Multiple gene segments are associated with enhanced virulence of clade 2.3.4.4 H5N8 highly pathogenic avian influenza virus in mallards

Christina M. Leyson\textsuperscript{a}, Sungsu Youk\textsuperscript{a}, Helena L. Ferreira\textsuperscript{a,b}, David L. Suarez\textsuperscript{a}, and Mary Pantin-Jackwood\textsuperscript{a}#

\textsuperscript{a} Southeast Poultry Research Laboratory, United States National Poultry Research Center, Agricultural Research Service, United States Department of Agriculture, Athens, Georgia, USA

\textsuperscript{b} Department of Veterinary Medicine, Faculty of Animal Science and Food Engineering, University of São Paulo (FZEA-USP), Pirassununga, São Paulo, Brazil

#Address correspondence to Mary Pantin-Jackwood, mary.pantin-jackwood@usda.gov

Running head: Virulence determinants of H5N8 HPAI virus in mallards

Abstract word count: (<250) 249

Importance word count: (<150) 145

Text word count: 14,137
Abstract

Highly pathogenic avian influenza (HPAI) viruses from the H5Nx Goose/Guangdong/96 lineage continue to cause outbreaks in domestic and wild bird populations. Two distinct genetic groups of H5N8 HPAI viruses, HA clades 2.3.4.4A and 2.3.4.4B, caused intercontinental outbreaks in 2014-2015 and 2016-2017 respectively. Experimental infections using viruses from these outbreaks demonstrated a marked difference in virulence in mallards, with the H5N8 virus from 2014 causing mild clinical disease and the 2016 H5N8 virus causing high mortality. To assess which gene segments are associated with enhanced virulence of H5N8 HPAI viruses in mallards, we generated reassortant viruses with 2014 and 2016 viruses. For single-segment reassortants in the genetic backbone of the 2016 virus, pathogenesis experiments in mallards revealed that morbidity and mortality were reduced for all eight single-segment reassortants compared to the parental 2016 virus, with significant reductions in mortality observed with the PB2, NP, and M reassortants. No differences in morbidity and mortality were observed with reassortants that either have the polymerase complex segments or the HA and NA segments of the 2016 virus in the genetic backbone of the 2014 virus. In vitro assays showed that the NP and PA segments of the 2014 virus lowered polymerase activity when combined with the polymerase complex segments of the 2016 virus. Furthermore, the M segment of the 2016 H5N8 virus was linked to filamentous virion morphology. Phylogenetic analyses demonstrated that gene segments related to the more virulent 2016 H5N8 virus have persisted in the contemporary H5Nx HPAI gene pool until 2020.
Importance

Outbreaks of H5Nx highly pathogenic avian influenza (HPAI) viruses from the Goose/Guangdong/96 lineage continue to occur in many countries and have resulted in substantial impact on wild birds and poultry. Epidemiological evidence has shown that wild waterfowl play a major role in the spread of these viruses. While HPAI virus infection in gallinaceous species causes high mortality, a wide range of disease outcomes have been observed in waterfowl species. In this study, we examined which gene segments contribute to severe disease in mallards infected with H5N8 HPAI viruses. No virus gene was solely responsible for attenuating the high virulence of a 2016 H5N8 virus, but the PB2, NP, and M segments significantly reduced mortality. The findings herein advance our knowledge on the pathobiology of avian influenza viruses in waterfowl and has potential implications in the ecology and epidemiology of H5Nx HPAI in wild bird populations.

Keywords: highly pathogenic avian influenza, H5N8, clade 2.3.4.4, mallards, pathogenicity, virulence.

INTRODUCTION

Avian influenza (AI) viruses are members of the genus Influenza virus A of the family Orthomyxoviridae and are classified into subtypes according to two surface proteins, hemagglutinin (HA) and neuraminidase (NA) (1). HA subtypes H1-16 and NA subtypes N1-9 are all found in wild waterfowl and thus, these species are regarded as natural reservoirs of AI viruses (2). AI viruses are classified as highly pathogenic avian influenza (HPAI) or low
pathogenic avian influenza (LPAI) virus based on either pathogenicity testing in chickens or sequence determination of the HA cleavage site of H5 and H7 subtypes (3). The H5 and H7 subtypes, which are predominantly low pathogenic in the wild bird host reservoir, can acquire a multi-basic cleavage site in the HA protein after circulating in gallinaceous species thereby resulting in an HPAI phenotype. These H5 and H7 HPAI viruses typically cause no or mild disease in mallards under experimental conditions and are rarely transmitted from poultry back to wild birds (4, 5). The major exception to this rule are the H5 subtype viruses from the Goose/Guangdong/96 (Gs/GD) lineage, which have spread back to many wild bird species, with some producing clinical signs and mortality in waterfowl (5). Viruses from this lineage initially emerged in China in 1996 and have caused multiple intercontinental outbreaks in domestic and wild bird populations through dissemination by migratory waterfowl (6).

The Gs/GD lineage H5 HA gene has diversified into multiple phylogenetic clades, with the clade 2 lineage becoming predominant (6). Two major intercontinental outbreaks caused by H5Nx HPAI clade 2.3.4.4 viruses of the Gs/GD lineage occurred in 2014-2015 and 2016-2017 (6-8). Two distinct genetic groups from this clade were involved in these outbreaks, Group A (named 2.3.4.4A) and Group B (named 2.3.4.4B). In the first wave (2014-2015), the clade 2.3.4.4A H5N8 HPAI virus emerged in East Asia and spread to Western Europe and North America (9). Experimental infections using the clade 2.3.4.4A H5N8 viruses showed that these viruses cause mild clinical disease with no or little mortality in domestic and wild ducks (4, 8, 10-17), which is typical of HPAI viruses (1, 18). The second wave of H5N8 outbreaks (2016-2017) involved the clade 2.3.4.4B H5N8 HPAI viruses, which reassorted with Eurasian LPAI viruses (8), generating many reassortants including H5N5 and H5N6 HPAI viruses (19-22). Notably, the
magnitude and breadth of the 2016-2017 outbreak was much larger than the previous 2014-2015 outbreak (7). A total of 48 countries were affected by the end of June 2017 and there were 80 times more poultry outbreaks in Europe for the 2016-2017 season compared to the previous 2014-2015 season (23).

The first detections of the clade 2.3.4.4B H5N8 HPAI viruses occurred near lakes in Russia and China (24-26), similar to earlier intercontinental outbreaks of Gs/GD H5Nx viruses (6, 8). In contrast to the first wave, mass die-offs of wild birds were reported in many affected countries and involved a wide range of species including diving ducks, swans, geese, gulls, terns, penguins, and other water birds (7, 26-29). Consistent with field observations, experimental studies using clade 2.3.4.4B 2016-2017 European H5N8 HPAI viruses resulted in high virus replication and mortality in ducks (30-32).

Investigations on the epidemiology and phylodynamics of the clade 2.3.4.4B viruses revealed that wild birds played a major role in the spread of the H5Nx viruses across various countries (7, 23, 29, 33, 34). Indeed, the potential for efficient transmission of clade 2.3.4.4A and B viruses by ducks to gallinaceous poultry species has been experimentally demonstrated (17, 32). In years following the second wave, continued circulation of clade 2.3.4.4B H5N8 viruses in domestic poultry was observed in Bulgaria (35) and other parts of Eurasia (36). The clade 2.3.4.4B H5N8 and H5N6 viruses were sporadically detected in Europe, East Asia, and Africa from 2018 to early 2020 (7, 33, 37-39). Furthermore, a surge in H5N8 HPAI cases occurred in Europe and Asia in the fall of 2020 (40-42), and outbreaks continue to be reported.

Given the scale of the 2016-2017 H5Nx HPAI outbreak, the continued reemergence of H5Nx viruses, and the role of wild waterfowl in the spread of these viruses, it is important to
identify gene segments that contribute to increased virulence of HSNx viruses in infected waterfowl. In this study, we rescued reassortant viruses using two H5N8 HPAI clade 2.3.4.4 viruses previously shown to have clear differences in virulence in mallard ducks. We subsequently characterized these reassortant viruses in mallards to determine which viral genes are associated with differences in virulence.

RESULTS

M, NP, and PB2 significantly decreased mortality caused by the 2016 H5N8 HPAI virus in mallards. To determine the gene segments that potentially contribute to increased virulence of H5N8 HPAI viruses in mallards, twelve reverse genetics (rg) viruses were made with the A/Tufted duck/Denmark/11740-LWPL/2016 H5N8 HPAI virus (TD16) from the European 2016-2017 outbreak (clade 2.3.4.4B) and the A/Gyrfalcon/Washington/40188-6/2014 H5N8 HPAI virus (GF14) from the North American 2014-2015 outbreak (clade 2.3.4.4A) (Figure 1). The TD16 virus caused 89% mortality in our 2-week-old mallard duck infection model (30), whereas the GF14 virus produced no mortality (4, 43). Eight single-segment reassortant viruses were generated with genes from GF14 in the TD16 backbone to determine which segment/s from GF14 would attenuate the virulent phenotype of TD16 in mallards (Figure 1). Two reassortant viruses containing the TD16 polymerase complex or HA and NA segments in the GF14 backbone were also examined.

No mortality was observed in mallards inoculated with rgGF14 (Table 1, Figure 2), consistent with our previous studies using the parental GF14 virus (4, 43). No mortality was also observed with rgGF14 reassortant viruses with the TD16 polymerase complex, or the HA and
NA segments. On the other hand, inoculation of mallards with rgTD16 resulted in 100% mortality with a mean death time (MDT) of 3.9 days. These findings are similar to our previous study with the parental TD16 virus in mallards, which caused death in 8 out of 9 ducks or 89% mortality, with an MDT of 3.4 days (30); the difference probably due to minor experimental variations since the pathogenicity of the two viruses were tested at different times and using different batches of mallards, and/or to the highly clonal nature of the rescued rgTD16 virus.

