Virology 443 (2013) 150–160

Contents lists available at SciVerse ScienceDirect

Virology

journal homepage: www.elsevier.com/locate/yviro

Frequency and patterns of reassortment in natural influenza A virus infection in a reservoir host

Michelle Wille [a], Conny Tolf [a], Alexis Avril [a,b], Neus Latorre-Margalef [a], Sofie Wallerström [c], Björn Olsen [d], Jonas Waldenström [a,*]

[a] Centre for Ecology and Evolution in Microbial Model Systems, Linnaeus University, Kalmar SE-391 82, Sweden
[b] CIRAD, Campus international de Baillarguet, Montpellier 34398, France
[c] Karolinska Institutet, Department of Microbiology, Tumor and Cell Biology (MTC), Box 280, Stockholm SE-171 77, Sweden
[d] Section of Infectious Diseases, Department of Medical Sciences, Uppsala University, Uppsala SE-751 85, Sweden

ARTICLE INFO

Article history:
Received 4 January 2013
Returned to author for revisions
4 February 2013
Accepted 3 May 2013
Available online 28 May 2013

Keywords:
Anas platyrhynchos
Evolution
Influenza A virus
Mallard
Reassortment
Sentinel ducks

ABSTRACT

Influenza A viruses (IAV) can dramatically alter both genotype and phenotype at a rapid rate as a product of co-infection and reassortment. Avian IAV exhibit high levels of phylogenetic incongruence, suggesting high levels of reassortment in the virus reservoir. Using a natural-experimental system, we reconstructed relationships amongst 92 viruses across 15 subtypes from 10 Mallards in an autumn season. Phylogenetic analyses estimated that 56% of the isolated viruses were reassorted. Network analysis demonstrated different patterns of reassortment and limited exchange of segments between primary and secondary infections. No clear patterns of linkage between segments were found, and patterns within a season were likely the consequence of continued introduction of new constellations, high viral load and diversity in the wild bird reservoir, and co-infections. This is the first IAV study to implement multiple tools available for elucidating factors governing reassortment patterns in naturally infected Mallards.

Introduction

Influenza A viruses (IAV) belong to the family *Orthomyxoviridae*, and are enveloped viruses with a genome consisting of eight segments of negative-sense single-stranded RNA (Kawaoka et al., 2005; Webster et al., 1992). IAV exhibit high mutation rates due to an error prone RNA-dependent RNA polymerase and, consequently, a rapid rate of change in antigenic proteins. Avian IAV evolves at the same rate as mammalian influenza viruses (Chen and Holmes, 2006, 2009), and the eight segments have multiple evolving lineages (Bahl et al., 2009; Chen and Holmes, 2009, 2010; Dugan et al., 2008). Reassortment, the process in which new virus variants can arise through exchange of the RNA segments in co-infected cells, results in the propensity for IAV to dramatically alter both genotype and phenotype by combining segments of different origin in new combinations (Rambaut et al., 2008; Salomon and Webster, 2009; Webster et al., 1992). By reassortment, novel viral strains can appear rapidly, and if these introductions contain genomic segments determining antigenicity (particularly the surface proteins hemagglutinin [HA] and neuraminidase [NA]) from distantly related strains, it may provide new strains the opportunity to proliferate among immunologically naïve hosts (Webster et al., 1992). IAV have a broad host range, and viruses infecting humans, pigs, horses and wild birds form different phylogenetic clades across all segments (Olsen et al., 2006; Webster et al., 1992). In avian lineages, additional genetic distinction is evident, whereby viruses isolated from North America, Eurasia and South America have different genetic and antigenic lineages (Olsen et al., 2006; Pereda et al., 2008). Although these distinctions exist, it has been shown that the PB1, PA and H6 segments currently circulating in North American wild birds have an Eurasian origin (Bahl et al., 2009; zu Donha et al., 2009). Not only can IAV exchange segments within a host group, but host-specific viruses can also reassort following co-infection, providing the capacity to alter host range. For instance, reassortment between human, avian and swine lineages have resulted in viruses with pandemic potential, and have been implicated in the last three human IAV pandemics (Lindstrom et al., 2004; Rabadan et al., 2006; Scholtissek et al., 1978; Taubenberger et al., 2005).

Wild waterfowl and shorebirds (including gulls) are the main avian reservoir hosts for IAV. In these hosts there are many co-circulating HA/NA subtypes, high levels of co-infection in single animals (Sharp et al., 1993) and high level of reassortment (Bahl et al., 2009; Chen and Holmes, 2006, 2010; Dugan et al., 2008).

* Corresponding author.
E-mail address: jonas.waldenstrom@lnu.se (J. Waldenström).

0042-6822/$ - see front matter © 2013 Elsevier Inc. All rights reserved.
http://dx.doi.org/10.1016/j.virol.2013.05.004

However, neither the specific pattern, nor the scope of co-infection and reassortment has been adequately assessed. One hurdle has been the lack of automated tools and computational limitations to deal with high levels of phylogenetic complexity across eight segments. A second problem is the lack of robust data. Indeed, to date, studies assessing reassortment in wild bird populations largely utilize multiyear data, often collected across large spatial scales and across multiple host species (Bahl et al., 2009; Chen and Holmes, 2006, 2010; Dugan et al., 2008). Multiyear data does not allow the detailing of specific patterns of reassortment in the short term, nor does it illustrate which factors are mechanistically important. It has been shown that the likelihood of segments being incorporated through reassortment varies, and that genetic linkages between different segments do occur in human and swine IAV (Downie, 2004; Khiabanian et al., 2009; Varich et al., 2008). Unlike IAV dynamics in these hosts, very high levels of incongruence have been demonstrated in analyses of low pathogenic avian IAVs (Bokhari and Janies, 2010; Dugan et al., 2008; Lam et al., 2011). Further, despite investigations utilizing sequence data from field isolates combined with laboratory assessments using controlled infections of cell lines, results pertaining to the non-random nature of reassortment are not consistent in avian IAV.

In the present article, we address the process of IAV reassortment on a short time scale in natural infections in a reservoir host, the Mallard (*Anas platyrhynchos*). We studied the IAV infection patterns in 10 immunologically naïve sentinel Mallards in a wild setting in Sweden over the course of a full autumn season. The birds were kept in contact with wild waterfowl; the sentinels shared water with, and were exposed to aerosols generated by wild Mallards entering and feeding within a duck trap. Hence, these 10 ducks represent a probe to illustrate the major patterns of introduction, extinction, reassortment, and maintenance in the viral reservoir. Using this experimental set up, with a daily sampling regime, we aimed to assess the (a) the order of infections and reinfections, (b) overall magnitude and scope of reassortment, (c) the carry-over of segments and viruses between primary and secondary infections, (d) disentangle the involvement of antigenic HA and NA combinations, (e) linkage between different segment combinations, and (f) the putative role of co-infection in our dataset.

