Genetic and antigenic characterization of an expanding H3 Cluster IV-A influenza A virus clade in US swine visualized by Nextstrain

Megan N. Neveaua,b, Michael A. Zellera, Bryan S. Kaplanc, Carine K. Souzac, Phillip C. Gaugera, Amy L. Vincentc, Tavis K. Andersonc#

aDepartment of Veterinary Diagnostic and Production Animal Medicine, College of Veterinary Medicine, Iowa State University, Ames, Iowa, USA

bBioinformatics and Computational Biology Program, Iowa State University, Ames, Iowa, USA

cVirus and Prion Research Unit, National Animal Disease Center, USDA-ARS, Ames, Iowa, USA

Running Head: Expanding C-IVA influenza A virus clade in US swine

#Address correspondence to Tavis K. Anderson, tavis.anderson@usda.gov

**Abstract**

The genetic and antigenic diversity of influenza A virus in swine (IAV-S) is shaped by transmission, establishment, and persistence of human IAV within pig populations. In 2019, a phylogenetic clade of the H3 Clade IV-A hemagglutinin (HA) lineage that circulated for 20 years in US swine was detected with increasing frequency. To determine putative mechanisms associated with increased detection, we paired comprehensive phylogenetic analyses with a phenotypic assessment of representative strains from the clade. To visualize and track the emergence, spatial spread, and genetic evolution of H3 IAV-S we developed and deployed a Nextstrain web application. These data identified two C-IVA clades that emerged concurrently and cocirculated within multiple US states. Phylodynamic analysis of the HA gene documented low relative genetic diversity from 2017 to present, suggesting clonal expansion of genetically similar viruses. The clade with the majority of detections was associated with an N156H amino acid substitution; 156 was previously defined as part of the antigenic motif associated with major impacts on antigenic phenotype. Using hemagglutination inhibition (HI) assays, we demonstrated that the 156N or 156H mutation did not result in significant antigenic drift. Genome constellation diversity was also quantified: the minor clade was paired with N2-2002B in ancestral strains, but through reassortment acquired an N2-2002A in 2016; in the major clade, the nucleoprotein (NP) from the North American swine lineage was replaced with an NP of the pandemic lineage by 2018. These data demonstrate that increased detection of the 1990.4.a clade was not associated with an increase in relative genetic or antigenic diversity of the HA gene; but may be associated with novel genome constellations impacting replication and transmission. The expansion of the H3 Clade IV-A clade warrants vaccine strain consideration based on regional prevalence in swine populations. Defining the factors driving spatial and temporal patterns in IAV-S genetic diversity in swine is essential to informing efficacious vaccine strategies and to reducing the chance of a swine-origin IAV emerging that has zoonotic potential. (Word count = 327)

Keywords: influenza A virus, H3N2, swine, reassortment, surveillance, vaccines

**Importance**

Efforts to control influenza A virus in swine (IAV-S) are undermined by the existence of many genetically distinct groups of viruses, called clades. Often, vaccines are produced against the most common clades in a farm or region. In 2019, we detected an increase in detection frequency of a clade, H3 C-IVA, in U.S. swine, which was previously circulating only at low levels. Our study attempted to identify genetic and antigenic factors contributing to its resurgence by leveraging the Nextstrain visualization platform, Bayesian statistical analysis, and wet-lab experiments. We found that the contemporary C-IVA viruses did not have increased genetic diversity nor significant antigenic changes that would allow them to escape the pig’s immune response. Many of the contemporary viruses do have a different nucleoprotein (NP) gene segment that has been swapped with the historic one and we suggest that this could have contributed to the clade’s success.  (Word count = 147)