The rgTD16 single segment reassortant viruses caused a wide range of mortality (25% - 100%, Table 1). Except for rgTD16/GF14-NA, the groups of mallards inoculated with these viruses had MDTs between 5-6.7 days. Notably, while the MDT for the rgTD16/GF14-NA group (3.8 days) is closest to that of the rgTD16 group (3.9 days), two ducks in this group survived. Using the Mantel-Cox log-rank test, significant differences in mortality were observed with the rgTD16/GF14-PB2, -NP, and -M groups compared to the rgTD16 group, whereas the survival curves of rgTD16/GF14-PB1, -PA, -HA, -NA, and -NS were not significantly different from the rgTD16 group (Table 1, Figure 2). Groups with at least one survivor at the end of the study were rgTD16/GF14-PB2, -PB1, -NP, -NA, and -M. In addition to comparing survival curves, we also compared time to death between groups (Figure 3). Mallards that survived at the end of the study were given 11 dpi as a value for statistical purposes. We found that compared to the rgTD16 group, significantly longer time to death was observed in rgTD16/GF14-NP and -M, and rgGF14 and its reassortants.

The clinical ante mortem signs were similar to what were previously observed in lethal H5Nx HPAI virus infections in ducks (5, 18), which included recumbency, decrease in feed and water consumption, and neurological signs such as tremors, incoordination, and torticollis.
Ducks that survived showed no clinical signs. Body weights at 2-, 4-, and 11-days post inoculation (dpi) from mallards inoculated with rgTD16 and the single segment reassortants were overall lower compared to the rgGF14 group, but only the rgTD16, rgTD16/GF14-PB2, -PA, -HA, and -NA groups were found to be significantly lower at 2 dpi (Figure 4A). No difference in body weights was found among rgGF14, rgGF14 reassortants, and the non-inoculated groups at any time point (Figure 4A-4C). Infection with rgTD16 and the rgTD16 single segment reassortants resulted in an overall increase in body temperatures at 2 dpi when compared to rgGF14, rgGF14 reassortants, and non-inoculated controls (Figure 5); but this increase was only significant in the rgTD16/GF14-PB2 group compared to rgGF14 and non-inoculated controls. Significant differences in body temperatures were also observed between the non-inoculated controls and the rgTD16/GF14-PB2, PB1, PA, and NP groups. There were no differences in both body weights and temperatures when doing pairwise comparisons between the rgTD16 and the rgTD16 reassortant groups and between rgGF14 and the rgGF14 reassortant groups.

Although the PB2, NP, and M reassortants had the highest impact on mortality, the results suggest that multiple genes are involved in the pathogenicity of H5N8 HPAI virus in mallards, with no single segment substitution completely abrogating the mortality caused by rgTD16 in mallards. Moreover, replacing the polymerase complex or the HA and NA segments of GF14 with that of TD16 was not sufficient to increase morbidity and mortality of rgGF14.

Systemic virus replication and cloacal shedding are associated with increased mortality in mallards. To measure viral shedding, oropharyngeal (OP) and cloacal (CL) swabs were collected from virus-inoculated mallards at different timepoints. All virus-inoculated mallards, including...
those that survived infection, shed virus by the OP and CL routes. Regardless of the virus used,
all mallards shed higher levels of viral RNA copies by the OP route than the CL route (Figure 6),
consistent with previous observations with other H5Nx Gs/GD lineage viruses (4, 15, 16, 43).
Mallards inoculated with rgTD16 shed significantly higher levels of viral RNA than mallards
inoculated with rgGF14 by the OP route at 2 dpi (Figure 6). Furthermore, rgTD16/GF14-PB2, -PA,
and -HA, rgGF14, and rgGF14/TD16-pol groups shed significantly lower OP virus RNA levels
at 2 dpi compared to the rgTD16 group. Through the CL route, rgTD16/GF14-PB2, -HA, and -NP,
and rgGF14/TD16-pol groups were found to shed significantly lower viral RNA at 2 dpi
compared to the rgTD16 group. Since only two mallards remained in the rgTD16 group at 4 dpi,
statistical analyses were not performed at subsequent time points.

Examining the trends in viral RNA levels across all groups, a wider variation in viral RNA
levels were observed in CL swabs than in OP swabs. A weak inverse correlation was found
between viral RNA levels in CL swabs and with time of death ($R^2=0.11$, $P<0.05$ using Pearson
correlation). For the surviving mallards, low levels of virus shedding were detected until the
termination of the experiment at 11 dpi (Figure 6).

Prior studies have shown that viral RNA and viral antigen can be detected in the organs
of waterfowl species when an H5Nx Gs/GD lineage virus infection is accompanied by high
mortality (4, 44). To assess virus presence in organs, brain, heart, liver, lung, muscle, and spleen
were collected from two mallards from each virus-inoculated group at 3 dpi. Necropsied
mallards were chosen randomly, unless mallards were exhibiting clinical signs that qualified for
euthanasia. Most necropsied mallards had no clinical signs, except for two mallards inoculated
with rgTD16, rgTD16/GF14-PA, and -NA and one of the mallards inoculated with rgTD16/GF14-
PB1, which showed mild ataxia or were lethargic. Similar gross lesions including enlarged spleen, mottled liver, and empty intestines were observed in all the mallards examined, except in the non-inoculated controls and mallards from the rgGF14, rgGF14/TD16-pol, and rgGF14/TD16-HA.NA groups. Overall, viral RNA levels in tissues from mallards infected with rgTD16 and rgTD16 reассortants were higher than those from mallards infected with rgGF14 and rgGF14 reассortants (Table 2). Although only two ducks were examined for each group, an inverse correlation between viral RNA levels in tissues and MDT of rgTD16 and rgTD16 reассortants was also found ($R^2=0.79$, $P<0.01$ using Pearson correlation), especially for the liver ($R^2=0.59$), spleen ($R^2=0.56$), and lung ($R^2=0.69$).

To further characterize viral spread within the host, immunohistochemistry was performed on a full set of tissues collected from the same necropsied mallards to detect AI virus antigen staining (Table 3). With all virus groups, moderate to high viral antigen staining was present in the epithelium of the nasal turbinates and, with some exceptions, in the trachea, corroborating the tropism of the viruses for the upper respiratory tract. Widespread viral antigen staining was also observed in many other tissues from mallards infected with rgTD16 and rgTD16/GF14-PB1, and -NA. Tissues from mallards infected with rgGF14, rgTD16/GF14-PB2, -HA, -NP, and -M had the lowest overall immunohistochemistry staining/score. These observations further suggest that disease severity is associated with multiple organ failure caused by systemic viral replication.
Filamentous particles are associated with the TD16 matrix segment. The matrix (M) segment is known to be an important determinant of virion morphology (45, 46). Since changing the M segment of TD16 to that of GF14 was implicated in reducing mortality in mallards, we further investigated if there are differences in the virion morphology between TD16 and GF14 viruses. To this end, we performed negative staining and electron microscopy examination on pelleted virus preparations of the rgTD16 and rgGF14 viruses as well as their reciprocal reassortants with the M gene segment switched. The rgTD16 virions had both spherical and filamentous shapes up to 1.6 µm (Figure 7A), whereas mostly spherical particles with a mean diameter of 0.14 µm were found with rgGF14 (Figure 7B). The rgGF14/TD16-M virions were also present in a mixture of spherical and filamentous particles (Figure 7C), but the filamentous virions were markedly shorter, about 0.5-1 µm for most, than those found with the parental rgTD16 virus (Figure 7E). The reciprocal reassortant, rgTD16/GF14-M, resulted in primarily spherical particles with a mean diameter of 0.15 µm (Figure 7D). Virions longer than 0.5 µm were only found in viruses where TD16 M segment was present (Figure 7E). Transmission electron microscopy of liver tissue from a mallard inoculated with rgTD16 likewise showed filamentous particles budding from infected cells (Figure 8). Due to technical difficulties and lower replication in liver, we were unable to visualize virions from the rgGF14-inoculated mallards in this organ.

NP and PA modulate the polymerase activity in the H5N8 viruses. Since the PB2 and NP of GF14 significantly reduced the virulence of the TD16 virus in mallards, and the PB1 and PA of GF14 increased MDT, we used a luciferase reporter system to assess whether the differences in virulence can be attributed to differences in polymerase activity. A previous study has shown
that polymerase activities of European H5N8 virus isolates from 2014-2015 and 2016-2017 were consistent across cells from several avian species including chicken, quail, duck, and geese (47). For this reason, polymerase activity was assayed in chicken fibroblast cells (DF-1) since this cell line is commonly used to examine polymerase activity in avian species (48, 49). The polymerase complex of TD16 had a significantly higher polymerase activity compared to that of GF14 (Figure 9). Significant decreases in polymerase activity were observed when PA or NP segments of GF14 are reassorted into the rest of TD16 polymerase complex, but no effect was found with PB1 and PB2. Moreover, a numerical increase in polymerase activity was observed when the PA and NP segments of TD16 were reassorted into the rest of the GF14 polymerase complex, but the increase was only statistically significant with the PA segment. Interestingly, these differences in polymerase activity do not seem to affect virus growth kinetics in vitro, since the original TD16 and GF14 isolates showed similar growth curves in duck fibroblast cells (Figure 10). Since no large differences in growth kinetics were observed between original TD16 and GF14 isolates, in vitro growth kinetics for reassortant viruses were not performed. Similarly, the same study mentioned above showed that European H5N8 virus isolates from the 2014-2015 and 2016-2017 outbreaks did not have significant differences in growth kinetics in vitro (47).

TD16-like gene segments were found in the genomes of 2017-2020 H5Nx HPAI viruses from Europe and Asia. Previous studies have shown that the TD16-like and GF14-like viruses have distinct genetic origins because of the reassortment among clade 2.3.4.4B viruses and Eurasian LPAI segments (8, 19, 30). Indeed, the sequence identities between the eight gene segments of
TD16 and GF14 were relatively low, ranging from 86.6-95.8% (Table 4), which complicates identifying specific changes associated with increased virulence in mallards. However, some of the changes between the two viruses have been previously reported as having a biological impact in avian or mammalian species (Table 5).