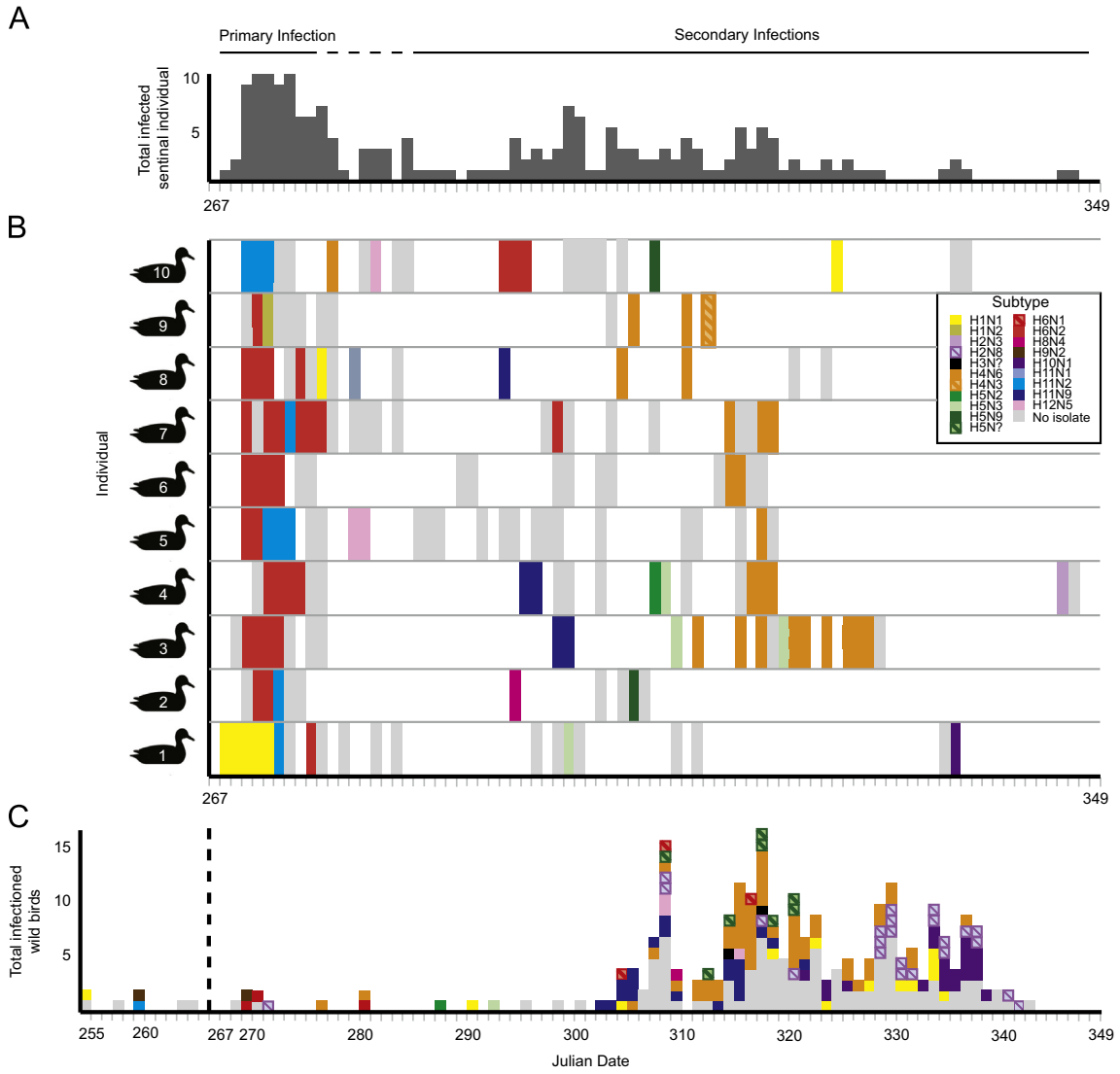

Fig. 1. Influenza A viruses detected and isolated from Mallards between 24 September and 15 December 2009 (Julian day 267–349). (A) Total number of infections per day by determined by rRT-PCR in the 10 sentinel ducks placed in the trap. (B) Detailed infection history for each sentinel duck. (C) Total number of viruses detected and isolated from the wild ducks during the period of study. Days where no infection was detected are in white, where infection was detected by rRT-PCR but no virus was isolated in culture are in gray, and where virus was detected and isolates grew in culture are colored by subtype.

Results

Infection profiles and diversity of IAV

Across the 83 days of study, IAV prevalence (assessed by rRT-PCR) had a bimodal distribution, with an initial peak corresponding to the primary infection, and a longer, flat secondary peak corresponding to secondary infections (Fig. 1A). We defined primary infection as the first infection detected by rRT-PCR, complimented with subtype information. The end of primary infection was defined by the first rRT-PCR negative day, which allowed for co-infections with multiple subtypes. In a single instance (individual 8) primary infection ended after the second rRT-PCR negative occasion, as sequence information indicated that the same virus was shed both before and after the intermittent rRT-PCR negative occasion on day 7. The threshold between primary and secondary infection differed across individuals (Fig. 1B). All individuals acquired detectable infections within the first 5 days, and there were 11–24 infection positive days, hence on average individuals were positive for 23.1% of the study period (Table S1). Specific infection patterns of RT-PCR and serology data for these individuals are more thoroughly explored in (Tolf et al., 2013). Ninety-two viruses were successfully cultured, and following subtyping, a detailed infection profile was generated for each individual (Fig. 1B). A total of 15 different subtypes were identified in the sentinel ducks, with H6N2 ($n=32$), H4N6 ($n=23$) and H11N2 ($n=9$) being the most common (Figs. 1 and 2, Table 1).

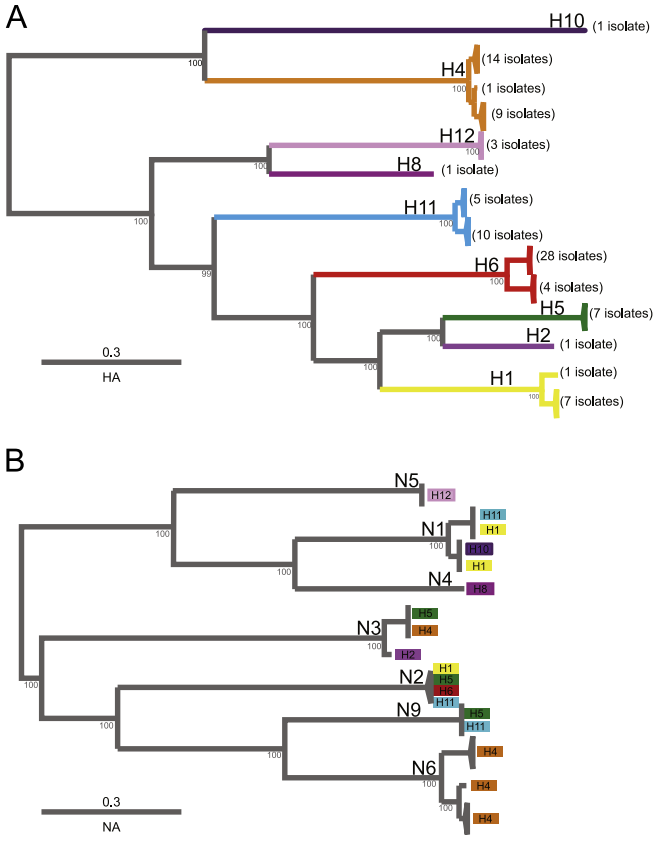

Fig. 2. Bayesian phylogenetic trees of external glycoprotein segments of 92 viruses isolated from sentinel ducks in 2009. (A) Hemagluttinin tree comprising of nine HA subtypes and associated number of isolates per branch. (B) Neuraminidase tree comprising of seven NA subtypes, and associated HA subtypes in boxes to illustrate subtype combinations present in the sentinel ducks. Branch colors in A and box color in B correspond to Fig. 1. Branches have been collapsed if pairwise similarity of the group is 99% or more. Trees have been rooted at the midpoint for clarity. Scale bar indicates the number of expected changes per site.

Table 1
Analysis of reassortment frequency of influenza A viruses by the GiRaF program.

Subtypes	Primary or Secondary	Number of isolates	Number of infections	% Reassorted GiRaF
H1N1	Primary	5	1	100
H1N2	Primary	1	1	100
H6N2	Primary	28	9	3.6
H11N2	Primary	9	5	66.7
H1N1	Secondary	2	2	100
H2N3	Secondary	1	1	0.0
H4N3	Secondary	1	1	100
H4N6	Secondary	23	8	95.7
H5N2	Secondary	1	1	100
H5N3	Secondary	4	3	100
H5N9	Secondary	2	2	100
H6N2	Secondary	4	2	100
H8N4	Secondary	1	1	100
H10N1	Secondary	1	1	0.0
H11N1	Secondary	1	1	100
H11N9	Secondary	5	3	100
H12N5	Secondary	3	2	0.0

However, it is important to note that these variants represent only those viruses that were successfully cultured (mean per individual=9, overall isolation success=48%; Table S1, Fig. 1B). Ducks were infected with two to six different subtypes, and only a single individual was infected with the same HA subtype in two different infections, where individual 7 was infected with H6N2 viruses on two separate occasions. Primary infection consisted predominately of H1N1, H6N2 and H11N2 viruses, and secondary infection comprised mainly of H11N9, H4N6, H6N2, as well as a number of H5 variants, with other subtypes detected only 1–3 times (Fig. 1B, Table 1). Some rare subtypes, such as H1N2, H11N1, and H4N3 were only detected toward the end of infections, after a number of days of shedding viruses with more common subtypes, such as H1N1, H6N2, or H4N6 (Fig. 1B).

