JVI Accepts, published online ahead of print on 15 June 2011
J. Virol. doi:10.1128/JVI.00219-11

Phylogeography and evolutionary history of reassortant H9N2 viruses with potential human health implications

Alice Fusaro[1*], Isabella Monne[1], Annalisa Salviato[1], Viviana Valastro[1], Alessia Schivo[1], Nadim Mukhles Amarin[2], Carlos Gonzalez[3], Mahmoud Moussa Ismail[4,5], Abdu-Rahman Al-ankari[4], Mohamed Hamad Al-Blowi[6], Owais Ahmed Khan[7], Ali Safar Maken Ali[8], Afshin Hedayati[9], Juan Garcia Garcia[10], Ghulam M. Ziay[11], Abdolhamid Shoushtari[12], Kassem Nasser Al Qahtani[13], Ilaria Capua[1], Edward C. Holmes[14,15], Giovanni Cattoli[1]

[1] Istituto Zooprofilattico Sperimentale delle Venezie, Research & Development Department, OIE/FAO and National Reference Laboratory for Newcastle Disease and Avian Influenza, OIE Collaborating Centre for Diseases at the Human-Animal Interface. Viale dell'Università 10, 35020, Legnaro, Padova, Italy.

[2] Boehringer Ingelheim, regional technical manager, Dubai – UAE

[3] Boehringer Ingelheim, S.A de C.V., Mexico

[4] College of Veterinary Medicine and Animal Resources, King Faisal University, Al-Ahsa, Kingdom of Saudi Arabia

[5] College of Veterinary Medicine, department of Poultry Diseases, Kafrelsheikh University, Egypt

[6] Ministry of Agriculture, Riyadh, Kingdom of Saudi Arabia

[7] College of Veterinary Medicine, Texas A&M University, College Station Texas - 77843, USA

[8] Iran FAO Office Animal Health Consultant on AI & Other TADs, Tehran, Iran

[9] Iran Veterinary Organization (IVO), Tehran, Iran

[10] FAO-ECTAD-TCEO CTA for Afghanistan & Pakistan

[11] Central Laboratory Director CVDRL-Afghanistan

[12] Department of Avian Diseases, Razi Vaccine and Serum Research Institute, Karaj, Iran

[13] Department of animal resources, Ministry of Environment, Doha, Qatar

[14] Center for Infectious Disease Dynamics, Department of Biology, The Pennsylvania State University, Mueller Laboratory, University Park, Pennsylvania 16802

[15] Fogarty International Center, National Institutes of Health, Bethesda, Maryland

1 **Running title:** Molecular evolution of H9N2 Influenza A viruses

2 **Abstract word count:** 213

3 **Text word count:** 5614

4

5 *Corresponding author: Alice Fusaro. Mailing address: Istituto Zooprofilattico Sperimentale delle
6 Venezie, Research & Development Department, OIE/FAO and National Reference Laboratory for
7 Newcastle Disease and Avian Influenza, OIE Collaborating Centre for Diseases at the Human-
8 Animal Interface. Viale dell'Università 10, 35020, Legnaro, Padova, Italy. Phone: 0039 049
9 8084381. Fax.: 0039 049 8084360. E-mail: afusaro@izsvenezie.it

ABSTRACT

Avian influenza viruses of the H9N2 subtype have seriously affected the poultry industry of the Far and Middle East since the mid-1990's and are considered one of the most likely candidates to cause a new influenza pandemic in humans. To understand the genesis and the epidemiology of these viruses we investigated the spatial and evolutionary dynamics of complete genome sequences of H9N2 viruses circulating in nine Middle Eastern and Central Asian countries from 1998 to 2010. We identified four distinct and co-circulating groups (A, B, C and D), each of which has undergone widespread inter- and intra- subtype reassortment, leading to the generation of viruses with unknown biological properties. Our analysis also suggested that Eastern Asia served as the major source for H9N2 gene segments in the Middle East and Central Asia, and that in this geographic region within-country evolution played a more important role in shaping viral genetic diversity than between-country migration. The genetic variability identified among the H9N2 viruses was associated with specific amino acid substitutions that are believed to result in increased transmissibility in mammals as well as resistance to antiviral drugs. Our study highlights the need to constantly monitor the evolution of H9N2 viruses in poultry to better understand the potential risk for human health posed by these viruses.

INTRODUCTION

Avian influenza viruses of the H9N2 subtype are endemic in poultry populations across Asia and the Middle East. These viruses fall into a number of genetically defined lineages, with the majority of viruses circulating in Asia belonging to two lineages – G1 and Y280 – represented by the prototype viruses A/quail/Hong Kong/G1/97 and A/duck/Hong Kong/Y280/97, respectively, and which both became established in domestic poultry during the mid-1990s (16, 50, 33). Although H9N2 viruses are classified as low pathogenic avian influenza (LPAI), they sometimes cause significant disease in poultry, occasionally accompanied by high mortality and a marked reduction in egg production (4). To reduce the economic impact of the H9N2 infection in poultry, vaccination programs are commonly undertaken in several Asian countries (2, 4, 24).

Since 1997 H9N2 viruses have been reported in multiple avian species throughout Asia, the Middle East, Europe and Africa (1, 2, 6, 14, 15, 16, 17, 18, 23, 24, 29, 44, 50), and have occasionally been transmitted from poultry to mammalian species including humans and pigs (5, 25, 33). Further evidence of an expanded mammalian host range includes efficient replication of H9N2 in experimental mice without adaptation (9). Numerous recent H9N2 isolates contain an amino acid leucine at position 226 (L226) in their hemagglutinin (HA) receptor-binding site (RBS), which exhibits preferential binding to human-like α2-6 linked sialic acid (SAα2-6) receptors, and regarded as one of the key elements for the successful infection of humans (14, 18, 25, 45, 50). Recent research using ferrets demonstrated that L226-containing H9N2 viruses were more likely to transmit, although no aerosol transmission was observed (44). However, a chimeric H9N2 influenza virus carrying the surface glycoproteins of an avian H9N2 and the six internal genes of a human H3N2 virus was found to possess an increased transmissibility via respiratory droplets (42).

It is well established that reassortment between isolates from different host species can generate viruses with pandemic potential. The detection of avian influenza H9N2 viruses in domestic pigs and humans clearly creates opportunities for such reassortment events (5, 49). In addition, the extensive co-circulation of H9N2 viruses with other avian influenza viruses is likely to

generate appropriate conditions for the development of reassortant viruses. Indeed, inter-subtype reassortment has been detected between co-circulating H9N2 and highly pathogenic H5N1 or H7N3 viruses in China (16, 51) and Pakistan (6, 18). The circulation of H9N2 viruses throughout Eurasia, along with their ability to infect mammals and humans and the potential for future reassortment, clearly raises concern about their pandemic potential.