There were several gene constellations that arose during the European 2016-2017 outbreak (21, 22, 50). In our previous study, we determined that all gene segments of TD16 (30) are closely related to the A/wild duck/Poland/82A/2016 H5N8 HPAI virus (21), and belong to the Netherlands cluster I (19) and to the Reassortant 3 group as described elsewhere (22). This genetic lineage of H5N8 viruses was predominant in the 2016-2017 outbreak in Europe (19, 21, 22), though variations within and among other genetic lineages exist. To determine if TD16-like genes continued to circulate in AI viruses found in wild birds and domestic poultry after the 2016-2017 outbreak, we generated phylogenetic trees with a sample of sequences from approximately 18,600 AI virus isolates of any subtype from 2007-2020 and 24 selected clade 2.3.4.4 H5N8 and H5N6 viruses from 2014-2020 (Supplementary Figure 1). Examination of the phylogenetic trees shows that, as expected, certain 2017 H5N8 viruses from European countries such as the Netherlands and Germany clustered with TD16 (19, 21, 50). Later 2018-2020, H5N8 isolates from Europe also clustered with TD16-like viruses and maintained high sequence identity with the TD16 sequence (Figure 11), suggesting that there was continuous circulation of TD16-like viruses from 2016-2020 as previously described in Russia and Bulgaria (35, 51). Notably, a particular virus that was isolated in 2020 from an organ of a domestic goose in Russia had high sequence similarity (≥99%) to TD16 in all segments, suggesting that TD16-like viruses continued to persist in Russia with few genomic changes. Apart from the NA segment, it
was also found that TD16-like gene segments persisted in 2017-2018 H5N6 viruses from the Netherlands, Japan, and Korea; except for an H5N6 virus from Netherlands that had reassortment at the PB2 segment (37, 39, 52). Furthermore, we found that the NP M105V amino acid change found between TD16 and GF14, or a similar change (M105I), were observed at increasing frequency from 2014-2019 (Figure 12). These findings collectively demonstrate the TD16-like viruses, and their gene segments continue to persist and predominate in avian species from Europe and East Asia.

DISCUSSION

A marked difference in pathogenicity in waterfowl has been observed between H5N8 HPAI clade 2.3.4.4 viruses from the 2014-2015 and the 2016-2017 outbreaks. Epidemiological and experimental data have shown that the H5N8 viruses from 2016-2017 caused high mortality in waterfowl, whereas viruses from 2014-2015 caused little to no clinical disease (4, 7, 30, 31). In this study, we sought to determine which virus gene segments are associated with the increased virulence observed with the 2016-2017 H5N8 viruses. We showed that the rgTD16 virus caused 100% mortality in mallards, whereas rgGF14 caused no mortality, as expected based on the pathogenicity of the parental viruses. No mortality was also observed with rgGF14 reassortants containing the TD16 polymerase complex or HA and NA segments. The single segment reassortants with GF14 genes in the rgTD16 backbone also caused mortality in mallards but at different rates (25%-100%), with a significant reduction in mortality observed with the rgTD16/GF14-PB2, -NP, and -M reassortants. Additionally, significant differences in time to death were observed in the rgTD16/GF14-NP, and -M groups compared to the rgTD16
Although infection with seven out of eight the reassortant viruses resulted in longer MDTs compared to rgTD16, none resulted in 0% mortality. These findings suggest that the high virulence in mallards caused by TD16 involves more than one gene segment since none of the reassortants reduced virulence comparable to the 2014 H5N8 virus, which caused no mortality in mallards. Furthermore, the lack of mortality in rgGF14/TD16-pol and rgGF14/TD16-HA.NA reassortants demonstrates that the TD16 polymerase complex or the TD16 HA and NA segments are not sufficient to confer enhanced virulence to GF14 and that the inclusion of other segments would be required. Since the rgGF14/TD16-pol includes TD16 PB2 and NP segments, which have been associated with significant reductions in mortality, the minimum requirements to elicit mortality in the rgGF14 background could presumably involve M and other segments. Previous studies with 2005 Gs/GD H5N1 HPAI viruses have yielded similar results where PB2, PA, HA, NP, and NS were found to be major contributors to high mortality in Pekin ducks, but only the combination of these five segments lead to high mortality similar to the parental H5N1 strain (53). Polymerase complex genes have been shown to be important contributors to virulence of H5N1 viruses (49, 53-62), similar to our findings. Because of limits in the number of experimental groups that could be tested at the same time, we did not examine more reassortants with other gene segment combinations. The identification of the minimum combination of segments to result in enhanced virulence will be an area of future investigation.

While AI virus infections typically cause no or mild disease in waterfowl, viruses from select clades of the Gs/GD H5Nx HPAI lineage occasionally cause high mortality in waterfowl species (18). Examples are clade 2.2 H5N1 viruses from the early 2000s which have
experimentally caused systemic disease and high mortality in several duck species (5). These H5N1 viruses showed extensive viral replication in many tissues and caused neurological signs (18, 31, 63, 64), similar to TD16 and other 2016-2017 H5Nx viruses from Europe (30-32, 65).

Likewise, many cases were reported in which wild clade 2.3.4.4B H5Nx viruses caused death in wild waterfowl (40). It is important to note however, that the clinical presentations during experimental studies are not necessarily the same as what is observed in the field.

Pathogenesis studies are usually performed with birds that are naïve to AI virus infection, and disease outcomes may be different from wild bird counterparts. Indeed, it has been shown that mallards previously infected with an LPAI virus had mild clinical signs and no mortality after an infection with a clade 2.3.4.4B H5N8 virus (63) genetically close to TD16 virus at 98.7-99.9% sequence identity across all segments. Similar cross-protection was also shown with a H5N1 virus infection of LPAI-exposed Canada geese (66). This cross-protection by previous exposure to AI viruses could explain why widespread deaths of mallards were not reported in the 2016-2017 outbreaks. In this scenario, the mallards could have still been infected and subsequently spread virus to susceptible birds.

Examination of tissues from mallards infected with the rgTD16 showed high viral RNA levels in many organs and extensive AI virus antigen staining, whereas tissues from rgGF14 had low viral RNA levels and less viral staining. This extensive systemic virus replication during TD16 infection is consistent with findings of previous pathogenesis studies with the parental TD16 virus (4, 30). Moreover, it has been shown that a 2016 H5N8 virus related to TD16 has higher virion binding to both respiratory and intestinal tissues of wild birds compared to a 2014 H5N8 virus (67), suggesting that 2016-2017 H5N8 viruses appear to have a wide breadth of tissue
tropism. Lower viral RNA levels in tissues were also observed for reassortants with GF14 segments PB2, HA, NP, and M. Altogether, the results show that mallards infected with the rgTD16 reassortants with GF14 segments PB2, NP, and M had the lowest virus antigen staining and viral RNA in tissues, the lowest viral RNA levels in CL swabs, and the lowest mortality. As for the rgTD16/GF14-HA group, the two ducks examined also had low viral RNA and staining in tissues, and mallards in this group also shed significantly lower levels of viral RNA, and had a longer MDT, than rgTD16 group. However, survival between the two groups was not significantly different, both showing 100% mortality. The use of more birds per group could have resulted in a clearer difference between these two groups, but because of experimental constrains, this was not possible. This highlights the importance of measuring more than one indicator of virulence, such as clinical signs, mortality, MDT, virus shedding, and presence of virus in tissues.

The acquisition of a polybasic cleavage site at the hemagglutinin (HA) gene is the most well-known genetic change that leads to systemic replication in gallinaceous species (68). However, the genetic determinants that lead to systemic replication in waterfowl are less known, and it appears that, in addition to the HA polybasic cleavage site, multiple genes are involved (53). In our study, the PB2, NP, and M segments had the most impact on virulence, although all segments had an attenuating effect on the rgTD16 virulent phenotype. Indeed, several point mutations that were previously reported to be associated with shifts in virulence or are in important regions of the viral protein (69), were found in many genes of GF14 (Table 5). Some of these genetic changes include those associated with increased adaptation or virulence in gallinaceous species, which could affect virus adaptation and virulence in mallards
Other changes appear to equally affect the virulence of the virus in chickens and ducks (49, 53, 62, 70, 78-80), and in some cases, the results in ducks were contrary to ours (53, 55). The significance of these genetic changes remains to be determined in consideration of co-occurring changes in other genes.

To uncover the potential mechanisms leading to increased virulence, we also examined the virion morphology of TD16, GF14, and their M segment reassortants and found that the TD16 matrix segment is associated with a more filamentous morphology (Figure 7). The M segment is known to be primarily responsible for determining virion morphology of influenza viruses, as it is involved in virus assembly and budding (45, 46, 81-85). It has also long been observed that laboratory-adapted strains tend to exhibit spherical morphology, whereas filamentous virions are primarily found in wild type viruses isolated from infected animals or during consecutive passages in their natural host. These observations suggest that filamentous virion morphology provides some advantage in vivo (86-89). Indeed, it has been shown that the M segment of the 2009 pandemic H1N1 virus is associated with a filamentous morphology, increased transmission, and increased NA activity (90, 91). A study tracking individual virions has shown that the HA and NA are enriched at opposite poles of filamentous virions (92). This arrangement of surface glycoproteins allows for directional movement of the virions in a sialic-acid rich environment and thus provides a potential selective advantage over spherical particles that randomly move in their environment (92). Another recent study comparing virions with different morphologies has shown that the filamentous virions are more resistant to inhibition of membrane fusion and antibody-mediated virus neutralization (93). Moreover, the investigators of this study proposed that the presence of more HA proteins on filamentous
virions increases the amount of antibody required to neutralize the virus and thereby, providing means to escape pre-existing virus neutralizing antibodies and provide a selective advantage to filamentous virion morphology (93). It may be possible that in the wild bird populations, this mechanism of escape from pre-existing virus neutralizing antibodies in reservoir species drives selection for increased filamentous morphology. In any case, further studies are needed to determine how the changes in the matrix segment impact TD16 virulence and whether pre-existing virus neutralizing antibodies drive selection for filamentous virions in wild bird populations.

Increased virus replication in tissues and increased virulence have been associated to changes in viral polymerase activity (49, 53). We found significant reductions in polymerase activity when the GF14 PA and NP segments were reassorted into the TD16 polymerase complex. On the other hand, we found only significant increase in the reciprocal exchange of TD16 PA into the GF14 polymerase complex, and only a non-significant, numerical increase for the TD16 NP segment. Since a significant reduction in mortality was observed in the rgTD16/GF14-NP group, these results suggest that the reduction in polymerase activity due to the GF14 NP segment could be associated with a reduction in mortality. However, while GF14 PA reduced polymerase activity, the rgTD16/GF14-PA group had 100% mortality, although a longer MDT than rgTD16. Furthermore, we found that, while the rgTD16/GF14-PB2 group had a significant reduction in mortality, the GF14 PB2 segment did not have an impact in the polymerase activity of TD16. These mixed observations suggest that polymerase activity may not be the only factor that contributes to the high virulence of TD16, and that other factors such as host protein interactions and packaging of viral ribonucleoprotein complexes possibly
play a role (94, 95). In a previous study by Vigeveno et al., polymerase assays were also conducted to compare European H5N8 viruses from 2014-2015 and 2016-2017. In contrast to the result shown here, the European 2014-2015 viruses generally had higher polymerase activities than 2016-2017 viruses though both had similar in vitro growth kinetics (47). A study examining the origins of GF14 showed that part of its segments came from Eurasian LPAI viruses and are distinct from certain lineages of Gs/GD H5Nx viruses circulating in Europe during the same time (96); thus, this may account for differences and variations of polymerase activity between GF14 and European 2014-2015 H5N8 viruses, which were circulating at the same time. This observation is consistent with growth curves generated using the original TD16 and GF14 viruses (Figure 10), further suggesting that polymerase activity and virus replication kinetics in vitro may not always be linked.