As part of an ongoing surveillance scheme, 3842 samples were collected from wild, migratory Mallards between August and December 2009 in the same trap as the sentinel ducks were housed. Four hundred and eighty-two samples were positive by rRT-PCR, and 175 viruses were isolated, corresponding to 23 HA–NA combination subtypes and nine with only the HA subtypes determined (Table S2). All the subtypes found in wild ducks captured between September and December were also isolated from the sentinel ducks, with the exception of the H3 and H9 subtypes (Table S2). Unlike patterns observed in the sentinel ducks, a large number of H10 viruses were detected in the wild ducks in November, where only a single isolate was detected in the sentinel ducks (Fig. 1). The highly proliferating H11N2 and H6N2 viruses that dominated the primary infections of the sentinel ducks were detected also in wild ducks shortly before the introduction of the sentinel ducks. Few viruses were detected in the wild ducks prior to Julian day 300, which was attributed to the low number of birds actually visiting the trap, and consequently the number of samples taken during this period (Table S2, Fig. 1).

Phylogenetic analysis of reassortment events

On a rudimentary level, different HA and NA subtype combinations illustrate reassortment. When including genetic diversity, more complex reassortment patterns became evident. Indeed, despite 15 subtype combinations, identical HA sequences had numerous different NA subtypes, e.g. H5N2, H5N3, H5N9 (Fig. 2). All viruses were genotyped and exhibited the G, G, D, 4, H, 6, F, 1E/2A pattern (Lu et al., 2007), most frequently found in Eurasian ducks. Small genetic distances were evident in the PB2, PA, NP and M segments with larger distances and more variance across the

HA, NA and NS segments. The latter was likely due to the inclusion of multiple subtypes (nine HA subtypes, seven NA subtypes and two NS subtypes, respectively; Fig. S1). Within the HA and NA segments, as indicated by inferred phylogenetic relationships, the highest diversity was observed for H4 (24 isolates), followed by H11 (15 isolates) and H6 (32 isolates), and N6 (Fig. 2). Comparisons of distances across segments revealed high levels of incongruence, reflected in low Pearson's correlation coefficients (r) (Table S3).

Phylogenetic topology comparisons revealed high levels of reassortment across all segments and subtypes within the dataset (Table 1, Fig. S2). All subtypes that occurred in the primary infections, except H6N2 viruses, were identified as reassorted viruses. Infection profiles suggested the presence of co-infections during primary infections, as there were no rRT-PCR negative days between infections, from which viruses with different subtype combinations were isolated. This was further supported by phylogenetic topological incongruence. The H6N2 and H11N2 viruses had segments with shared history and H11N2 viruses were found in clades dominated by the H6N2 viruses on several occasions (Fig. S2). The most common subtypes detected as secondary infections were H11N9, H4N6, and H5 viruses, which were all identified as reassorted (Table 1). Unlike the primary infections, there were rRT-PCR negative days between infections. However, these highly reassorted secondary infections occurred during a period with high IAV prevalence among wild ducks and where multiple subtypes were present simultaneously (Fig. 1C).

Only three viruses occurring during secondary infections had independent lineages for all segments. Among these, H10N1 and H2N3 appeared as late, short-lived, secondary infections, while the H12N5 viruses were isolated earlier and from two different individuals (Table 1). Unlike the H10N1 and H2N3 viruses, which also appeared only once, the H8N4 virus was identified as being reassorted. However, this virus occurred much earlier than the latter unique introductions (Fig. 1B). Unlike these rare subtypes that appear in low numbers each autumn at Ottenby, other rare subtypes that were not maintained in the population, such as H1N2, H11N1 and H4N3 were the product of reassortment. Manual inspection of phylogenetic trees suggested that these viruses shared segments with more common subtypes shed during the same time period, e.g. the rare subtypes H1N2 and H11N1 shared segments with the H1N1 and H11N2 virus of the primary infection (Fig. S2).

Reassortment of genome constellations

Limited transfer of segments between primary and secondary infections

In terms of single segments, few segments present in the highly proliferating primary infections were detected again in the secondary infections, with the exception of the putatively reassorted H11N1 virus and the secondary H1N1 infections (Fig. 3, Tables S4 and S5). The reoccurrence of internal (PB2, PB1, PA, NP and M) segments was infrequent, as was transfer within the NS segment. However, at an allele level, both the NS A and B alleles were found in the primary and secondary infections. There was more transfer of NA segments between isolates, with little genetic differentiation between NA segments detected in primary and secondary infections (Figs. 2 and 3). While, there was greater discrimination within the HA subtypes, it remains unclear whether these differences were both genetic and antigenic.

Network analyses, illustrating the relationships between constellations (identified in Fig. 3), demonstrated little carry-over of segments between primary and secondary infections (Fig. 4A). Adjusting the level of stringency in the analyses graphically demonstrated linkage and dissociation of linkage between constellations and overall network patterns (Fig. S4). Thus, at the lowest level of stringency, the entire network is connected (Fig. S4E), owing to the limited transfer of some segments, such as the NS segment (Fig. 3). However, by increasing the stringency slightly, the connection between primary and secondary infections immediately breaks down (Fig. S4D).

To test the level of association between primary and secondary infections, two-way tables for each set of infections were constructed (Fig. 4, Tables S4 and S5). Utilizing these tables, and simulating the reassortment process, we could determine whether the available segments observed during the primary infections could produce the constellations observed in the secondary infections. Of 2916 possible segment constellations that could have occurred in the secondary infections (based on the primary infection), only one case was observed (in an H1N1 virus). Based on these analyses, the diversity of segments in the secondary infections seem better related to the arrival of new viruses in the population, than due to transfer of segments between infections (Tables S4 and S5).

Patterns of reassortment across HA/NA combinations

Further investigating the patterns of the networks at different stringency levels revealed different patterns of shared segments between subtypes. At the highest level of stringency (Fig. S4A) there were few shared and related constellations; only between the H11N2 and H6N2 viruses in primary infections (a likely result of co-infections), and the H4N6, H4N3, and H5 viruses in secondary infections. If constellation constraints were lifted, whereby a smaller number of segments were artificially linked, shared histories of the H5N3, H5N2, H5N9, H11N9, H4N3 and some H4N6 viruses became more evident (Fig. 4A, Fig. S4). The H12N5, H10N1, H8N4, secondary H1N1, H2N3, and secondary H6N2 constellations were distantly connected to the remainder of the network, with high link levels. Moreover, as the level of stringency was increased, these constellations rapidly dissociated from the network (Fig. S4), suggesting they have few shared segments with the other viruses in the dataset (Fig. 4A). Single circles, or constellations, with two different colors, or HA/NA subtypes, suggested that the same internal segment constellations (PB2, PB1, PA, NP, M, NS) were shared across different subtypes, as was the case with one H11N2 virus and H6N2 viruses in the primary infection; H5N2 and H5N3 viruses, and H4N3 and one H4N6 virus constellation in the secondary infections (Fig. 3, Fig. 4A). Further, different constellations with the same color suggest more than one constellation for each subtype. This was clearly evident for the H4N6 viruses, where there were two main constellations, which were not linked to each other (Fig. 4A). One of H4N6 constellations (constellation 8) shared numerous segments with the H5N2, H5N3, and H4N3 viruses, where the other H4N6 constellation (constellation 2) was independent of these putative reassortants (Fig. 4A).