While the genetic and antigenic evolution of H9N2 viruses in China is well documented, information on the genetic properties of those H9N2 viruses circulating in Central Asia and the Middle East is limited to studies of few strains collected in a single country (1, 15, 17, 18, 27, 34, 34, 41). To explore the genetic characteristics, as well as the spatial and evolutionary dynamics of the H9N2 lineages that co-circulate in Central Asia and the Middle East, we conducted a phylogenetic analysis of whole genome sequences from H9N2 viruses sampled between 1998 to 2010 from nine Asian countries: Iran, Iraq, Israel, Jordan, the United Arab Emirates (UAE), Pakistan, Afghanistan, Saudi Arabia and Qatar. The central aim of our study was to reveal the extent of inter- and intra- subtypic reassortment, as well as the frequency and pattern of viral gene flow in order to identify those geographical regions that are likely to serve as key source populations for H9N2 viruses. Additionally, we examined the emergence of amino acid mutations and their significance in terms of drug-resistance, adaptation to different host species, and pandemic potential.

MATERIALS AND METHODS

Viruses included in this study

We sequenced the complete genomes of 29 avian influenza H9N2 viruses isolated from poultry in Afghanistan, Jordan, Saudi Arabia, Iraq-Kurdistan, Iran, UAE from 2004 to 2010, as well as the partial genomic sequences of five additional H9N2 viruses from Iraq (HA and NA), Qatar (PB2, HA, NP, NA, M and NS), Jordan (PB2, PB1, HA, NP, NA, M and NS), UAE (PB2, PB1, PA, HA,

NA, M and NS), and Saudi Arabia (PA, HA, NP, NA, M and NS). These sequences were analyzed together with all publicly available sequences data of H9N2 viruses isolated between 1998 and 2008 from Central Asian and the Middle Eastern countries, including Israel, Iran, Saudi Arabia, UAE, Pakistan and Jordan. Overall, a total of 107, 109, 111, 178, 109, 138, 103 and 119 full-length H9N2 sequences were analyzed for the PB2, PB1, PA, HA, NP, NA, M, and NS segments, respectively.

Nucleotide sequencing

For the isolates sequenced in this study, viral RNA was extracted from the infective allantoic fluid of SPF fowls' eggs using the Nucleospin RNA II Kit (Machery-Nagel, Duren, Germany) and was reverse transcribed with the SuperScript III Reverse Transcriptase kit (Invitrogen, Carlsbad, CA - USA). PCR amplifications were performed by using specific primers (primer sequences available on request). The complete coding sequences were generated using the Big Dye Terminator v3.1 cycle sequencing kit (Applied Biosystem, Foster City, CA - USA). The products of the sequencing reactions were cleaned-up using PERFORMA DTR Ultra 96-Well kit (Edge BioSystems, Gaithersburg, MD - USA) and sequenced in a 16-capillary ABI PRISM 3130xl Genetic Analyzer (Applied Biosystem, Foster City, CA - USA). Sequence data were assembled and edited with SeqScape software v2.5 (Applied Biosystem). Sequences from all eight gene segments were aligned and compared with all publicly available H9N2 sequences of viruses from Central Asian and Middle Eastern countries, and with Eurasian H9N2 sequences representative of distinct H9N2 lineages. Sequences of the six internal genes were compared also with representative sequences of viruses of different subtypes available in GenBank (table 1).

Nucleotide sequence accession numbers

The nucleotide sequences obtained in this study are available from GISAID under accession numbers EPI301452 to EPI301655 and EPI223068 to EPI223123.

Phylogenetic analysis

For each of the eight genome segments, maximum likelihood (ML) trees were estimated using the best-fit general time-reversible (GTR)+I+Γ_4 model of nucleotide substitution using PAUP* (47). Parameter values for the GTR substitution matrix, base composition, gamma distribution of among-site rate variation (with four rate categories, Γ_4), and proportion of invariant sites (I) were estimated directly from the data using MODELTEST (37). A bootstrap resampling process (1000 replications) using the neighbor-joining (NJ) method was used to assess the robustness of individual nodes on the phylogeny, incorporating the ML substitution model defined above.

Substitution rates and times to common ancestry

Rates of nucleotide substitution per site, per year and the time to the Most Recent Common Ancestor (tMRCA) of the sampled data were estimated using the BEAST program version 1.5.3 (11) which employs a Bayesian Markov chain Monte Carlo (MCMC) approach. For each analysis the Bayesian skyline coalescent tree prior was used, as this is clearly the best descriptor of the complex population dynamics of influenza A virus (12). Two molecular clock models – strict (constant) and uncorrelated lognormal (UCLN) relaxed clock – were compared by analyzing values of the coefficient of variation (CoV) in Tracer (11), such that CoV values >0 are evidence of non-clock-like evolutionary behaviour. We also compared two nucleotide substitution models: a GTR + Γ_4 model similar to that described above and a codon-based SRD06 model. In all cases the SRD06 model performed better than the GTR + Γ_4 model, as previously demonstrated for a large set of RNA viruses (39). In all cases, uncertainty in the data is reflected in the 95% highest probability density (HPD) values for each parameter estimate, and in each case chain lengths were run for sufficient time to achieve coverage as assessed using the Tracer v1.5 program (11). Finally, Maximum Clade Credibility (MCC) phylogenetic trees were estimated from the posterior distribution of trees generated by BEAST using the program TreeAnnotator v1.5.3 (11) after the

removal of an appropriate burn-in (10% of the samples). The MCC trees were visualized using the program FigTree v1.3.1 (http://tree.bio.ed.ac.uk/software/figtree/), which allowed us to estimate the tMRCA of each individual node on the trees.

Phylogeography of H9N2 in Central Asia and the Middle East

To determine the extent and pattern of geographical structure in the H9N2 viruses sampled here we first grouped the sequences of viruses isolated in Central Asia and the Middle East into eight geographic regions, representing one or two adjacent countries. These geographic regions were as follows: Saudi Arabia/Qatar (the latter represented by a single sequence), Jordan, Israel, Iran, Iraq, Afghanistan, Pakistan, United Arab Emirates (UAE), Eastern Asia and Europe. To assess the overall degree of geographical structure among H9N2 viruses sampled from the Central Asian and Middle Eastern regions analyzed here, we used the BaTS program (32) to estimate values of the association index (AI) and parsimony score (PS) statistics of phylogeny-trait association, with the traits (the geographical regions) defined above. Importantly, this method uses the posterior distribution of trees obtained from the BEAST analysis described above, and thereby accounts for phylogenetic uncertainty in the data. The BaTS program also allowed us to assess the level of clustering in individual locations using the monophyletic clade (MC) size statistic (32). In all cases, 1000 random permutations of tip locations were undertaken to create a null distribution for each statistic

Analysis of selection pressures

Gene and site-specific selection pressures for all segments of Central Asian and the Middle Eastern H9N2 viruses were measured as the ratio of nonsynonymous (d_N) to synonymous (d_S) nucleotide substitutions per site. In all cases d_N/d_S ratios were estimated using the Single Likelihood Ancestor Counting (SLAC) and Fixed Effects Likelihood (FEL) maximum likelihood methods available at

the Datamonkey online version of the Hy-Phy package (21). All analyses utilized the GTR model of nucleotide substitution and employed neighbor-joining phylogenetic trees.

