Genetic characterization of a new candidate hemagglutinin subtype of influenza A viruses

Sasan Fereidouni, Elke Starick, Kobey Karamendin, Cecilia Di Genova, Simon D. Scott, Yelizaveta Khan, Timm Harder & Aidyn Kydyrmanov

View supplementary material ⃗

Accepted author version posted online: 19 Jun 2023.

Submit your article to this journal ⃗

View related articles ⃗

View Crossmark data ⃗

Check for updates

Genetic characterization of a new candidate hemagglutinin subtype of influenza A viruses

Sasan Fereidouni, Elke Starick, Kobey Karamendin, Cecilia Di Genova, Simon D. Scott,

Yelizaveta Khan, Timm Harder, and Aidyn Kydyrmanov

Author affiliations: University of Veterinary Medicine Vienna, Vienna, Austria (S.

Fereidouni); Friedrich Loeffler Institute, Insel Riems, Germany (E. Starick, T. Harder);

Research and Production Center for Microbiology and Virology, Almaty, Kazakhstan (K.

Karamendin, Y. Khan, A. Kydyrmanov); Universities of Kent & Greenwich, Chatham

Maritime, Kent, UK (C. Di Genova, S. D. Scott).

Corresponding authors: Sasan Fereidouni, Research Institute of Wildlife Ecology,

University of Veterinary Medicine Vienna, Savoyenstrasse 1, Vienna, Austria, Tel. +43 1

250777154, Fax +43 1 25077947180: Email: sasan.fereidouni@vetmeduni.ac.at, Aidyn

Kydyrmanov, Research and Production Center for Microbiology and Virology, Bogenbai

batyr st. 105, Almaty, Kazakhstan Tel. +7 727 2 91 84 97, Fax +7 727 2 91 84 96 , Email:

aidyn.kydyrmanov@gmail.com

Abstract

Avian influenza viruses (AIV) have been classified on the basis of 16 subtypes of

hemagglutinin (HA) and 9 subtypes of neuraminidase. Here we describe genomic evidence

for a new candidate HA subtype, nominally H19, with a large genetic distance to all previously described AIV subtypes, derived from a cloacal swab sample of a Common Pochard (*Aythya ferina*) in Kazakhstan, in 2008. Avian influenza monitoring in wild birds especially in migratory hotspots such as central Asia is an important approach to gain information about the circulation of known and novel influenza viruses. Genetically, the novel HA coding sequence exhibits only 68.2% nucleotide and 68.5% amino acid identity with its nearest relation in the H9 (N2) subtype. The new HA sequence should be considered in current genomic diagnostic AI assays to facilitate its detection and eventual isolation enabling further study and antigenic classification.

Keywords: Influenza A viruses, influenza subtypes, One-Health, Wild birds, Kazakhstan

Introduction

Metapopulations of wild aquatic birds (waterfowl and shorebirds) represent the primary natural reservoir of avian influenza A viruses (AIVs) of which 16 hemagglutinin (HA) and nine neuraminidase (NA) subtypes are currently known. Two further subtypes, H17 and H18, were recently discovered in Little yellow-shouldered (*Sturnia lilium*) and a flat-faced fruit bats (*Artibeus planirostris*) respectively, but not in any avian species or human to date [1-2]. The most recent IAV subtype identified in an avian species was the H16 subtype which was isolated from a Black-headed gull in 2005 [3]. An important part of our knowledge regarding low pathogenicity avian influenza viruses (LPAIVs) has been acquired via wildlife monitoring studies. Such wild bird studies have been carried out over the past few decades, especially after emergence and geographical expansion of high pathogenicity AIVs (HPAIVs) of the H5Nx subtype since 2005. Yet, no evidence for another new AIV subtype has been reported since, until now [4-12].

We have examined approximately 8000 oropharyngeal and cloacal samples in the framework of an AIV monitoring program in wild birds in Kazakhstan from 2006 to 2011. Molecular screening of samples and subsequent sequencing of AIV positive samples revealed a number of known AIV subtypes but a single influenza A virus HA sequence could not be classified within the existing subtypes. Here we describe sequence analyses of this sample providing genomic evidence for a novel HA subtype of AIV from a Common Pochard (*Aythya ferina*). With more than 30% genetic distance [3,13], based on the HA coding region nucleotide sequence, to its closest relatives in the H9 subtype, the novel sequence represents a potential candidate new subtype, nominally H19.

Materials & Methods

Study design

In the framework of an AIV wild bird monitoring program in Kazakhstan during 2002 to 2009, approximately 4800 cloacal and oropharyngeal swab samples from 155 different species of wild birds were collected and analysed [11]. The study was previously described comprehensively [11], in summary, samples were tested using virus culture in embryonated chicken eggs, conventional RT-PCR, and a generic quantitative real-time RT-PCR (RT-qPCR) targeting the AIV nucleoprotein (NP) gene [14]. In total, 95 AIV from different subtypes, especially H3, H4 and H13 subtypes, were isolated. In addition, 99 positive samples were identified using conventional RT-PCR and quantitative RT-PCR (RT-qPCR). A cloacal swab sample with sample ID of Kz52, that was collected on 23.11.2008 from a Common Pochard (*Aythya ferina*) was among the RT-qPCR positives. Positive RT-qPCR samples, including Kz52 were used for re-inoculation in embryonated chicken SPF (specific pathogen free) eggs, MDCK II (Madin-Darby canine kidney) cell culture cultivation, genome amplification and sequencing.

Virus isolation and characterization

Virus isolation was attempted in 11-day-old embryonated SPF chicken eggs and MDCK II cell cultures based on standard procedures [15]. The samples were passaged two or three times before being considered negative based on lack of hemagglutination activity and negative generic RT-qPCR tests.

RNA extraction

RNA was extracted from the original swab sample using the QIAamp Viral RNA kit (Qiagen) for the conventional RT-PCR or sequencing, according to the manufacturer's instructions. RNA for RT-qPCR analysis were extracted directly from the samples by an automated RNA extraction robot (Freedom Evo 3000, Tecan) using the NucleoSpin 96 Virus Core kit (Macherey & Nagel).

Quantitative Real-time RT-PCR (RT-qPCR)

A previously validated generic quantitative real-time RT-PCR (RT-qPCR), using forward and reverse primers (1448-F and 1543-R) and probe (1473-FAM), targeting the AIV NP gene was used for screening of the samples or evaluation of cultivated cell cultures, and viral load in original sample was estimated [14,16].