There are only five amino acid differences between the NP proteins of TD16 and GF14, with two of them (M105V and K452R) previously shown to be associated with pathogenesis in avian species (62, 71, 76, 78-80)(Table 5). To determine if the amino acids at these positions have changed through time, we examined NP protein sequences of Gs/GD clade 2.3.4.4 viruses from 2014-2019. We found that the amino acid usages were stable during this period, except for position 105 and 452. Additional studies are needed to determine the impact of the M105V and K452R changes on virus replication and/or virulence in mallards and other avian species.

To determine if the constellation of TD16-like gene segments continued to circulate in the AI virus gene pool, we conducted phylogenetic analyses using influenza A viruses of all subtypes from 2007-2020 that were isolated from avian species. We found that the TD16 gene constellation was still maintained in recent H5N8 isolates from Russia and other European
countries. Furthermore, the TD16 constellation remained consistently intact, although reassortment was present in some isolates. These findings are consistent with previous studies of isolates from Russia, Iraq, Kazakhstan (36), and Bulgaria (35) where clade 2.3.4.4 H5N8 viruses were found to be maintained with very little evidence for reassortment. Likewise, in 2017-2018 H5N6 viruses, TD16-like segments were found in some isolates from the Netherlands, Korea, and Japan, with reassortment at the N6 segment and, for the Netherlands isolate, also at the PB2 segment. Previous studies have in fact shown that H5N6 virus isolates from the Netherlands, Nigeria, Japan, Korea, and Taiwan were a result of reassortments of European LPAI and European 2016-2017 H5N8 virus strains (34, 52, 97-100). Many of the H5N6 isolates from the Netherlands (52), Korea (101), and Japan (97) were isolated from either domestic ducks or wild birds. Moreover, a phylodynamic study proposed that domestic ducks were the primary drivers of sustained transmission of H5N8 viruses in Bulgaria (35). Since TD16 was shown in our previous study to be well-adapted to mallards compared to chickens (30), the persistence of TD16-like segments in the contemporary H5 HPAI outbreaks may be due to the high adaptation of the virus to domestic and wild ducks, all the while remaining highly pathogenic to gallinaceous species. These observations also suggest that TD16-like segments have some selective advantage in certain avian species, particularly in ducks. Indeed, a large phylogeographic analysis of clade 2.3.4.4 H5Nx 2016-2017 in Eurasia showed that among the seven main reassortant viruses that originated from the precursor H5Nx clade 2.3.4.4B virus, the M and NS segments were maintained alongside the HA segment, whereas other segments were reassorted more than once (22). These studies further support the hypothesis that the M
segments of TD16-like viruses, in conjunction with the H5 segment, appear to have a complementary selective advantage from an evolutionary perspective.

In summary, we have used a reverse genetics approach to determine which gene segments are associated with the severe clinical disease caused by a European clade 2.3.4.4B H5N8 virus from 2016. We have shown that multiple gene segments are involved in the mortality caused by TD16 infection and that high mortality is correlated to extensive systemic viral replication. The severe disease outcome can be greatly attenuated by substitution of the PB2, NP, and M segments, but an effect on virus replication was also observed with the HA segment, and on polymerase activity with the PA segment. However, the TD16 HA and NA segments, and polymerase complex, were not sufficient to confer an increase in virulence to rgGF14. We also found that the TD16 NP and PA segments can affect polymerase activity in avian cells, and that the TD16 M segment can confer a filamentous virion shape. The precise mechanisms that cause the increased virulence in mallards of the European clade 2.3.4.4B H5N8 viruses from 2016-2017 are still widely unknown and warrant further investigation. Although the pathogenesis of the H5N8 HPAI viruses will vary between different wild bird species, results obtained in this study in mallards may explain the increased mortality observed in other wild bird species with the 2016-2017 outbreaks and with the currently circulating H5Nx HPAI viruses. Given that gene segments from European H5N8 viruses from 2016-2017 were found in the gene pool of succeeding H5Nx viruses in 2017-2020, it is important to further examine the unique properties of the Gs/GD H5Nx viruses containing these segments, which potentially could help predict disease outcomes through the examination of contemporary viral
Uncovering the mechanisms behind disease progression will also further our understanding of avian influenza in its reservoir hosts and its ecology on a larger scale.

**MATERIALS AND METHODS**

**Generation of reassortant viruses.** The viruses used for this study were generated using an eight plasmid system with bi-directional RNA polymerase promoters (pCMV-BDRG) (49). We used two H5N8 HPAI viruses that were previously shown to have clear differences in pathogenicity in different duck species (4, 30, 43). Recombinant viruses were generated using the A/Tufted duck/Denmark/11740-LWPL/2016 H5N8 HPAI virus (TD16) from the European 2016-2017 clade 2.3.4.4B and the A/Gyrfalcon/Washington/40188-6/2014 H5N8 HPAI virus (GF14) from the clade 2.3.4.4A in the North American 2014-2015 outbreak. TD16 was originally isolated from a dead tufted duck at a lake in central Copenhagen, Denmark, on November 2016 (30), while GF14 was isolated from a dead captive-reared gyrfalcon that was fed American wigeon in the State of Washington, United States, on December 2014 (102). These two viruses are representative of H5N8 viruses circulating in wild birds at the time. Reassortants using TD16 and GF14 segments were also rescued, namely, eight single segment H5N8 reassortants in the TD16 backbone, and two reassortants with TD16 HA and NA segments or TD16 polymerase complex in the GF14 backbone (Figure 1). Briefly, RNA from two viruses was extracted and used for reverse transcriptase PCR (RT-PCR) using modified gene-specific influenza primers adapted from Hoffman et al, 2002 (103). Primer sequences are available upon request. Whole genome segments were cloned into pCMV-BDRG using NEBuilder® HiFi DNA Assembly kit (New England Biolabs, Ipswich, Massachusetts). NEB® Stable competent *E. coli* (New England Biolabs) was
used for transformation and grown at 30°C for gene segments that exhibited genetic instability in *E. coli*. All sequences were verified using Sanger sequencing and were all consistent with previously published sequences of the TD16 and GF14 viruses, except for GF14 PB1 gene which had three synonymous mutations C36T, C48T, and G54A. GenBank accession numbers corresponding to the sequences of the viruses used are MN708198 - MN708205 for TD16 and KP307981- KP307988 for GF14. Eight plasmids corresponding to the gene segments in Figure 1 were then transfected into HEK 293 cells using TransIT®-LT1 Transfection reagent (Mirus Bio, Madison, Wisconsin). One day post-transfection 10⁶ DF1 cells were added to the transfection reaction. At 48 hours post-transfection, conditioned cell culture media was inoculated into 5-6 embryonating chicken eggs (ECE). Embryo mortality was observed for up to 4 days post-inoculation (dpi). Chorioallantoic fluid was collected from each ECE and tested for the presence of influenza A virus using a hemagglutination test (104). Allantoic fluid from all HA positive eggs were pooled, filter-sterilized, aliquoted, and stored at -80°C until further use. A second passage of rgTD16, rgGF14, rgTD16/GF14-M, and rgGF14/TD16-M viruses was also prepared to obtain enough volume for electron microscopy. All experiments with the viable viruses were performed in biosafety level-3 enhanced facilities in accordance with procedures approved by the United States National Poultry Research Center (USNPRC) Institutional Biosecurity Committee.

**Animal experiments.** To examine the pathogenicity of the reassortant viruses we followed our mallard duck model (4, 5, 105). Mallards (*Anas platyrhynchos*) are used to examine avian influenza infections in wild waterfowl since they are one of the species most commonly infected and have been demonstrated to have an important role in virus spread (4, 106-109). Mallard
ducklings were purchased at day-of-age from a commercial vendor and reared at the SEPRL, USNPRC, for two weeks. Sera were collected from 15 randomly selected ducks before virus challenge and subsequently screened for influenza A antibodies using FlockCheck Avian Influenza MultiS-Screen Antibody Test® ELISA kit (IDEXX Laboratories, Westbrook, ME, USA). After virus challenge, each experimental group of mallards was housed in individual isolation units with temperature control, negative air pressure, and HEPA-filtered inlet and exhaust in a BSL-3E facility. This study and all associated procedures were reviewed and approved by the USNPRC Institutional Animal Care and Use Committee.

Animal experiment design. Two recombinant parental viruses and 8 reassortant viruses were examined in this study (Figure 1). The viruses were diluted in brain heart infusion (BHI) media to achieve $10^6$ EID$_{50}$ (50% egg infectious dose) in 0.1 mL. Ten mallards from each group were inoculated with 0.1 mL of the appropriate inoculum via the choanal cleft. A control sham group was inoculated with 0.1 mL of the appropriate inoculum via the choanal cleft. A control sham group was inoculated with 0.1 mL of 1:300 (v/v) sterile allantoic fluid: BHI broth mixture. Mallards were observed for clinical signs until 11 dpi. Cloacal temperature was measured at 2 dpi, and body weights at 2, 4 and 11 dpi. Oropharyngeal and cloacal swabs were taken at 2, 4, 7, and 11 dpi. Each swab was placed into 1.5 mL of BHI media supplemented with penicillin (2000 units/mL; Sigma Aldrich, St. Louis, MO, USA), gentamicin (200 μg/mL) and amphotericin B (5 μg/mL). All swabs were stored at -80°C until further processing. At 3 dpi, two mallards per group were euthanized and necropsied, and brain, heart, liver, lung, muscle, and spleen tissues were collected and stored at -80°C for further processing. A full set of tissues were also collected from the same two mallards in 10% neutral-buffered formalin for histopathological examination and immunohistochemical virus antigen detection as previously described (110).
Ducks that exhibited severe neurological signs, did not eat or drink, and/or remained recumbent were euthanized. At 11 dpi, all surviving mallards were weighed, bled, and euthanized.