RESULTS

Phylogenetic relationships among H9N2 viruses in Central Asia and the Middle East from 1998 to 2010: multiple introductions and extensive reassortment events

To investigate the genetic diversity of avian influenza H9N2 viruses in Central Asia and the Middle East we analysed the eight genome segments of all the H9N2 viruses from this geographic region that are available on GenBank (the number of available sequences ranged from 71 for the M gene to 145 for the HA gene) combined with 34 genome sequences newly generated in this study. Phylogenetic relationships were determined between these viruses and representative sequences from European and Asian countries, along with the G1 and Y280 established Eurasian H9N2 lineages. The topologies of all the eight maximum likelihood phylogenetic trees revealed substantial genetic diversity, exemplified by multiple co-circulating H9N2 lineages and frequent reassortment events, which could be schematically identified as either inter-subtype (between different HA subtypes) or intra-subtype (between different H9 genetic clusters) exchanges.

Analysis of the HA phylogeny identifies four monophyletic groups of H9N2 viruses in Central Asia and the Middle East – denoted A, B, C and D (table 2) – each of which is defined by high bootstrap (>70%) values and average percentage pairwise nucleotide distances within and between groups of <5% and >5%, respectively (figure 1). The phylogenetic groups identified in the HA tree (A to D) are maintained in most of the remaining gene segments (figure 2 and figures S1 to S6 in the supplemental material). Importantly, the topologies of the eight segment phylogenies show that each group (A to D) is characterized by a unique genomic constellation, indicative of the extensive mixing of gene segments from multiple avian influenza subtypes.

Of particular note, our phylogenetic analysis suggests that the HA, NA and M genome segments of the four genetic groups (A to D), as well as the NP gene of group B, derive from the G1-lineage, while the remaining gene segments have been replaced through inter-subtype reassortment and do not show any relationship with the previously identified G1 and Y280 Eurasian H9N2 lineages, generating many different genotypes (figures 1 and 2 and figures S1 to S6 in the supplemental material). As recently described by Iqbal et al. (2009), some H9N2 strains collected in Pakistan from 2005 to 2006 have acquired the NS gene segment from HPAI H5N1 (clade 2.2) and Pakistani HPAI H7N3 viruses (18). However, in most cases the origin of the H9N2 segments cannot be established.

Viruses possessing the gene segment constellation A or B appear to have circulated extensively in Central Asia and the Middle East (table 2), while groups C and D were only detected in the UAE between 2000 and 2002, and in Iran from 1998 to 2007, respectively. Group A contains viruses isolated in Israel (Israel-1 and Israel-2 group in the phylogenetic trees and in table 2) and Jordan (Jordan-1) between 2000 and 2007, and A/quail/Saudi Arabia/08vir3489-46-as104/2006. Of note, A/ck/Saudi Arabia/532/1999 contains the same gene composition, and could be considered the progenitor of these viruses as it is located near the root of cluster A in all the phylogenetic trees (figures 1 and 2 and figures S1 to S6 in the supplemental material), except the HA, where it does not belong to any identified cluster.

Viruses from Pakistan (2006-2008), Afghanistan (2008-2009), A/ck/Iran/10vir854-5/2008 and A/ck/Iran/10vir854-3/2009 were identified as group B. Our phylogenetic analysis suggests that this group arose after acquiring the HA, NA, M and NP gene segments from viruses isolated in Pakistan in 1999 (entirely belonging to the G1 lineage), as Pakistani strains tended to cluster near the root of the cluster B in the four phylogenies. In contrast, the remaining genes have been obtained by inter-subtype reassortment events.

Cluster D consists only of viruses from Iran (Iran-1) that, based on the HA phylogeny, circulated from 1998 to 2007. Due to the lack of sequence data for the internal genes of most of the

Iranian viruses previously deposited in GenBank, this group is only represented by A/ck/Iran/11T/99 and A/ck/Iran/10vir854-1/02 in the PB2, PA and NP phylogenies. Interestingly, the NP gene appears to cluster with A/duck/Hong Kong/Y439/97, suggesting that the Y439 H9N2 lineage circulating in Asia may be the source of this gene.

The occurrence of inter-subtype reassortment is evident also for three H9N2 viruses from Iraq-Kurdistan and Iran. Interestingly, four gene segments (HA, NA, M and NS) of two 2009 Kurdistani strains and five gene segments (HA, NA, M, NP and NS) of A/chicken/Iran/10vir854-4/2009 fall within the genotype B, while the remaining genes do not belong to any H9N2 groups (A to D) and show low sequence similarity with H9N2 Asian viruses, indicative of inter-subtype gene mixing (table 2).

Besides these inter-subtype reassortment events, extensive intra-subtype reassortments among H9N2 viruses belonging to the four genetic groups (A to D) seem to have played an important role in the genesis of viruses with unique genome constellations (table 2). In particular, reassortant strains containing a mix of genes from groups A and B have been isolated in Israel (Israel-3) from 2006 to 2009, in Jordan (Jordan-2, 2010) and in Saudi Arabia in 2007 (table 2). Interestingly, these viruses show three different type of reassortments: Israel-3 and Jordan-2 contain four genes from group A and four from group B, two Saudi Arabian isolates possess five genes of group A and three of group B, and one virus from Saudi Arabia has acquired another gene from cluster A.

Distinct reassortments between groups B and C are identified in the UAE in 2003 and 2008, in Qatar in 2008, and in Saudi Arabia in 2005 and 2006 (table 2). In addition, two different triple reassortant viruses, containing gene segments belonging to the clusters A, B and C, have been detected in Saudi Arabia from 2005 to 2007. The 2007 isolate was probably generated by further reassortment events, involving the PB2 and the NA genes, of the Saudi Arabian strains circulating in 2005-2006.

More generally, these data suggest that a unique reassortant virus became predominant in Israel and Jordan and spread in these two countries between 2006 and 2010. In the other countries, reassortant viruses circulated for shorter periods and underwent further reassortment events, generating strains with distinct gene compositions. For example, in Saudi Arabia six distinct reassortant viruses were detected: two in 2005, two in 2006 and two in 2007 (table 2).

Rates of nucleotide substitution and time-scale of viral evolution in Central Asia and the Middle East

Rates of nucleotide substitution and tMRCAs were estimated for each segment of the entire population of H9N2 viruses from Central Asia and the Middle East. For all genes, the lower 95% HPDs of CoV values of the relaxed (uncorrelated lognormal) molecular clock were >0, so a relaxed molecular clock model, which allows for the rate variation across lineages, was used in preference to a strict molecular clock.