RT-PCR and Sequencing

Reverse transcription-PCR (RT-PCR) assays were performed via one-step protocols using the Qiagen RT-PCR Kit according to the manufacturers' instructions. Hoffmann primers were used for full length amplification of genomic RNA and further sequencing of different viral segments, including polymerase basic 2 (PB2), polymerase basic 1 (PB1), polymerase acidic (PA), HA, NP, NA, matrix (M), and non-structural protein (NS) genes [17]. PCR products of the anticipated size range were purified using the QIAquick Gel

Extraction Kit (Qiagen, Germany). Purified DNA fragments were cycle sequenced in both directions using the same primers as for RT-PCR. The Prism Big Dye Terminator v1.1 cycle sequencing kit (Applied Biosystems) was utilized and amplicons were analyzed on an automatic sequencer (ABI-377, Applied Biosystems). Assembled nucleotide sequences were then used in BlastN2 database searches for subtype specification.

A second round of sequencing focused specifically on the HA gene of sample Kz52 and was conducted in collaboration with Virology Research Services (Sandwich, UK) and was achieved via two approaches. Firstly, the sequence of the N-terminal two-thirds (~1150bp) was obtained by one strategy, followed by the C-terminal one third (~600bp) by an alternative approach, both described below. For the former, total extracted RNA was retro-transcribed using random hexamer and influenza HA-specific primers (details below) to initiate first round cDNA synthesis via Superscript Reverse Transcriptase (ThermoFisher) priming at 65°C for 5 minutes (min), reverse transcribed at 55°C for 10 min, then deactivated at 80°C for 10 min. HA sequences from this cDNA were then amplified by polymerase chain reaction (PCR) using HA-specific primers (details below) and Phusion Hi-Fidelity DNA Polymerase (ThermoFisher) according to manufacturer's instructions, using the thermal cycling; initially 98°C for 30 secs, then 25 cycles of 98°C for 10 secs + 52°C for 30 secs + 72°C for 30 secs, with final extension at 72°C for 5 min. PCR products (~300-400bp) were purified by gel electrophoresis and extraction. These products were Sanger sequenced (Source Bioscience, Cambridge, UK) and used to design novel primers to produce further nested PCR products for sequencing. These sequences were assembled into the ~1150bp consensus. Despite several attempts, it was not possible to extend this sequence further by the same methodology, possibly due to target degradation. Consequently, a second 'gene walking' approach was employed. This involved using a PCR primer based on a known sequence adjacent to unknown region (5'-AAAGAGACACCGCTCAAG-3' using above cycle) to produce ssDNAs of varying

lengths from the cDNA above, followed by non-specific binding of reverse primers to generate dsDNAs (single cycle of 98℃ for 30 secs + 45℃ for 30 secs + 72℃ for 3 min). A second round of nested PCR was conducted (30 cycles, as above) to produce more amplicons that were gel purified and sequenced. Lastly, the N and C-terminal sequences were assembled as a single 1686 nucleotide contiguous open reading frame.

Sequences of primers utilized in above methodologies:

Primers used for RNA retro-transcription reaction I:
Rev2: 5′-AAAGGTACTCTTGTTCTGATT-3′
Rev1: 5′-TCAGATACAAATGGTGCATTTG-3′

Primers used for RNA retro-transcription reaction II:
Rev1: 5′-TCAGATACAAATGGTGCATTTG-3′

Primers used for RNA retro-transcription reaction III:
Rev3: 5′-CGAAGTTATCTTATCAATGGCC-3′

Primers used for PCR reactions:

HAFor: 5′-ATGTGGAAACTAGCATTAGTAACG-3′
HARev: 5′-AAAGGTACTCTTGTTCTGATT-3′
HA1For: 5′-TTAGTAACGACTTTTTTGATGC-3′
HA1Rev: 5′-ACTTGAGGAAGAGAATAAGAC-3′
HA2For: 5′-AAAGAGACACCGCTCAAG-3′
HA2Rev: 5′-ATTTTTAAAGTCTGATTGGGC-3′
HA3For: 5′-GAGTTTCAAGTCAACATTCGG-3′
HA3Rev: 5′-CGAAGTTATCTTATCAATGGCC-3′
HA4For: 5′-AATTCAGAGGGAACAGGAATGG-3′
HA4Rev: 5′-TCAGATACAAATGGTGCATTTG-3′

Estimates of Evolutionary Divergence between Sequences

The number of base substitutions per site between sequences were calculated using the Tajima-Nei model implemented within MEGA 11 software. The rate variation among sites was modeled with a gamma distribution. Codon positions included were 1st+2nd+3rd+Noncoding. All positions with less than 95% site coverage were eliminated and ambiguous bases were allowed at any position (partial deletion option). There were a

total of 1680 positions in the final dataset. Evolutionary analyses were conducted in MEGA 11.

Prediction of Potential N-glycosylation sites:

Potential N-glycosylation sites (PGS) of the HA were predicted using the NetNGlyc 1.0 Server (https://services.healthtech.dtu.dk/services/NetNGlyc-1.0).

Phylogenetic analyses:

Phylogenetic analyses were carried out for the complete open reading frame (ORF) of the HA gene of A/Common Pochard/Kazakhstan/Kz52/2008 (Kz52 in brief) AIV and representatives of other influenza A virus subtypes. Ninety-one representative sequences of all AIV and bat-origin HA subtypes for phylogenetic analysis were retrieved from the NCBI Influenza Database, aligned and trimmed to equal lengths using BioEdit 7.2.5.

The trees were generated using a maximum-likelihood (ML) approach based on a General Time Reversible (GTR) model with rate heterogeneity and a proportion of invariant sites with 1000 bootstrap replicates implemented within MEGA 11 software. In addition, all trees were reconstructed and confirmed using the neighbor-joining algorithm based on the Tamura 3-parameter model, rate heterogeneity and 1000 bootstrap replicates.

Three-dimensional (3D) structural modelling of HA:

A 3D model of the Kz52 HA polypeptide was predicted using the free software Phyre2 [18; www.sbg.bio.ic.ac.uk/phyre2]. Modelling compared the novel HA sequence with that of other HAs for which structures have been determined. The best structural alignment was with H16 (A/black-headed gull/Sweden/2/99; NCBI accession number AY684888.1). Phyre 2 'normal mode' default settings were used to predict the 3D models

of HA. The Phyre 2 PDB file was then imported into the PyMOL program (www.pymol.org) in order to highlight specific features on the HA structure.

Results

One cloacal swab sample, Kz52, that was collected from a Common Pochard was considered positive in a generic AIV-specific RT-qPCR with low viral load of about 10^2-10^3 viral genome copies per 100 µl of cloacal swab suspension (Ct value equal to 34.12 in original swab material). No virus could be propagated either in embryonated chicken SPF eggs, or MDCK cell cultures, possibly because the sample had already experienced several cycles of freezing and thawing following initial analysis, storage, shipment and RNA extraction phases, before any attempts at virus isolation were made and the sample included low virus load. RNA extracted from the residual sample was only adequate for complete sequencing of the hemagglutinin and short fragments of three other segments.