**Viral RNA quantification for swabs and tissues.** Oropharyngeal and cloacal swab samples were processed using the MagMAX™-96 AI/NDV Viral RNA Isolation Kit® (Ambion Inc/Thermo Fisher Scientific; Grand Island, NY, USA) according to the manufacturer’s recommendations. Tissue samples were homogenized in BHI to achieve a 10% (w/v) homogenate using appropriate sterile nuclease-free beads (OPS Diagnostics, Lebanon, NJ) and the Fast Prep-24™ dissociation instrument (MP Biomedicals, Irvine, CA, USA). RNA was extracted from the tissue homogenates using Trizol LS Reagent (Thermo Scientific) and chloroform as per manufacturer protocol. The RNA in aqueous phase was purified using RNA Clean and Concentrator kit (Zymo, Irvine, CA, USA), quantified by spectrophotometry, and subsequently normalized to 50 ng/µL. Virus quantification from both swab and tissue samples was performed by quantitative real-time reverse transcriptase PCR (qRRT-PCR) following a standard influenza A matrix protocol (111). AgPath-ID one-step RT-PCR kit (Ambion/Thermo Fisher) was used for qRRT-PCR reactions. All primer sequences and qRRT-PCR protocol are available upon request. An internal control, which consists of a primer-probe and RNA template set, was also added to determine if the qRRT-PCR reaction was free from inhibitors. The viral copy numbers in swab and tissue samples were interpolated by using a standard curve correlating with a known number of matrix gene copies of the bidirectional plasmid used for the virus rescue. Results are reported as log_{10} copy number per mL for swabs and log_{10} copy number per g for tissues. For statistical purposes, samples that
did not contain detectable viral RNA were given a value of 1 log_{10} copy number per mL/g, which 
was the limit of detection for the test.

**Viral ribonucleoprotein (vRNP) reconstitution assay.** A reporter system for assaying 
polymerase activity was performed as previously described (49, 112). Briefly, chicken fibroblast 
cells (ATCC DF1) were transfected using TransIT®-LT1 Transfection reagent (Mirus Bio) with the 
equal amounts of the following plasmids: 1) pCMV-BDRG encoding polymerase genes PB2, PB1, 
PA, and NP as indicated in figure legends, 2) pMACkv-Gluc (NS) reporter plasmids, and 3) pCMV-
SEAP transfection control plasmid. At 24 hours post-transfection, conditioned media was 
harvested from transfected cells and assayed for luciferase activity using Pierce™ Gaussia 
Luciferase Flash Assay Kit (ThermoFisher) and secreted alkaline phosphatase activity using 
NovaBright™ Phospha-Light™ EXP Assay Kit for SEAP Reporter Gene Detection (Thermo Fisher) 
with Synergy HT reader (Bio-Tek, VT, USA) as per the manufacturers' recommended protocols. 
Relative luciferase units are reported as the luciferase signal divided by the positive control 
signal from secreted alkaline phosphatase (SEAP) activity. Polymerase assays were done for 
three independent experiments with eight replicates per experiment. Statistics were calculated 
for each independent experiment.

**Virus electron microscopy examination.** Fourteen milliliters of ECE chorioallantoic fluid were 
harvested for each of the following viruses: rgGF14, rgTD16, rgTD16/GF14-M, and 
rgGF14/TD16-M. A pre-centrifugation step was performed at 170 x g for 30 minutes at 4°C. The 
samples were then placed into an ultracentrifugation tube and bulked to full with cold sterile 
phosphate-buffered saline (PBS). Ultracentrifugation was then carried out at 72,000 x g for 1.5 
hours at 4°C. The supernatant was discarded, and the pellet was resuspended in 400 µL of cold
sterile PBS. Formalin was then added to the sample at a final concentration of 0.02%. The
samples were then incubated for 18 hours at 37°C and tested for complete inactivation via
inoculation of ECE. The samples were then treated with formaldehyde and glutaraldehyde both
at a final concentration of 2% and stored at 4°C overnight. Preparations were subsequently
processed for electron microscopy using a standard negative staining protocol with OsO₄ (113)
at the Georgia Electron Microscopy Laboratory in Athens, GA. At the same laboratory,
immunohistochemistry slides were used as a guide to identify areas for transmission electron
microscopy. Areas which have viral antigen staining were cored out of the formalin-fixed
paraffin embedded blocks from the cut surface. The thin sections were then subjected to heat
and xylene to remove paraffin, and rehydration through a series of ethanol (95-30%) washes.
The sections were then stained with 1% OsO₄ and subjected to embedding on Epon-Araldite
resin mixture using standard procedures. Samples were viewed with a JEOL JEM1011 (JEOL
USA, Inc, Peabody, MA) transmission electron microscope at HV= 100kV. To compare virion
morphologies, the length of virions along the longest axis were measured using ImageJ (114).
The number of virions used were \( \text{rgGF14}=106, \text{rgTD16}=96, \text{rgTD16/GF14-M}=101, \text{rgGF14/TD16-M}=152 \)

**Virus growth kinetics.** Duck embryo fibroblasts (ATCC CCL-141) were counted using a
hemocytometer, and replicate wells were infected with TD16 or GF14 viruses at a multiplicity of
infection of 0.01. Viruses were absorbed for 45 minutes and subsequently washed with fresh
cell culture media. Conditioned cell culture media was harvested at 6-,12-,24-, and 72- hours
post-inoculation. Viral titers were measured using the plaque assay method with chicken
fibroblast cells (ATCC DF-1) using a semi-solid overlay method as previously described (115).

Virus growth assays were done in triplicate wells.

**Statistical analyses.** All statistical tests and figures with numerical data were generated using

Graphpad Prism (San Diego, CA, USA). The statistical tests performed were 1) for survivorship,

Log-rank (Mantel-Cox) test with a Bonferroni correction for multiple comparisons; 2) for body

weight, body temperature, time to death, and virus shedding, Kruskal-Wallis test with Dunn’s

multiple comparisons test; 3) for polymerase assays and virion lengths, one-way ANOVA with

Tukey’s multiple comparisons test. A $p$ value of 0.05 was used for all statistical calculations.

EID$_{50}$ was calculated using the Reed-Muench method (116). For calculations of percent

mortality, time to death, and mean death time, the two mallards that were euthanized for

necropsy at 3 dpi were excluded. For viral shedding, mallards that did not shed detectable viral

RNA were given the value 1 log$_{10}$ EID$_{50}$, which was the limit of detection for the qRRT-PCR assay.

**Phylogenetic analyses.** To test whether TD16-like virus gene sequences persisted in the AI virus

gene pool after 2016, phylogenetic trees were constructed with sequences of isolates from

2007-2020 using an adapted method (34, 97). Influenza A viruses isolated from avian species on

January 1, 2007 to December 31, 2020 from any location were obtained from GISAID database

(https://www.gisaid.org/). This search resulted in over 18,600 isolates with 148,800 segments.

Selected Gs/GD lineage H5 viruses from 2005 and 2014-2016 were also included in the

phylogenetic analysis (Table 6). All strain names and accession numbers are found in

Supplementary Table 1. Nucleotide sequences that had more than 1% of ambiguous bases

and/or were less than 80% of the full-length segment were excluded from the analyses. For

each segment, sequences were down-sampled using CD-HIT-EST (117) with 90% identity
threshold to reduce the dataset to 30-50 representative sequences. To capture the more recent
strains, the dataset for H5N8 viruses in 2018-2020 was down-sampled at 99% identity using CD-
HIT-EST. These two down-sampled datasets and selected clade 2.3.4.4 H5N8 viruses from 2014-
2016 (Supplementary Table 1, Strains of interest) were aligned using MAFFT (118).

Subsequently, Maximum Likelihood phylogenetic trees were created using RaxML (119) as
implemented within Geneious Prime 2019.2.3 (Biomatters Ltd; Auckland, New Zealand) and
CIPRES Science Gateway (120). Rapid bootstrapping was utilized and halted automatically by
RaxML. Phylogenetic trees were visualized and annotated using Geneious. Pairwise distances
were calculated between TD16 and H5N8 isolates from 2018-2020 using MAFFT.

Amino acid usage in NP protein sequence. TD16 and GF14 differ at five amino acid positions in
NP. To determine if the amino acid present in these positions changed over time, we collected
NP protein sequences of Gs/GD clade 2.3.4.4 viruses in 2014-2019 from the Influenza Research
were downloaded for 2014-2019 from each year, respectively. Thus, there is a total of 953 NP
sequences for this analysis. At the time of writing this manuscript, only two full-genome
sequences from 2020 were available and were thus not included in the analysis. An alignment
on all sequences was then performed using MUSCLE as implemented on the IRD multiple
sequence alignment (MSA) service (121). The amino acids present at positions 105, 286, 353,
452, and 492 were then tabulated and the percentage of usage were calculated using the total
number of sequences examined and the number of sequences that carry the indicated amino
acid at each position.
Data availability. Sequences of the reassortant viruses generated in this study were confirmed by high throughput sequencing as previously described (30). Each segment of the reassortant virus was identical to the original virus, except for GF14 PB1 gene, which had three synonymous mutations, as disclosed in GenBank database (Accession number: MN708198 - MN708205 for TD16; and KP307981- KP307988 for GF14).