The mean substitution rates for all segments were found to be within the range typically observed for avian influenza viruses (8). The highest rate of evolution was observed in the NA segment (4.26×10^{-3} substitutions/site/year; 95% HPD, 3.70×10^{-3} to 4.80×10^{-3}), while the lowest rate was observed in the M segment (2.22×10^{-3} substitutions/site/year; 95% HPD, 1.78×10^{-3} to 2.69×10^{-3}) (table 3).

To establish when H9N2 virus genes were introduced in Central Asia and the Middle East, we estimated the tMRCA for the entire population of H9N2 viruses circulating in this area (the time-scaled MCC trees for each gene segment, showing nodal divergence times, are provided in figure S7 in the Supplementary Information). As expected given the different origins of the internal gene segments, the tMRCAs calculated for the PB2, PB1, PA, NP and NS genes were relatively distant in time (mean tMRCAs ranged from 1968 to 1981, table 3), while the mean tMRCAs estimated for the segments belonging to the G1 lineage (HA, NA and M) ranged from 1989 to 1994. Notably, the tMRCAs estimated for the HA and NA genes were very similar – 1993 (1990-1996,

95% HPD) and 1994 (1991-1996, 95% HPD), respectively – and approximately five and four years before the detection of H9N2 viruses in Central Asia and the Middle East (table 3).

Dating the time of emergence of each gene segment of lineage A and B showed different tMRCAs, supporting the results of the phylogenetic analysis depicting multiple reassortment events among gene segments derived from different sources. In particular, the mean tMRCAs estimated for the genes of lineage A ranged from 1987 to 1997, while the tMRCAs obtained for the genes of lineage B ranged from 1991 to 2001, suggesting that this lineage may have emerged in 2001 when H9N2 viruses were already circulating in Central Asia and the Middle East.

Phylogeography of H9N2 in Central Asia and the Middle East

Our phylogenetic analysis also revealed that most of the H9N2 genes circulating in Central Asia and the Middle East clustered with viruses from Eastern Asia, suggesting that birds in this area may be an important source of H9N2 genetic diversity. In addition, in all phylogenetic trees the Jordanian and the Afghani H9N2 viruses showed a close relationship to the H9N2 sequences from Israel and Pakistan, respectively, and which may reflect a considerable trade in commercial poultry between these countries.

Although our phylogenetic analysis revealed some mixing of H9N2 sequences among localities, most notably Saudi Arabia, which is indicative of at least some widespread of viral gene flow, most of the viruses sampled from individual countries tended to cluster together, which is strongly suggestive of major geographic subdivision among H9N2 viruses. To examine the extent and pattern of geographical structure in these data in a more quantitatively rigorous manner, we used a series of phylogeny-trait association tests, in which each H9N2 virus was assigned to a different geographic region. Accordingly, our Bayesian MCMC analysis of geographical association revealed a very strong geographic clustering of H9N2 strains by country of origin (P=0 for both AI and PS statistics in all gene segments). Similarly, when the extent of phylogenetic clustering of individual countries was tested, significant population subdivision was observed in all

countries. Such localized clustering was especially strong in Jordan and Afghanistan, in which the MC statistic was strongly significant in all gene segments (P<0.0009; table S1 in the supplemental material), while in Pakistan and Iran lower significance values were observed for four and three genes, respectively, (*P*>0.001 table S1 in the supplemental material) indicative of some gene flow involving birds from these localities. Notably, the MC statistic was significant for every country analyzed, indicating that H9N2 genetic diversity is shaped primarily by within-country evolution rather then extensive between-country migration.

Diversifying selection in the H9N2 genes and identification of potential additional glycosylation sites

Our analysis of selection pressures revealed that most codons in all genes were subject to purifying selection (mean d_N/d_S ratios varied between 0.07 and 0.61; table 4). Despite this, we identified several individual codons in the HA, NA, PB2, PA, NS1 and M1 genes that may be subject to positive selection (table 4). Specifically, three positively selected residues in the HA gene at positions 168, 198 and 234 (160, 190, 226 in H3 numbering) are located at the receptor binding site or on the tip of the HA (26). Substitutions in these positions may affect virus interactions with cell surface receptors. Furthermore, site 234 (226 in H3 numbering) falls within the epitope II of H9 and plays a role in the receptor binding specificity of HA (20). In addition, additional potential glycosylation sites (AGS) are detected at positions 145 (137 in H3 numbering) in two viruses from Israel (2007-2008), at position 206 (198 in H3 numbering) in the HA protein of A/quail/Saudi Arabia/08vir3489-46-as104/06 and A/chicken/Iran/SS1/1998, and at position 218 (210 in H3 numbering) in four viruses from the UAE (2000-2003), three from Pakistan (1999) and three from Iran (1998-2007). Interestingly, sites 137 and 198 have previously been demonstrated to be antigenically relevant amino acid positions in H9 hemagglutinin (20, 36) and glycosylation in these two sites could play a role in antigenic variation. Similarly, in the NA gene we identified seven

residues under positive selection pressure, six of which were located in the NA globular head (26). Mutations at these positions may influence the antigenic specificity of these viruses.

Molecular characterization

Amino acid mutations with possible implications for pandemic emergence

Recent studies have demonstrated that the amino acid leucine (L) at position 226 in the HA receptor binding site (RBS) plays a key role in human virus-like receptor specificity, and promotes transmission of H9N2 in ferrets (46). Sorrel et al. (2009) showed that the combination of four key amino acid residues at the RBS of the HA molecule (H183, A189, E190 and L226) is essential for respiratory droplets transmission of a reassortant virus carrying the surface proteins of an avian H9N2 in a human H3N2 backbone (42). Our analysis of H9N2 from Central Asia and the Middle East reveals that 23/42 (54.8%) viruses of lineage A, 91/97 (93.8%) viruses of lineage B, 6/7 (85.7%) viruses of lineage C and 16/29 (55.1%) viruses of lineage D contain the amino acid leucine at position 226 (H3 numbering) at the RBS. Moreover, seven viruses show three of the four key amino acid residues at the RBS. In particular, A/ostrich/Eshkol/619/02, A/turkey/Israel/619/02, A/chicken/Iran/SS1/98, A/chicken/Iran/TH77/98, A/quail/Dubai/301/00 and A/quail/Dubai/302/00 viruses possess H183, E190 and L226, while A/quail/Saudi Arabia/08vir3489-46-as104/06 contains H183, A189, and E190. Interestingly, an alanine at position 189 is rarely observed among H9N2 viruses (only 6 of 847 Eurasian H9 sequences available in GenBank show A189), and the Saudi Arabian isolate is the only virus from the G1 lineage that carries this substitution. In addition, position 189, located on the tip of the globular head of HA1, is considered a critical antigenic site and has been implicated in H9 escape mutants (42).