**1. Introduction**

Influenza A virus (IAV) is an economically important pathogen of swine that has the ability to evolve and evade the host immune response which, therefore, presents a challenge to current disease control strategies. The negative-sense, single-stranded RNA genome consists of eight non-contiguous gene segments that are known to encode between 10 and 17 proteins (Gamblin and Skehel 2010; Muramoto et al. 2013; Yamayoshi et al. 2016). Mutation and selection increase diversity at the gene level. The segmented genome structure creates opportunity for reassortment when two or more IAV strains concurrently infect the same host, resulting in novel gene combinations and increased diversity at the genomic level (Essere et al. 2013; Marshall et al. 2013). As swine have both α2,6- and α2,3-Gal-linked sialic acid on the surface of their respiratory epithelial cells; therefore in addition to swine-to-swine transmission, they are susceptible to infection from both human and avian origin IAVs as well (Ito et al. 1998; Nicholls et al. 2008). Consequently, observed IAV diversity in swine is further increased by the transmission, occasional establishment, and evolution of avian and human IAV in swine populations.

The genetic diversity of IAV is paired with a similarly large breadth of antigenic diversity. The accumulation of amino acid substitutions from polymerase mutation can result in changes of the antigenic phenotype of IAV (Bedford et al. 2014; Smith et al. 2004). For the H3 subtype, a small number of amino acid residues have a disproportionate effect on antigenic phenotype in both humans and swine (Koel et al. 2013; Lewis et al. 2014; Santos et al. 2019). In swine, six of these amino acid positions (145,155,156,158,159 and 189; H3 mature peptide numbering (Burke and Smith 2014)) are referred to as the H3 antigenic motif (Abente et al. 2016). These six residues are located on the globular head of the HA protein and are adjacent to the receptor binding site. There may be limited substitution flexibility at these positions due to necessary conservation of receptor binding functionality (Santos et al. 2019). These data suggest that minimal genetic change may result in significant antigenic change. Substitutions at these positions could reduce the efficacy of current IAV-S vaccines in minimizing clinical disease and transmission (Vincent et al. 2017). Vaccination with whole inactivated virus (WIV) with oil in water adjuvant is common in swine in the United States (U.S.). However, vaccines have only been proven to induce robust antibody responses when the vaccine and challenge strains were closely related (Kitikoon et al. 2006; Loving et al. 2013). Thus, an important factor in vaccine strain selection is surveillance of the circulating strains to inform understanding of contemporary genetic and antigenic IAV-S diversity at a national and regional level.

In 1998, investigations into severe respiratory disease in swine in the U.S. lead to the first recognition of the H3N2 subtype of IAV in North American swine. The H3N2 that persisted was a triple reassortant virus with HA, NA, and PB1 gene segments derived from human seasonal H3N2; PB2 and PA gene segments from avian IAV; and NP, M, and NS gene segments from classical swine H1N1(Anderson et al. 2020; Webby et al. 2000; Zhou et al. 1999). The HA gene from this introduction evolved into genetically distinct clades, establishing H3 cluster IV in the U.S. (Anderson et al. 2020). This lineage continued to circulate in the U.S. since 2005 and genetically diversified into clades A through F (Kitikoon et al. 2013). Cluster IV-A (C-IVA) began to increase in detection frequency beginning in 2010 and was the predominant H3 clade until 2016, when it was surpassed by H3 2010.1, a more recent human-seasonal incursion that established as a swine lineage (Anderson et al. 2020; Rajao et al. 2015; Zeller et al. 2018). In 2019, passive IAV-S surveillance conducted by the USDA indicated a resurgence in C-IVA sequence detection, as well as a relative decrease in detection of H3 2010.1 that required further investigation.

In this study, we quantified genetic and antigenic characteristics associated with the recent increased detection frequency of the H3 C-IVA clade. Concurrently, we adapted the Nextstrain platform (Hadfield et al. 2018) to IAV-S to provide near real-time phylogenetic visualization of surveillance data for the H3 subtype. Collectively, these analyses provide insight into the factors contributing to the expansion of the clade and improve our ability to predict mechanisms that IAV-S employs to evade current control measures.