ACKNOWLEDGEMENTS

We gratefully acknowledge the technical assistance from Scott Lee, Jesse Gallagher, Nikolai Lee, Ricky Zoller, Dr. Karen Segovia, and Suzanne Deblois, and animal care assistance from Seth Lee, Roger Brock, and Tarrah Bigler. We would also like to thank the technical assistance of Mary Ard for electron microscopy, and Dr. Matthew Angel at the Laboratory of Viral Diseases, NIAID, NIH, Bethesda, MD, and Dr. Daniel Perez at the Poultry Diagnostic and Research Center in the University of Georgia, Athens, GA for providing the luciferase reporter plasmids. We would also like to acknowledge the GISAID contributors of the sequences used in this study and their names and affiliations are found in Supplementary Table 1. This research was supported by the United States Department of Agriculture (USDA), Agricultural Research Service (ARS) Project 6612-32000-066-00D and by the USDA/ARS- Animal and Plant Health Inspection Service (APHIS) Interagency Agreement #60-6040-6-005. This research was also supported in part by an appointment to the ARS Research Participation Program administered by the Oak Ridge Institute for Science and Education (ORISE) through an interagency agreement between the U.S. Department of Energy (DOE) and USDA. ORISE is managed by ORAU under DOE contract number DE-SC0014664. All opinions expressed in this paper are the authors’ and do not necessarily reflect the policies and views of USDA, DOE, or ORAU/ORISE. Mention of trade
names or commercial products in this publication is solely for the purpose of providing specific information and does not imply recommendation or endorsement by the USDA. The USDA is an equal opportunity provider and employer.


92. Vahey MD, Fletcher DA. 2019. Influenza A virus surface proteins are organized to help penetrate host mucus. eLife 8:e43764.


influenza viruses isolated from wild birds and a chicken in Japan during winter 2017-2018.


Figure legends

Figure 1. Viruses generated by reverse genetics using two H5N8 clade 2.3.4.4 viruses, the
A/Tufted duck/Denmark/11740-LWPL/2016 H5N8 HPAI virus (TD16) and the

Figure 2. Survival curves for mallards inoculated with the rgH5N8 viruses. Shown are rgTD16
reassortant groups that had (A) 100% mortality, (B) 75-88% mortality, and (C) 25-75% mortality.
Three separate plots for survival data were created for clarity. All groups were statistically
compared to rgTD16 group using the log-rank Mantel Cox test. An asterisk (*) beside the group
name indicates $P < 0.05$. The rgTD16, rgGF14, rgGF14-TD16/pol, and rgGF14-TD16/HA.NA
groups are shown in all plots for reference.

Figure 3. Time to death for mallards inoculated with rgH5N8 viruses. Mallards that survived by
the end of the study (11 dpi) were given a value of 11 days. Groups that share at least one letter
superscript indicate that no significant statistical differences were found in pairwise
comparisons ($P \geq 0.05$).

Figure 4. Body weights at (A) 2, (B) 4, and (C) 11 days post-inoculation of mallards inoculated
with the rgH5N8 viruses. Groups that share the same superscript letter are not significantly
different ($P < 0.05$). No statistical comparisons were made at 4 and 11 dpi due to the number of
ducks left in the rgTD16 group. Hash symbol (#) indicates that no ducks were left in the group at
the time of data collection.
Figure 5. Body temperatures at two days post-inoculation of mallards inoculated with the rgH5N8 viruses. Groups that share at least one letter superscript indicate that no significant statistical differences were found in pairwise comparisons ($P \geq 0.05$).

Figure 6. Oropharyngeal and cloacal virus shedding detected by qRRT-PCR from mallards inoculated with the rgH5N8 viruses. Swab type and days post-inoculation (dpi) are indicated on top of each graph. Error bars are omitted for groups with less than three mallards left. Groups that share at least one letter superscript indicate that no significant statistical differences were found in pairwise comparisons ($P \geq 0.05$). For 4 dpi and onwards, there were not enough ducks left in the rgTD16 group to compare to other groups. The hash symbol (#) indicates that no mallards were left at the indicated group and time point. The limit of detection (1 log$_{10}$ EID$_{50}$), which is indicated by the dotted line, was given to mallards that shed no detectable virus.

Figure 7. Electron microscopy photographs of negatively stained rgTD16 and rgGF14 virions (A and B), as well as their M reassortants (C and D). The black scale bar represents 200 nm. The lengths of about 100 virions from each virus were measured from electron micrographs (E). Groups that share a letter superscript indicate that no statistical differences were found ($P \geq 0.05$).

Figure 8. Transmission electron microscopy of virions from the liver of a mallard infected with rgTD16 at 6,000x (A) and 10,000x (B). The black bar represents 600 nm. Black arrows indicate virions budding from the surface of an infected cell.

Figure 9. Polymerase activity of rgH5N8 viruses. Single segment reassortments in the polymerase complex were tested as indicated in the legend below the x-axis. Blue boxes
represent segments from TD16 while red boxes represent segments from GF14. Assays were performed on chicken DF1 cells using the reporter plasmids with an avian RNA polymerase I promoter. Groups that share at least one letter superscript indicates that no significant statistical differences were found ($P \geq 0.05$). Shown here are representative data from three independent experiments.

**Figure 10.** Growth curves of TD16 and GF14 in duck cells. Cells were infected at an MOI of 0.01.

**Figure 11.** Pairwise sequence identity of H5N8 HPAI viruses from 2018 to 2020 and the TD16 virus.

**Figure 12.** Nucleoprotein amino acid usage in Gs/GD clade 2.3.4.4. viruses from 2014-2019 at positions (A) 105, (B) 286, (C) 353, (D) 452, and (E) 492. TD16 and GF14 were found to differ at these five positions. Blue bars represent amino acid residues found in TD16, whereas red bars represent those in GF14. Light blue bars represent amino acid residues that are similar to those from TD16. Green and purple bars represent amino acid residues that are not similar to those found in either TD16 or GF14. For B, the S286 is at low frequency (0.41% in 2014) and thus cannot be visualized well.
Table 1. Mortality and mean death time for mallards inoculated with the indicated H5N8 HPAI viruses. Mallards that were necropsied at 3 dpi were excluded from calculations for % mortality and mean death time. *Indicates significant differences when comparing reassortant groups to the rgTD16 group. Log-rank (Mantel-Cox) test with Bonferroni corrected P-value threshold for multiple comparisons. n/a = not applicable.

<table>
<thead>
<tr>
<th>Virus name</th>
<th>Mean death time (days)</th>
<th>%mortality</th>
<th>Adjusted P value</th>
</tr>
</thead>
<tbody>
<tr>
<td>rgTD16</td>
<td>3.9</td>
<td>100%</td>
<td>n/a</td>
</tr>
<tr>
<td>rgTD16/GF14-PB2</td>
<td>5.5</td>
<td>75%</td>
<td>0.0152*</td>
</tr>
<tr>
<td>rgTD16/GF14-PB1</td>
<td>5.1</td>
<td>88%</td>
<td>0.1776</td>
</tr>
<tr>
<td>rgTD16/GF14-PA</td>
<td>5.7</td>
<td>100%</td>
<td>0.4696</td>
</tr>
<tr>
<td>rgTD16/GF14-HA</td>
<td>6.0</td>
<td>100%</td>
<td>0.3120</td>
</tr>
<tr>
<td>rgTD16/GF14-NP</td>
<td>6.7</td>
<td>38%</td>
<td>&lt;0.0008*</td>
</tr>
<tr>
<td>rgTD16/GF14-NA</td>
<td>3.8</td>
<td>75%</td>
<td>1.000</td>
</tr>
<tr>
<td>rgTD16/GF14-M</td>
<td>6.0</td>
<td>25%</td>
<td>&lt;0.0008*</td>
</tr>
<tr>
<td>rgTD16/GF14-NS</td>
<td>5.7</td>
<td>100%</td>
<td>0.4536</td>
</tr>
<tr>
<td>rgGF14</td>
<td>n/a</td>
<td>0%</td>
<td>n/a</td>
</tr>
<tr>
<td>rgGF14/TD16-pol</td>
<td>n/a</td>
<td>0%</td>
<td>n/a</td>
</tr>
<tr>
<td>rgGF14/TD16-HA.NA</td>
<td>n/a</td>
<td>0%</td>
<td>n/a</td>
</tr>
</tbody>
</table>
Table 2. Viral RNA levels in tissues of mallards inoculated with the indicated rgH5N8 virus. Data from two birds per group necropsied at 3 days post-inoculation. Virus RNA levels were determined by qRRT-PCR and are expressed as log_{10} copy numbers per gram of tissue. *Tissue unavailable. Cells shaded in darker green indicates higher viral RNA levels than cells in lighter color.