In order to address the role of subtype in reassortment of this dataset, we used a two step approach. First, the identification of putative parental constellations and second, comparison of observed and expected frequencies of these constellations under the hypothesis of random reassortment (Fig. 5, Tables S6–S8). Theoretically, the parental constellations should be the most likely to appear when recombining the available lineages. Using this approach, statistically plausible parental constellations were identified in H6N2, H11N2, H5N2, H5N3 and H4N6 constellations (Table S6). Assuming that each of the five identified constellations were parental, we computed the observed reassortment frequency for each of them (Table S7) and we compared the observed reassortment frequencies to those expected under the null hypothesis of random combinations of the clades available for each segment. Overall, the observed frequencies for all parental constellations were less than expected under the null hypothesis (Table S7). The observed reassortment frequencies of H6N2 and

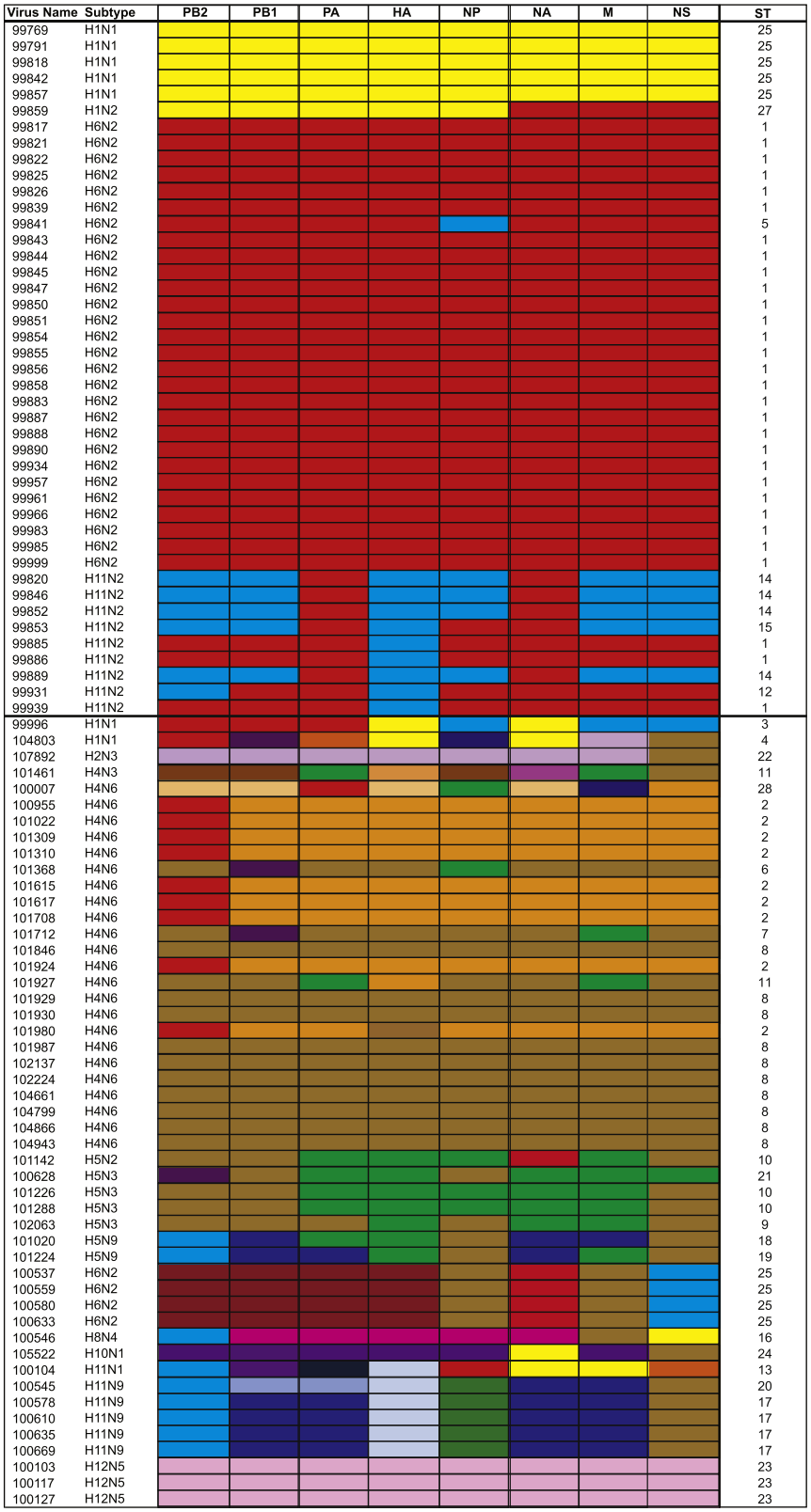

Fig. 3. Genome constellations of the 92 viruses isolated derived from phylogenetic trees and pairwise alignments. The isolate name and subtype are listed on the left of the genome constellations, and are ordered by infection status, with the primary infections first followed by the secondary infections and separated by a line, and further arranged by subtype. Each segment is represented by a box, and ordered by segment size from left to right (PB2, PB1, PA, HA, NP, NA, M, and NS). Segments with a shared color (in a vertical fashion) are identical. Each unique constellation of internal segments (PB2, PB1, PA, NP, M, NS) has been given a unique constellation number (ST), and corresponds to constellation numbers in Fig. 4. Phylogenetic trees on which genome constellations are based can be found in Fig. 2 and Fig. S2, and pairwise identity charts in Fig. S1.

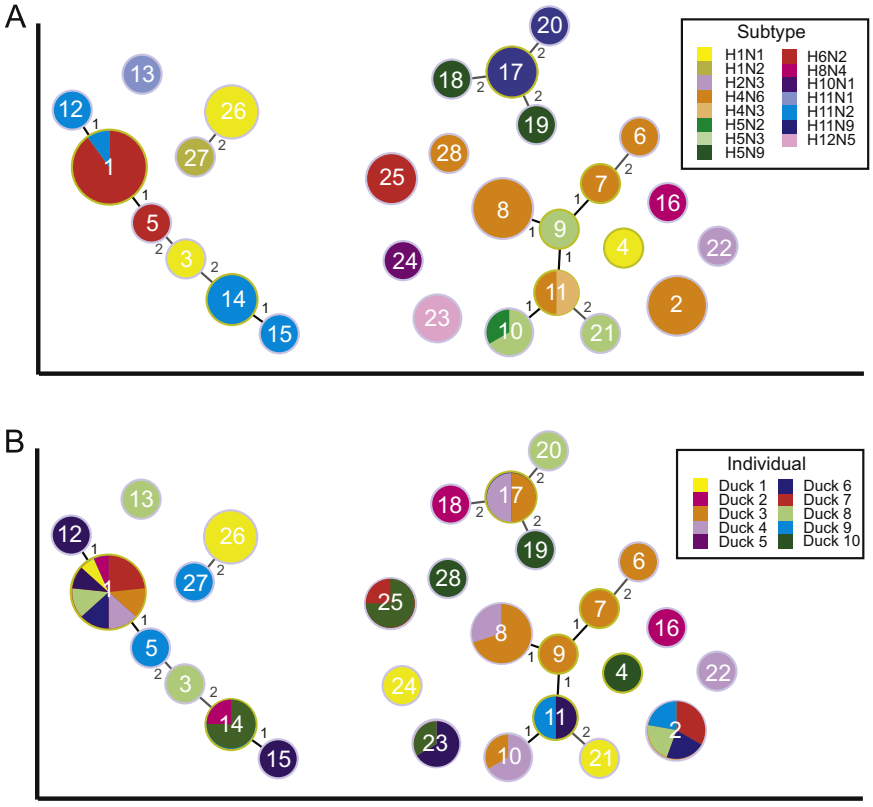

Fig. 4. Network analysis of genome constellations from 92 virus isolates whereby any four of six segments needed to be shared between constellations. Each circle represents a distinct genome constellation, and numbered text within each circle identifies the constellation number (ST). The size of the circle is proportional to the number of times the ST was present in the dataset. Lines between circles suggest that STs have partly shared genomes, where black lines indicate large similarities between constellations and gray lines less similarities. The levels of connection between constellations are included on the connectors, where the higher the level the more differences there are between constellations. (A) Constellations colored by virus subtype and, (B) constellations colored by individual. Circles with more than one color are constellations that are found in more than one subtype or individual, respectively. Circles are placed along the *x*-axis in approximate position of infection, with STs in the left cluster representing primary infections, followed by secondary infection clusters to the right. Full network analysis by subtype can be found in Fig. S4.

one H4N6 virus constellation were lower than the other parental constellations tested, which reflects the results of the GiRaF analysis (for the H6N2) and observation of limited reassortment in constellation 2 (H4N6). Overall, we failed to find significant difference in reassortment frequencies between each identified parental HA lineage (Table S8), thus in this analysis subtype was not a driver of reassortment.