The sequencing of the HA gene via different complimentary approaches and protocols finally revealed a complete sequence of the HA open reading frame (ORF) comprising 1686 nucleotides (including start and stop codons) and 561 amino acids (Figure 1). Extensive database searches and estimation of evolutionary divergence revealed that the HA gene shared only 68.2% nucleotide and 68.5% amino acid sequence identity with its closest related sequences, all found among strains of the H9 subtype isolated from wild and domestic birds and bats (https://blast.ncbi.nlm.nih.gov/Blast.cgi), such as A/duck/NZL/76/1984(H9N1), A/chicken/Korea/GH2/2007(H9N2) and A/Bat/Egypt/381OP/2017(H9N2), respectively. Kz52 is placed as a sister clade to the H9 subtype with subtypes H8 and H12 being closely related to the sister clade (Figures 2a & 2b).

Due to the low viral load of the specimen, the gene specific RT-PCR assays did not reveal any positive PCR product for PB2, PA, NA, and M genes of Kz52; and partial

sequences were generated only for PB1 (620 of 2370 nt), NP (310 of 1594 nt), and NS (330 of 919 nt) genes.

Analysis of partial sequences of the NP gene show 98.4% sequence identity with A/mallard/Sweden/124418/2010 (H4N6), NS with 93% identity with A/mallard/Sweden/104690/2009 (H4N6), and PB1 with 93.7% identity with A/mallard/Netherlands/10/1999 (H1N8) revealing much higher sequence identity to known naturally occurring wild bird influenza viruses of the Eurasian lineage compared to the HA. However, the phylogenetic analyses indicated that Kz52 NS and PB1 sequences form a separate branch even compared to closely related correspondent sequences (Figures S1-S3). The HA sequence has been deposited in Genbank database with accession number ON637239.

The Kz52 HA cleavage site was defined as PIKETR↓GLF (with G at position 340). Despite the fact that Kz52 HA showing the highest sequence identity and 3D predicted HA model similarity (Figure 3) with H9 subtypes, its cleavage site pattern was identical to that of the bat influenza strain A(H18N11) isolated in South America [2]. The cleavage motif of Kz52 has two basic amino acids that are seen often in low pathogenic H5 AIV strains. Historically, amino acid substitutions at positions E190D and G225D for H1N1 [19] and Q226L with G228S in the case of H2N2 and H3N2 (H3 numbering) [20] within the HA receptor-binding site (RBS) of all influenza viruses have been shown to play a key role in transition from avian-like to human-like terminal sialic acid receptor recognition. The Kz52 HA RBS has typical avian residues at the corresponding positions: E190, G225 and G228 but has a substitution Q226I that differed from the closest H9 subtype sequences that have leucine or glutamine at this position. In addition, the HA protein sequence of Kz52 has tyrosine 98 (Y98), serine 136 (S136), tryptophan 153 (W153), and histidine 183 (H183) forming the base of the RBS which is conserved among all HA subtypes [21].

Six potential N-glycosylation sites (PGSs) at positions 30, 142, 193, 249, 299 and 306 in HA1, and one 493 in HA2 were identified. Two of those PGSs, at positions 193 and 249, have not been observed in avian and bat-origin H9 viruses [22]. None of these PGS affected the RBSs.

Discussion

Our knowledge regarding the circulation of AIVs in wild bird populations as potential donors of previous pandemic causing influenza viruses, the relative prevalence of various subtypes in different species, susceptibility or resistance of different bird species to different genetic lineages of AIV subtypes, mainly stems from monitoring and surveillance studies in an extensive range of countries [4-11,23]. Thus, continued influenza monitoring of wild bird populations is essential to assess potential disease impact in these species, domestic birds, and mammals, including man.

As a result of such activities, the genomic sequence of a potentially novel HA subtype of AIVs was identified from a cloacal swab sample of a Common Pochard (*Aythya ferina*), A/ Common Pochard/Kazakhstan/Kz52/2008. Unfortunately, due to the low virus load in the bird, the virus itself could not be isolated. Despite successful sequencing of the entire HA ORF, it was not possible to obtain the complete sequence of the Kz52 genome or further complete ORFs with the remaining viral genomic RNA. Nevertheless, partial sequences of other genes (NP, NS and PB1) were obtained, providing some further genetic information.

The HA of this virus is genetically significantly divergent from all known AIVs, and appears to represent a separate monophyletic group established as a sister clade to the H9 phylogenetic branch. The H9 HA sequence obtained from bats in Egypt is likewise close to the root of H9, although it is reported to represent a true H9 subtype with an

identity of 73% to avian H9 viruses [22]. The full HA coding region of Kz52 showed only 68.2% sequence homology with closest related H9 subtypes viruses. In addition, as the genetic distance on the basis of amino acid sequences from other known subtypes was more than 30%, in accordance with current AIV subtype classification [3,13], thus we suggest the Kz52 as a potential candidate for a new subtype, although antigenic classification is essential to confirm this, as described in WHO guidelines [24]. Nevertheless, this represents the first new avian HA sequence as substantially distinct from all other subtypes, as of the most recent additions, H16 AIV in 2005 [3], and two bat-origin influenza viruses identified a decade ago [H17, H18: 1-2].

Since this new HA subtype remained undetected despite extensive AIV monitoring in wild birds, it is possible that the corresponding AIV may only propagate in very limited numbers of avian species, or there may be a strict geographic restriction. The intensified monitoring of wild birds at the location where this sample has been obtained would be a logical strategy to obtain further material for analysis of this new subtype candidate, including isolating the virus itself. A specifically designed RT-qPCR would aid in high throughput screening of such new samples and existing sample banks [25]. The sample was collected in the Delta of Ural River and the North Caspian region, one of the important areas for stopover, accumulation and nesting of migratory birds, where the Black Sea-Mediterranean and East African-West Asian migratory flyways intersect. The retrospective confirmation of circulation of H16 subtype AIVs among gulls in the Northern Caspian since 1976 [26] and first isolation of the H14 subtype from this region [13] indicates that the Northern Caspian is one of the potential epicenters for emerging novel and rare influenza A virus subtypes.

The close phylogenetic association of Kz52 with bat-origin H9 sequences [22] is strikingly paralleled by its HA cleavage site which was similar to bat H18N11 influenza viruses [2] but has a typical avian RBS (at positions of E190, G225 and G228). However,

the unique substitution Q226I was observed at the RSB key position of Kz52, while H9 subtype viruses mainly contain leucine (L) or glutamine (Q) at the position 226 [27]. Although the novel subtype hemagglutinin possesses avian-like RBS, experimental studies are necessary to determine the role of its RBS in host specificity.

