

Evolutionary features of a prolific subtype of avian influenza A virus in European waterfowl

Michelle Wille,^{1,†,‡} Conny Tolf,¹ Neus Latorre-Margalef,¹ Ron A. M. Fouchier,² Rebecca A. Halpin,³ David E. Wentworth,³ Jayna Ragwani,⁴ Oliver G. Pybus,^{4,5,*} Björn Olsen,⁶ and Jonas Waldenström^{1,††}

¹Centre for Ecology and Evolution in Microbial Model Systems, Linnaeus University, Universitetsplatsen 1, Kalmar SE-39231, Sweden, ²Department of Virology, Erasmus Medical Centre, Rotterdam, The Netherlands, ³J. Craig Venter Institute, Rockville, MD, USA, ⁴Department of Zoology, University of Oxford, South Parks Road, Oxford OX1 3SY, UK, ⁵Department of Pathobiology and Population Sciences, The Royal Veterinary College, London AL9 7TA, UK and ⁶Section of Infectious Diseases, Department of Medical Sciences, Uppsala University, Uppsala SE751 85, Sweden

[†]Present address: Sydney Institute for Infectious Diseases, School of Life and Environmental Sciences and School of Medical Sciences, The University of Sydney, Sydney, New South Wales, Australia.

[‡]<https://orcid.org/0000-0002-5629-0196>

[§]<https://orcid.org/0000-0001-8538-4995>

^{||}<https://orcid.org/0000-0002-8797-2667>

^{††}<https://orcid.org/0000-0002-1152-4235>

*Corresponding authors: E-mail: michelle.wille@sydney.edu.au; jonas.waldenstrom@lnu.se

Abstract

Avian influenza A virus (AIV) is ubiquitous in waterfowl and is detected annually at high prevalence in waterfowl during the Northern Hemisphere autumn. Some AIV subtypes are globally common in waterfowl, such as H3N8, H4N6, and H6N2, and are detected in the same populations at a high frequency, annually. In order to investigate genetic features associated to the long-term maintenance of common subtypes in migratory ducks, we sequenced 248 H4 viruses isolated across 8 years (2002–9) from mallards (*Anas platyrhynchos*) sampled in southeast Sweden. Phylogenetic analyses showed that both H4 and N6 sequences fell into three distinct lineages, structured by year of isolation. Specifically, across the 8 years of the study, we observed lineage replacement, whereby a different HA lineage circulated in the population each year. Analysis of deduced amino acid sequences of the HA lineages illustrated key differences in regions of the globular head of hemagglutinin that overlap with established antigenic sites in homologous hemagglutinin H3, suggesting the possibility of antigenic differences among these HA lineages. Beyond HA, lineage replacement was common to all segments, such that novel genome constellations were detected across years. A dominant genome constellation would rapidly amplify in the duck population, followed by unlinking of gene segments as a result of reassortment within 2–3 weeks following introduction. These data help reveal the evolutionary dynamics exhibited by AIV on both annual and decadal scales in an important reservoir host.

Key words: avian influenza; evolution; influenza A virus; phylodynamics; reassortment.

Introduction

Influenza A viruses have a significant impact on human and animal health worldwide. These viruses cause seasonal outbreaks and pandemics in humans but also cause disease in domestic animals such as horses, pigs and poultry, and wild animals including birds and seals (Webster et al. 1992; Daly et al. 2011; Anthony et al. 2012; Vincent et al. 2014; Zohari et al. 2014). Central to the epidemiology of influenza A viruses are wild birds, particularly the *Anseriformes* (ducks, geese, and swans) and *Charadriiformes* (shorebirds and gulls) (Webster et al. 1992; Olsen et al. 2006). Within this wild bird reservoir, influenza A viruses have been isolated from more than 105 species across twenty-six different families. Wild birds maintain a large diversity of avian influenza A virus (AIV) subtypes; sixteen of the possible eighteen haemagglutinin (HA) subtypes and nine of the eleven neuraminidase (NA) subtypes are found in wild birds (Olsen et al. 2006; Olson et al. 2014).

Of importance in the context of emerging and re-emerging disease is that this large diversity of AIV in waterfowl frequently spillover into poultry (Verhagen et al. 2017), which may facilitate the introduction of AIVs into future human pandemic viruses (Smith et al. 2009; Wille and Holmes 2020).

In the Northern Hemisphere, waterfowl have a high AIV prevalence in the autumn, linked to the congregation of migrating birds, including large numbers of immunologically naïve juveniles (Latorre-Margalef et al. 2014; van Dijk et al. 2014). A large diversity of HA subtypes are maintained in waterfowl; however, certain HA–NA subtype combinations, such as H3N8, H4N6, and H6N2, are over-represented at waterfowl surveillance sites and may comprise the majority of viruses detected and isolated. Conversely, a number of HA–NA subtype combinations are uncommon or absent from these study sites (Munster and Fouchier 2009; Wilcox et al. 2011; Latorre-Margalef et al. 2014; Wille et al. 2018).

The consistent detection of diverse subtypes, combined with low rates of evolutionary change of amino acid sequences, led early studies to postulate that avian AIV was in evolutionary stasis (Webster et al. 1992; Sharp et al. 1997; Hatchette et al. 2004). However, high rates of molecular evolution ($\sim 10^{-3}$ substitutions per site per year) and inference of positive selection (dN/dS) in AIV genomes have firmly refuted this hypothesis (Chen and Holmes 2006).

In the last decade, a number of interconnected hypotheses of AIV evolution have been put forward based on observed AIV genetic structure. Central to these hypotheses are the evolutionary features of antigenic drift and antigenic shift. Antigenic drift is the fixation, by natural selection, of mutations in the HA and NA that enable the virus to evade the host immune response. These mutations arise through an error-prone RNA-dependant RNA polymerase (Gething et al. 1980; Chen and Holmes 2006). Antigenic shift or the process of reassortment following coinfection allows for the generation of novel genome constellations with altered antigenic properties (Steel and Lowen 2014; Lowen 2017). Together, these processes result in the capacity for rapid genetic and antigenic change, and while the same subtypes may appear annually in a population, these subtypes consist of multiple genetic lineages, which accumulate mutations, resulting in high levels of genetic diversity in the viral population over time.

In this study, we aimed to reveal the genetic and evolutionary features that allow for the long-term maintenance of AIV subtypes in waterfowl populations. We collected and sequenced 248 H4 isolates across 8 years from migratory mallards (*Anas platyrhynchos*) at a stopover site in southeastern Sweden. H4 is the most common AIV in not only European ducks (Munster et al. 2007; Latorre-Margalef et al. 2014) but also Asian wild and domestic ducks (Cheng et al. 2010; Wisedchanwet et al. 2011; Zhang, Chen, and Chen 2012; Kang et al. 2013). We used mallards as a model as they are considered to be an important reservoir for AIV diversity and have formed the basis of the current understanding of spatial and temporal avian AIV dynamics in nature (Olsen et al. 2006; Latorre-Margalef et al. 2014). Specifically, from these data, we aimed to reveal the role of different genetic lineages, examine how this genetic variation may relate to antigenic differences, and address the role of reassortment in both within-year and among-year maintenance of this virus subtype. We suggest that a complex combination of natural selection, genetic hitchhiking, and reassortment is responsible for the evolutionary patterns of AIV observed in wild birds.

Methods

Ethics statement

All trapping and handling of mallards were done in accordance with regulations provided by the Swedish Board of Agriculture under permits from the Linköping Animal Research Ethics Board (permit numbers 8-06, 34-06, 80-07, 111-11, and 112-11).

Study site and virus collection

Wild mallards were captured as part of an ongoing long-term AIV surveillance scheme (Latorre-Margalef et al. 2014) at Ottenby Bird Observatory, Sweden (56° 12'N, 16° 24'E). Details on trapping, AIV surveillance, and virus isolation have been published earlier (Latorre-Margalef et al. 2014, 2016). Briefly, for each captured duck, either a cloacal swab or a faecal sample was collected and placed in a virus transport medium (VTM). All samples were stored in an 80°C freezer within 2–6 h of collection. Viral RNA was extracted

from the VTM samples and assayed for a short fragment of the matrix gene using reverse-transcription real-time PCR (RT-PCR) (Spackman et al. 2002). Samples positive for AIV were inoculated into 10- to 12-day-old specific-pathogen-free embryonated chicken eggs by the allantoic route. HA subtypes of positive samples were determined using haemagglutination inhibition, and NA subtypes were determined by sequencing (Latorre-Margalef et al. 2014, 2016).

Sequence dataset

Full genomes were sequenced as part of the Influenza Genome Project (<http://gcid.jcvi.org/projects/gsc/influenza/index.php>), an initiative by the National Institute of Allergies and Infectious Diseases, as previously described (Wille et al. 2018). Briefly, AIV RNA was extracted and the entire genome was amplified using a multi-segment RT-PCR strategy (Zhou et al. 2009; Zhou and Wentworth 2012). The amplicons were sequenced using the Ion Torrent PGM (Thermo Fisher Scientific, Waltham, MA, USA) and/or the Illumina MiSeq v2 (Illumina, Inc., San Diego, CA, USA) instruments. When sequencing data from both platforms were available, the data were merged and assembled together; the resulting consensus sequences were supported by reads from both technologies.