<table>
<thead>
<tr>
<th>Virus name</th>
<th>Brain 1</th>
<th>Brain 2</th>
<th>Heart 1</th>
<th>Heart 2</th>
<th>Liver 1</th>
<th>Liver 2</th>
<th>Lung 1</th>
<th>Lung 2</th>
<th>Muscle 1</th>
<th>Muscle 2</th>
<th>Spleen 1</th>
<th>Spleen 2</th>
</tr>
</thead>
<tbody>
<tr>
<td>rgTD16</td>
<td>8.3</td>
<td>8.3</td>
<td>8.2</td>
<td>7.3</td>
<td>7.1</td>
<td>8.4</td>
<td>7.9</td>
<td>7.3</td>
<td>7.9</td>
<td>6.9</td>
<td>8.0</td>
<td>5.5</td>
</tr>
<tr>
<td>rgTD16/GF14-PB2</td>
<td>7.8</td>
<td>6.0</td>
<td>4.0</td>
<td>3.7</td>
<td>3.6</td>
<td>4.8</td>
<td>6.6</td>
<td>4.6</td>
<td>5.4</td>
<td>4.8</td>
<td>3.7</td>
<td>3.4</td>
</tr>
<tr>
<td>rgTD16/GF14-PB1</td>
<td>6.6</td>
<td>8.2</td>
<td>7.1</td>
<td>7.3</td>
<td>8.0</td>
<td>8.3</td>
<td>7.2</td>
<td>7.7</td>
<td>6.6</td>
<td>6.7</td>
<td>6.3</td>
<td>7.1</td>
</tr>
<tr>
<td>rgTD16/GF14-PA</td>
<td>8.3</td>
<td>8.1</td>
<td>7.0</td>
<td>5.8</td>
<td>6.7</td>
<td>4.4</td>
<td>6.8</td>
<td>6.4</td>
<td>6.3</td>
<td>5.2</td>
<td>6.3</td>
<td>4.0</td>
</tr>
<tr>
<td>rgTD16/GF14-HA</td>
<td>*</td>
<td>4.9</td>
<td>5.6</td>
<td>5.6</td>
<td>4.9</td>
<td>4.6</td>
<td>5.0</td>
<td>4.7</td>
<td>6.1</td>
<td>5.8</td>
<td>3.9</td>
<td>3.0</td>
</tr>
<tr>
<td>rgTD16/GF14-NP</td>
<td>7.6</td>
<td>8.1</td>
<td>5.6</td>
<td>4.2</td>
<td>5.2</td>
<td>5.5</td>
<td>5.0</td>
<td>5.9</td>
<td>6.2</td>
<td>6.3</td>
<td>5.0</td>
<td>4.9</td>
</tr>
<tr>
<td>rgTD16/GF14-NA</td>
<td>7.1</td>
<td>7.6</td>
<td>7.2</td>
<td>7.4</td>
<td>8.2</td>
<td>8.0</td>
<td>7.8</td>
<td>7.7</td>
<td>6.0</td>
<td>6.6</td>
<td>7.0</td>
<td>6.9</td>
</tr>
<tr>
<td>rgTD16/GF14-M</td>
<td>6.5</td>
<td>6.0</td>
<td>6.8</td>
<td>7.2</td>
<td>4.9</td>
<td>5.1</td>
<td>6.2</td>
<td>6.4</td>
<td>6.8</td>
<td>6.6</td>
<td>4.2</td>
<td>5.0</td>
</tr>
<tr>
<td>rgTD16/GF14-NS</td>
<td>7.5</td>
<td>7.8</td>
<td>5.0</td>
<td>5.7</td>
<td>4.7</td>
<td>3.0</td>
<td>6.6</td>
<td>6.2</td>
<td>6.5</td>
<td>6.5</td>
<td>3.9</td>
<td>4.1</td>
</tr>
<tr>
<td>rgGF14</td>
<td>2.8</td>
<td>3.6</td>
<td>3.5</td>
<td>3.6</td>
<td>5.4</td>
<td>3.8</td>
<td>4.9</td>
<td>5.1</td>
<td>2.7</td>
<td>2.8</td>
<td>4.8</td>
<td>3.9</td>
</tr>
<tr>
<td>rgGF14/TD16-pol</td>
<td>3.4</td>
<td>4.3</td>
<td>3.8</td>
<td>4.4</td>
<td>2.1</td>
<td>3.7</td>
<td>4.8</td>
<td>4.7</td>
<td>4.3</td>
<td>5.4</td>
<td>3.9</td>
<td>5.5</td>
</tr>
<tr>
<td>rgGF14/TD16-HA.NA</td>
<td>3.7</td>
<td>3.4</td>
<td>5.6</td>
<td>3.7</td>
<td>3.4</td>
<td>3.1</td>
<td>6.6</td>
<td>5.0</td>
<td>3.8</td>
<td>4.6</td>
<td>4.4</td>
<td>4.8</td>
</tr>
</tbody>
</table>
Table 3. Immunohistochemistry staining for avian influenza virus on tissues collected from mallards inoculated with the indicated rgHSN8 viruses. Legend: (+++) widespread staining, (++) moderate staining, (+) few stained cells, (-) no staining.

<table>
<thead>
<tr>
<th>Virus name</th>
<th>Nasal epithelium</th>
<th>Trachea</th>
<th>Lung</th>
<th>Heart</th>
<th>Brain</th>
<th>Liver</th>
<th>Spleen</th>
<th>Kidney</th>
<th>Muscle</th>
<th>Intestine</th>
<th>Pancreas</th>
<th>Proventriculus</th>
<th>Thymus</th>
<th>Bursa</th>
</tr>
</thead>
<tbody>
<tr>
<td>rgTD16</td>
<td>+++/+++</td>
<td>+++/+</td>
<td>+++/++</td>
<td>+++/++</td>
<td>+++/++</td>
<td>+++/++</td>
<td>+/++</td>
<td>+++/+</td>
<td>++/+</td>
<td>+++/++</td>
<td>++/++</td>
<td>+++/++</td>
<td>+++/+</td>
<td>+/++</td>
</tr>
<tr>
<td>rgTD16/GF14-PB2</td>
<td>++/++</td>
<td>+/+++</td>
<td>+/‐</td>
<td>‐/‐</td>
<td>++/+</td>
<td>‐/+</td>
<td>‐/‐</td>
<td>+/‐</td>
<td>+/‐</td>
<td>++/++</td>
<td>+/‐</td>
<td>+/‐</td>
<td>+/‐</td>
<td>+/++</td>
</tr>
<tr>
<td>rgTD16/GF14-PB2</td>
<td>+++/+</td>
<td>+/+</td>
<td>+++/++</td>
<td>+++/++</td>
<td>+++/++</td>
<td>+++/++</td>
<td>+++/++</td>
<td>++/+</td>
<td>++/+</td>
<td>+++/++</td>
<td>+++/++</td>
<td>+++/++</td>
<td>+++/++</td>
<td>+++/++</td>
</tr>
<tr>
<td>rgTD16/GF14-PA</td>
<td>+++/+</td>
<td>+/+</td>
<td>+++/++</td>
<td>+++/++</td>
<td>+++/++</td>
<td>+++/++</td>
<td>+++/++</td>
<td>++/+</td>
<td>++/+</td>
<td>+++/++</td>
<td>+++/++</td>
<td>+++/++</td>
<td>+++/++</td>
<td>+++/++</td>
</tr>
<tr>
<td>rgTD16/GF14-NA</td>
<td>+++/+</td>
<td>+/+</td>
<td>+++/++</td>
<td>+++/++</td>
<td>+++/++</td>
<td>+++/++</td>
<td>+++/++</td>
<td>++/+</td>
<td>++/+</td>
<td>+++/++</td>
<td>+++/++</td>
<td>+++/++</td>
<td>+++/++</td>
<td>+++/++</td>
</tr>
<tr>
<td>rgTD16/GF14-NA</td>
<td>+++/+</td>
<td>+/+</td>
<td>+++/++</td>
<td>+++/++</td>
<td>+++/++</td>
<td>+++/++</td>
<td>+++/++</td>
<td>++/+</td>
<td>++/+</td>
<td>+++/++</td>
<td>+++/++</td>
<td>+++/++</td>
<td>+++/++</td>
<td>+++/++</td>
</tr>
<tr>
<td>rgTD16/GF14-M</td>
<td>+++/+</td>
<td>+/+</td>
<td>+++/++</td>
<td>+++/++</td>
<td>+++/++</td>
<td>+++/++</td>
<td>+++/++</td>
<td>++/+</td>
<td>++/+</td>
<td>+++/++</td>
<td>+++/++</td>
<td>+++/++</td>
<td>+++/++</td>
<td>+++/++</td>
</tr>
<tr>
<td>rgTD16/GF14-N5</td>
<td>+++/+</td>
<td>+/+</td>
<td>+++/++</td>
<td>+++/++</td>
<td>+++/++</td>
<td>+++/++</td>
<td>+++/++</td>
<td>++/+</td>
<td>++/+</td>
<td>+++/++</td>
<td>+++/++</td>
<td>+++/++</td>
<td>+++/++</td>
<td>+++/++</td>
</tr>
<tr>
<td>rgGF14</td>
<td>+++/+</td>
<td>+/+</td>
<td>+++/++</td>
<td>+++/++</td>
<td>+++/++</td>
<td>+++/++</td>
<td>+++/++</td>
<td>++/+</td>
<td>++/+</td>
<td>+++/++</td>
<td>+++/++</td>
<td>+++/++</td>
<td>+++/++</td>
<td>+++/++</td>
</tr>
<tr>
<td>rgGF14/TD16-po</td>
<td>++/+</td>
<td>+/+</td>
<td>+++/++</td>
<td>+++/++</td>
<td>+++/++</td>
<td>+++/++</td>
<td>+++/++</td>
<td>++/+</td>
<td>++/+</td>
<td>+++/++</td>
<td>+++/++</td>
<td>+++/++</td>
<td>+++/++</td>
<td>+++/++</td>
</tr>
<tr>
<td>rgGF14/TD16-HA.NA</td>
<td>+++/+</td>
<td>+/+</td>
<td>+++/++</td>
<td>+++/++</td>
<td>+++/++</td>
<td>+++/++</td>
<td>+++/++</td>
<td>++/+</td>
<td>++/+</td>
<td>+++/++</td>
<td>+++/++</td>
<td>+++/++</td>
<td>+++/++</td>
<td>+++/++</td>
</tr>
</tbody>
</table>
Table 4. Pairwise genome and protein sequence comparison between two clade 2.3.4.4 H5N8 HPAI viruses, TD16 and GF14.

<table>
<thead>
<tr>
<th>Gene</th>
<th>Nucleotide identity</th>
<th># of changes</th>
<th>Protein</th>
<th>Amino acid identity</th>
<th># of changes</th>
</tr>
</thead>
<tbody>
<tr>
<td>PB2</td>
<td>86.6%</td>
<td>310</td>
<td>PB2</td>
<td>97.6%</td>
<td>18</td>
</tr>
<tr>
<td>PB1</td>
<td>90.8%</td>
<td>215</td>
<td>PB1</td>
<td>98.2%</td>
<td>14</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>PB1-F2</td>
<td>76.9%</td>
<td>12</td>
</tr>
<tr>
<td>PA</td>
<td>91.5%</td>
<td>190</td>
<td>PA</td>
<td>96.4%</td>
<td>26</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>PA-X</td>
<td>93.7%</td>
<td>16</td>
</tr>
<tr>
<td>HA</td>
<td>94.1%</td>
<td>105</td>
<td>HA0</td>
<td>96.3%</td>
<td>21</td>
</tr>
<tr>
<td>NP</td>
<td>92.7%</td>
<td>114</td>
<td>NP</td>
<td>99.0%</td>
<td>5</td>
</tr>
<tr>
<td>NA</td>
<td>95.8%</td>
<td>62</td>
<td>NA</td>
<td>96.4%</td>
<td>17</td>
</tr>
<tr>
<td>M</td>
<td>91.9%</td>
<td>83</td>
<td>M1</td>
<td>94.8%</td>
<td>13</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>M2</td>
<td>90.7%</td>
<td>9</td>
</tr>
<tr>
<td>NS</td>
<td>93.1%</td>
<td>61</td>
<td>NS1</td>
<td>90.5%</td>
<td>9</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>NEP</td>
<td>92.6%</td>
<td>22</td>
</tr>
</tbody>
</table>
Table 5. Amino acid differences between GF14 and TD16 previously reported as associated with biological changes.