We were not able to amplify the Neuraminidase (NA) gene using a range of sequencing primers previously used for all other known NAs. This may indicate that NA was also genetically different from previously identified NAs. Among three partial sequences of other viral segments, only NP showed very high genetic identity (98%) with NP of other circulating Eurasian AIVs in wild birds. NS and PB1 fragments showed approximately 93% genetic relatedness with other NS and PB1 of AIVs. Although short fragment sequencing of other segments such as NP, NS, and PB1 revealed different levels of genetic relatedness with commonly circulating genes in wild birds, without access to the virus or complete genome of the virus, a definitive conclusion about the genetic background of the virus is not possible. A small amount of original genomic RNA was preserved for future investigations, when genome sequencing techniques might permit further refined analysis, prior to reconstructing a replication-competent virus by reverse genetics. There are several Next Generation Sequencing approaches, which might be applied, such as the multisegment reverse transcription PCR method [28], or more recent Oxford Nanopore minION direct RNA sequencing platform [29], however, the current scarcity of sample material is a barrier to progress. This highlights the need for further wildlife surveillance to provide new material, combined with improved detection with novel PCR primers. Nevertheless, we have cloned the HA gene into expression vectors and are currently devising serological tests for antigenic subtyping and potential subtype delineation [unpublished data].

With the current intensified AIV surveillance in wild and domestic animals, it is probable that further viruses closely related to Kz52 will be detected, and indeed further

AIV subtypes. The development of diagnostic kits and updating the primers for AIV screening using the results of current study may facilitate further identification of related influenza viruses and aid in specific monitoring in a range of avian species [30]. Increasing usage of whole genome sequencing (WGS)/metagenomics as a diagnostic tool could be considered more with wild bird samples, which will help identify rare and divergent subtypes. Identification of a such divergent HA sequence of AIV in a Common Pochard indicates the continuing emergence of influenza viruses and emphasizes the intriguing variability of these viruses. It seems, the full spectrum of variability of AIV remains to be unraveled yet.

Disclosure statement

The authors report there are no competing interests to declare.

References:

[1] Tong S, Li Y, Rivailler P, Conrardy C, et al. A distinct lineage of influenza A virus from bats. Proc Natl Acad Sci U S A. 2012 Mar 13;109(11):4269-74. doi: 10.1073/pnas.1116200109.

[2] Tong S, Zhu X, Li Y, et al. New world bats harbor diverse influenza A viruses. PLoS Pathog. 2013;9(10):e1003657. doi: 10.1371/journal.ppat.1003657.

[3] Fouchier RA, Munster V, Wallensten A, et al. Characterization of a novel influenza A virus hemagglutinin subtype (H16) obtained from black-headed gulls. J Virol. 2005;79(5):2814-22.

[4] Baek YH, Pascua PN, Song MS, et al. Surveillance and characterization of low pathogenic H5 avian influenza viruses isolated from wild migratory birds in Korea. Virus Res. 2010; 150(1-2):119-28. doi: 10.1016/j.virusres.2010.03.002.

[5] Brown IH. Summary of avian influenza activity in Europe, Asia, and Africa, 2006-2009. Avian Dis. 2010; 54(1 Suppl):187-93. doi: 10.1637/8949-053109-Reg.1.

[6] Busquets N, Alba A, Napp S, et al. Influenza A virus subtypes in wild birds in North-Eastern Spain (Catalonia). Virus Res. 2010; 149(1):10-8. doi: 10.1016/j.virusres.2009.12.005.

[7] Fereidouni SR, Werner O, Starick E, et al. Avian influenza virus monitoring in wintering waterbirds in Iran, 2003-2007. Virol J. 2010b; 7:43. doi: 10.1186/1743-422X-7-43.

[8] Fouchier RA, Olsen B, Bestebroer TM, et al. Influenza A virus surveillance in wild birds in Northern Europe in 1999 and 2000. Avian Dis. 2003; 47(3 Suppl):857-60. doi: 10.1637/0005-2086-47.s3.857.

[9] Henriques AM, Fagulha T, Barros SC, et al. Multiyear surveillance of influenza A virus in wild birds in Portugal. Avian Pathol. 2011;40(6):597-602. doi:10.1080/03079457.2011.618943.

[10] Hoye BJ, Munster VJ, Nishiura H, et al. Surveillance of wild birds for avian influenza virus. Emerg Infect Dis. 2010; 16:1827-34.

[11] Kydyrmanov A, Sayatov M, Karamendin K, et al. Monitoring of influenza A viruses in wild bird populations in Kazakhstan in 2002-2009. Arch Virol. 2017 Jan;162(1):147-155. doi: 10.1007/s00705-016-3076-4.

[12] Miller BJ. Why unprecedented bird flu outbreaks sweeping the world are concerning scientists. Nature. 2022 Jun;606(7912):18-19. doi: 10.1038/d41586-022-01338-2.

[13] Kawaoka Y, Yamnikova S, Chambers TM, et al. Molecular characterization of a new hemagglutinin, subtype H14, of influenza A virus. Virology. 1990;179(2):759-67. doi: 10.1016/0042-6822(90)90143-f.

[14] Fereidouni SR, Harder TC, Gaidet N, et al. Saving resources: Avian influenza surveillance using pooled swab samples and reduced reaction volumes in real-time RT-PCR. J Virol Meth. 2012;186:119-25.

[15] EC (European Commission) (2019) Commission Decision 2020/689/EC of 4 August 2006 approving a Diagnostic Manual for avian influenza as provided for in Council Directive 2005/94/EC. Available at: https://eur-lex.europa.eu/legal-content/EN/TXT/PDF/?uri=CELEX:32020R0689&from=en.

[16] Fereidouni SR, Starick E, Beer M, et al. Highly pathogenic avian influenza virus infection of mallards with homo- and heterosubtypic immunity induced by low pathogenic avian influenza viruses. PLoS One. 2009 Aug 20;4(8):e6706. doi: 10.1371/journal.pone.0006706.

[17] Hoffmann E, Stech J, Guan Y, et al. Universal primer set for the full-length amplification of all influenza A viruses. Arch Virol 2001, 146:2275-2289.

[18] Kelley LA, Mezulis S, Yates CM, et al. The Phyre2 web portal for protein modeling, prediction and analysis. Nat Protoc. 2015 Jun;10(6):845-58. doi: 10.1038/nprot.2015.053.

[19] Gamblin SJ, Haire LF, Russell RJ, et al. The structure and receptor binding properties of the 1918 influenza hemagglutinin. Science. 2004 19;303(5665):1838-42. doi: 10.1126/science.1093155.

[20] Liu J, Stevens DJ, Haire LF, et al. Structures of receptor complexes formed by hemagglutinins from the Asian Influenza pandemic of 1957. Proc Natl Acad Sci U S A. 2009 Oct 6;106(40):17175-80. doi: 10.1073/pnas.0906849106.

[21] Martín J, Wharton SA, Lin YP, et al. Studies of the binding properties of influenza hemagglutinin receptor-site mutants. Virology. 1998 Feb 1;241(1):101-11. doi: 10.1006/viro.1997.8958.