To test for specific patterns of linkage throughout the dataset, we compared the observed and expected counts of linkage if the phenomenon was random. Overall, most of segment pairs were less frequent than expected if linkages occurred randomly (Table S9). Together, these results show little statistical support for specific and repeated patterns of linkage between segments.

Influence of co-infections

Assessing the distribution of constellations across individuals provides information regarding the role of co-infection, rather than reassortment in the dataset. The infection profile of individual 3 (Fig. 1) suggests prolonged shedding of an H4N6 infection, in addition to intermittent shedding of H5N3 viruses. Network analysis demonstrated that linked constellations 6, 7, 9, 8 and 10 (comprised of H4N6 and H5N3 subtypes) were shed by the same individual (Fig. 4B). This strongly suggests the occurrence co-infection in this individual, but the role of reassortment remains unclear. As previously alluded to, due to the lack of RT-PCR negative days, and the relatedness between constellations, we predict that patterns seen in the H6N2–H11N2 viruses in the primary infection are the result of co-infection.

Discussion

Patterns of co-infection and reassortment within an autumn season

Exploiting the standard approach of sentinels to monitor population health, we were able to investigate reassortment from longitudinal infection histories at the individual level amongst immunologically naïve ducks across an autumn season. We could assess frequency and patterns of reassortment utilizing both new and standard phylogenetic tools and a network approach. Regardless of the methodology utilized, all analyses found a high frequency of reassortment in the sentinel Mallard viruses. This was illustrated by poor correlations of pairwise distances, phylogenetic topologies, and genome constellations between the eight segments. Overall, 56% of collected viruses were estimated to be reassorted, however, this is reflective of the entire dataset and is biased by numerous putative viral clones from the same infection, parallel infections across multiple individuals, and putative co-infections. Even with the removal of candidate co-infections (i.e. primary infection, and all H4N6, H5N3 viruses from individual 3), and controlling for multiple isolates per individual in the statistical analysis, the estimate of reassortment remains high.

Within a season, evidence for multiple successive viral imports into the sentinel ducks were supported by the limited carry-over of segments and lack of entire viruses between primary and secondary infections. Primary infections consisted of fewer subtypes, were shed for a shorter period, and detected reassortment occurred between few subtypes (H6N2–H11N2, H1N1–H11N1–H1N2). Due to the lack of RT-PCR positive days between different infections in the primary infections, we are likely detecting co-infection in the sentinel ducks.

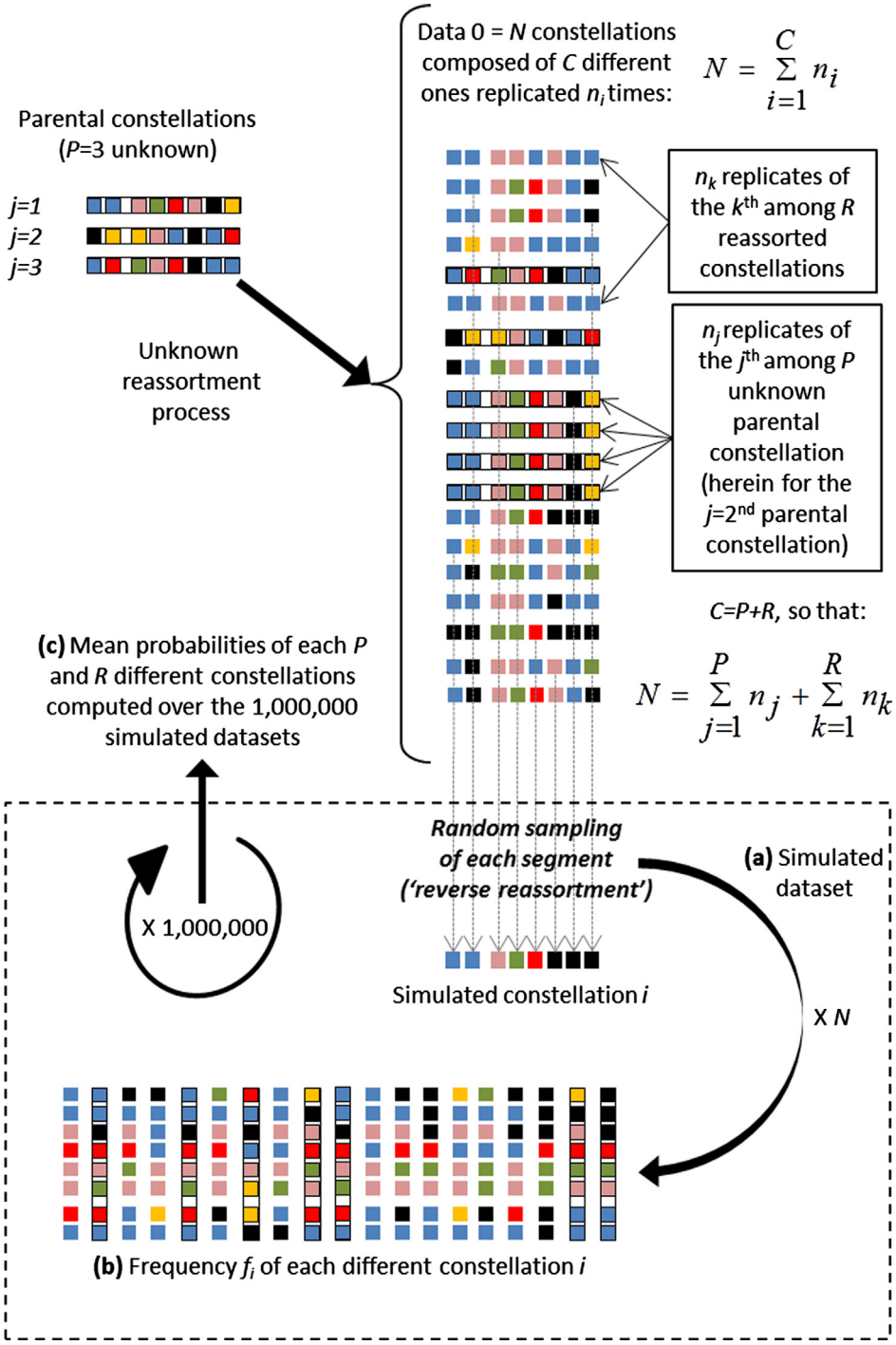

Fig. 5. Random sampling scheme used to identify the parental constellation. We assumed that the parental constellations were still present at the time of sampling, even at low frequency. Hence, the observed data (data 0) is composed of *C* individually different constellations *i*. Among the C individually different constellations, we supposed *P* parental constellations and *R* reassorted ones, with n_j and n_k the number of replicates for the *j*th unknown parental constellation and *k*th unknown reassorted constellation respectively. We assumed that the parental constellations should have the highest probability of occurrence if we reverse the reassortment process. To have a measure of the probability that each *i* different segment constellation is parental, we simulated a reverse reassortment process (dashed box): (a) for each constellation *i* we permuted the lineages for each of the eight segments available in data 0 to obtain *N* simulated constellations (i.e. a simulated dataset "before reassortment"), (b) we computed the simulated frequency of the *i*th constellation over the simulated dataset, and (c) we repeated this procedure 1,000,000 to obtain the mean probability for each *i* constellation. We used this rational to test for (i) random reassortment and (ii) linkages between segments. In (i), the observed reassortment frequencies in data 0 were compared to those expected under the hypothesis of random reassortment, i.e. against the distribution of the reassortment frequency obtained by permuting only the NA-PB2/PB1/PA/NP/M/NS segments for each HA subtype belonging to a parental constellation. In (ii), we compared the observed linkage frequencies of each segment lineages pairs or triplets against the expected distribution of the linkage frequencies obtained by permutations of the segment lineages available in data 0. A more exhaustive explanation of the methods can be found in the supplemental materials Tables S6–S9.