**2. Materials and Methods**

2.1 Data Collection

All available U.S. swine H3 nucleotide sequences (n=3395) detected between January 2010 and March 2021 deposited into GenBank (Benson et al. 2013) were downloaded from the Influenza Research Database (IRD) (Zhang et al. 2017). Duplicate strains were removed from the dataset. Sequences were then classified using octoFLU (Chang et al. 2019), and those that were from the C-IVA clade (n=1376) were retained for further analysis. All available corresponding gene segments for viruses with whole genome sequences (WGS, n=545) were collated and classified into lineage. H3 clade detection frequency was derived from octoFLU classification of public sequence data and validated with the private regional surveillance data housed in the Iowa State University Veterinary Diagnostic Lab (ISU VDL) visualized on ISU FLUture (Zeller et al. 2018).

2.2 Estimation of relative genetic diversity

To generate a computationally tractable dataset, we generated a random subset (n=500) of H3 C-IVA sequences via smof v2.21.0 (Arendsee et al. 2018). Sequences were aligned with mafft v7.450 (Katoh and Standley 2013) and a maximum-likelihood phylogenetic tree was inferred using the generalized time-reversible model (GTR) of nucleotide substitution in FastTree v2.1.11 (Price et al. 2010). This tree was used in a root-to-tip regression analysis in TempEst v1.5.3 (Rambaut et al. 2016) to assess temporal signal and detect genes with incongruous genetic divergence and sampling dates. The final dataset (n=493) was then analyzed with BEAST v1.8.4 (Drummond and Rambaut 2007) to estimate effective population size of the C-IVA lineage over time. We applied the GMRF Bayesian Skyride coalescent model (Minin et al. 2008) with a GTR substitution model with gamma-distributed rate variation, an uncorrelated relaxed clock, and a MCMC chain length of 100,000,000 with sampling every 10,000 iterations. Demographic reconstruction was performed using the GMRF skyride reconstruction in Tracer v1.7.1 (Rambaut et al. 2018) and a maximum clade credibility (MCC) tree was generated using TreeAnnotator v1.8.4 (Rambaut and Drummond 2015).

2.3 Deployment of Nextstrain for H3 IAV in swine

The Nextstrain (Hadfield et al. 2018) platform was adapted for H3 IAV-S. A time-scaled tree was estimated for all H3 swine IAV HA genes, and a focused C-IVA HA nucleotide sequence dataset using the “refine” Augur command (Huddleston et al. 2021). A separate time-scaled tree was estimated for paired NA nucleotide sequences and the two trees were then compared using the Auspice visualization platform (https://github.com/nextstrain/auspice). Amino acid substitutions were annotated on the backbone of the tree using the “ancestral” and “translate” commands. The H3 antigenic motif was visualized by combining the “Color By Genotype” function for positions 145, 155, 156, 158, 159, and 189. The lineages determined through octoFLU of the other six gene segments were mapped onto the HA tree using the “traits” command. The “traits” command also integrated geographic information at the U.S. state level and computed putative transmission between states. These data were exported as JSON files that are interactively visualized on the web at https://flu-crew.org on an AWS server using USDA-ARS SCInet with all files provided at https://github.com/flu-crew/.

2.4 Antigenic Characterization

We identified two C-IVA genetic clades co-circulating in the U.S. from January 2019 to Mar 2021; one clade formed the majority of detections (89.3% of 326 sequences) and the other was minor, but persistent. To identify a representative sequence for each clade, we generated an HA1 consensus sequence in Geneious Prime 2020.2.3 and selected the best matching field strain from the USDA-APHIS IAV in swine repository at the National Veterinary Services Laboratories. For the major clade, we identified an amino acid substitution at position 156 on the backbone of the phylogeny using Nextstrain. We selected an “ancestral” strain and a “contemporary” strain to reflect the substitution at 156. The three field strains most similar to the consensus sequences (A/swine/Oklahoma/A01770191/2015 – C-IVA major/ancestral, A/swine/North Carolina/A02245294/2019 – C-IVA major/contemporary, and A/swine/Minnesota/A02266068/2018 – C-IVA minor) were selected to be antigenically characterized.

