Molecular analysis of highly pathogenic avian influenza virus of subtype H5N1 isolated from wild birds and mammals in northern Germany

Siegfried Weber,1 Timm Harder,2 Elke Starick,2 Martin Beer,2 Ortrud Werner,2 Bernd Hoffmann,2 Thomas C. Mettenleiter1 and Egbert Mundt1†

1Institute of Molecular Biology, Friedrich-Loeffler-Institut, 17493 Greifswald-Insel Riems, Germany
2Institute of Diagnostic Virology, Friedrich-Loeffler-Institut, 17493 Greifswald-Insel Riems, Germany

Analysis of the full-length sequences of all eight segments of the German wild-bird H5N1 highly pathogenic avian influenza virus index isolate, A/Cygnus cygnus/Germany/R65/2006, and an H5N1 isolate from a cat (A/cat/Germany/R606/2006) obtained during an outbreak in February 2006 revealed a very high similarity between these two sequences. One amino acid substitution in the PA gene, encoding a protein involved in virus RNA replication, and one amino acid substitution in the haemagglutinin (HA) protein were observed. Phylogenetic analyses of the HA and neuraminidase nucleotide sequences showed that avian influenza H5N1 isolates from the Astrakhan region located in southern Russia were the closest relatives. Reassortment events could be excluded in comparison with other ‘Qinghai-like’ H5N1 viruses. In addition, an H5N1 isolate originating from a single outbreak in poultry in Germany was found to be related closely to the H5N1 viruses circulating at that time in the wild-bird population.

Enveloped particles of influenza A viruses harbour eight segments of single-stranded genomic RNA of negative polarity. Two of the eight segments encode the envelope glycoproteins haemagglutinin (HA) and neuraminidase (NA), whose antigenic properties are used to distinguish influenza virus subtypes. Until now, 16 HA and nine NA subtypes have been described. All HA and NA subtypes co-exist with their natural host, wild aquatic birds, in a meticulously balanced equilibrium (Fouchier et al., 2005). Whereas serious diseases in mammals have in the past been caused predominantly by subtypes H1, H2 and H3, only viruses of subtypes H5 or H7 are known to exist in two different pathotypes: low-pathogenic avian influenza viruses (LPAIV) remain in evolutionary stasis in their natural hosts, whereas highly pathogenic avian influenza viruses (HPAIV) may arise unpredictably from their LPAIV H5 or H7 progenitors only after transmission to susceptible poultry species. A major molecular determinant for pathogenicity of H5 and H7 viruses is the amino acid sequence specifying the proteolytic-cleavage site of HA. In LPAIV, single basic residues at the cleavage site restrict proteolytic activation of HA to the respiratory and intestinal tracts. In contrast, insertional mutations at the genomic locus encoding the endoproteolytic-cleavage site resulting in the presence of a polybasic site render it accessible for ubiquitous proteases of the subtilisin type, resulting in severe, systemic infections (Rott et al., 1995). In addition, other non-H5/H7 subtypes may also cause serious illness in chickens, but only in combination with other pathogens and factors. In the past, HPAI outbreaks have been eliminated by culling the affected birds and by application of strict hygienic measures, restrictions in transportation and appropriate surveillance programmes. An exception is the H5N1 HPAIV, which has re-emerged at least twice since 1997 in South-East Asia. From 2003 onwards, this virus not only caused large-scale outbreaks in poultry in several countries in Asia, but was also transmitted to wild-bird populations, causing significant mortalities, e.g. among migratory birds at Lake Qinghai in north-western China in 2005 (Chen et al., 2006). In 2005, a progression of H5N1 HPAIV towards western Eurasia was observed in poultry and wild birds. Most recently, within 1 week in February 2006, H5N1 HPAIV surfaced almost simultaneously in several countries in central Europe, among them Greece, Italy, Bulgaria, Slovenia, Austria and Germany. Although the vectors of this expansion have not been identified unambiguously, the patterns of progression strongly implicate an important role of migratory birds for long-distance transmission.
However, H5N1 HPAIV of Asian lineage is not confined to birds, and a slowly but steadily increasing cumulative number of confirmed human infections leads to growing concerns about an imminent pandemic caused by this strain. The molecular properties of influenza viruses that govern host tropism are still not well understood. Distinct receptor specificities only partly explain the species barrier. Therefore, analysis and comparison of full-length genomic sequences not only allow insights into the molecular epidemiology of H5N1 HPAIV, but also may provide data for the identification of molecular markers for host tropism and pathogenicity. For these purposes, we determined the full-length sequences [except for the highly conserved extreme ends (Hoffmann et al., 2001), which were used as primer-binding sites] from all eight segments of the German wild-bird H5N1 HPAIV index isolate A/Cygnus cygnus/ Germany/R65/2006 (later called R65/06), originating from a dead whooper swan found in early February 2006 on the island of Rügen, located in the south-western part of the Baltic Sea. An H5N1 isolate was also obtained 14 days later from a cat (A/cat/Germany/606/2006; later called R606/06) found dead in close proximity to the site where the first H5N1 cases in wild birds were observed in Germany.