To understand the pandemic risk posed by the H9N2 viruses we also investigated the species-associated signature positions thought to be characteristic of human influenza A viruses. Residues located at the avian-human signature positions previously described by Chen et al. (2006), Finkelstein et al. (2007), and Pan et al. (2010) and identified in the H9N2 genes are listed in table

S2 (7, 13, 31). Although the predominant amino acid found is consistent with avian influenza viruses at most marker locations, in a small proportion of H9N2 sequences the amino acid prevalent in human-hosted viruses has been acquired. Accordingly, we found 21 host markers in the PB2, PB1-F2, PA, NP, M2 and NS2 proteins (table S2). However, no single H9N2 isolate contains more than four of these 21 sites.

Amino acid mutations associated with virulence

Our analysis of the NS1 protein revealed the existence of two mutations, at positions 92 and 227, which have previously been demonstrated to modulate pathogenicity of avian influenza. Mutation D92E is observed in three viruses isolated in Pakistan in 1999, while E227K is found in 47 of 119 H9N2 viruses analysed here. The glutamic acid residue at position 92 of the NS gene appears to be implicated in influenza virus resistance to host antiviral activity exerted by interferons and TNF-α (38).

Mutation E227R/K is a persistent host marker (table S2) located in the extreme C- terminal region of the NS1 (7, 13, 31), and has been shown to modulate pathogenicity of avian influenza through mechanisms not yet completely clarified (19, 30). In particular, the four C- terminal residues of the NS1 gene correspond to a PDZ ligand domain (PL), which is a protein–protein recognition module that organizes diverse cell-signalling assemblies (40). Viruses containing the C-terminal four residue PL sequence from the 1918 H1N1 (KSEV) and from H5N1 highly pathogenic avian influenza (HPAI) viruses (ESEV or EPEV) showed a significantly increased virulence and pathogenicity in infected mice (19). Interestingly, 82 of the 119 H9N2 viruses analysed in this study contain one of these motifs. In addition, all H9N2 viruses belonging to group A contained the substitution N66S in the PB1-F2 protein. This mutation, also found in the Hong Kong 1997 H5N1 viruses and in the 1918 H1N1 pandemic strain, has been shown to be associated with high pathogenic phenotypes in mice (10).

Amino acids substitutions associated with resistance to antiviral drugs

Sequence analysis of the M2 ion channel protein revealed that 22 of the H9N2 viruses analyzed here contained amino acid changes at two positions previously associated with adamantane resistance (43). Specifically, twelve isolates (one from Iran, six from UAE and five from Israel) had the substitution V27A, while ten viruses (six from Iran, three from the UAE and one from Qatar) possessed the mutation S31N, as already described by Aamir et al. (2006) in seven isolates from UAE (1).

DISCUSSION

The H9N2 influenza viruses currently endemic in Asian poultry populations are considered one of the most likely candidates for a new pandemic influenza in humans (3). The detection of increasingly large numbers of H9N2 strains showing human-like receptor specificity, combined with the growing evidence for reassortment involving this subtype, emphasizes its potential to emerge and establish itself in the human population.

This study represents the largest evolutionary analysis of H9N2 viruses undertaken to date. We show that H9N2 influenza viruses isolated from 1998 to 2010 in Central Asia and the Middle East comprise four distinct and co-circulating groups (A, B, C and D), each of which has undergone widespread inter- and intra- subtype reassortment. Unlike groups C and D that were detected only in one country, groups A and B have circulated extensively in Central Asia and the Middle East and have been identified in six Asian countries from 1999 to the present day. Although it is difficult to determine the precise origin of the H9N2 lineages, our phylogenetic analysis revealed that the HA, NA and M genes of H9N2 viruses from Central Asia and the Middle East share the same progenitor, as they are all included within the G1-lineage. This indicates that these gene segments have ever undergone inter-subtype reassortment events since 1998, when the first isolates of this lineage were detected in this region. Of note, the G1-lineage was responsible for human infection

cases in Hong Kong in 1999 (25) and at the end of 2009 (48), suggesting that this variant is still potentially infectious for humans.

Despite the frequent reassortment events and the identification of multiple co-circulating viruses in each country, we also observed a statistically significant geographical subdivision of viral strains. Hence, after entering a specific country, H9N2 viruses circulate among poultry populations in that locality, giving rise to well defined genetic groups. This was particularly apparent for most of the viruses isolated from Israel, Jordan, UAE and Afghanistan. In contrast, Saudi Arabian H9N2 sequences are dispersed throughout the phylogenetic trees and are generated by multiple reassortment events between lineage A, B and C. The Arabian peninsula is one of the most import markets for poultry in the world, and during the last ten years has seen a rapid increase of imports of live poultry (28). Birds from this locality may act as an ecological sink, receiving viruses from different areas, but also as a regional source population. Indeed, our results show that the precursor of group A, which includes viruses from Israel, Jordan and Iraq, is a Saudi Arabian strain (A/chicken/Saudi Arabia/532/99).

Our phylogeny-based analysis of geographical association identifies avian populations from Pakistan and Iran as two possible sources for the H9N2 virus in Central Asia and the Middle East, although across the data set as a whole by far the strongest signal was for population subdivision. This is consistent with our phylogenetic analysis that identifies Pakistani viruses as the main source of genes for group B, while Iranian strains appear to be a source of genes (PB1 and M) for UAE and Saudi Arabian viruses. However, all inferences regarding the geographical source of H9N2 will clearly need to be confirmed with a larger and less biased sample of viruses. Indeed, the long branches in all the phylogenetic trees that divide viruses collected from different countries confirm that surveillance and sampling of Central Asian and the Middle Eastern is currently inadequate.

The trade in poultry and poultry products seems to play the major role in the spread of H9N2 viruses among Central Asian and Middle Eastern countries, as all H9N2 isolates to date have been collected from domestic birds, mainly chickens. Unfortunately, our data does not shed light on the

means by which the virus was introduced into Central Asia and the Middle East, as neither of the two main means of spread – through wild birds or trade of poultry – can be definitely excluded on these data. However, our phylogenetic analysis does suggest that Eastern Asia plays the main role in the introduction of H9N2 viruses into this geographic region. To date, the mixing between HPAI H5N1 and H9N2 has only been detected in China (16) and Pakistan (18). However the extensive co-circulation of these two subtypes in some countries could lead to further reassortment events, generating viruses with a range of phenotypic properties.

Also of note was our identification of amino acid residues under positive selection in the antigenic sites of the HA molecule, such that some H9N2 viruses could differ substantially in antigenicity. This finding is compatible with the idea that the mass vaccination of poultry adopted in many of these countries is a major factor driving the evolution of the H9N2 viruses. Among the amino acid substitutions in H9 HA, 76.4% of the viruses analyzed contain the amino acid leucine at position 226 that is responsible for human virus-like receptor specificity and critical for replication and direct transmission of H9N2 viruses in ferrets (46). Furthermore, seven viruses contain three of the four residues at the RBS of the HA molecule that are considered critical for human-to-human transmission (42). In particular, A/quail/Saudi Arabia/08vir3489-46-as104/2006 contains three of these point mutations essential for respiratory droplet transmission and, more importantly, it possesses the change T189A that has been shown to dramatically alter the antigenicity of viruses (42).