<table>
<thead>
<tr>
<th>Protein</th>
<th>Residue in GF14</th>
<th>Residue in TD16</th>
<th>Position</th>
<th>Change</th>
<th>Citation</th>
<th>Comments</th>
</tr>
</thead>
<tbody>
<tr>
<td>PB2</td>
<td>R</td>
<td>K</td>
<td>702</td>
<td>R702K</td>
<td>(73)</td>
<td>K702R change associated with increased polymerase activity in mammalian cells.</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>(48)</td>
<td>Position 702 is known to modulate host range where K is found in most Al viruses and R is found in most human influenza viruses.</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>(56)</td>
<td>K702R associated with lower mouse 50% lethal dose in TY93/HSN1 (clade 2.3.4).</td>
</tr>
<tr>
<td>HA</td>
<td>S</td>
<td>P</td>
<td>141</td>
<td>S141P</td>
<td>(74)</td>
<td>S141P change is involved in antigenic drift of the H5N1 clade 2.2.1 virus in Egypt.</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>(49)</td>
<td>S141P change is associated with increased virus replication and transmissibility of Gs/GD H5N2 virus in chickens.</td>
</tr>
<tr>
<td>NP</td>
<td>M</td>
<td>V</td>
<td>105</td>
<td>M105V</td>
<td>(62, 78-80)</td>
<td>M105V change observed in viruses from chickens and other gallinaceous species infected with Gs/GD H5 viruses; also known to increase virulence and replication in chickens.</td>
</tr>
<tr>
<td>NP</td>
<td>K</td>
<td>R</td>
<td>452</td>
<td>K452R</td>
<td>(75, 76)</td>
<td>At position 452, K is found predominantly in mammalian viruses whereas an R is found in most avian viruses.</td>
</tr>
<tr>
<td>M1</td>
<td>K</td>
<td>R</td>
<td>95</td>
<td>K95R</td>
<td>(81)</td>
<td>R95K (reciprocal change) leads to a reduction in filamentous virions in A/WSN/33 with A/Udorn/72 M segment.</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>(69)</td>
<td>Position at the M1 monomer interface.</td>
</tr>
<tr>
<td>M1</td>
<td>F</td>
<td>L</td>
<td>144</td>
<td>F144L</td>
<td>(55) L144F change along with E156D associated with increased mortality with Gs/GD H5N1 virus in ducks.* In this study, these M1 changes were sufficient to confer a partial increase in mortality.</td>
<td></td>
</tr>
<tr>
<td>M1</td>
<td>K</td>
<td>R</td>
<td>101</td>
<td>K101R</td>
<td>(69) Position at the M1 monomer interface. (85) R101G change leads to filamentous virions in PR8. (71) R101K change is associated with increased mortality with Gs/GD H5N1 virus in chickens.**</td>
<td></td>
</tr>
<tr>
<td>M1</td>
<td>N</td>
<td>S</td>
<td>224</td>
<td>N224S</td>
<td>(72) N224 became predominant in H9N2 viruses from China in 2010-2015. Position 224 is located at the vRNP binding domain. 224N is associated with enhanced early replication in vitro and extrapulmonary spread in chickens. (71) S224N change is associated with increased mortality with Gs/GD H5N1 virus in chickens.** (70) S244N occurred when aH7N3 HP virus was passaged in quail.</td>
<td></td>
</tr>
<tr>
<td>M1</td>
<td>N</td>
<td>K</td>
<td>242</td>
<td>N242K</td>
<td>(72) N242 became predominant in H9N2 viruses from China in 2010-2015 and is associated with enhanced early replication in vitro and extrapulmonary spread in chickens.</td>
<td></td>
</tr>
<tr>
<td>M2</td>
<td>N</td>
<td>S</td>
<td>31</td>
<td>N31S</td>
<td>(70) S31N change occurred as H5N1 clade 2.2 viruses circulated in poultry in Egypt.</td>
<td></td>
</tr>
</tbody>
</table>
M2  E  G  66  E66G  (77)  A66E change lead to increased replication and dominance over other H5N1 viruses in ECE.

*Results contrary to what found in our study.
**Changes related to increased adaptation or virulence in chickens which could be contrary to increased adaptation or virulence in ducks.
Table 6. Strain names of clade 2.3.4.4 and 2.2 HS viruses from 2005 and 2014-2016 that were previously studied. The sequences from these strains were used for phylogenetic analysis in addition to other sequences that were downsampled from the 2007-2020 datasets.

<table>
<thead>
<tr>
<th>Strain</th>
<th>Serotype</th>
<th>Continent</th>
<th>Clade</th>
<th>Group</th>
<th>Reference</th>
</tr>
</thead>
<tbody>
<tr>
<td>A/baikal_teal/Korea/Donglim3/2014</td>
<td>H5N8</td>
<td>Asia</td>
<td>2.3.4.4</td>
<td>A</td>
<td>(122, 123)</td>
</tr>
<tr>
<td>A/breeder_duck/Korea/Gochang1/2014</td>
<td>H5N8</td>
<td>Asia</td>
<td>2.3.4.4</td>
<td>B</td>
<td>(122, 123)</td>
</tr>
<tr>
<td>A/broiler_duck/Korea/Buan2/2014</td>
<td>H5N8</td>
<td>Asia</td>
<td>2.3.4.4</td>
<td>A</td>
<td>(122, 123)</td>
</tr>
<tr>
<td>A/chicken/Kagawa/1T-1/2018</td>
<td>H5N6</td>
<td>Asia</td>
<td>2.3.4.4</td>
<td>B</td>
<td>(97)</td>
</tr>
<tr>
<td>A/duck/India/10CA01/2016</td>
<td>H5N8</td>
<td>Asia</td>
<td>2.3.4.4</td>
<td>B</td>
<td>(124)</td>
</tr>
<tr>
<td>A/duck/Korea/HD1/2017</td>
<td>H5N6</td>
<td>Asia</td>
<td>2.3.4.4</td>
<td>B</td>
<td>(99)</td>
</tr>
<tr>
<td>A/Eurasian_Wigeon/Netherlands/4/2016</td>
<td>H5N8</td>
<td>Europe</td>
<td>2.3.4.4</td>
<td>B</td>
<td>(28)</td>
</tr>
<tr>
<td>A/goose/Zhejiang/925037/2014</td>
<td>H5N8</td>
<td>Asia</td>
<td>2.3.4.4</td>
<td>B</td>
<td>(125)</td>
</tr>
<tr>
<td>A/great_crested_grebe/Tyva/34/2016</td>
<td>H5N8</td>
<td>Europe</td>
<td>2.3.4.4</td>
<td>B</td>
<td>(126)</td>
</tr>
<tr>
<td>A/great_crested_grebe/Uvs-Nuur_Lake/341/2016</td>
<td>H5N8</td>
<td>Europe</td>
<td>2.3.4.4</td>
<td>B</td>
<td>(24)</td>
</tr>
<tr>
<td>A/grey_heron/Korea/W779/2017</td>
<td>H5N8</td>
<td>Asia</td>
<td>2.3.4.4</td>
<td>B</td>
<td>(127)</td>
</tr>
<tr>
<td>A/gyrfalcon/Washington/41088/6/2014</td>
<td>H5N8</td>
<td>North America</td>
<td>2.3.4.4</td>
<td>A</td>
<td>(4, 43, 102)</td>
</tr>
<tr>
<td>A/Jungle_crow/Hyogo/2803E023C/2018</td>
<td>H5N6</td>
<td>Asia</td>
<td>2.3.4.4</td>
<td>B</td>
<td>(97)</td>
</tr>
<tr>
<td>A/mallard/Korea/Jeju-H24/2017</td>
<td>H5N6</td>
<td>Asia</td>
<td>2.3.4.4</td>
<td>B</td>
<td>(99, 101)</td>
</tr>
<tr>
<td>A/mallard/Korea/W452/2014</td>
<td>H5N8</td>
<td>Asia</td>
<td>2.3.4.4</td>
<td>A</td>
<td>(12)</td>
</tr>
<tr>
<td>A/Mallard/Netherlands/18012508-017/2018</td>
<td>H5N6</td>
<td>Europe</td>
<td>2.3.4.4</td>
<td>B</td>
<td>(128)</td>
</tr>
<tr>
<td>A/mute_swan/Krasnodar/25/2017</td>
<td>H5N8</td>
<td>Europe</td>
<td>2.3.4.4</td>
<td>B</td>
<td>(25)</td>
</tr>
<tr>
<td>A/Northern_Goshawk/Tokyo/1301B003/2018</td>
<td>H5N6</td>
<td>Asia</td>
<td>2.3.4.4</td>
<td>B</td>
<td>(97)</td>
</tr>
<tr>
<td>A/Northern_pintail/Washington/40964/2014</td>
<td>H5N2</td>
<td>North America</td>
<td>2.3.4.4</td>
<td>A</td>
<td>(4, 43, 102)</td>
</tr>
<tr>
<td>Strain</td>
<td>HSN</td>
<td>Region</td>
<td>Pathogenicity</td>
<td>Reference</td>
<td></td>
</tr>
<tr>
<td>------------------------</td>
<td>------</td>
<td>----------</td>
<td>----------------</td>
<td>-----------</td>
<td></td>
</tr>
<tr>
<td>A/painted_stork/India/10CA03/2016</td>
<td>H5N8</td>
<td>Asia</td>
<td>2.3.4.4</td>
<td>B</td>
<td>(21, 124)</td>
</tr>
<tr>
<td>A/Tufted_duck/Denmark/11740/LWPL/2016</td>
<td>H5N8</td>
<td>Europe</td>
<td>2.3.4.4</td>
<td>B</td>
<td>(30)</td>
</tr>
<tr>
<td>A/tufted_duck/Germany/AR8444-L01987/2016</td>
<td>H5N8</td>
<td>Europe</td>
<td>2.3.4.4</td>
<td>B</td>
<td>(28)</td>
</tr>
<tr>
<td>A/whooper_swam/Mongolia/244/2005</td>
<td>H5N1</td>
<td>Asia</td>
<td>2.2</td>
<td>-</td>
<td>(4, 129)</td>
</tr>
<tr>
<td>A/wild_duck/Poland/82A/2016</td>
<td>H5N8</td>
<td>Europe</td>
<td>2.3.4.4</td>
<td>B</td>
<td>(21)</td>
</tr>
</tbody>
</table>
A. 100% mortality for rgTD16 reassortants

B. 75-88% mortality for rgTD16 reassortants

C. 25-75% mortality for rgTD16 reassortants