We attempted sequencing 289 viruses, and 248 were successfully sequenced. Viruses that were not successfully sequenced did not pass QC at the sequencing facility, and we did not attempt resubmission of samples or resequencing.

All the sequences generated in this study have been deposited in GenBank (accession numbers CY164135–CY166103).

Phylogenetic analysis

Sequences generated in this study, in addition to all Eurasian reference sequences mined from the Influenza Research Database (<http://www.fludb.org/>), were aligned using MAFFT (Katoh, Asimenos, and Toh 2009). Maximum likelihood trees were used to explore the temporal signal and clock-like behaviour of each dataset by performing linear regressions of root-to-tip distances against the year of sampling, using TempEst (Rambaut et al. 2016). Using BEAST v10.4 or v1.8, time-stamped data were analysed under the uncorrelated lognormal relaxed molecular clock (Li and Drummond 2012), the SRD06 codon-structured nucleotide substitution model (Shapiro, Rambaut, and Drummond 2006), and the Bayesian skyline coalescent tree prior. Three independent analyses of 100 million states were performed, which were then combined in LogCombiner v1.8 following the removal of a burnin of 10 per cent. Convergence was assessed using Tracer v1.6 (<http://tree.bio.ed.ac.uk/software/tracer/>). Maximum credibility phylogenies were generated using TreeAnnotator v1.8 and visualised in FigTree v1.4.

The N2 phylogeny was estimated using MrBayes 3.2.1 (Ronquist et al. 2012). All trees were run until convergence, which was assessed using Tracer v1.4. The tree was visualised using FigTree v1.4.

Phylogenetic lineages were identified through a combination of pairwise identity plots (summarised in Fig. S1) and phylogenetic tree shapes. From these analyses, 95 per cent sequencing identity was established as a cut-off for genetic lineage delineation, similar to previous assessments (Reeves et al. 2011; Wille et al. 2013; Huang et al. 2014). Lineage proportion plots were plotted using the *ggplot2* library in R v. 3.5.1 integrated into RStudio v. 1.0.143.

Discrete trait analysis was performed using a symmetric trait evolution model, and social networks were inferred with Bayesian Stochastic Search Variable Selection (Lemey et al. 2009).

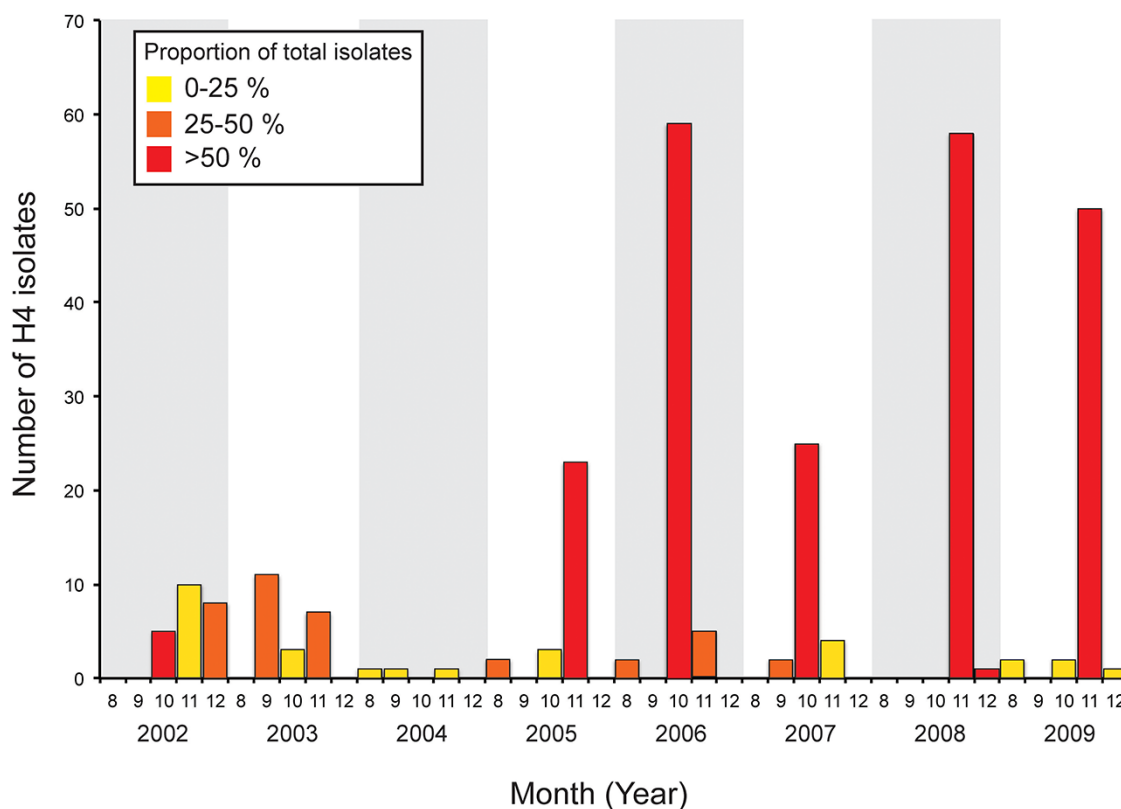


Figure 1. H4 AIVs are isolated at high frequency and constitute a substantial proportion of viruses from mallards sampled during the autumn migration at Ottenby Bird Observatory.

Connectivity among locations was determined using Bayes Factor analysis as implemented in Spread3 (Bielejec et al. 2016). We considered Bayes Factors of greater than 10 to be strong support and greater than 100 to be decisive support (Jeffreys 1961; Hill et al. 2021). Number of strains included in each location category is found in Table S1.

Initial computations were performed on resources provided by SNIC through the Uppsala Multidisciplinary Center for Advanced Computational Science (UPPMAX) under project b2013122.

Comparison of H4 HA1 peptide sequences of Swedish H4N6 virus

In order to evaluate possible antigenic properties governing between different lineages of H4, HA amino acid sequences of each lineage were aligned using MAFFT, and the consensus sequence for each lineage was determined using Geneious Prime 2020.2.4. Consensus sequences from different years were aligned to identify substitutions that had been introduced between years. These substitutions were mapped in the 3D structure of the HA1 monomer derived from the crystal structure of AIV H4 HA (PDB ID: 5XL1), using Geneious Prime 2020.2.4. In addition, structural regions in the H4 structure corresponding to established antigenic sites (A–E) in the closely related and well-studied H3 HA1 were indicated as reference (Broecker et al. 2018).

Results

H4 appears annually, at high frequency, in migratory mallards

As part of a long-term AIV surveillance scheme, 18,643 cloacal/faecal samples from migratory mallards were collected

between 2002 and 2009 at Ottenby Bird Observatory in southeast Sweden, as outlined in Latorre-Margalef et al. (2014). From these samples, 289 H4 viruses were successfully isolated, representing 27 per cent of all viruses isolated from mallards in the study period (Fig. 1).

For 2005–9, the years for which we have the most data, we find ‘outbreak-like’ patterns of H4 in the population, wherein during October or November, this subtype comprised >50 per cent of all isolated viruses (Fig. 1). In 2002 and 2003, years in which we have less data, we find a bimodal H4 prevalence peak, with high prevalence in September and November, respectively. No prevalence pattern could be observed in 2004 due to a limited number of isolates (Fig. 1). Moreover, for H4 viruses isolated from 2002 to 2009, all NA subtypes except N4 and N7 were detected. The majority (79 per cent), however, were H4N6 viruses (Fig. S2A).

Long-term phylodynamics of H4N6

A total of 246 H4 viruses isolated from 2002 to 2009 were successfully sequenced. Reconstruction of the H4 phylogeny of Eurasian sequences revealed two main lineages: one lineage comprising sequences from Asia and Australia (1980–present) (Fig. 2A, collapsed node) and a second lineage dominated by European sequences, most of which were generated in this study, in addition to a smaller lineage of Asian sequences (Fig. 2A). The estimated time of the most recent common ancestor (tMRCA) of all Eurasian sequences, including both the Asian/Australian and European lineages, is 1926–57 (95 per cent highest posterior density [HPD], mean 1946).