For virus isolation, tracheal swab material from the swan was inoculated into embryonated chicken eggs and RNA was purified from allantoic fluid by using a QIAamp Viral RNA mini kit (Qiagen). For the analysis of all eight segments of the H5N1 cat virus, RNA was purified directly from brain tissue of the animal. In addition, the HA sequence of the feline H5N1 virus was analysed after one passage of a tissue of the animal. In order to gain more detailed information on the origin of R65/06, different phylogenetic analyses were performed. Alignments of the HA and NA genes were generated using CLUSTAL_W (http://www.ebi.ac.uk/clustalw). The alignments were edited by hand to exclude missing data at the 5’ and 3’ ends. Finally, sequences extending from nt 51 to 1700 of the HA gene (GenBank accession no. DQ464354), nt 21 to 1347 of the NA gene (DQ464355), nt 46 to 1019 of the NP gene (DQ464359) and nt 27 to 821 of the NS gene (DQ464358) of R65/06 were included in the analysis, using minimum evolution (ME) implemented in MEGA 3.1 (Kimura, 1980; Kumar et al., 2004) and maximum-likelihood methods (ML, Puzzle, implemented in the HUSAR clone of GCG; Strimmer & von Haeseler, 1997). Distance matrices for the ME approach were calculated by the Kimura two-parameter model. Gaps were excluded from pairwise-distance calculations. Robustness of nodes was evaluated by bootstrapping (500–1000 replications).

Phylogenetic analyses either by the ME or ML approach produced similar topologies that were supported by

| Table 1. Comparison of amino acid sequences of three isolates of the highly pathogenic avian influenza virus H5N1 |
|-----------------|-----------------|-----------------|-----------------|
| GenBank access no. | Length (nt) | Length (aa) | GenBank access no. |
| PB1 | DQ464361 | 2341 | 757 |
| PB2 | DQ464357 | 2341 | 759 |
| PA | DQ464360 | 2233 | 716 |
| HA | DQ464354 | 1779 | 568 |
| NP | DQ464359 | 1565 | 498 |
| NA | DQ464355 | 1398 | 449 |
| M | DQ464356 | 1027 | 252 (M1) |
| NS | DQ464358 | 875 | 225 (NS1) |

*Amino acid substitutions are shown using the appropriate amino acid sequence of A/Cygnus cygnus/Germany/R65/2006 (H5N1) as reference.
†Percentage identity of the nucleotide sequence to the sequence of A/Cygnus cygnus/Germany/R65/2006 (H5N1).
Fig. 1. Phylogenetic relationship of HA and NA nucleotide sequences of the H5N1 HPAI isolates A/Cygnus cygnus/Germany/R65/2006 and A/cat/Germany/R606/2006 compared with other representative H5 subtype (HA) and H5 or H7 (NA) avian influenza A viruses. The inset (HA inset) shows a separate analysis of HA sequences of an H5N1 HPAI isolate from a turkey obtained during an outbreak in Saxonia/Germany. The trees were generated by using an ME approach and have been drawn to scale. Bootstrap values >70% are shown.
comparable bootstrap values. Fig. 1 shows the results obtained with the ME method. The selected genes of swan R65/06 and cat R606/2006 encode proteins of the envelope (HA and NA), an inner protein (NP) and a protein involved in the regulation of transcription and modulation of the host innate immune response (NS). As all analysed sequences revealed similar topologies, only those for HA and NA are shown (Fig. 1). They cluster together in a group of ‘Qinghai-like’ H5N1 viruses, which are clearly separated from other HPAIV H5N1 viruses of South-East Asian lineages. Our result confirms the finding of Shestopalov et al. (2006), demonstrating that HA genes of the isolates from western Siberia, including the Astrakhan region, are distinct from isolates of South-East Asian lineage. Highest similarity rates in all segments were observed between the swan and cat viruses (> 99 %), followed by sequences of a H5N1 virus from Astrakhan (Fig. 1). With respect to the HA phylogeny, there was a tendency, based on bootstrap values, of a further separation within the Qinghai cluster, with sequences from Mediterranean (A/mallard/Italy/332/2006, A/chicken/Egypt/960N3-004/2006, A/turkey/Turkey/1/2005), African (A/chicken/Nigeria/641/2006) and inner Asiatic (A/bar-headed goose/Mongolia/1/2005, A/whooper swan/ Mongolia/3/2005) locations forming different subgroups. The very close relationship of all four genes between the Ruegen whooper swan and cat isolates is indicative of a direct epidemiological link between the viruses. It is highly likely that the cat became infected through direct contact with H5N1-infected wild birds. Our phylogenetic analyses showing that the closest relative is a virus found in southern Russia are in line with the hypothesis of an introduction of H5N1 HPAI into eastern Europe via wild (migratory) birds and that a sudden westward movement of wild birds caused by spills of Arctic cold in late January/early February 2006 led to the almost simultaneous occurrence of H5N1 viruses along a single longitude, spanning several thousands of kilometres from south Sweden across eastern Germany, Austria and Slovenia to southern Italy. As the HA sequences of the Italian and Ruegen isolates can be distinguished phylogenetically (Fig. 1), these viruses may have been derived from closely related, but distinct progenitors.