[22] Kandeil A, Gomaa MR, Shehata MM, et al. Isolation and Characterization of a Distinct Influenza A Virus from Egyptian Bats. J Virol. 2019 Jan 4;93(2):e01059-18. doi: 10.1128/JVI.01059-18.

[23] Lewis NS, Javakhishvili Z, Russell CA, et al. Avian influenza virus surveillance in wild birds in Georgia: 2009-2011. PLoS One. 2013;8(3):e58534. doi: 10.1371/journal.pone.0058534.

[24] World Health Organization (WHO). A revision of the system of nomenclature for influenza viruses: a WHO memorandum. Bull World Health Organ. 1980;58(4):585-91.

[25] Shriner SA, Root JJ, Lutman MW, et al. Surveillance for highly pathogenic H5 avian influenza virus in synanthropic wildlife associated with poultry farms during an acute outbreak. Sci Rep. 2016 Nov 4;6:36237. doi: 10.1038/srep36237.

[26] Yamnikova SS, Gambarian AS, Aristova VA, et al. [A/H13 and A/H16 influenza viruses: different lines of one precursors]. Vopr Virusol. 2009;54(4):10-17. Russian. Abstract in English.

[27] Wan H, Perez DR. Amino acid 226 in the hemagglutinin of H9N2 influenza viruses determines cell tropism and replication in human airway epithelial cells. J Virol. 2007 May;81(10):5181-91. doi: 10.1128/JVI.02827-06.

[28] Zhou B, Donnelly ME, Scholes DT, et al. Single-reaction genomic amplification accelerates sequencing and vaccine production for classical and Swine origin human influenza a viruses. J Virol. 2009;83(19):10309-13. doi: 10.1128/JVI.01109-09.

[29] Keller MW, Rambo-Martin BL, Wilson MM, et al. Direct RNA Sequencing of the Coding Complete Influenza A Virus Genome. Sci Rep. 2018;8(1):14408. doi: 10.1038/s41598-018-32615-8.

[30] Panzarin V, Marciano S, Fortin A, et al. Redesign and Validation of a Real-Time RT-PCR to Improve Surveillance for Avian Influenza Viruses of the H9 Subtype. Viruses. 2022 Jun 10;14(6):1263. doi: 10.3390/v14061263.
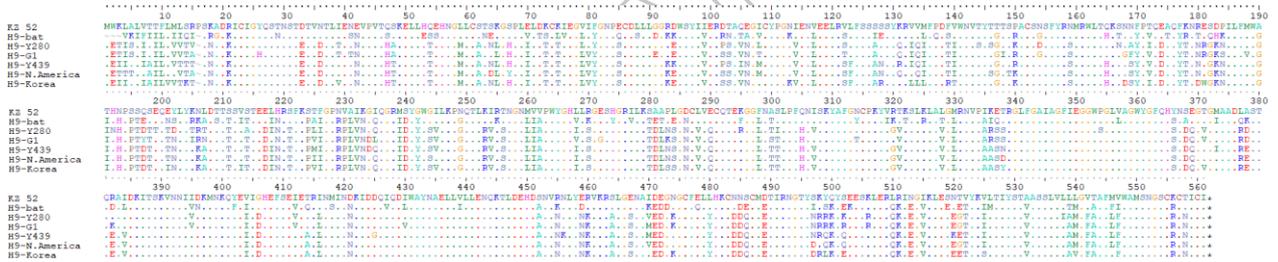
Figure 1. The complete open reading frame (ORF) of the HA gene of A/ Common Pochard/Kazakhstan/Kz52/2008 influenza virus (new subtype), compared to closely related H9 subtype isolates. Representative isolates for each lineage included: H9-bat: A/Bat/Egypt/381OP/2017(H9N2), H9-Y280: A/Chicken/Beijing/1/97(H9N2), H9-G1: A/quail/Hong Kong/G1/97(H9N2), H9-Y439: A/Duck/Hong Kong/Y439/97(H9N2), H9-N. America: A/Mallard/California/D1713634/2017(H9N2), H9-Korea: A/chicken/Korea/GH2/2007(H9N2).

0.26 HA 2
0.13
0.27 HA 5
0.07
0.42 HA 1
0.09
0.51 HA 6
0.12
0.43 HA 17
0.56
0.46 HA 18
0.06
0.33 HA 11
0.19
0.26 HA 13
0.36
0.22 HA 16
0.14
0.39 ● KZ 52
0.14
0.28 HA 9
0.25
0.37 HA 8
0.15
0.35 HA 12
0.23 HA 4
0.16
0.31 HA 14
0.34
0.55 HA 3
0.45
0.38 HA 10
0.37
0.27 HA 7
0.26
0.20 HA 15

0.20

A/Berlin/3/1964(H2N2)L11126
A/Japan/305/1957(H2N2)CY014976
A/England/1/1961(H2N2)CY125854
A/Georgia/1/1967(H2N2)CY033980
H2

A/duck/Minnesota/1525/1981(H5N1)CY014726
A/duck/France/02166/2002(H5N3)AJ632268
A/duck/Singapore/3/1997(H5N3)EF619972
A/mallard/Bavaria/1/2005(H5N2)DQ387854
H5

A/Iowa/CEID23/2005(H1N1)DQ889689
A/India/6263/1980(H1N1)CY020453
A/Denmark/3/2005(H1N1)EU097950
A/New Caledonia/20/1999(H1N1)CY031336
H1

A/chicken/Taiwan/0204/05(H6N1)DQ376652
A/mallard/Netherlands/16/99(H6N5)AY684892
A/pintail/Ohio/226/1998(H6N2)CY013255
A/sanderling/Delaware/1258/1986(H6N6)CY014607
H6

A/little yellow-shouldered bat/Guatemala/153/2009(H17N10)(CY103876)
A/little yellow-shouldered bat/Guatemala/164/2009(H17N10)(CY103884)
A/little yellow-shouldered bat/Guatemala/060/2010(H17N10)(CY103892)
H17

A/dark fruit-eating bat/Bolivia/PBV780-781/2011(H18N11)(KR077932)
A/flat-faced bat/Peru/033/2010(H18N11)(CY125945)
A/Artibeus lituratus/Brazil/2344/2012(H18N11)(MH682211)
A/Artibeus lituratus/Brazil/2301/2012(H18N11)(MH682203)
H18

A/duck/Washington/663/1997(H11N9)EF599120
A/mallard/ALB/124/1991(H11N2)CY006003
A/mallard duck/ALB/797/1983(H11N3)CY006002
A/environment/Delaware/232/2005(H11N8)CY021149
H11

A/black-headed gull/Astrakhan/227/84(H13N6)M26089
A/black-headed gull/Sweden/1/1999(H13N6)AY684887
A/black-headed gull/Republic of Georgia/7/2011(H13N6)CY185489
H13