The data generated in this study provide the most comprehensive insight into the epidemiology and evolution of H9N2 virus in Central Asia and the Middle East. Despite these efforts, there remains a lack of information concerning H9N2 in this very large geographic region. Additional study and timely surveillance of H9N2 is clearly needed to identify any increments in viral adaptation to humans, and to constantly monitor the mixing of avian H9N2 with other influenza virus subtypes, including those circulating in humans, that may favor the emergence of influenza viruses with pandemic potential.

ACKNOWLEDGEMENTS

The authors wish to thank Boehringer Ingelheim Vetmedica for its assistance in facilitating exchange of information, data sharing, sample submissions and financial support. Part of this work was financially supported by the Italian Ministry of Health through the RC IZS VE 14/09, the Food & Agricultural Organization of the United Nations (LoA PR n. 41097) and has been conducted in the framework of the EU project FLUTRAIN. Edward C. Holmes was funded in part by NIH grant GM080533-04.

REFERENCES

1. **Aamir, U. B., U. Wernery, N. Ilyushina, and R. G. Webster.** 2006. Characterization of avian H9N2 influenza viruses from United Arab Emirates 2000 to 2003. Virology. **361:**45-55.

2. **Alexander, D.J.** 2003. Report on avian influenza in the Eastern Hemisphere during 1997-2002. Avian Dis. 47(3 Suppl):792-7.

3. **Alexander, P. E., P. De, S. Rave.** 2009. Is H9N2 avian influenza virus a pandemic potential? Can J Infect Dis Med Microbiol.20:e35-6.

4. **Brown, I. H., J. Banks, R. J. Manvell, S. C. Essen, W. Shell, M. Slomka, B. Londt, and D. J. Alexander**. 2006. Recent epidemiology and ecology of influenza A viruses in avian species in Europe and the Middle East. Dev Biol (Basel). **124:**45-50.

5. **Butt, K. M., G. J. Smith, H. Chen, L. J. Zhang, Y. H. Leung, K. M. Xu, W. Lim, R. G. Webster, K. Y. Yuen, J. S. Peiris, and Y. Guan**. 2005. Human infection with an avian H9N2 influenza A virus in Hong Kong in 2003. J Clin Microbiol. **43:**5760-7.

6. **Cameron, K. R., V. Gregory, J. Banks, I. H. Brown, D. J. Alexander, A. J. Hay, and Y. P. Lin.** 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.

7. **Chen, G. W., S. C. Chang, C. K. Mok, Y. L. Lo, Y. N. Kung, J. H. Huang, Y. H. Shih, J.**

Y. Wang, C. Chiang, C. J. Chen, and S. R. Shih. 2006. Genomic signatures of human versus avian influenza A viruses. Emerg Infect Dis. **12:**1353-60.

8. **Chen, R., and E. C. Holmes.** 2006. Avian influenza virus exhibits rapid evolutionary dynamics. Mol. Biol. Evol. **7:**214

9. **Choi, Y. K., H. Ozaki, R. J. Webby, R. G. Webster, J. S. Peiris, L. Poon, C. Butt, Y. H. Leung, and Y Guan**. 2004. Continuing evolution of H9N2 influenza viruses in Southeastern China. J Virol. **78:**8609-14.

10. **Conenello, G. M., D. Zamarin, L. A. Perrone, T. Tumpey, P. Palese.** 2007. A single mutation in the PB1-F2 of H5N1 (HK/97) and 1918 influenza A viruses contributes to increased virulence. PLoS Pathog. **3:**1414-21

11. **Drummond A. J., Rambaut A.** 2007. BEAST: Bayesian evolutionary analysis by sampling trees.BMC Evolutionary Biology **7:**214

12. **Drummond, A. J., A. Rambaut, B. Shapiro, and O. G. Pybus.** 2005. Bayesian coalescent inference of past population dynamics from molecular sequences. Mol. Biol. Evol. **22:**1185–1192.

13. **Finkelstein, D.B., S. Mukatira, P. K. Mehta, J. C. Obenauer, X. Su, R. G. Webster, and C. W. Naeve.** 2007. Persistent host markers in pandemic and H5N1 influenza viruses. J Virol. **81:**10292-9

14. **Ge, F.F., J.P. Zhou, J. Liu, J. Wang, W.Y. Zhang, L.P. Sheng, F. Xu, H. B. Ju, Q. Y.**

Sun, and P. H. Liu. 2009. Genetic evolution of H9 subtype influenza viruses from live poultry markets in Shanghai, China. J Clin Microbiol. **47:**3294-300.

15. **Golender, N., A. Panshin, C. Banet-Noach, S. Nagar, S. Pokamunski, M. Pirak, Y. Tendler, I. Davidson, M. García, and S. Perk**. 2008. Genetic characterization of avian influenza viruses isolated in Israel during 2000-2006. Virus Genes. **37:**289-97.

16. **Guan, Y., K. F. Shortridge, S. Krauss, P. S. Chin, K. C. Dyrting, T. M. Ellis, R. G. Webster, and M. Peiris**. 2000. H9N2 influenza viruses possessing H5N1-like internal genomes continue to circulate in poultry in southeastern China. J Virol. **74:**9372-80.

17. **Homayounimehr, A. R , H. Dadras, A. Shoushtari, and S. A. Pourbakhsh**. 2010. Sequence and phylogenetic analysis of the haemagglutinin genes of H9N2 avian influenza viruses isolated from commercial chickens in Iran. Trop Anim Health Prod. **6:**1291-1297, DOI: 10.1007/s11250-010-9565-7

18. **Iqbal, M., T. Yaqub, K. Reddy, and J. W. McCauley.** 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(6): e5788. doi:10.1371/journal.pone.0005788

19. **Jackson, D., M. J.Hossain, D. Hickman, D. R. Perez, and R. A. Lamb.** 2008. A new influenza virus virulence determinant: the NS1 protein four C-terminal residues modulate pathogenicity. Proc Natl Acad Sci U S A. **105:**4381–4386.

20. **Kaverin, N. V., I. A. Rudneva, N. A. Ilyushina, A. S. Lipatov, S. Krauss, and R. G.**

Webster. Structural differences among hemagglutinins of influenza A virus subtypes are reflected in their antigenic architecture: analysis of H9 escape mutants. J Virol. **78:**240-9.

21. **Kosakovsky Pond, S. L., and S. D. W Frost**. 2005. Datamonkey: Rapid detection of selective pressure on individual sites of codon alignments. Bioinformatics **21:** 2531–253

22. **Kosakovsky Pond, S. L., and D. W. S. Frost** 2005. Not So Different After All: A Comparison of Methods for Detecting Amino Acid Sites Under Selection. Molecular Biology and Evolution. **22:**1208-1222

23. **Lee, Y. J., J. Y. Shin, M. S. Song, Y. M. Lee, J. G. Choi, E. K. Lee, O. M. Jeong, H. W. Sung, J. H. Kim, Y. K. Kwon, J. H. Kwon, C. J. Kim, R. J. Webby, R. G. Webster, and Y. K. Choi**. 2007. Continuing evolution of H9 influenza viruses in Korean poultry. Virology. **359:**313-23.