Given that two-thirds of H4 sequences from Eurasia between 2000 and 2009 are sequences generated in this study, the European lineage of the H4 phylogenetic tree is largely defined by the

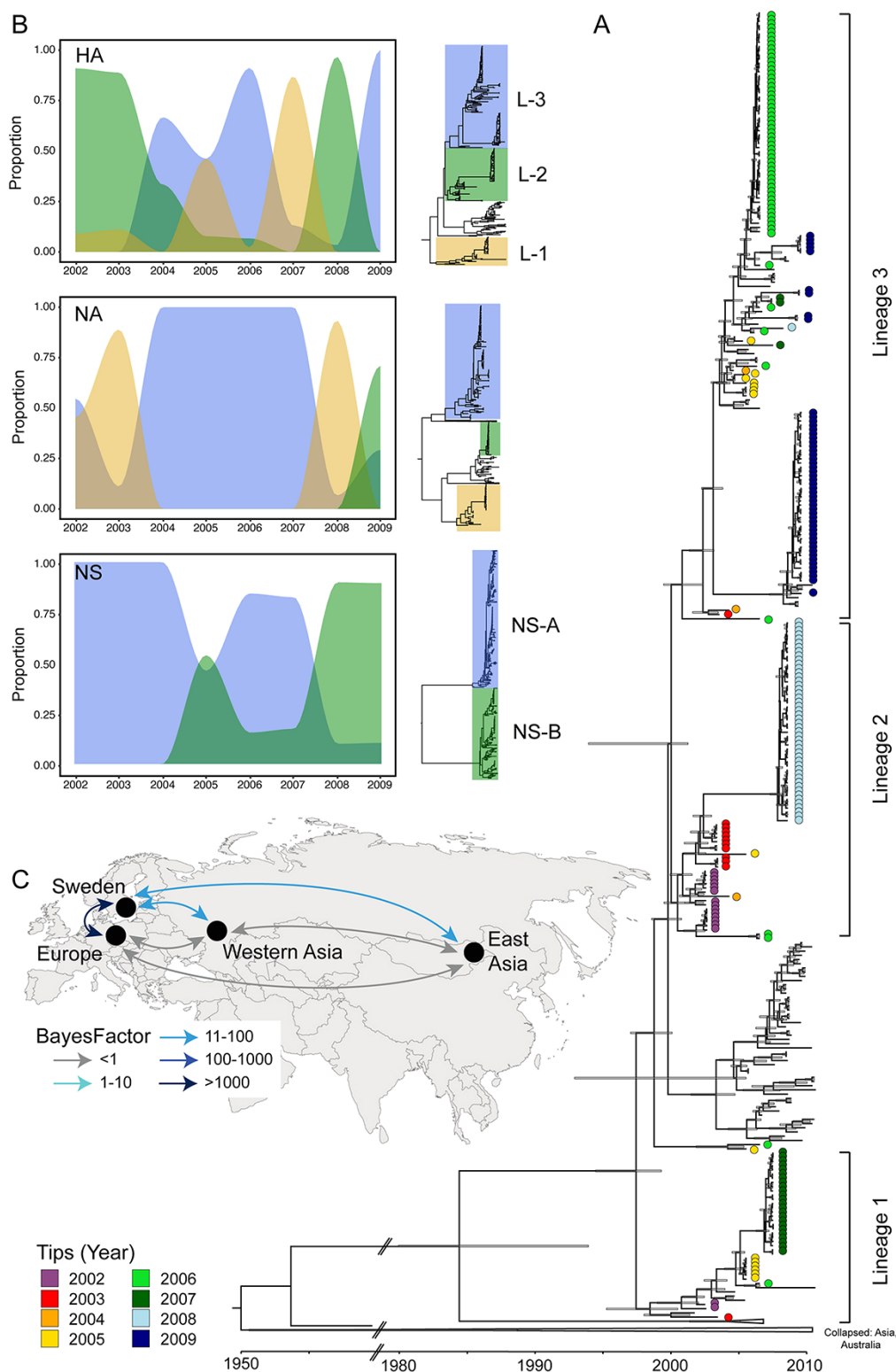


Figure 2. Features of the long-term evolution of H4. (A) Time-scaled phylogeny of the Eurasian lineage of H4. Phylogeny includes all H4 sequences in GenBank from 1950 to 2010. Sequences generated in this study are denoted by a filled circle, with different colours referring to the year of sample collection. Branch nodes correspond to the 95 per cent HPD of node height. (B) Proportion of lineages of each H4, N6, and NS detected in Sweden between 2002 and 2010. For each segment, the phylogeny with highlighted lineages is shown. For NS, the lineages correspond to the A and B alleles. In both (A) and (B), lineages circulating in Sweden are referred to as Lineages (L) 1–3 for clarity only. (C) Viral diffusion between geographic locations. Sequences included in Europe include the Netherlands, the Czech Republic, and UK; Western Asia includes western Russia and the Republic of Georgia; Eastern Asia includes China, Japan, Korea, Mongolia, and Russian Siberia. Results of a symmetrical model in which all sequences from Asia (western and eastern Asia) have been included into a single trait category are found in Fig. S3. Number of sequences in each location category is found in Table S1. Arrow colour refers to the strength of support for inclusion in a phylogeographic model (i.e., Bayes Factors), where values greater than 10 represent strong support, and values greater than 100 are considered decisive support.

sequences from our study site. In addition to these, there are a smaller number of sequences from the Netherlands and the Czech Republic in GenBank. All H4 sequences from Sweden are within 95 per cent sequence identity (Fig. S1) and the European lineages comprising the sequences generated in this study diverged in 1997.7 (tMRCA, 95 per cent HPD 1994–9) (Fig. 2A). We find the sequences generated in this study to fall into three lineages with the exception of four sequences (including two sequences from 2010) that group closely with Asian sequences nested within this European lineage. We refer to those lineages detected in Swedish mallards as Lineages 1–3 for clarity only (Fig. 2A).

There are three general observations regarding the distribution of sequences across the three lineages. First, sequences from 2002 and 2003 fall into the same lineage (Lineage 2). Second, unlike all other years in this dataset, sequences from 2005 fall into two distinct lineages with equal proportion: the lineage comprising sequences from 2006 (Lineage 3) and that of 2007 (Lineage 1). Finally, starting in 2006 (for which we have the most data), sequences from within each year fall into a single lineage with ~99 per cent nucleotide similarity, and these lineages are distinct from the dominating lineage in the prior and subsequent years. Indeed, the lineages containing sequences in 2006 and 2009 (Lineage 3) do not contain more than three sequences from intermediate years. Taken together, we find that lineage replacement in H4 is high between years (Fig. 2B).

To attempt to better elucidate the source of annual viral introduction into Sweden, we undertook a phylogeographic analysis. Based on the available sequences from 1980 to 2010, there is strong evidence for virus movement between Sweden and other European locations (Bayes factor, 30,532.9; posterior probability, 1), between Sweden and western Asia (here western Russia [Moscow] and the Republic of Georgia) (Bayes factor, 59; posterior probability, 0.98), and between Sweden and eastern parts of Asia (Bayes factor, 52.4; posterior probability, 0.97). In this analysis, virus movement outside Sweden was poorly supported (Bayes factor, <10; posterior probability, <0.5), most likely because of the low number of sequences from other regions (Fig. 2C). Results of a symmetrical model in which all sequences from Asia (western and eastern Asia) have been included into a single trait category are shown in Fig. S3. Overall, due to poor sequence coverage from Eurasian locations outside of Sweden, it is challenging to infer the role of different geographic locations in the introduction of viruses into Swedish-caught mallards.

Similar to the H4 tree, sequences from this study play an important role in defining the shape of the Eurasian N6 phylogenetic tree (Fig. S2). The N6 lineages are more distantly related, as compared to H4 lineages, where the three distinct N6 lineages detected in Ottenby are ~93 per cent identical and postulated to have evolved from their MRCA ranging from 1977 to 1987 (95 per cent HPD). The second most common NA subtype identified was N2, and as with N6, Swedish N2 sequences were similar to those circulating in Europe, most often Sweden, during the same time period. N2 sequences from 2005 and 2008 formed discrete lineages sharing 99 per cent sequence identity, whereas sequence variation was greater in 2002 and 2007 (Fig. S4). In addition to N6 and N2, the dataset included N1 ($n=1$), N3 ($n=3$), N5 ($n=1$), and N9 ($n=3$). In all instances, these sequences were most closely related to other European AIV sequences, and the top 10 Blast matches often included other sequences from Sweden (Table S2).

Globally, non-structural protein (NS) sequences fall into two major lineages, termed allele A (including viruses with broad host range) and allele B (viruses specific to birds). NS sequences of viruses sequenced in this study fell into both alleles A and B and

were most similar to NS sequences of AIV from Eurasian wild birds (Fig. S5). Starting in 2015, both alleles were found in the population each year, although with differing proportions. The shapes of the phylogenetic trees for the 'internal' segments (PB2, PB1, PA, NP, and M) were similar to those previously characterised (Fig. S6). Within the 'internal' segments, we found a pattern similar to that of HA and NA, wherein sequences from 2006 to 2009 fell into discrete year-specific lineages.

Amino acid sequence variation in the HA of Swedish H4N6 virus

Amino acid sequence variation predominately occurred in the HA1, and more specifically in, or in close proximity to, regions corresponding to antigenic sites identified in the H3 HA1 (Broecker et al. 2018). While antigenic sites for H4 have not been mapped, H3 and H4 belong to the same lineage within group 2 HA subtypes (Latorre-Margalef et al. 2013), and H3 therefore constitutes the best available proxy for initial evaluation of the possible effect of identified substitutions. However, in order to conclusively determine whether identified substitutions in H4 result in antigenic changes, detailed experiments, corresponding to those made for H3, are needed.