A/common gull/Norway/10 1617/2006(H16N3)FM179755
A/environment/California/1242V/2012(H16N3)CY176997
A/glaucous gull/Alaska/44198-027/2006(H16N3)HM059998
A/glaucous-winged gull/Southcentral Alaska/9JR0783R0/2009(H16N3)CY070890
H16

A/pintail duck/Alaska/102/2005(H12N5)CY017733
A/pintail duck/Alberta/49/2003(H12N5)CY005920
A/mallard/Ohio/409/1988(H12N5)CY016419
A/mallard/Sweden/100103/2009(H12N5)JX566034
H12

A/common teal/Netherlands/1/2005(H8N4)CY041258
A/blue-winged teal/Guatemala/CIP049-07/2010(H8N4)CY096696
A/duck/Alaska/702/1991(H8N2)CY015173
H8

KZ 52
H19

A/Bat/Egypt/381OP/2017(H9N2)(MH376902)
Bat-origin H9

A/mallard/Finland/13353/2010(H9N2)KF18363
A/mallard/Austria/WV1090234/2007(H9N2)GU19448
A/gadwall/Netherlands/1/2006(H9N2)CY043864
A/Duck/Hong Kong/Y439/97(H9N2)KF188265
A/duck/Germany/113/1995(H9N2)HE802066
Y439 lineage

A/chicken/Korea/GH2/2007(H9N2)HQ871933
A/swine/Korea/S452/2004(H9N2)AY790305
A/mallard/South Korea/N11-087/2011(H9N2)MN530493
Korean lineage

A/turkey/New York/11275/1999(H9N2)LC431370
A/Mallard/California/D1713634/2017(H9N2)MK995793
A/ruddy turnstone/New Jersey/UGAI17-2707/2017(H9N2)MH068463
North American line

A/chicken/Japan/AQ-HE61/2015(H9N2)LC208476
A/chicken/Hubei/H2168/2017(H9N2)MK553893
A/chicken/Shandong/1784/2019(H9N2)MN765086
A/Japanese Quail/Vietnam/4/2009(H9N2)AB753192
A/Duck/Hong Kong/Y280/97(H9N2)AF156376
A/Chicken/Beijing/1/97(H9N2)AF461530
Y280 lineage

A/quail/hong kong/G1/97(H9N2)KY785896
A/parakeet/Narita/92A/98(H9N2)AB049160
A/chicken/Pakistan/2/1999(H9N2)KF188299
A/quail/Bangladesh/33376/2017(H9N2)MH135567
A/chicken/Iran/N101/2011(H9N2)JQ970436
A/chicken/Egypt/A15074/2018(H9N2)MN038193
G1 lineage

H9

A/Netherlands/233/1982(H3N2)CY077834
A/Udorn/307/1972(H3N2)M54895
A/Netherlands/241/1993(H3N2)CY112613
A/Sichuan/2/1987(H3N2)CY121293
A/Netherlands/213/2003(H3N2)HQ166052
A/Vienna/28/2006(H3N2)JF340085
H3

A/duck/Hong Kong/365/1978(H4N6)CY006027
A/duck/Nanchang/4-165/2000(H4N6)CY006017
A/duck/New Zealand/31/1976(H4N6)M25286
H4

A/herring gull/Astrakhan/267/1982(H14N5)FJ975075
A/long-tailed duck/Wisconsin/10OS3912/2010(H14N6)JN696314
A/long-tailed duck/Wisconsin/10OS4225/2010(H14N6)JN696316
H14

A/quail/Italy/1117/1965(H10N8)CY014644
A/swan/Shimane/1331/1981(H10N6)AB289339
A/red knot/Delaware/1269/2000(H10N7)GU050906
A/ruddy turnstone/NJ/1600/2001(H10N7)EU743418
H10

A/chicken/Pakistan/447/1995(H7N3)AF202226
A/duck/Hokkaido/143/2003(H7N1)AB269694
A/chicken/Rostock/8/1934(H7N1)M24457
A/chicken/New York/12273-11/1999(H7N3)AY240892
A/chicken/New York/19499/2005(H7N2)CY034246
H7

A/teal/Chany/7119/2008(H15N4)CY098540
A/duck/Australia/341/1983(H15N8)CY006009
A/Australian shelduck/Western Australia/1756/1983(H15N2)CY006032
A/wedge-tailed shearwater/Western Australia/2327/1983(H15N9)CY006034
A/wedge-tailed shearwater/Western Australia/2576/1979(H15N9)CY006010
H15