The shedding of reassorted subtypes not detected annually at our field site (e.g. H1N1, H1N2) may be the product of reassortment, rather than a product of the process of egg isolation. These viruses appear following the shedding of more common subtypes, and are a feature observed in other field studies. (Lebarbenchon et al., 2012; Wilcox et al., 2011). In contrast, during secondary infections there

were more RRT-PCR negative days between different subtypes and constellations, suggesting a decreased role of co-infection, with the exception of individual 3. The different constellations detected in the secondary infections displayed a number of trends, including different constellations with the same HA/NA subtype combination, and shared constellations across different HA/NA subtypes, suggesting reassortment occurring at this stage. The lack of segment exchange between primary and secondary infections suggests that reassortment occurs between viruses circulating simultaneously, particularly during periods of heavy viral load in the host reservoir population. Hence, there are strong temporal trends within a season. The final infections (H1N1, H2N3 and H10N1) were unique and not reassorted with the other viruses during the secondary infections; indeed these short infections should probably be classified as tertiary infections.

Estimating reassortment, and the role of co-infections

Quantifying reassortment has proven to be challenging to estimate in wild bird IAV due to the combination of multiple segments, and multiple lineages in each segment, high mutation rates, and unequal rates of change across segments. In contrast, human and swine IAV have only a few HA/NA subtypes, and internal segment lineages (Khiabanian et al., 2009; Rambaut et al., 2008). Further, the process of vaccination is a large driver in limiting only 1–2 subtypes to circulate annually in humans, thereby limiting reassortment. Therefore, more work has been done to quantify and describe reassortment in human and swine IAV (Ghedin et al., 2009; Khiabanian et al., 2009; Lam et al., 2011; Rambaut et al., 2008).

Incongruence methodology and phylogenetic analysis based on large multi-year datasets have revealed high levels of both reassortment, but also transient genetic linkage in avian IAV (Chen and Holmes, 2006, 2010; Dugan et al., 2008). Existing phylogeny estimates of reassortment compare the fit of data for each segment to each maximum likelihood tree (Dugan et al., 2008). Phylogenetic estimators of reassortment, including GiRaF, while providing an estimate of reassortment frequency are plagued by a number of problems such as influence of drift, different numbers of circulating lineages across the different segments, inefficiencies in heuristic searches, and a lack of consistency in defining a threshold to determine when trees or lineages are different enough to imply reassortment (Bokhari and Janies, 2010). The characterization of lineages from trees, or by pairwise clustering (e.g. FluGenome) has been utilized to compare viral genome constellations, or specific groups of lineages. In segmented viruses, this type method has been used both in IAV as well as rotaviruses (Dugan et al., 2008; McDonald et al., 2009), and network approaches based on constellations have been utilized in poultry and swine IAV (Bataille et al., 2011; Bokhari and Janies, 2010; Bokhari et al., 2012). Networks do have an advantage as they demonstrate how segment lineages are shared across different viruses, allowing us to better explain reassortment patterns ecologically/epidemiologically and alleviating some of the biases of phylogenetic trees. In our study, we demonstrate how different constellations are related, or unrelated, by network analysis. And while we have an initial estimate of within season reassortment, the ability to discern specific patterns, and test whether these patterns occur due to random chance provides insight into the process.

Reassortment is a result of co-infections, and aside from a study by Sharp et al. (1997), no one has assessed the level of co-infection of IAV in wild birds (Sharp et al., 1997). Utilizing standard techniques of sample collection, screening, and culture we are describing the most common variant in the host, resulting in a consistent methodological bias: egg adaptation and/or egg based reassortment. This bias has recently been identified in studies of human influenza by comparing the result of pipeline capillary sequencing to pyrosequencing (Ghedin et al., 2009). Due to the high resolution of our sampling scheme, putative co-infections are more evident as overlapping infections and can be identified in the infection histories and numerous constellations of the same subtype are shed by the same individual. Therefore, co-infections and egg-reassortment may have biased the overall GiRaF estimate of reassortment, suggesting the importance of network analysis over absolute tree-based estimates of reassortment. Co-infections and egg-based reassortment are an inherent problem in all IAV studies utilizing field data, as usually infection history data are not available, hence the identification of co-infections versus reassortants are not possible. Our analysis highlights the need for an assessment of co-infection in IAV studies, a phenomenon which can be better accounted for by moving away from egg isolation and towards deep sequencing of original material (Hoper et al., 2011; Ramakrishnan et al., 2009).

Perspectives

Long term studies of IAV dynamics have demonstrated a number of subtype and genome constellation patterns: some subtypes are common and some are rare, some are annually maintained and others are infrequently detected; some genome constellations are linked and appear across many years and some are novel, and/or detected only once (Chen and Holmes, 2010; Dugan et al., 2008; Wallensten et al., 2007; Wilcox et al., 2011). Our dataset encompasses viruses representing all these variants, and indeed, different patterns of reassortment (and co-infections) within an autumn season are likely the driving force for this phenomenon. We hypothesize that the independence of primary, secondary, and tertiary viruses is the result of immunity in the ducks, where they are unlikely to be infected by the same or similar subtype and genome constellation within an autumn season (Latorre-Margalef et al., 2013; Tolf et al., 2013). High rates of reassortment, particular in the secondary infections are likely the effect of high virus load in the wild ducks resulting in the circulation of multiple subtypes, and subsequently a higher likelihood of co-infections in both the sentinel ducks and the wild ducks. Our initial estimate of within season reassortment is higher than anticipated, demonstrating a discrepancy between annual and inter-annual patterns of reassortment. These high rates of reassortment, resulting in the continual production of new subtypes and constellation combinations should result in total panmixia of the virus population. The limited carry-over of putatively reassorted viruses (Lebarbenchon et al., 2012) suggests an evolutionary barrier between seasons, and lower fitness of these viruses. However, in experiments putative reassortants do not have lower survival compared with viruses that are seen annually in surveillance programs (Lebarbenchon et al., 2012). We hypothesize that the limited maintenance of reassortants across seasons may maintain genetic structure in the virus reservoir despite high seasonal reassortment rates, and it is these high seasonal rates that produce the overall diversity detected. In addition to frequency, we illustrate the patterns of virus introductions and reassortment within a resident duck population. Further, the large multitude of subtypes found during periods of peak influenza prevalence have highly connected genomes, and there is no clear bias towards reassortment of certain segments, or across specific HA/NA subtypes. This is the first study to dedicatedly assess the phenomenon of reassortment in wild birds, and consequently demonstrates the sheer capacity of IAV reassortment in a duck population, and the role that resident ducks have in the propagation of new viruses following infection from their migratory counter parts.

Materials and Methods

Sample collection

Ten immunologically naïve, first year, female Mallards from a commercial hunt-and-release farm were placed into an outdoor duck trap located near Ottenby Bird Observatory, Sweden (56°12′N 16°24′ E). The trap is primarily used for long-term monitoring of IAV and sampling methods for wild birds have been described in previous publications (Latorre-Margalef et al., 2009; Wallensten et al., 2007). The sentinel individuals were negative for IAV nucleoprotein (NP) antibodies and by rRT-PCR (see below) prior to placement in the trap. They were housed in a separate compartment of the trap, but physical contact through mesh was possible and individuals were exposed any droplets generated by wild ducks within the trap, and shared the same water as wild individuals. Cloacal samples or freshly deposited feces were collected daily, from all 10 individuals, between September 24 and December 15, 2009, using a sterile tipped applicator and placed into viral transport media. Samples were stored at −80 °C within 1–5 h of collection.

IAV screening

RNA extraction, real-time reverse transcriptase PCR (rRT-PCR) and isolation were performed as previously described (Spackman et al., 2002), with slight modifications. Briefly, RNA was extracted using the Viral Mini Kit and M48 Robot (Qiagen, Hilden, Germany) following the manufacturer's instructions. Subsequently, samples were assayed for the presence of the IAV matrix gene by rRT-PCR using a TaqMan assay with a threshold cut-off (Ct) of 40 cycles (Spackman et al., 2002). Positive samples were inoculated via the allantoic route into 11 day-old SPF embryonated chicken eggs (Valo, Germany). The allantoic fluid was harvested 2 days after inoculation and assayed for the presence of IAV using hemaglutiniation assays (HA) (Fouchier et al., 2005).