When assessing the amino acid changes between consecutive years, we identified between 6 and 9 amino acid substitutions in the HA1 region. Specifically, between 2006 and 2007 (Lineage 3 → Lineage 1, Fig. 3A), we found seven substitutions; between 2007 and 2008 (Lineage 1 → Lineage 2, Fig. 3B), we found nine substitutions; and between 2008 and 2009 (Lineage 2 → Lineage 3, Fig. 3C), we found six substitutions. When comparing the sequences from 2006 compared to 2007, amino acid substitutions were located around the top of the globular head of HA1, close to secondary structures likely forming the receptor-binding site, including the 130 loop, the 190 helix, and the 220 loop (Fig. 3A). It is unclear whether these sites are important for H4, but antigenic information from H3 provides us with a useful framework, while recognising that confirmation work is required. Based on scores given by the BLOSUM62 substitution matrix (positive and negative values are considered favourable and unfavourable, respectively, from a functional standpoint for a given protein), the majority of these substitutions are functionally favourable (Table S3), with the exception of K186T and T208I. In the transition from 2007 to 2008, four out of nine substitutions are found in loop structures closer to the stem region of HA1, including the 260 loop. These include Asn88Ser and Ser259Asn in a region corresponding to antigenic site E in H3 as well as I270T, K283R, and I295V adjacent to site C (Fig. 3B). All substitutions except I208T and I270T are functionally favourable. It is noteworthy that four substitutions in positions 124, 204, 208, 259, and 270 are reversions to amino acid residues expressed in the 2006 consensus sequence. Finally, in the transition between 2008 and 2009, the only unique substitution is Q42H, which is neutral from a functional standpoint. In addition, five of six substitutions found in the consensus sequence from 2009 compared to that from 2008 represent reversions to amino acids expressed in the 2007 sequence. These include amino acid at positions 88, 166, 259, 283, and 295 (Fig. 3C).

Lineage replacement and no perpetuation of genome constellations across years

Analysis of temporal changes of H4N6 genome constellations revealed frequent lineage replacement across years (Fig. 2B). Plots of lineage proportions of the glycoproteins and the NS alleles

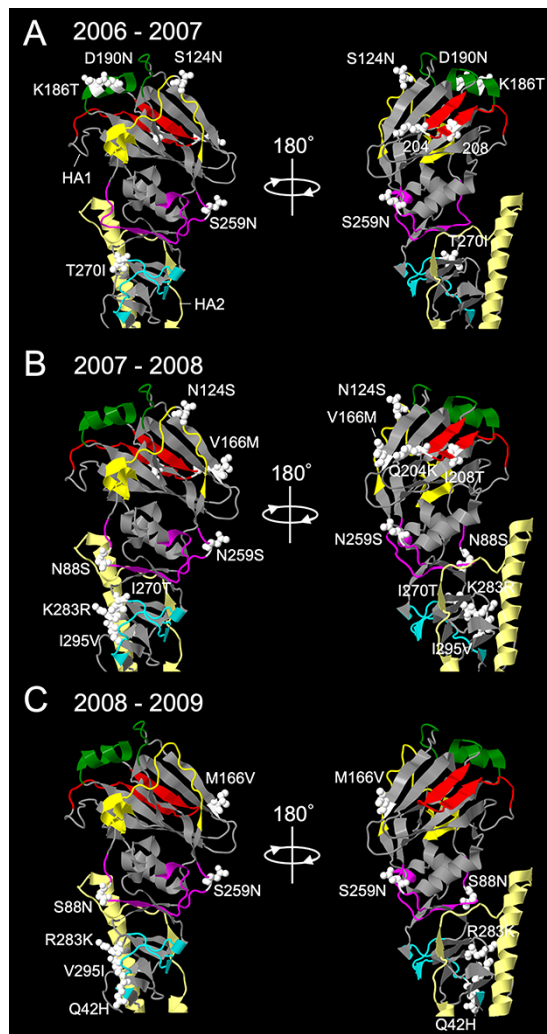


Figure 3. Amno acid substitutions across years between distinct H4 NA1 lineages of Swedish H4N6 strains. Specifically, amino acid substitutions between sequences from (A) 2006–7, (B) 2007–8, and (C) 2008–9. HA1 (grey) and HA2 (light yellow) peptide monomers are indicated in the left structure of the 2006–7 panel. Structural regions corresponding to antigenic site A (yellow), B (green), C (cyan), D (red), and E (magenta) in the HA1 peptide of H3 are indicated as a reference (Broecker et al. 2018).

illustrate that, while lineage replacement is frequent for HA, concomitant replacement among multiple segments occurred less frequently (Fig. 2B). For example, between 2003–4 and 2008–9, there was a replacement of both HA and NA lineages, and between 2007 and 2008, there was a simultaneous replacement of HA, NA, and NS (Fig. 2B). The replacement between 2007 and 2008 was not limited to only segments with established antigenic properties (HA and NA), but rather new lineages were introduced for all segments in addition to the replacement from NS allele A to allele B (Figs 2B, S3, S5, and S6). This global replacement also resulted in a reintroduction of lineages in 2008 that were circulating among sampled birds 6 years earlier, in 2003.

Bringing all lineage data together, we may assess whether genome constellations remain in the population across multiple years (Figs 4 and S7). Importantly, despite not detecting annual global replacements, there was sufficient lineage replacement across the different segments for novel constellations to proliferate each year. That is, genome constellations identified in 1 year were not found again in the population the following

year (Figs 4 and S7). Using 2005–6 as an example, no constellations identified in 2005 were detected again in 2006, despite the fact that some H4 sequences from 2005 were found in the same phylogenetic lineage as sequences from 2006. In this case, the lineages of four gene segments were different between the dominant genome constellation in 2005 compared to the dominant constellation from 2006 (Fig. 4). This example is not an outlier, rather is consistent with findings between all years of this dataset (Fig. S7).

Rapid unlinking of genome constellations within a year

Within each year, particularly between 2006 and 2009, we identified a number of different genome constellations. Despite the diversity, more than 50 per cent of viruses characterised within each year comprised a single, dominant constellation (Table S4). Data from 2006 to 2009 further indicate that the first genome constellations detected each year were not the most frequently observed constellations overall that year. Following the introduction of the dominant constellation into the population, it was consistently detected and isolated from mallard samples for 1–2 weeks, followed by a period (1–2 weeks) wherein this constellation rapidly unlinked and a diversity of new genome constellations were detected in the population. These novel and diverse constellations most likely contained gene lineages originating from other co-circulating viruses with different HA–NA subtype combinations and formed through the process of reassortment (Figs 4, S7, and S8).

Using 2006 as an example, the first virus constellations were detected in Week 23, followed by a detection in Week 32. These constellations were not detected again and were replaced by a frequently detected constellation that accounted for 73 per cent of isolates (Fig. 4A) in Week 41. This constellation initially appeared at low numbers in Week 41 ($n = 2$ sequenced isolates); however, it was the only H4 genome constellation detected in Week 42 ($n = 18$ sequenced H4 isolates) and remained at a high proportion relative to other constellations in Week 43 ($n = 13$ sequenced H4 isolates). Seven other genome constellations were also identified in Week 43, and one additional novel constellation was detected in Week 44 (Fig. 4). Data from 2005 revealed a genetic variation that was different compared to other years, wherein isolates comprised high levels of constellation diversity, presumably due to the presence of two different HA lineages and NS alleles (Fig. S7). In 2006, we detected fifteen ‘mixed’ viruses—these are viruses containing more than one sequence for any one segment. These mixed infections are likely reflective of coinfection in the mallard host, but as we sequenced viruses propagated in eggs, we cannot make strong inferences about these viruses (Fig. 4 and Table S5). These mixed viruses spanned the periods of highest prevalence and HA/NA diversity during the autumn. The year 2006 accounted for a large proportion of mixed viruses in our dataset, and this year, there was substantial co-circulation of other viral subtypes in the Ottenby duck population (e.g. H1N1, H3N6, and H6N2; Latorre-Margalef et al. 2014) (Fig. 4 and Table S5).

Discussion

How are subtypes maintained in waterfowl on a decadal scale?