24. **Li, C, K. Yu, G. Tian, D. Yu, L. Liu, B. Jing, J. Ping, and Chen H.** 2005. Evolution of H9N2 influenza viruses from domestic poultry in Mainland China. Virology. **340:**70-83.

25. **Lin, Y. P., M. Shaw, V. Gregory, K. Cameron, W. Lim, A. Klimov, K. Subbarao, Y. Guan, S. Krauss, K. Shortridge, R. Webster, N. Cox, and A. Hay.** 2000. Avian-to-human transmission of H9N2 subtype influenza A viruses: relationship between H9N2 and H5N1 human isolates. Proc Natl Acad Sci U S A. **97:**9654-8.

26. **Matrosovich, M. N., S. Krauss, and R. G. Webster.** 2001. H9N2 influenza A viruses from poultry in Asia have human virus-like receptor specificity. Virology. **281:**156-62.

27. **Monne, I., G. Cattoli, E. Mazzacan, N. M. Amarin, H. M. Al Maaitah, M. Q. Al-Natour, and I. Capua.** 2007. Genetic comparison of H9N2 AI viruses isolated in Jordan in 2003. Avian Dis51(1 Suppl):451-4.

28. **Nicita, A.** Avian influenza and the poultry trade. Policy research working paper. Available at: http://siteresources.worldbank.org/INTTOPAVIFLU/Resources/AI_PoultryTrade_WorkingPaper.pdf

29. **Nili, H., and K. Asasi.** 2003. Avian influenza (H9N2) outbreak in Iran. Avian Dis. 47(3 Suppl):828-31.

30. **Obenauer, J. C., J. Denson, P. K. Mehta, X. Su, S. Mukatira, D. B. Finkelstein, X. Xu, J. Wang, J. Ma, Y. Fan, K. M. Rakestraw, R. G. Webster, E. Hoffmann, S. Krauss, J. Zheng, Z. Zhang, and C. W. Naeve.** 2006. Large-scale sequence analysis of avian influenza isolates. Science. **311:**1576-80.

31. **Pan, C., B. Cheung, S. Tan, C. Li, L. Li, S. Liu, and S. Jiang**. 2010. Genomic signature and mutation trend analysis of pandemic (H1N1) 2009 influenza A virus. PLoS One. 8;5(3):e9549. doi: 10.1371/journal.pone.0009549.

32. **Parker, J., A. Rambaut, and O. G. Pybus.** 2008. Correlating viral phenotypes with phylogeny: accounting for phylogenetic uncertainty. Infect. Genet. Evol. **8:**239–246.

33. **Peiris, J. S., Y. Guan, D. Markwell, P. Ghose, R. G. Webster, and K. F. Shortridge**. 2001. Cocirculation of avian H9N2 and contemporary "human" H3N2 influenza A viruses

in pigs in southeastern China: potential for genetic reassortment? J Virol. **70:**9679-86.

34. **Perk,, S., C. Banet-Noach, E. Shihmanter, S. Pokamunski, M. Pirak, M. Lipkind, and A. Panshin.** 2006. Genetic characterization of the H9N2 influenza viruses circulated in the poultry population in Israel. Comp Immunol Microbiol Infect Dis. **29:**207-23.

35. **Perk, S, N. Golender, C. Banet-Noach, E. Shihmanter, S. Pokamunsky, M. Pirak, Y. Tendler, M. Lipkind, and A. Panshin**. 2009. Phylogenetic analysis of hemagglutinin, neuraminidase, and nucleoprotein genes of H9N2 avian influenza viruses isolated in Israel during the 2000-2005 epizootic. Comp Immunol Microbiol Infect Dis. **32:**221-38

36. **Ping, J, C Li, G Deng, Y Jiang, G Tian, S Zhang, Z Bu, and H Chen**. 2008. Single-amino-acid mutation in the HA alters the recognition of H9N2 influenza virus by a monoclonal antibody. Biochem Biophys Res Commun. **371:**168-71.

37. **Posada, D. and K. A. Crandall.** 1998. MODELTEST: testing the model of DNA substitution. Bioinformatics. **14:**817-818

38. **Seo, S. H., E. Hoffmann, R. G. Webster**. 2002. Lethal H5N1 influenza viruses escape host anti-viral cytokine responses. Nat Med. **8:**950-4

39. **Shapiro, B., A. Rambaut, A. J. Drummond**. 2006. Choosing appropriate substitution models for the phylogenetic analysis of protein-coding sequences. Mol Biol Evol. **23:**7-9

40. **Sheng, M., and C. Sala.** 2001. PDZ domains and the organization of supramolecular complexes. Annu Rev Neurosci **24:**1–29.

41. **Soltaniavar, M, H. Shoushtari, M. Bozorgmehrifard, S. Charkhkar, and Eshratabadi.** 2010. Molecular characterization of Hemagglutinin and Neuraminidase genes of H9N2 avian influenza viruses isolated from commercial broiler chicken in Iran. J. Biol. Sci. **10:**145-150

42. **Sorrell, E. M., H. Wan, Y. Araya, H. Song, D. R. Perez.** 2008. Minimal molecular constraints for respiratory droplet transmission of an avian-human H9N2 influenza A virus. Proc Natl Acad Sci U S A. 106:7565-70.

43. **Suzuki H, Saito R, Masuda H, Oshitani H, Sato M, Sato I.** 2003. Emergence of amantadine-resistant influenza A viruses: epidemiological study. J InfectChemother. **9**(3):195-200.

44. **Toroghi, R., and R. Momayez.** 2006. Biological and molecular characterization of Avianinfluenza virus (H9N2) isolates from Iran. Acta Virol. 50:163-8.

45. **Wan, H., and D.R. Perez.** 2007. Amino acid 226 in the hemagglutinin of H9N2 influenza viruses determines cell tropism and replication in human airway epithelial cells. J Virol. 81:5181-91.

46. **Wan, H., E. M. Sorrell, H. Song, M. J. Hossain, G. Ramirez-Nieto, I. Monne, J. Stevens, G. Cattoli, I. Capua, L. M. Chen, R. O. Donis, J. Busch, J. C. Paulson, C. Brockwell, R. Webby, J. Blanco, M. Q. Al-Natour, and D. R. Perez.** 2008. Replication and transmission of H9N2 influenza viruses in ferrets: evaluation of pandemic potential. PLoS One. 3(8): e2923. doi: 10.1371/journal.pone.0002923.