Virus characterization

RNA was extracted from the allantoic fluid using the HighPure RNA isolation kit (Roche, Mannheim, Germany) and cDNA was synthesized using the UNI12 primer (Chan et al., 2006) and the Superscript III First Strand System for Reverse Transcriptase PCR (Invitrogen, Carlsbad, CA). HA and NA subtypes of isolated viruses were determined by PCR (Hoffmann et al., 2001; Orozovic et al., 2010). Viruses were further characterized by PCR using previously published primers (Orozovic et al., 2010; Wille et al., 2011). The amplification products were purified by PEG precipitation; briefly, 60 μl 10% PEG (Sigma, Steinheim, Germany) containing 2.5 M NaCl was added to 20 μl PCR product and centrifuged for 45 min at 4000 rpm. The supernatant was discarded and the pellet was then washed with 60 μl of 70% EtOH (room temperature), followed by centrifugation at 4000 rpm. The supernatant was removed and the pellet was air-dried and resuspended in 30 μl 5 mM TrisHCl (pH 9). HA and NA sequences were gel purified using the Wizard® SV Gel and PCR Clean-Up System (Promega, Madison, WI), and cloned using the pGEM-T Vector System (Promega). Capillary sequencing of purified PCR products, of at least three colonies containing inserts, was carried out at Eurofins MWG Operon (Ebersberg, Germany). Full length sequence of the HA, NP, NA, M, and NS segments were generated, and a shorter fragment of a variable region of each of the PB2 (1–1200 bp), PB1 (533–1262 bp) and PA (747–1919 bp) segments was amplified and subsequently sequenced. To ensure that these short regions of the polymerase segments were representative of full length sequence, pairwise distances and diversity scores were compared using 500 previously published sequences from the Influenza Virus Resource Database (Bao et al., 2008).

Sequence analysis

Sequences were assembled, and MUSCLE (Edgar, 2004) and/or MAFFT (Katoh et al., 2009) alignments generated, within Geneious Pro v5.5.6 (Biomatters, New Zealand). Distance estimates were calculated using MEGA 5.0 (Tamura et al., 2011) and all viruses were genotyped using FluGenome (Lu et al., 2007). Diversity was measured under the Tamura-Nei substitution model for all segments. Phylogenetic models were determined in MEGA 5.0 and/or JModelTest 1.0 (Posada, 2008), and phylogenetic analyses were completed using Mr. Bayes v3.1.2. (Ronquist and Huelsenbeck, 2003). All trees were run for the same number of generations, at which point the most complex tree converged. Consensus trees were then displayed using FigTree v1.3.1 (Rambaut, 2008). The topologies of phylogenetic trees of all 8 segments (Fig. S1, Fig. 2) were compared in every combination using the GiRaF software (Nagarajan and Kingsford, 2010), whereby each segment tree was compared with every other tree, resulting in 28 comparisons. Based on validation experiments with IAV, reassorted viruses are those where isolates are identified in seven of 28 comparisons. The algorithm identifies sets of taxa with different phylogenetic placement across the trees (Nagarajan and Kingsford, 2010). Automated and manual comparisons of phylogenetic trees, while providing information about the presence of reassortment, does not illustrate the patterns of reassortment between viruses, nor adequately identify which viruses are affecting our estimate of reassortment the most. Thus we utilized network analysis to allow for comparisons of constellations, or the combination of sublineages in the segments of interest. Sublineages for genome constellation comparisons were generated by utilizing tree topology generated from bayesian trees, complimented with a differentiation method utilizing pairwise distance estimates and a cut-off value of 97% identity (Fig. S3) (Reeves et al., 2011). Unrooted Minimum Spanning Trees (MSTs) were built for network analysis and used to illustrate relationships between genome constellations, where we artificially linked the internal segments (PB2, PB1, PA, NP, M, NS) to create different unique constellations. Each constellation (ST) is represented by a circle, and the size of the circle corresponds to the number of isolates with that constellation. Constellations are then linked, and a level is computed for each link. Black connectors indicate large similarities between constellations and have low link levels, and gray connectors have fewer similarities with large link levels. The algorithm allows for any possible distance of links, thus at the lowest levels of stringency all constellations will be linked. Stringency settings can be adjusted; at the highest level of stringency, any five of six segments had to be shared across constellations, and at the lowest level of stringency any 1 of the 6 segments had to be shared. Network analysis and figures were generated utilizing the goeBURST v1.2.1 algorithm within PhyloViz v1.0 (Francisco et al., 2012). All sequences generated in this study have been deposited in GenBank under the accession numbers JX565713 to JX566448.

Identifying parental segment constellations, testing for non-random reassortment and linkage between segments

We relied on simulations by permutations of the lineages recorded for each segment, for each observed constellation, to identify the parental constellation (Fig. 5). We hypothesized that the parental combinations should have the highest probability to occur if we permute the lineages of all constellations, simulating a reverse reassortment process. To investigate whether particular reassortment and/or linkage occurred between the different segments in our data, we used a null model approach (Gotelli, 2001; Gotelli and Graves, 1996). This approach relies on random sampling from a known or hypothetical distribution. Elements of the

data are held constant, while others are allowed to vary stochastically to create new assemblage patterns that would occur without a particular ecological mechanism. Utilizing lineage counts derived from genome constellation data, we generated all possible segment combinations under a null hypothesis, and we repeated this operation a sufficient number of times to explore all the possible outcomes and hence to obtain the distribution of the statistic of interest (e.g. reassortment frequency) under the null hypothesis of random reassortment. The observed statistic (e.g. observed reassortment frequency) was then compared to a $1-\alpha$ confidence interval of the statistic under the null hypothesis of random reassortment. The main findings of each test are presented in the main text; the rationale, statistical test, and null hypothesis of each analysis are fully described in the supplemental materials. With the exception of absolute frequencies presented in Tables S3 and S4, we did not include repetitions of a given constellation during a given infection for the analysis. All statistical analyses were completed in the R 2.12 software (R Development Core Team, 2013).

Role of the funding source

This work was supported by the Swedish Research Councils VR (2010-3067, 2010-5399, 2011-3568), FORMAS (2009-1220), and the EC FP6 program NewFlubird. MW was supported by a PGS-D2 fellowship from the Natural Sciences and Engineering Research Council of Canada (www.nserccrsng.gc.ca). The funding sources had no role in study design, in the collection, analysis, and interpretation of data, in the writing of the report, or in the decision to submit the paper for publication.

Acknowledgments

Ethical approval for trapping, sampling, and keeping birds was obtained from the Swedish Animal Research Ethics Board ("Linköpings djurförsöksetiska nämnd", reference number 46-09). We specifically want to thank F. Johnsson and S. Andersson for assistance in sample collection, and A. Jawad for assistance in virus isolation. We thank J. Chapman and anonymous reviewers for insightful comments on the manuscript. This is publication no. 273 from Ottenby Bird Observatory.

Appendix A. Supporting information

Supplementary data associated with this article can be found in the online version at http://dx.doi.org/10.1016/j.virol.2013.05.004.