Low pathogenic AIV H4N6 is one of the most abundant virus subtypes isolated from waterfowl in Europe. Indeed, these viruses have been circulating at high frequency in migratory mallards at our study site, in southeast Sweden, since the start of AIV

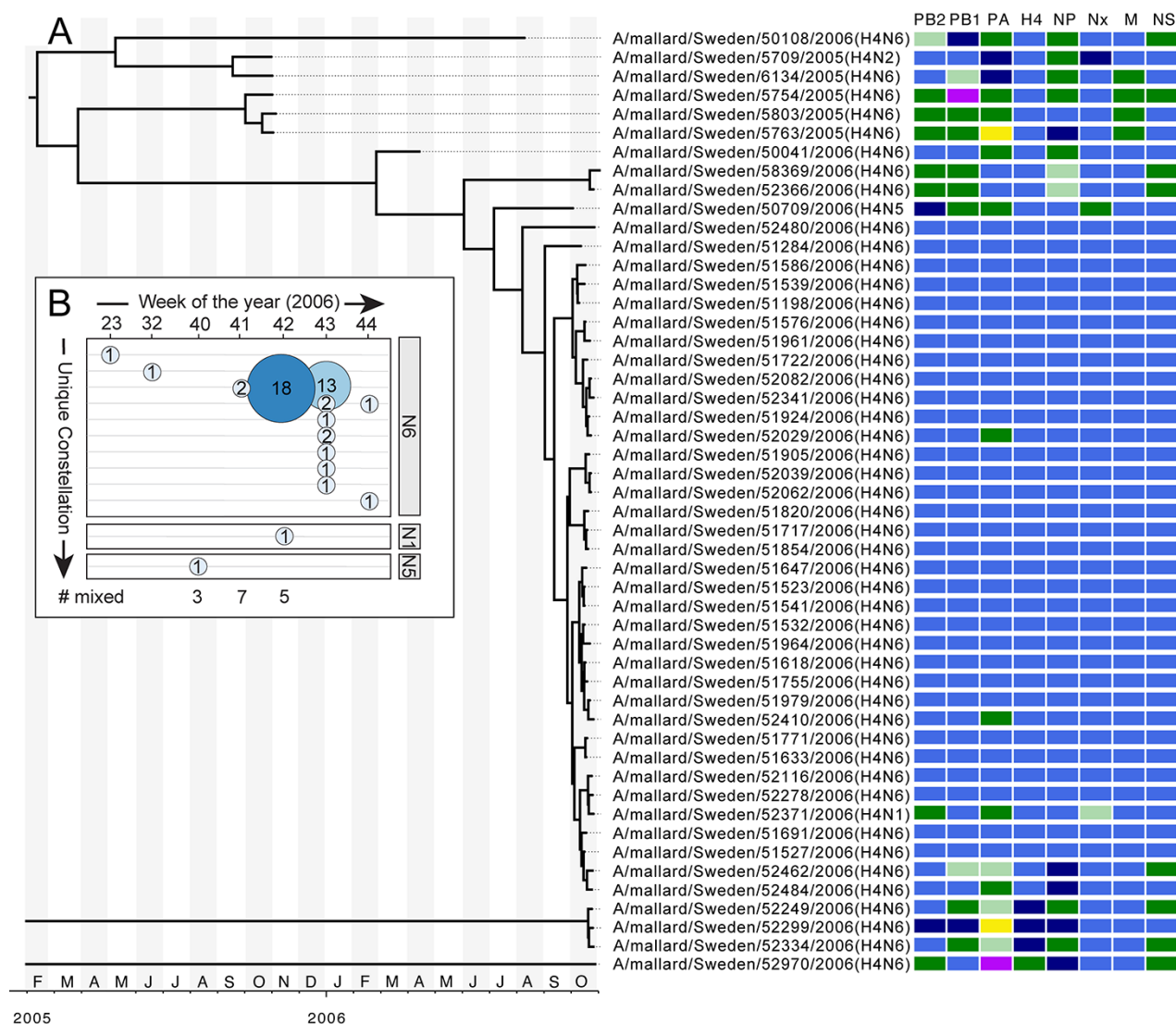


Figure 4. Within-year reassortment. (A) Time-structured phylogeny of H4 sequences from 2005 and 2006 comprising Lineage 1. Scale bar indicates time, with year and month indicated. Four H4 sequences from viruses detected in 2006 did not fall into this lineage, and therefore, tMRCAs are not shown in order to preserve resolution. Each tile column represents a segment, ordered by size: PB2, PB1, PA, H4, NP, Nx, M, and NS, and each row comprises the genome constellation of a virus. Different colours pertain to different viral lineages. The tile colours do not correspond to Fig. 2 or S7, rather were selected for aesthetic purposes. Genome constellations for all viruses are presented in Fig. S7. (B) Genome constellation for each week of 2006. Week number is presented on the X axis, and each unique constellation is on the Y axis. Bubble size and colour refer to the number of times each genome constellation was detected. Different NA subtypes are in separate panels. At the bottom, the number of mixed H4 infections detected each week is indicated. Change of genome constellations over time for all years is presented in Fig. S8.

surveillance in 2002 (Munster et al. 2007; Wallensten et al. 2007; Latorre-Margalef et al. 2014), and continue to circulate at high frequency today (Wille et al. 2015; Verhagen et al. 2017; Venkatesh et al. 2018; Bergervoet et al. 2019). An important question, however, is how this subtype is maintained in a mallard population across many years, given that high rates of infection should confer immunity in the population. Through the analysis of a dataset comprising viruses from the same species and location, but across multiple years, we were able to reveal key genetic features likely playing a key role in this phenomenon.

Unlike human influenza A virus, AIV HA segments do not exhibit a classic 'ladder-shaped' phylogeny (Fitch et al. 1997). These ladder-shaped phylogenies in human influenza A viruses are the result of strong positive selection on the HA, leading to antigenic drift, or the fixation of mutations in the HA (and NA) that enables the virus to evade the immune system (Koel et al. 2013;

Vijaykrishna et al. 2015; Wille and Holmes 2020). Rather, multiple lineages co-circulate in wild bird populations (Chen and Holmes 2006, 2010). Herein, we detect three lineages of H4 in the mallard population. These lineages are genetically similar (within 95 per cent nucleotide similarity), although none are present in all years. We find a phenomenon whereby there is a different HA lineage circulating each year, and within years 2006–9, one lineage comprises ~90 per cent of all isolates sequenced. Indeed, this is reflective of antagonistic co-evolution in a sympatric model of circulating AIV, wherein host immune pressure will drive phylogenetic branching of viruses into discrete antigenic lineages and eventually, subtypes, with limited overlap in antigenic space (Recker et al. 2007). Alternatively, or in concert, these patterns could rather be driven by migration-mediated metapopulation dynamics, which frequently re-introduces lineages to the site.

Our data is consistent with the hypothesis proposed by [Chen and Holmes \(2010\)](#), whereby the genetic structure of AIV is shaped by a combination of occasional selective sweeps in the HA and NA segments, coupled with transient genetic linkage to the internal gene segments. Indeed, we identified annual replacement of HA lineages, although simultaneous replacement of HA, NA, and NS lineages was less frequent. While NS is less studied as a target for cellular and humoral immune responses, there are several reports on the inhibitory effect of NS1 on the innate immunity and fitness differences of the different alleles ([Gack et al. 2009](#); [Rajsbaum et al. 2012](#); [Adams et al. 2013](#); [Yodsheewan et al. 2013](#); [Nogales et al. 2018](#)). In 2008, we detected lineage replacement in all segments, which coincided with the shift of the NS A to B allele in the virus population. This event does support the transient linkage hypothesis, where a change from NS A to NS B may have conferred some type of fitness advantage, promoting the genome constellation with the new NS allele to become fixed and resulting in a local sweep. Despite lineage replacement, previously circulating lineages do not go extinct, but rather have a marked decrease in frequency in the system, which may be explained by negative frequency-dependent selection ([Gandon and Michalakis 2002](#)). That is, the fitness of a phenotype or genotype decreases as it becomes more common, likely due to increased immunity against these lineages in the population. In addition to competitive interactions between lineages, the migratory behaviour of avian hosts and their distinct geographic distribution may contribute to lineage maintenance in different host populations and to the patterns of replacement between lineages in one site. Given these lineages do not go extinct, it is likely they circulate in other dabbling duck populations, including mallard populations, in other geographic locations, including Asia, or in other less characterised host species, such as diving ducks, geese, or even shorebirds, which are not as intensively monitored in surveillance programmes.