0,50

Figure 2a & 2b. Phylogenetic analysis of Kz52 HA coding region nucleotide sequences and the representative sequences of other influenza subtypes using obtained from publicly available sequence datasets (GenBank). Figure 2a shows the distances and 2b the corresponding bootstrap values. Analysis was based on full-length coding sequence of the HA gene. Only bootstrap values >80 are shown. Bar represents 0.01 nt substitutions.
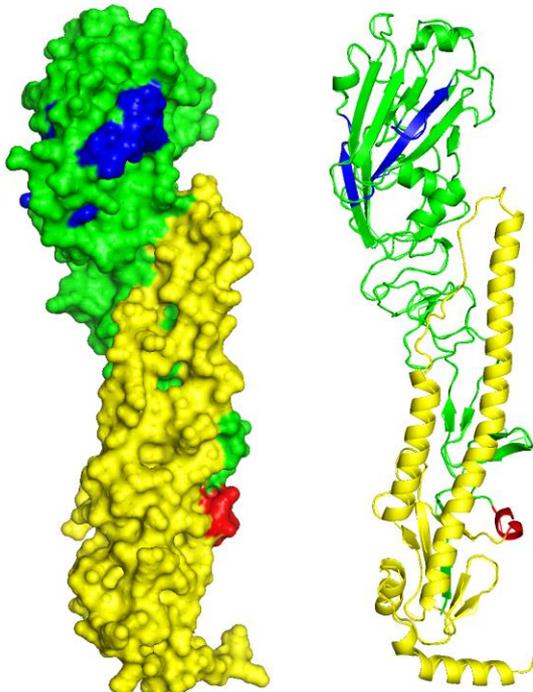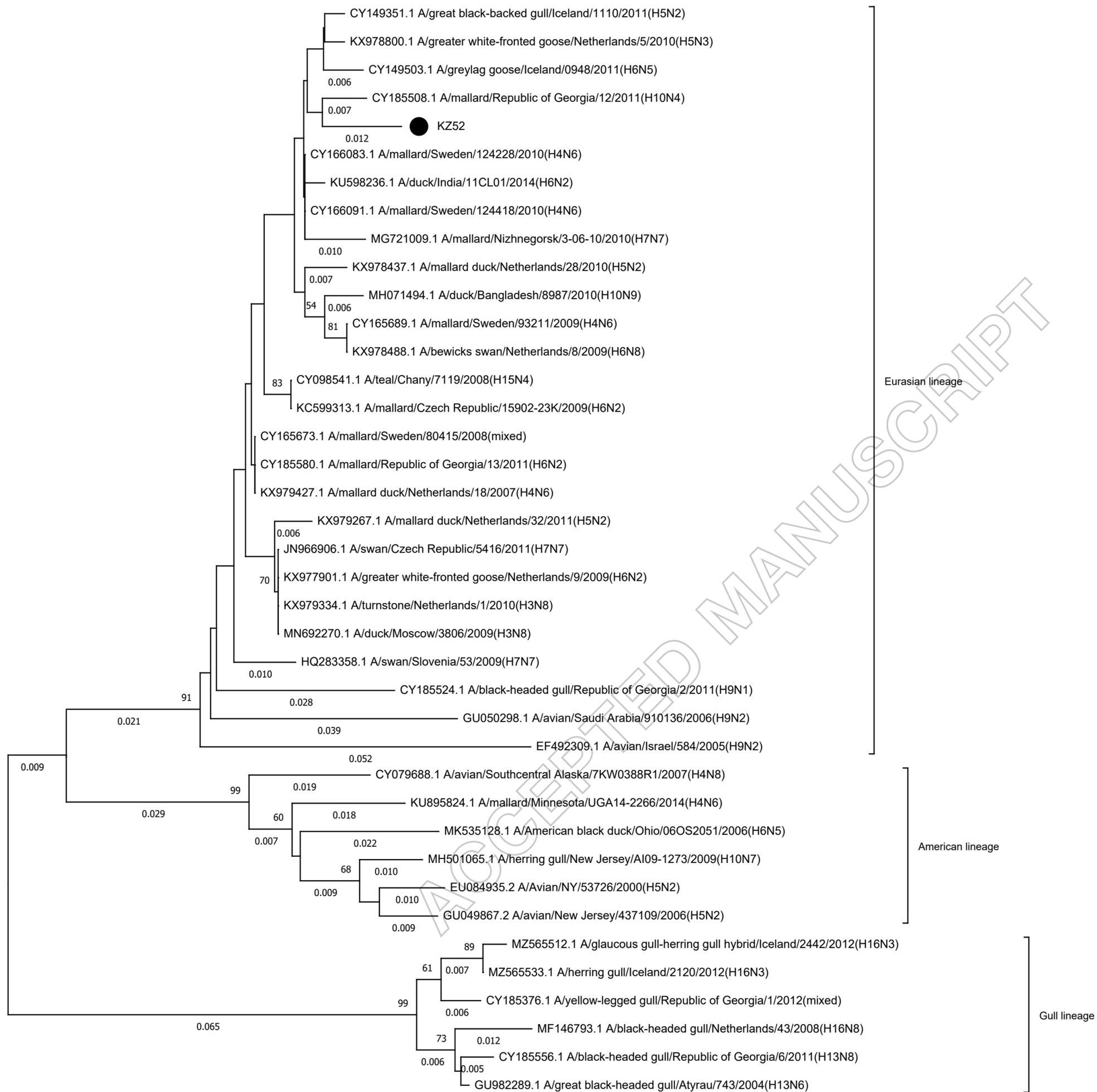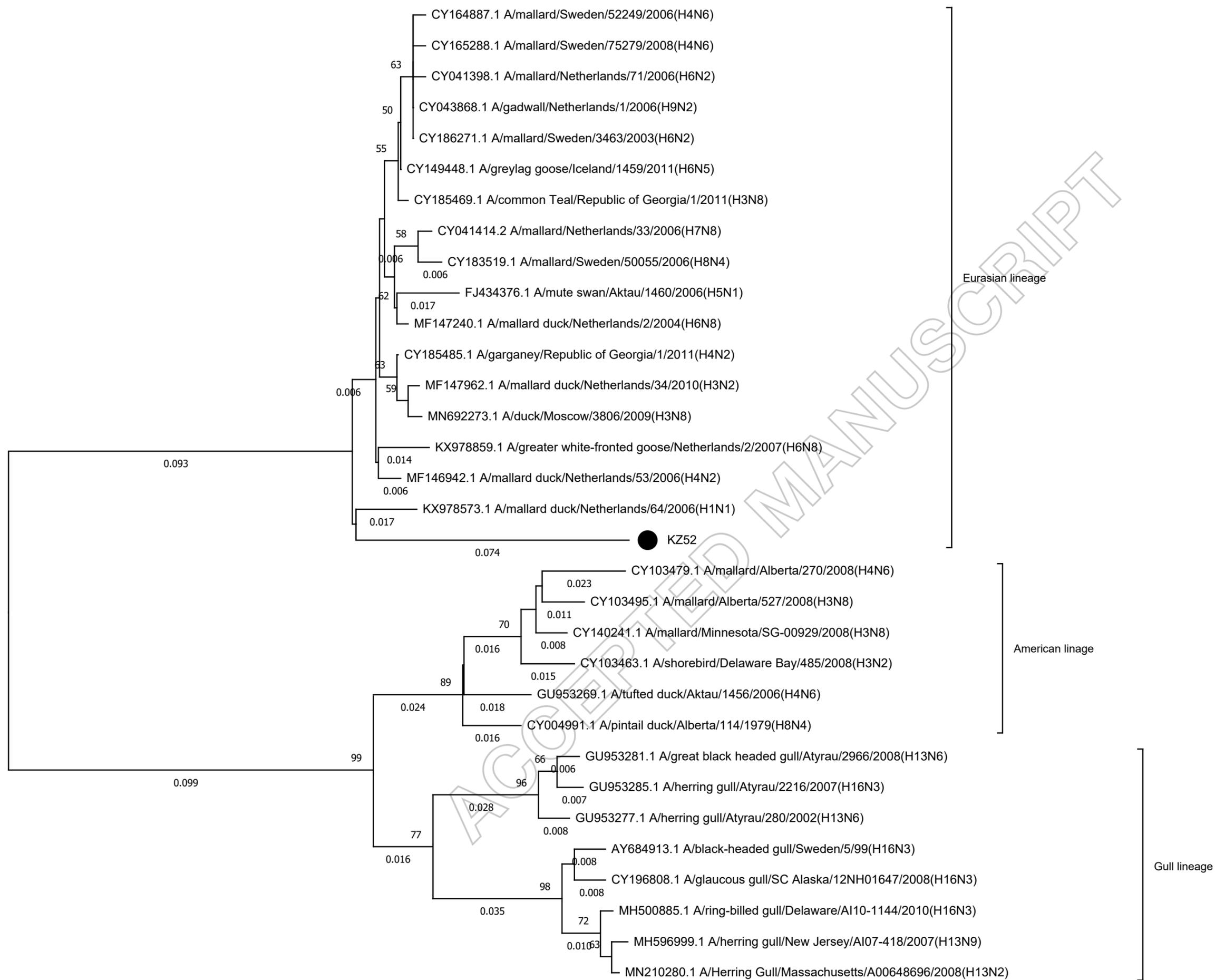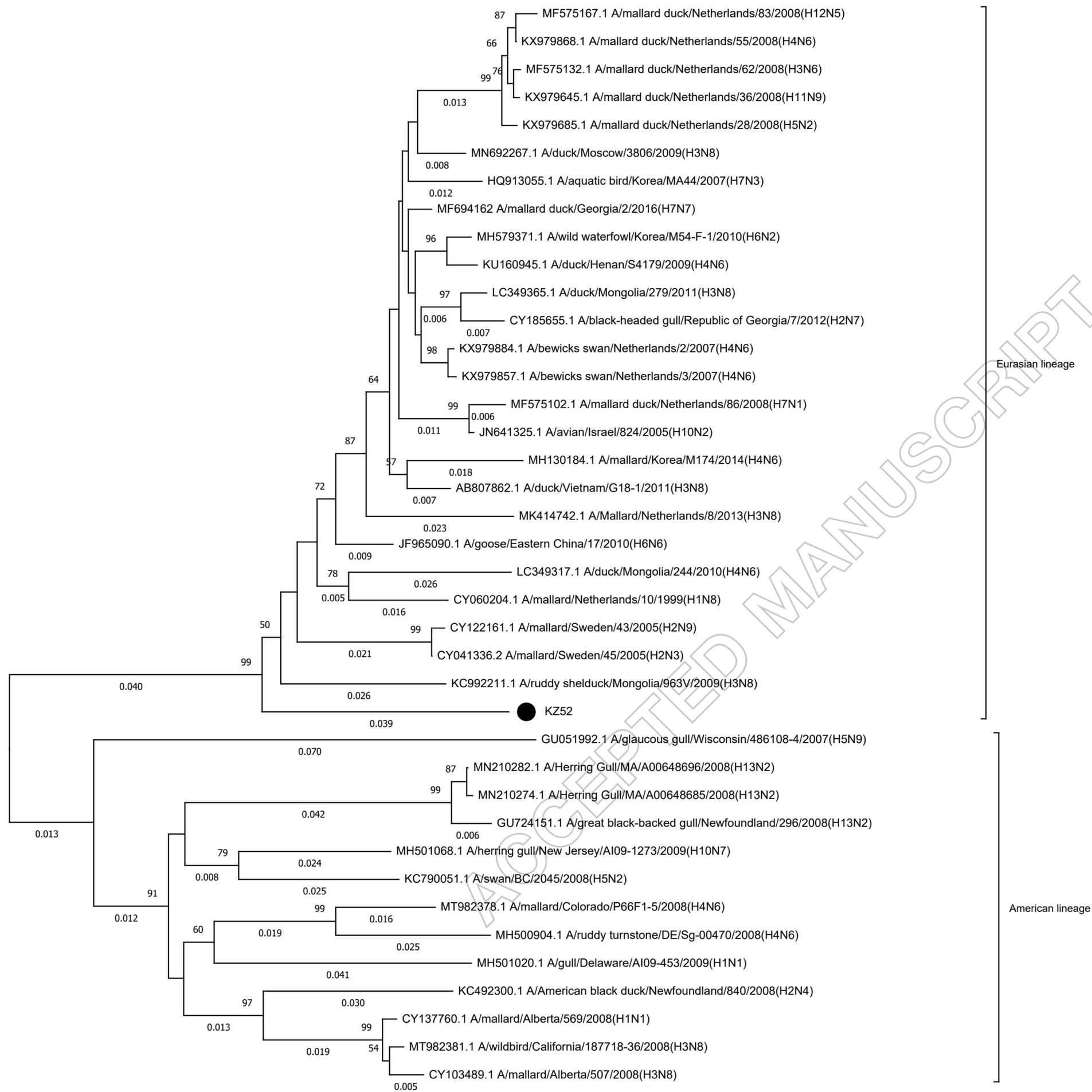