47. **Wilgenbusch, J. C. and D. Swofford.** 2003. Inferring evolutionary trees with PAUP*. Curr. Protoc. Bioinformatics Chapter 6:Unit 6.4.

48. **World Health Organization.** 2010. Antigenic and genetic characteristics of influenza A(H5N1) and influenza A(H9N2) viruses and candidate vaccine viruses developed for potential use in human vaccines. Available at http://www.who.int/csr/disease/avian_influenza/guidelines/201002_H5_H9_VaccineVirusUpdate.pdf

49. **Xu C., W. Fan, R. Wei, and H. Zhao**. 2004. Isolation and identification of swine influenza recombinant A/Swine/Shandong/1/2003(H9N2) virus. Microbes Infect. **6:**919-25.

50. **Xu, K. M., K. S. Li, G. J. Smith, J. W. Li, H. Tai, J. X. Zhang, R. G. Webster, J. S. Peiris, H. Chen, Y. Guan.** 2007. Evolution and molecular epidemiology of H9N2 influenza A viruses from quail in southern China, 2000 to 2005. J Virol. **81:**2635-45

51. **Xu, K. M., G. J. Smith, J. Bahl, L. Duan, H. Tai, D. Vijaykrishna, J. Wang, J. X. Zhang , K. S. Li, X. H. Fan, R. G. Webster, H. Chen, J. S. Peiris, and Y. Guan.** 2007. The genesis and evolution of H9N2 influenza viruses in poultry from southern China, 2000 to 2005. J Virol. **81:**10389-401.

Table 1. Influenza subtypes included in the phylogenetic analysis of internal gene segments.

SUBTYPE COMPOSITION												
H1N1	H2N2 H2N3 H2N9	H3N2 H3N3 H3N5 H3N6 H3N8	H4N2 H4N6	H5N1 H5N2 H5N3 H5N6 H5N9	H6N1 H6N2 H6N5 H6N8	H7N1 H7N2 H7N3 H7N7 H7N9	H8N4	H9N2	H10N5 H10N7	H11N2 H11N7 H11N8 H11N9	H12N2 H12N3 H12N5	H14N5

Table 2. Genome constellations of the H9N2 viruses collected in Central Asia and the Middle East.

VIRUS	GENETIC GROUPS							
	PB2	PB1	PA	HA	NP	NA	M	NS
Israel-1 (2000-03)	A	A	A	A	A	A	A	A
Israel-2 (2003-07)	A	A	A	A	A	A	A	A
Jordan-1 (2004-07)	A	A	A	A	A	A	A	A
Saudi Arabia 1999	A	A	A	A	A	A	A	A
quail/Saudi Arabia/08vir3489-46-as104/06	A	A	A	A	A	A	A	A
Iraq 2005	-	-	-	A	-	A	-	-
Afghanistan (2008-09)	B	B	B	B	B	B	B	B
Pakistan (2005-08)	B	B	B	B	B	B	B	B
Iran-2 (2003-2009)**	B	B	B	B	B	B	B	B
ck/Iran/10vir854-3/2009	B	B	B	B	B	B	B	B
UAE 2000-02	C	C	C	C	C	C	C	C
Iran-1 (1998-2007)*	D	D	D	D	D	D	D	D
Israel-3 (2006-09)	A	A	A	B	A	B	B	B
Jordan-2 (2010)	A	A	A	B	A	B	B	B
ck/Saudi Arabia/08vir3489-47-as215/07	A	A	A	B	A	A	B	A
ck/Saudi Arabia/08vir3489-50-as118/07	A	A	A	B	A	B	B	A
ck/Saudi Arabia/08vir3489-51-as366/07	A	A	A	B	A	B	B	A
UAE 2003	C	B	C	C	B	C	B	C
UAE 2008	C	C	C	B	B	C	B	C
Qatar 2008	C	-	-	B	B	C	B	C
ck/Saudi Arabia/08vir3489-44-as253/06	C	B	C	B	B	B	B	C
ck/Saudi Arabia/08vir 3489-38-as92/05	-	-	C	B	B	C	B	C
ck/Saudi Arabia/08vir 3489-33-as638/05	C	C	C	B	B	A	B	C
falcon/Saudi Arabia/08vir 3489-40-as977/06	C	C	C	B	B	A	B	C
ck/Saudi Arabia/08vir 3489-48-as997/07	A	C	C	B	B	B	B	C
Pakistan-1999	G1	G1	G1	G1	G1	G1	G1	G1
Iraq-Kurdistan 2009	?	?	?	B	?	B	B	B
ck/Iran/10vir854-4/2009	?	?	?	B	B	B	B	B

* cluster represented in all phylogenies by ck/Iran/11T/99; ck/Iran/854v10-1/02

** cluster represented in all phylogenies by ck/Iran/854v10-5/2008

- no sequences available

? genetic group different from A, B; C or D

1 Table 3. Estimates of the rate of nucleotide substitution and tMRCA for H9N2 viruses from Central

2 Asia and the Middle East.

3

Gene	Substitution/site/y x 10^{-3} (95% HPD)	Mean tMRCA (95% HPD)
PB2	3.20 (2.32-4.12)	1968 (1945-1985)
PB1	3.13 (2.35-4.02)	1976 (1957-1991)
PA	3.14 (2.31-3.95)	1974 (1952-1992)
HA	4.14 (3.54-4.75)	1993 (1990-1996)
NP	2.84 (2.24-3.46)	1981 (1970-1990)
NA	4.26 (3.70-4.80)	1994 (1991-1996)
NS	3.82 (2.76-4.87)	1979 (1960-1993)
M	2.22 (1.78-2.69)	1989 (1982-1995)

4

Table 4. Amino acid sites under putative positive selection and mean d_N/d_S ratios for each gene

Gene	HA*	NA	PB2	PB1	PA	NP	NS1	NS2	M1	M2
Sites under positive selection (p<0.05)	168 (160) 198 (190) 234 (226) 282 (274) 283 (275)	42 43 50 111 141 356 463	451	-	237 272	-	171 197 215	-	74	-
Mean d_N/d_S	0.23	0.26	0.08	0.07	0.08	0.07	0.38	0.31	0.13	0.61

* Position numbers are for the H9 HA, with those for H3 HA in parentheses.

FIGURE LEGENDS

Figure 1. Maximum likelihood (ML) phylogenetic tree for the HA gene segment of H9N2 avian influenza viruses from the Middle East and Asia. Sequences of H9N2 viruses analyzed in this study are highlighted with different colors according to the country of origin: purple for Iran, green for Israel, red for Saudi Arabia, grey for UAE, yellow for Afghanistan, orange for Pakistan, blue for Jordan, light blue for Iraq. Genetic groups are coloured as follows: group A is blue, group B is green, group C is red, and group D is yellow. The numbers at each node represent bootstrap values.

Figure 2. ML phylogenetic tree for the NA gene segment of H9N2 avian influenza viruses from the Middle East and Asia. The color scheme is the same as that used in figure 1.