Unlike in humans, reassortment of AIV (both inter and intra-subtypic reassortment) is prolific in avian systems ([Macken, Webby, and Bruno 2006](#); [Dugan et al. 2008](#); [Wille et al. 2013](#); [Steel and Lowen 2014](#)) and is likely critical in both the long-term (decadal scale) and short-term (within year) arms race of AIV against mallard immunity. Specifically, it is through reassortment that novel HA or NA lineages are introduced and also the mechanism by which entire genome constellations are formed within the population between years. Indeed, it is the generation of novel genome constellations that drives the emergence of pandemic influenza viruses in human populations ([Smith et al. 2009](#); [Wille and Holmes 2020](#)). In our study, the role of reassortment is most evident when assessing H4N6 genome constellations within a year. Specifically, during periods of high infection burden of H4 viruses in the mallard population, we may expect that the dominant genome constellation has high fitness, allowing for rapid proliferation in the mallard population. However, within ~2 weeks of entering/detection in the mallard population, the initially dominant H4 constellation rapidly unlinks through reassortment and is replaced by a variety of auxiliary genome constellations. We hypothesise that this is driven by increasing immunity in the host population, resulting in a decrease in fitness for this common constellation. While most antibodies are directed at the HA, and to a lesser extent the NA, during infection all influenza proteins are expressed in infected cells and can potentially induce an antibody response ([Krammer 2019](#)). However, some proteins are more accessible than others, including the NP and M segments ([Haaheim 1977](#); [Sukeno et al. 1979](#)). There is some evidence from monoclonal antibody isolation and from

antigenic fingerprinting that natural infection also induces antibodies against internal segment proteins, although the magnitude and quality of these responses are not well defined ([Thathaisong et al. 2008](#); [Krejnosova et al. 2009](#); [Yodsheewan et al. 2013](#); [Krammer 2019](#)). As such, it is likely that through the generation of novel genome constellations, the H4 subtype may be maintained in the population beyond the point at which there is an increase in HA-directed immunity and a subsequent decrease in viral fitness. Novel segments incorporated into novel genome constellations originate from other viruses co-circulating in the mallard population ([Wille et al. 2013](#)). Indeed, in years with the highest number of mixed infections, a number of other subtypes were co-circulating at a relatively high frequency: H1N1 in 2006 and H11N2 in 2008.

Do genetic patterns reflect antigenic patterns?

Both the phenomenon of lineage replacement and reassortment are important drivers in this system and would only be selected if they provided a fitness advantage. One such advantage would be escape from acquired population immunity. Following infection, mallards develop an immune response that is homosubtypic, with potential for heterosubtypic effects ([Latorre-Margalef et al. 2013, 2017](#)). Although the strength and duration of the immune response are largely unknown in ducks, it is hypothesised that immune memory may be weak and short-lived ([Magor 2011](#)). However, repeated sampling of individual sentinel mallards illustrates the maintenance of anti-NP antibodies for months following AIV infection. Furthermore, these same mallards were unlikely to be reinfected with the same subtypes in their second year ([Tolf et al. 2013](#)). Indeed, it is hypothesised that this pattern of both homo- and heterosubtypic immunity drives the temporal succession of different HA lineages across the sampling season ([Latorre-Margalef et al. 2014](#)).

Based on the phenomenon of lineage replacement or succession of H4 between years, we hypothesise that the different HA lineages observed must confer homosubtypic immune escape. Molecular and structural comparisons of sequences in each of the three HA lineages found in the mallard population had substitutions in regions of HA1, corresponding to antigenic sites in the homologous H3 peptide ([Broecker et al. 2018](#)). Unfortunately, dedicated work on identifying domains has not been completed for H4; however, H3 and H4 comprise the H3 Clade of Group 2 viruses ([Latorre-Margalef et al. 2013](#)), such that we predict the antigenic sites of H3 should correspond to H4 as well. We found that the six to nine amino acid differences delineating the H4 lineages described in this study partly map to different regions in the crystal structure of H4 HA1, where substitutions in the 2007 sequence were located at the top of the globular head and substitutions in 2008 and 2009 sequences were located closer to the stem region of HA1. This may suggest that a few changes in structural regions in putative antigenic sites are sufficient to evade immunity. Indeed, from work on human influenza A virus H3, it is clear that only a few substitutions are required for immune escape in humans, given that they are located in or adjacent to the receptor-binding domain ([Smith et al. 2004](#); [Koel et al. 2013](#)). In contrast to the HA of human influenza A (e.g. [Koel et al. 2013](#)), the antigenic diversity of avian HAs has not been well explored, with a few exceptions (e.g. [Koel et al. 2014](#); [Hill et al. 2016](#); [Verhagen et al. 2020](#)). However, a recent study investigating infection probability given previous exposure illustrated H3 vaccine escape despite >95 per cent sequence differences in ducks ([Wille et al. 2016](#)). In this study, ducks were vaccinated with an inactivated H3 virus,

isolated in 2010, and were challenged with naturally circulating AIVs in 2013. Despite a raised immune response due to the vaccination, ducks were not protected against H3 viruses, and it was hypothesized that the lack of immunity was associated to substitutions in immunogenic epitopes of more recent H3 strains (e.g. vaccine escape) (Wille et al. 2016). Despite the parallels, without performing a characterisation of H4 antigenic sites or through antigenic cartography, we are unable to confirm whether the different lineages (L1–L3) of H4 confer antigenic differences and/or whether reassortment confers partial escape from population immunity.

Understanding the evolutionary processes of AIV is critical. We generally have a limited understanding of the evolution of LPAIV in the wild bird reservoir (Chen and Holmes 2006, 2010; Dugan et al. 2008; Wille et al. 2013). But genetic diversity is shaped by virus-related factors, including antigenic drift and reassortment, as well as immunological and ecological features of the hosts, such as migratory behaviour (van Dijk et al. 2018). Low pathogenicity viruses circulating in wild birds play an important additional role by providing an increased gene pool, which can be transferred to other influenza A viruses. Within the avian reservoir, there is a key focus on HA subtypes that cause significant morbidity and mortality in food production birds (e.g. H5 and H7) or have zoonotic potential (e.g. H5, H7, and H9) (Chang et al. 2018, 2020; Mahmoud et al. 2019; Cui et al. 2020). The results of this study have important implications for our understanding of not only LPAIV but also HPAIV. Indeed, HPAIV H5Nx outbreaks are an interesting parallel. In 2014, a novel HPAIV H5 lineage (Goose/Guangdong Lineage 2.3.4.4) emerged and has subsequently caused substantial and repeated outbreaks in wild birds and poultry in Eurasia. Most likely, these viruses were maintained in Eurasian, wild bird populations, through the mechanisms we describe here: evolution including HA lineage replacements (e.g. a, b, and c lineages of 2.3.4.4) and reassortment (Global Consortium for H5N8 and Related Influenza Viruses 2016; Poen et al. 2018, 2019). In humans, the detailed interrogation of seasonal influenza viruses is key to vaccine selection and detection of antiviral resistance (Koel et al. 2013; Vijaykrishna et al. 2015; Van Poelvoorde et al. 2020; Wille and Holmes 2020). On a larger scale, it is through the process of reassortment that novel pandemic viruses are generated (Gething et al. 1980; Smith et al. 2009). For example, all human pandemic influenza A viruses have had at least one gene of avian origin (Scholtissek et al. 1978; Kawaoka, Krauss, and Webster 1989; Lindstrom, Cox, and Klimov 2004; Taubenberger et al. 2005; Rabadan, Levine, and Robins 2006; Runstadler et al. 2013), and HPAIV emerges following rapid evolution of LPAIV in poultry (Seekings et al. 2018). Thus, in order to understand evolutionary factors governing AIV dynamics, including host ecology affects, long-term evolution, and genetic structure, it is pivotal to characterise viruses circulating in the wild bird reservoir.

Data availability

All sequences generated in this study have been deposited in GenBank (accession numbers CY164135–CY166103).

Supplementary data

Supplementary data are available at *Virus Evolution* online.

Acknowledgements

We thank all the staff at Ottenby Bird Observatory for trapping and sampling mallards, and all laboratory staff for processing samples over the years.

This is manuscript 325 from Ottenby Bird Observatory.

Funding

This study was supported by grants from the Swedish Environmental Protection Agency (V-124-01 and V-98-04), the Swedish Research Council VR (2008-58, 2010-3067, and 2011-48), the Swedish Research Council Formas (2007-297 and 2009-1220), and the Sparbanksstiftelsen Kronan. The surveillance at Ottenby was part of the European Union wild bird surveillance and has received support from the Swedish Board of Agriculture and from the EU FP6-funded New Flubird project. Sequencing has been funded in whole or part with federal funds from the National Institute of Allergy and Infectious Diseases, National Institutes of Health, Department of Health and Human Services under Contract Number HHSN272200900007C. M.W. was supported by a PGS-D2 from the Natural Sciences and Engineering Research Council (2011–13) and is currently supported by an Australian Research Council Discovery Early Career Research Award (DE200100977). O.G.P. and J.R. were supported by the GCRF One Health Poultry Hub (Grant No. BB/S011269/1). This manuscript was prepared while DEW was employed at the J. Craig Venter Institute. The opinions expressed in this article are the author's own and do not reflect the view of the Centers for Disease Control, the Department of Health and Human Services, or the US government.

Conflict of interest: None declared.