Figure 3. Schematics showing predicted 3D models of the Kz52 HA Ribbon (right) and surface (left) formats generated using Phyre2 software. The images show the HA1 and HA2 subunits (green and yellow respectively), the cleavage site (red) and receptor binding sites (blue).

CY164887.1 A/mallard/Sweden/52249/2006(H4N6)
CY165288.1 A/mallard/Sweden/75279/2008(H4N6)
CY041398.1 A/mallard/Netherlands/71/2006(H6N2)
CY043868.1 A/gadwall/Netherlands/1/2006(H9N2)
CY186271.1 A/mallard/Sweden/3463/2003(H6N2)
CY149448.1 A/greylag goose/Iceland/1459/2011(H6N5)
CY185469.1 A/common Teal/Republic of Georgia/1/2011(H3N8)
CY041414.2 A/mallard/Netherlands/33/2006(H7N8)
CY183519.1 A/mallard/Sweden/50055/2006(H8N4)
FJ434376.1 A/mute swan/Aktau/1460/2006(H5N1)
MF147240.1 A/mallard duck/Netherlands/2/2004(H6N8)
CY185485.1 A/garganey/Republic of Georgia/1/2011(H4N2)
MF147962.1 A/mallard duck/Netherlands/34/2010(H3N2)
MN692273.1 A/duck/Moscow/3806/2009(H3N8)
KX978859.1 A/greater white-fronted goose/Netherlands/2/2007(H6N8)
MF146942.1 A/mallard duck/Netherlands/53/2006(H4N2)
KX978573.1 A/mallard duck/Netherlands/64/2006(H1N1)
KZ52

Eurasian lineage

CY103479.1 A/mallard/Alberta/270/2008(H4N6)
CY103495.1 A/mallard/Alberta/527/2008(H3N8)
CY140241.1 A/mallard/Minnesota/SG-00929/2008(H3N8)
CY103463.1 A/shorebird/Delaware Bay/485/2008(H3N2)
GU953269.1 A/tufted duck/Aktau/1456/2006(H4N6)
CY004991.1 A/pintail duck/Alberta/114/1979(H8N4)

American linage

GU953281.1 A/great black headed gull/Atyrau/2966/2008(H13N6)
GU953285.1 A/herring gull/Atyrau/2216/2007(H16N3)
GU953277.1 A/herring gull/Atyrau/280/2002(H13N6)
AY684913.1 A/black-headed gull/Sweden/5/99(H16N3)
CY196808.1 A/glaucous gull/SC Alaska/12NH01647/2008(H16N3)
MH500885.1 A/ring-billed gull/Delaware/AI10-1144/2010(H16N3)
MH596999.1 A/herring gull/New Jersey/AI07-418/2007(H13N9)
MN210280.1 A/Herring Gull/Massachusetts/A00648696/2008(H13N2)

Gull lineage