References

Bahl, J., et al., 2009. Gene flow and competitive exclusion of avian influenza A virus in natual resevoir hosts. Virology 390, 289–297.
Bao, Y., et al., 2008. The Influenza virus resource at the National Centre for Biotechnology Institute for Biotechnology Information. J. Virol. 82, 596–601.
Bataille, A., et al., 2011. Evolutionary analysis of inter-farm transmission dynamics in a highly pathogenic avian influenza epidemic. PLoS Pathog. 7 (6), e1002094. http://dx.doi.org/10.1371/Journal.Ppat.1002094.
Bokhari, S.H., Janies, D.A., 2010. Reassortment networks for investigating evolution of segmented viruses. IEEE Comp. Biol. 7, 288–298.
Bokhari, S.H., et al., 2012. Reassortment networks and the evolution of pandemic H1N1 swine-origin influenza. IEEE/ACM Trans. Comput. Biol. Bioinformatics 9, 214–227.
Chan, C.H., et al., 2006. Amplification of the entire genome of influenza A virus H1N1 and H3N2 subtypes by reverse-transcriptase polymerase chain reaction. J. Virol. Methods 136, 38–43.
Chen, R., Holmes, E.C., 2006. Avian influenza virus exhibits rapid evolutionary dynamics. Mol. Biol. Evol. 23, 2336–2341.
Chen, R., Holmes, E.C., 2009. Frequent inter-species transmission and geographic subdivision in avian influenza viruses from wild birds. Virology 383, 156–161.
Chen, R., Holmes, E.C., 2010. Hitchhiking and the population genetic structure of avian influenza virus. J. Mol. Evol. 70, 98–105.
Downie, J.C., 2004. Reassortment of influeza A viruses genes linked to PB1 polymerase gene. Int. Congr. 1263, 714–718.
Dugan, V.G., et al., 2008. The evolutionary genetics and emergence of avian influenza A viruses in wild birds. PLoS Pathog. 4, e1000076, doi: 1000010.1001371/journal/ppat/1000076.
Edgar, R.C., 2004. MUSCLE: a multiple sequence alignment method with reduced time and space complexity. BMC Bioinformatics 5, 113–132.
Fouchier, R.A.M., et al., 2005. Characterization of a novel influenza A virus hemagglutinin subtype (H16) obtained from Black-headed Gulls. J. Virol. 79, 2814–2822.
Francisco, A.P., et al., 2012. PHYLOViZ: phylogenetic inference and data visualization for seuqnece based typing methods. BMC Bioinformatics 13, 87, http://dx.doi.org/10.1186/1471-2105-1113-1187.
Ghedin, E., et al., 2009. Mixed infection and the genesis of influenza virus diversity. J. Virol. 83, 8832–8841.
Gotelli, N.J., 2001. Research frontiers in null model analysis. Global Ecol. Biogeogr. 10, 337–343.
Gotelli, N.J., Graves, G.R., 1996. Null Models in Ecology. Smithsonian Institution Press, Washington, U.S.A.
Hoffmann, E., et al., 2001. Universal primer set for the full-length amplification of all influenza A viruses. Arch. Virol. 146, 2275–2289.
Hoper, D., et al., 2011. A comprehensive deep sequencing strategy for full-length genomes of influenza A. PLoS One 6, e19075, doi: 19010.11371/journal.pone.0019075.
Katoh, K., et al., 2009. Multiple alignment of DNA sequences with MAFFT. Methods Mol. Biol. 537, 39–64.
Kawaoka, Y., et al., 2005. Orthomyxoviridae. In: Fauquet, C.M., Mayo, M.A., Maniloff, J., Desselberger, U., Ball, L.A. (Eds.), Virus Taxonomy: Eighth Report of the International Committee for the Taxonomy of Viruses. Elsevier Academic Press, San Diego, USA, pp. 681–693.
Khiabanian, H., et al., 2009. Reassortment patterns in swine influenza viruses. PLoS ONE 4, e7366, doi: 7310.1371/journal.pone.007366.
Lam, T.T., et al., 2011. Reassortment events among swine influenza A viruses in China: implications for the origin of the 2009 influenza pandemic. J. Virol. 85, 10279–10285.
Latorre-Margalef, N., et al., 2013. Heterosubtypic immunity to influenza A virus infections in Mallards may explain existence of multpile virus subtypes. PLoS Pathog., In Press.
Latorre-Margalef, N., et al., 2009. Effects of influenza A virus infection on migrating mallard ducks. Proc. R. Soc. B 276, 1029–1036.
Lebarbenchon, C., et al., 2012. Reassortment influenza A viruses in wild duck populations: effects on viral shedding and persistence in water. Proc. R. Soc. B 279, 3967–3975.
Lindstrom, S.E., et al., 2004. Genetic analysis of human H2N2 and early H3N2 influenza viruses, 1957–1972: evidence for genetic divergence and multiple reassortment events. Virology 328, 101–119.
Lu, G., et al., 2007. FluGenome: a web tool for genotyping influenza A virus. Nucleic Acids Res. 35, 275–279.
McDonald, S.M., et al., 2009. Evolutionary dynamics of human rotaviruses: balancing reassortment with preferred genome constellations. PLoS Pathog. 5(10), e1000634. http://dx.doi.org/10.1371/journal.ppat.1000634.
Nagarajan, N., Kingsford, C., 2010. GiRaF: robust, computational identificatin of influenza reassortments via graph mining. Nucleic Acids Res. 39, e34.
Olsen, B., et al., 2006. Global patterns of influenza A virus in wild birds. Science 312, 384–388.
Orozovic, G., et al., 2010. Degenerate primers for PCR amplification and sequencing of the avian influenza A neuraminidase gene. J. Virol. Methods 170, 94–98.
Pereda, A.J., et al., 2008. Avian influenza virus isolated in wild waterfowl in Argentina: evience of a potentially unqie phylogenetic lineage in South America. Virology 378, 363–370.
Posada, D., 2008. jModelTest: phylogenetic model averaging. Mol. Biol. Evol. 25, 1253–1256.
Rabadan, R., et al., 2006. Comparison of avian and human influenza A viruses reveals a mutational bias on the viral genomes. J. Virol. 80, 11887–11891.
R Core Team, 2013. R: A Language and Environment for Statistical Computing. R Foundation for Statistical Computing, Austria. 〈http://www.R-project.org〉.
Ramakrishnan, M.A., et al., 2009. The feasibility of using high resolution genome sequencing of Influenza A viruses to detect mixed infections and quasispecies. PloS ONE 4, e7105, doi: 7110.1371/journal.pone.0007105.
Rambaut, A., 2008. FigTree v1.1.1: Tree figure drawing tool.
Rambaut, A., et al., 2008. The genomic and epidemiological dynamics of human influenza A virus. Nature 453, 615–619.
Reeves, A.B., et al., 2011. Interspecies transmission and limited persistence of low pathogenic avian influenza genomes among Alaska dabbling ducks. Infec. Genet. Evol. 11, 2004–2010.
Ronquist, F., Huelsenbeck, J.P., 2003. MRBAYES 3: Bayesian phylogenetic inference under mixed models. Bioinformatics 19, 1572–1574.
Salomon, R., Webster, R.G., 2009. The influenza virus enigma. Cell 136, 402–410.
Scholtissek, C., et al., 1978. On the origin of human influenza virus subtypes H2N2 and H3N2. Virology 87, 13–20.
Sharp, G.B., et al., 1997. Coinfection of wild ducks by influenza A viruses: distribution patterns and biological significance. J. Virol. 71, 6128–6135.

Sharp, G.B., et al., 1993. Wild ducks are the resevoir for only a limited number of influenza A subtypes. Epidemiol. Infect. 110, 161–176.
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. J Clin. Microbiol. 40, 3256–3260.
Tamura, K., et al., 2011. MEGA5: molecular evolutionary genetics analysis using maximum likelihood, evolutionary distance, and maximum parsimony methods. Mol. Biol. Evol. 28, 2731–2739.
Taubenberger, J.K., et al., 2005. Characterization of the 1918 influenza virus polymerase genes. Nature 437, 889–893.
Tolf, C., et al., 2013. Individual variation in influenza A virus infection histories and long-term immune responses in Mallards. PLoS ONE 8, e61201, doi:61210.61371/journal.pone.0061201.
Varich, N.L., et al., 2008. Deviation from the random distribution pattern of influenza A virus gene segments in reassortants produced under non-selective conditions. Arch. Virol. 153, 1149–1154.
Wallensten, A., et al., 2007. Surveillance of influenza A virus in migratory waterfowl in Northern Europe. Emerg. Infect. Dis. 13, 404–411.
Webster, R.G., et al., 1992. Evolution and ecology of influenza A viruses. Microbiol. Rev. 56, 152–179.
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, doi: 24010.21371/journal.pone.0024010.
Wille, M., et al., 2011. Extensive geographic mosaicism in avian influenza viruses from gulls in the northern hemisphere. PLoS ONE 6, e20664, doi:20610.21371/journal.pone.0020664.
zu Donha, H., et al., 2009. Invasions by Eurasian avian influenza virus H6 genes and replacement of its North American clade. Emerg. Infect. Dis. 15, 1040–1045.