A panel of swine antisera was constructed using sera previously produced by immunizing two pigs (Bolton et al. 2019; Lewis et al. 2014). Hemagglutination inhibition (HI) assays were performed on test antigens using turkey red blood cells and sera treated with Receptor Destroying Enzyme (II) (Hardy Diagnostics). Fold reduction in titer was calculated by dividing the log transformed homologous titer of each antisera by the log transformed heterologous titer of each test antigen. HI data with the selected H3 C-IVA strains were merged with a subset of previously generated H3 antigenic data and used to create three-dimensional antigenic maps via ACMACS (Bolton et al. 2019; Lewis et al. 2014; Smith et al. 2004).

An enzyme-linked lectin assay (ELLA) was used to determine neuraminidase inhibiting (NI) antibody titers using peanut agglutinin-horse radish peroxidase (PNA-HRP) (Sigma-Aldrich, St. Louis, MO) and 3,3’,5,5’-tetramethylbenzidine (TMB) (KPL Laboratories, Gaithersburg, MD), as previously described (Gao et al. 2016; Kaplan and Vincent 2020). The optical density (OD) of the plates was read at 650 nm and the titer was assigned as the reciprocal of the highest dilution resulting in at least 50% inhibition.

**3. Results**

3.1 Increased detection frequency followed by increased relative genetic diversity of H3 Clade IV-A

From 2011 to 2015 the C-IVA genetic clade was the most frequently detected H3 in U.S. swine (Figure 1a). The C-IVA clade showed a steep decline in detection frequency throughout 2016. In 2017 and 2018, C-IVA viruses represented less than 20% of H3 detections, with the 2010.1 clade instead accounting for the majority of H3 detections (2017: 81.9%; 2018: 60.4%). In 2019 and 2020, C-IVA detection frequency increased to 32.1% (101 detections) and 53% (187), respectively. Concurrently, the 2010.1 clade decreased to 53.3% (168 detections) of H3 detections in 2019 and 36.8% (130 detections) in 2020. The makeup of the H3 clade in the first three months of 2021 remains similar to that of 2020.

The median posterior rate of nucleotide substitution for the C-IVA clade estimated by Bayesian analysis was 4.265 x 10-3 (95%HPD: 3.956 x 10-3, 4.589 x 10-3).Relative genetic diversity of the HA gene was estimated by a Bayesian demographic reconstruction and demonstrated an almost linear increase from 2011 to 2015 and decrease from 2015 to 2019 (Figure 1b). Despite minimal genetic diversity in 2018 and 2019, detection frequency began to increase. The increase in detection frequency is followed by an increase in relative genetic diversity. The trends in relative genetic diversity are supported by the topology of a maximum-likelihood phylogeny (Supplemental Figure 1) with external branches that are shorter relative to branches on the interior of the tree.

3.2 Two co-circulating clades with onward transmission after 2018

The HA tree shows many co-circulating C-IVA genetic clades that corresponded with high levels of relative genetic diversity in 2013 and 2015 (Figure 1b; Figure 2). After 2018, only two distinct genetic clades of the C-IVA clade were apparent: a major clade representing 245 detections (71 in 2019, 174 in 2020) and a minor clade representing 24 detections (19 in 2019 and 6 in 2020). The major clade viruses were first detected in the Southwest region (Texas, Oklahoma, and Kansas) of the U.S. in 2017 and 2018, but were detected in the major pork producing states of the Midwest by January 2019. By late-2019, major clade viruses were detected in North Carolina and some less hog-dense states such as Michigan and Pennsylvania. Despite the broad geographic representation, the majority of detections of this clade were in Iowa and Indiana (56%). The minor clade was initially detected in the Midwest; it was rarely outside of these states from 2018 to present (5 detections).