Following the occurrence of HPAIV H5N1 in the wild-bird population of Germany, a single outbreak of HPAI was also noted in a mixed-poultry holding in the federal state of Saxony. For the analysis of HA sequences of an H5N1 HPAI isolate from a turkey obtained during this outbreak, a smaller part of the HA sequence (from nt 377 to 1613) was included (Fig. 1, inset). The turkey isolate fits phylogenetically between the Ruegen cat and whooper swan sequences. Thus, the origin of this outbreak in poultry was clearly related to viruses circulating in wild birds in Germany at that time.

The identity of the determined nucleotide sequences to those of A/Cygnus olor/Astrakhan/Ast05-2-3/2005 (H5N1) ranged between 99 % (PB1, PB2, PA, HA, NA, NP, NS) and 100 % (M) (Table 1). This excludes the occurrence of reassortment events of genomic segments when comparing the ‘Qinghai-like’ H5N1 viruses of inner Asiatic origin. Analysis of the amino acid sequences revealed only a few differences in all segments except M. Two amino acid differences were observed between R65/06 and A/Cygnus olor/Astrakhan/Ast05-2-3/2005 in HA (N403D, L514F) and NP (I270V, M371I), and one amino acid difference between both isolates in PB1 (R531K), PB2 (A274T), PA (V636A), NA (D316G), NS1 (I189V) and NS2 (E63G). No differences have been detected in the deduced amino acid sequences of PB1-F2, M1 and M2. In addition, the HA sequence encoded a highly basic amino acid sequence (here RRRKKR) at the HA1/HA2 junction, which is typical for HPAIV. Several amino acids in different influenza virus proteins have been associated with enhanced replication in mammalian hosts (reviewed by Harder & Werner, 2006). Examples were also found in the Ruegen swan isolate: proline at position 13 of PB1 (Gabriel et al., 2005) and lysine at position 627 of PB2 (Subbarao et al., 1993; Hatta et al., 2001; Shinya et al., 2004). However, all of these amino acids were already present in the H5N1 HPAIV sequence obtained from the Astrakhan swan.

In the course of the epidemic among wild birds on the island of Ruegen, three stray cats and one stone marten (Martes foina) were also found to be infected with HPAIV H5N1, presumably from scavenging on carcasses of dying or dead wild birds. Apparently, these animals succumbed to the infection. With the availability of complete sequences of the swan isolate R65/06, it was possible to analyse whether additional mutations had occurred, facilitating infection of mammals. To exclude the introduction of adaptive mutations during virus isolation in embryonated chicken eggs, RNA was obtained directly from central nervous tissues of one cat. Comparison of the obtained sequences with those of the Ruegen swan isolate revealed a total of two amino acid substitutions in two genomic segments (Table 1). One of them resulted in a phenylalanine to leucine substitution in the PA protein (F4L). The second substitution was observed in the HA1 part of the HA protein (N110D). None of them had been described previously as an adaptive mutation facilitating infection of mammals. In

![Fig. 2. Alignment of HA amino acid sequences of several H5N1 isolates obtained from mammals. The alignment was based on published amino acid sequences (GenBank accession numbers: H5N1 swan-Germ, DQ464354; HA-cat-Germ, DQ463982; HA-cat-Iraq, DQ435200; HA-hum-Iraq, DQ435202). Only regions with amino acid substitutions are shown. Stretches of identical amino acids are symbolized by a double dot and identical amino acids are marked by a dash.](http://virsmjournals.org)
addition, the identified mutation in HA has not been noted to be of importance in alterations of the receptor-binding properties by using the H3 numbering system (Matrosovich et al., 1997, 2000), but may have a function for this H5N1 isolate from the cat. By comparison of the amino acid sequence of the HA gene from R65/06 with H5N1 isolates obtained recently from other mammals (cat and human; Fig. 2), several amino acid substitutions were observed, but no distinct position within the HA gene appeared to be affected consistently. Matrosovich et al. (2004), Shinya et al. (2006) and van Riel et al. (2006) showed the presence of α2,3-linked sialic acids, thought to be necessary for efficient infection of cells with avian influenza A viruses, on cells in the lower respiratory tract of humans, cats and ferrets. In contrast, the upper respiratory tract of mammals predominantly harbours sialic acid of the α2,6 types. These data explain that infection of mammals with H5N1 of avian origin will occur provided that the virus reaches the lower respiratory tract. Whether the HA mutation identified by us in the cat isolate alters receptor specificity to facilitate infection of mammals has not yet been analysed. For an efficient excretion and spread of HPAIV H5N1 among mammals, additional mutations in the virus genome, influencing virus replication efficacy and receptor specificities, are probably required. An indication for such adaptive mutations may be seen in the described non-synonymous mutations distinguishing the Ruegen swan and cat isolates.

References