References

- Adams, S. et al. (2013) 'The Effect of Avian Influenza Virus NS1 Allele on Virus Replication and Innate Gene Expression in Avian Cells', *Molecular Immunology*, 56: 358–68.
- Anthony, S. J. et al. (2012) 'Emergence of Fatal Avian Influenza in New England Harbor Seals', *MBio*, 3: e00166–00112.
- Bergervoet, S. A. et al. (2019) 'Circulation of Low Pathogenic Avian Influenza (LPAI) Viruses in Wild Birds and Poultry in the Netherlands, 2006–2016', *Scientific Reports*, 9: 13681.
- Bielejec, F. et al. (2016) 'Spread3: Interactive Visualization of Spatiotemporal History and Trait Evolutionary Processes', *Molecular Biology and Evolution*, 33: 2167–9.
- Broecker, F. et al. (2018) 'Immunodominance of Antigenic Site B in the Hemagglutinin of the Current H3N2 Influenza Virus in Humans and Mice', *Journal of Virology*, 92: e01100–01118.
- Chang, P. et al. (2020) 'Immune Escape Adaptive Mutations in the H7N9 Avian Influenza Hemagglutinin Protein Increase Virus Replication Fitness and Decrease Pandemic Potential', *Journal of Virology*, 94: e00216–00220.
- et al. (2018) 'Amino Acid Residue 217 in the Hemagglutinin Glycoprotein Is a Key Mediator of Avian Influenza H7N9 Virus Antigenicity', *Journal of Virology*, 93: e01627–01618.
- Chen, R., and Holmes, E. C. (2006) 'Avian Influenza Virus Exhibits Rapid Evolutionary Dynamics', *Molecular Biology and Evolution*, 23: 2336–41.
- (2010) 'Hitchhiking and the Population Genetic Structure of Avian Influenza Virus', *Journal of Molecular Evolution*, 70: 98–105.

- Cheng, M. C. et al. (2010) 'Avian Influenza Monitoring in Migrating Birds in Taiwan during 1998-2007', *Avian Diseases*, 54: 109–14.
- Cui, Y. et al. (2020) 'Evolution and Extensive Reassortment of H5 Influenza Viruses Isolated from Wild Birds in China over the past Decade', *Emerging Microbes & Infections*, 9: 1793–803.
- Daly, J. M. et al. (2011) 'Equine Influenza: A Review of an Unpredictable Virus', *The Veterinary Journal*, 189: 7–14.
- Dugan, V. G. et al. (2008) 'The Evolutionary Genetics and Emergence of Avian Influenza A Viruses in Wild Birds', *PLoS Pathogens*, 4: e1000076.
- Fitch, W. M. et al. (1997) 'Long Term Trends in the Evolution of H(3) HA1 Human Influenza Type A', *Proceedings of the National Academy of Sciences*, 94: 7712–8.
- Gack, M. U. et al. (2009) 'Influenza A Virus NS1 Targets the Ubiquitin Ligase TRIM25 to Evade Recognition by the Host Viral RNA Sensor RIG-I', *Cell Host & Microbe*, 5: 439–49.
- Gandon, S., and Michalakakis, Y. (2002) 'Local Adaptation, Evolutionary Potential and Host-parasite Coevolution: Interactions between Migration, Mutation, Population Size and Generation Time', *Journal of Evolutionary Biology*, 15: 451–62.
- Gething, M. J. et al. (1980) 'Cloning and DNA Sequence of Double-stranded Copies of Hemagglutinin Genes from H2 and H3 Strains Elucidates Antigenic Shift and Drift in Human Influenza Virus', *Nature*, 287: 310–306.
- Global Consortium for H5N8 and Related Influenza Viruses. (2016) 'Role for Migratory Wild Birds in the Global Spread of Avian Influenza H5N8', *Science*, 354: 213–7.
- Haaheim, R. (1977) 'Single-Radial-Complement-Fixation: A New Immunodiffusion Technique. 2. Assay of the Antibody Response to the Internal Antigens (MP and NP) of Influenza A Virus in Human Sera after Vaccination and Infection', *Developments in Biological Standardization*, 39: 481–4.
- Hatchette, T. F. et al. (2004) 'Influenza A Viruses in Feral Canadian Ducks: Extensive Reassortment in Nature', *Journal of General Virology*, 85: 2327–37.
- Hill, N. J. et al. (2021) 'Co-habiting in a Disease Hotspot: Overlap between Wild and Domestic Birds in Egypt Impacts Transmission of Highly Pathogenic H5N1', *Virus Evolution*: veaa093.
- Hill, S. C. et al. (2016) 'Antibody Responses to Avian Influenza Viruses in Wild Birds Broaden with Age', *Proceedings of the Royal Society B*, 283: 20162159.
- Huang, Y. et al. (2014) 'Genetic Structure of Avian Influenza Viruses from Ducks of the Atlantic Flyway of North America', *PLoS One*, 9: e86999.
- Jeffreys, H. (1961) *The Theory of Probability*. Oxford: Oxford University Press.
- Kang, H. M. et al. (2013) 'Genetic and Antigenic Characteristics of H4 Subtype Avian Influenza Viruses in Korea and Their Pathogenicity in Quails, Domestic Ducks and Mice', *Journal of General Virology*, 94: 30–9.
- Katoh, K., Asimenos, G., and Toh, H. (2009) 'Multiple Alignment of DNA Sequences with MAFFT', *Methods in Molecular Biology* (Clifton, N.J.), 537: 39–64.
- Kawaoka, Y., Krauss, S., and Webster, R. G. (1989) 'Avian-to-Human Transmission of the PB1 Gene of Influenza A Viruses in the 1957 and 1968 Pandemics', *Journal of Virology*, 63: 4603–8.
- Koel, B. F. et al. (2013) 'Substitutions near the Receptor Binding Site Determine Major Antigenic Change during Influenza Virus Evolution', *Science*, 342: 976–9.
- et al. (2014) 'Antigenic Variation of Clade 2.1 H5N1 Virus Is Determined by a Few Amino Acid Substitutions Immediately Adjacent to the Receptor Binding Site', *mBio*, 5: 301070–14.
- Krammer, F. (2019) 'The Human Antibody Response to Influenza A Virus Infection and Vaccination', *Nature Reviews Immunology*, 19: 383–97.
- Krejnosova, I. et al. (2009) 'Antibodies to PB1-F2 Protein Are Induced in Response to Influenza A Virus Infection', *Archives of Virology*, 154: 1599–604.
- Latorre-Margalef, N. et al. (2016) 'How Does Sampling Methodology Influence Molecular Detection and Isolation Success in Influenza A Virus Field Studies?', *Applied and Environmental Microbiology*, 82: 1147–53.
- et al. (2017) 'Competition between Influenza A Virus Subtypes through Heterosubtypic Immunity Modulates Re-infection and Antibody Dynamics in the Mallard Duck', *PLOS Pathogens*, 13: e1006419.
- et al. (2013) 'Heterosubtypic Immunity to Influenza A Virus Infections in Mallards May Explain Existence of Multiple Virus Subtypes', *PLoS Pathogens*, 9: e1003443.
- et al. (2014) 'Long-term Variation in Influenza A Virus Prevalence and Subtype Diversity in A Migratory Mallards in Northern Europe', *Proceedings of the Royal Society B*, 281.
- Lemey, P. et al. (2009) 'Bayesian Phylogeography Finds Its Roots', *PLoS Computational Biology*, 5: e1000520.
- Li, W. L., and Drummond, A. J. (2012) 'Model Averaging and Bayes Factor Calculation of Relaxed Molecular Clocks in Bayesian Phylogenetics', *Molecular Biology and Evolution*, 29: 751–61.
- Lindstrom, S. E., Cox, N. J., and Klimov, A. I. (2004) 'Genetic Analysis of Human H2N2 and Early H3N2 Influenza Viruses, 1957-1972: Evidence for Genetic Divergence and Multiple Reassortment Events', *Virology*, 328: 101–19.
- Lowen, A. C. (2017) 'Constraints, Drivers, and Implications of Influenza A Virus Reassortment', *Annual Review of Virology*, 4: 105–21.
- Macken, C. A., Webby, R. J., and Bruno, W. J. (2006) 'Genotype Turnover by Reassortment of Replication Complex Genes from Avian Influenza A Virus', *Journal of General Virology*, 87: 2803–15.
- Magor, K. E. (2011) 'Immunoglobulin Genetics and Antibody Responses to Influenza in Ducks', *Developmental and Comparative Immunology*, 35: 1008–16.
- Mahmoud, S. H. et al. (2019) 'Evolution of H5-Type Avian Influenza A Virus Towards Mammalian Tropism in Egypt, 2014 to 2015', *Pathogens*, 8: 224.
- Munster, V., and Fouchier, R. A. M. (2009) 'Avian Influenza Virus: Of Virus and Bird Ecology', *Vaccine*, 27: 6340–4.
- Munster, V. J. et al. (2007) 'Spatial, Temporal, and Species Variation in Prevalence of Influenza A Viruses in Wild Migratory Birds', *PLoS Pathogens*, 3: e61.
- Nogales, A. et al. (2018) 'Modulation of Innate Immune Responses by the Influenza A NS1 and PA-X Proteins', *Viruses*, 10: 708.
- Olsen, B. et al. (2006) 'Global Patterns of Influenza A Virus in Wild Birds', *Science*, 312: 384–8.
- Olson, S. H. et al. (2014) 'Sampling Strategies and Biodiversity of Influenza A Subtypes in Wild Birds', *PLoS ONE*, 9: e90826.
- Poen, M. J. et al. (2018) 'Local Amplification of Highly Pathogenic Avian Influenza H5N8 Viruses in Wild Birds in the Netherlands, 2016 to 2017', *Eurosurveillance*, 23: 17–00449.
- et al. (2019) 'Co-circulation of Genetically Distinct Highly Pathogenic Avian Influenza A Clade 2.3.4.4 (H5N6) Viruses in Wild Waterfowl and Poultry in Europe and East Asia, 2017-18', *Virus Evolution*, 5: vez004.
- Rabadan, R., Levine, A. J., and Robins, H. (2006) 'Comparison of Avian and Human Influenza A Viruses Reveals A Mutational Bias on the Viral Genomes', *Journal of Virology*, 80: 11887–91.

- Rajsbaum, R. et al. (2012) 'Species-specific Inhibition of RIG-I Ubiquitination and IFN Induction by the Influenza A Virus NS1 Protein', *PLoS Pathogens*, 8: e1003059.
- Rambaut, A. et al. (2016) 'Exploring the Temporal Structure of Heterochronous Sequences Using TempEst (Formerly Path-O-Gen)', *Virus Evolution*, 2: vew007.
- Recker, M. et al. (2007) 'The Generation of Influenza Outbreaks by a Network of Host Immune Responses against a Limited Set of Antigenic Types', *Proceedings of the National Academy of Sciences of the United States of America*, 104: 7711–6.
- Reeves, A. B. et al. (2011) 'Interspecies Transmission and Limited Persistence of Low Pathogenic Avian Influenza Genomes among Alaska Dabbling Ducks', *Infection, Genetics and Evolution*, 11: 2004–10.
- Ronquist, F. et al. (2012) 'MrBayes 3.2: Efficient Bayesian Phylogenetic Inference and Model Choice across a Large Model Space', *Systematic Biology*, 61: 539–42.
- Runstadler, J. et al. (2013) 'Connecting the Study of Wild Influenza with the Potential for Pandemic Disease', *Infection, Genetics and Evolution*, 17: 162–87.
- Scholtissek, C. et al. (1978) 'On the Origin of Human Influenza Virus Subtypes H2N2 and H3N2', *Virology*, 87: 13–20.
- Seekings, A. H. et al. (2018) 'Direct Evidence of H7N7 Avian Influenza Virus Mutation from Low to High Virulence on a Single Poultry Premises during an Outbreak in Free Range Chickens in the UK, 2008', *Infection, Genetics and Evolution*, 64: 13–31.
- Shapiro, B., Rambaut, A., and Drummond, A. J. (2006) 'Choosing Appropriate Substitution Models for the Phylogenetic Analysis of Protein-coding Sequences', *Molecular Biology and Evolution*, 23: 7–9.
- Sharp, G. B. et al. (1997) 'Coinfection of Wild Ducks by Influenza A Viruses: Distribution Patterns and Biological Significance', *Journal of Virology*, 71: 6128–35.
- Smith, D. J. et al. (2004) 'Mapping the Antigenic and Genetic Evolution of Influenza Virus', *Science*, 305: 371–6.
- Smith, G. J. et al. (2009) 'Origins and Evolutionary Genomics of the 2009 Swine-Origin H1N1 Influenza A Epidemic', *Nature*, 459: 1122–5.
- Spackman, E. et al. (2002) 'Development of a Real-time Reverse Transcriptase PCR Assay for Type A Influenza Virus and the Avian H5 and H7 Hemagglutinin Subtypes', *Journal of Clinical Microbiology*, 40: 3256–60.
- Steel, J., and Lowen, A. C. (2014) 'Influenza a Virus Reassortment', *Current Topics in Microbiology and Immunology*, 385: 377–401.
- Sukeno, N. et al. (1979) 'Anti-nucleoprotein Antibody Response in Influenza A Infection', *The Tohoku Journal of Experimental Medicine*, 128: 241–9.
- Taubenberger, J. K. et al. (2005) 'Characterization of the 1918 Influenza Virus Polymerase Genes', *Nature*, 437: 889–93.
- Thathaisong, U. et al. (2008) 'Human Monoclonal Single Chain Antibodies (HuscFv) that Bind to the Polymerase Proteins of Influenza A Virus', *Asian Pacific Journal of Allergy and Immunology*, 26: 23–35.
- Tolf, C. et al. (2013) 'Individual Variation in Influenza A Virus Infection Histories and Long-term Immune Responses in Mallards', *PLoS ONE*, 8: e61201.
- van Dijk, J. G. B. et al. (2014) 'Juveniles and Migrants as Drivers for Seasonal Epizootics of Avian Influenza Virus', *Journal of Animal Ecology*, 83: 266–75.
- et al. (2018) 'Host and Virus Ecology as Determinants of Influenza A Virus Transmission in Wild Birds', *Current Opinion in Virology*, 28: 26–36.
- Van Poelvoorde, L. A. E. et al. (2020) 'Next-generation Sequencing: An Eye-opener for the Surveillance of Antiviral Resistance in Influenza', *Trends in Biotechnology*, 38: 360–7.
- Venkatesh, D. et al. (2018) 'Avian Influenza Viruses in Wild Birds: Virus Evolution in a Multihost Ecosystem', *Journal of Virology*, 92: e00433–18.
- Verhagen, J. H. et al. (2017) 'Discordant Detection of Avian Influenza Virus Subtypes in Time and Space between Poultry and Wild Birds: Towards Improvement of Surveillance Programs', *PLoS ONE*, 12: e0173470.
- et al. (2020) 'Phylogeography and Antigenic Diversity of Low-pathogenic Avian Influenza H13 and H16 Viruses', *Journal of Virology*, 94: e00537–00520.
- Vijaykrishna, D. et al. (2015) 'The Contrasting Phylodynamics of Human Influenza B Viruses', *Elife*, 4: e05055.
- Vincent, A. et al. (2014) 'Review of Influenza A Virus in Swine Worldwide: A Call for Increased Surveillance and Research', *Zoonoses and Public Health*, 61: 4–17.
- Wallensten, A. et al. (2007) 'Surveillance of Influenza A Virus in Migratory Waterfowl in Northern Europe', *Emerging Infectious Diseases*, 13: 404–11.
- Webster, R. G. et al. (1992) 'Evolution and Ecology of Influenza A Viruses', *Microbiological Reviews*, 56: 152–79.
- Wilcox, B. R. et al. (2011) 'Influenza A Viruses in Ducks in Northwestern Minnesota: Fine Scale Spatial and Temporal Variation in Prevalence and Subtype Diversity', *PLoS ONE*, 6: e24010.
- Wille, M. et al. (2015) 'Temporal Dynamics, Diversity, and Interplay in Three Components of the Viriodiversity of a Mallard Population: Influenza A Virus, Avian Paramyxovirus and Avian Coronavirus', *Infection, Genetics and Evolution*, 29: 129–37.
- Wille, M., and Holmes, E. C. (2020) 'The Ecology and Evolution of Influenza Viruses', *Cold Spring Harbor Perspectives in Medicine*, 10: a038489.
- Wille, M. et al. (2018) 'Where Do All the Subtypes Go? Temporal Dynamics of H8–H12 Influenza A Viruses in Waterfowl', *Virus Evolution*, 4: vey025.
- et al. (2016) 'No Evidence for Homo- or Heterosubtypic Immunity of Influenza H3 in Mallards following Vaccination in a Natural-experimental System', *Molecular Ecology*, 26: 1420–31.
- et al. (2013) 'Frequency and Patterns of Reassortment in Natural Influenza A Virus Infection in A Reservoir Host', *Virology*, 443: 150–60.
- Wisedchanwet, T. et al. (2011) 'Genetic Characterization of Avian Influenza Subtype H4N6 and H4N9 from Live Bird Market, Thailand', *Virology Journal*, 8: 131.
- Yodsheewan, R. et al. (2013) 'Human Monoclonal ScFv Specific to NS1 Protein Inhibits Replication of Influenza Viruses across Types and Subtypes', *Antiviral Research*, 100: 226–37.
- Zhang, H., Chen, Q., and Chen, Z. (2012) 'Characterization of an H4N2 Avian Influenza Virus Isolated from Domestic Duck in Dongting Lake Wetland in 2009', *Virus Genes*, 44: 24–31.
- Zhou, B. et al. (2009) 'Single-reaction Genomic Amplification Accelerates Sequencing and Vaccine Production for Classical and Swine Origin Human Influenza A Viruses', *Journal of Virology*, 83: 10309–13.
- Zhou, B., and Wentworth, D. E. (2012) 'Influenza A Virus Molecular Virology Techniques', *Methods in Molecular Biology (Clifton, N.J.)*, 865: 175–92.
- Zohari, S. et al. (2014) 'Avian Influenza A(H10N7) Virus Involvement in Mass Mortality of Harbour Seals (*Phoca vitulina*) in Sweden, March through October 2014', *Eurosurveillance*, 19: 20967.