3.4 Reassortment with the H3 IV-A and novel NA and NP gene segment pairings

The HA gene segments of the major C-IVA clade were consistently paired with N2 2002.B2 gene segments, matching the topology of the congruent NA phylogenetic tree (Figure 2). The ancestral viruses of the minor clade were paired with N2 2002.B2 from mid-2012 to 2016. In 2016 (95% CI: 2016-06-05, 2017-02-08), the minor clade showed evidence of reassortment with a genetically distinct N2 2002.A2 that no other C-IVA HA gene segment was paired with in the past decade. The lineages of the remaining six gene segments were annotated onto the HA tree for those viruses with WGS (Figure 3). Historically the C-IVA clade was paired with a nucleoprotein (NP) gene segment from the triple-reassortment H3N2 (TRIG) lineage. The major clade showed evidence of reassortment with an NP from the H1N1 pdm09 lineage beginning in 2017. All available WGS (n=58) from the major clade contained a pdm09 lineage NP after September 2018. There was no evidence of lineage replacement in the M, NS, PA, PB1, or PB2 gene segments.

3.5 N156H antigenic motif substitution

The time-scaled tree annotated with amino acid substitutions created with Nextstrain showed two amino acid substitutions associated with the expansion of the major clade, N156H and K368E, dating back to November 2016 (95% CI: 2016-06-04,2017-05-18). The 156 position was identified in the previously characterized antigenic motif (Abente et al. 2016; Koel et al. 2013; Lewis et al. 2014). This substitution was the only sustained mutation in the antigenic motif that occurred in either of the two contemporary clades. Fourteen other amino acid substitutions were detected in the major clade after December 2016. Ten of the fourteen occurred in antigenic regions of HA1 (N96S, V323I, A131T, R141K, T121N, S146G, T48A, N158D, V196A, I214V; H3 mature peptide numbering).

3.6 Hemagglutination Inhibition and Antigenic Cartography

Three representative strains were selected to represent three distinct groups of C-IVAs: the major clade prior to the N156H substitution (A/swine/Oklahoma/A01770191/2015; OK/15) with 156N; the major clade following the N156H substitution (A/swine/North Carolina/A02245294/2019; NC/19) with 156H; and the minor clade containing 156N (A/swine/Minnesota/A02266068/2018; MN/18). Significant antigenic drift defined by an 8-fold loss in HI cross-reactivity, which corresponds to 3 antigenic units (AU) between viruses. The three strains were estimated to be within 2 AU of each other through the use of antigenic cartography and extracted antigenic distances (Figure 4a). Most C-IVA viruses, excluding those with an S or K substitution at position 145, were within 3 AU of the three test strains. Antigens from cluster IV-B, IV-C, or human vaccines were the only viruses tested to be greater than 3 AU from the three test strains. The CIV-B virus MN/13 with an NYHNYK antigenic motif was less than 3 AU away from all three test viruses and, of the three test viruses, was closest to the one with the same motif, NC/19.

3.7 Neuraminidase Inhibition with Enzyme-Linked Lectin Assay (ELLA)

The NI titers of the same three test strains were assessed against reference antisera of the swine N2 clades. The A/swine/Oklahoma/A01770191/2015 had a N2-2002X gene, the A/swine/North Carolina/A02245294/2019 had a N2-2002X gene; and the minor clade A/swine/Minnesota/A02266068/2018 had a N2-2002X gene. There were differences in NI cross-reactivity within the N2 2002 lineage, further divided into 2002.A.1, 2002.A.2, 2002.B.1 and 2002.B.2 (Zeller et al. 2020) (Zeller et al.; Kaplan, personal communication). In particular, the MN/18 virus from the 2002.A.2 clade showed a large antigenic distance from the 2002.B.2 representative virus (Figure 4b). The 2002.A.2 and 2002.B.2 both had at least two antigen representatives and demonstrated no significant loss in NI cross-reactivity within the respective clades.

**4. Discussion**

In this study, we investigated possible factors to explain the recent increase in detection of H3 C-IVA viruses in US swine. As the number of C-IVA detections increased since 2018, the relative genetic diversity within the clade did not increase, but instead continued to decrease through 2020. This pattern is distinct from a prior increase in detection frequency that was paired with simultaneous diversification from 2013 to 2015. The current clonal expansion of C-IVA with low diversity suggests that a selective sweep occurred in the population. Sweep-related changes were identified in human seasonal H3N2 IAV and most often detected at amino acid sites located on the HA (Klingen et al. 2018; Rambaut et al. 2008). The possible correlation between EPS and viral expansion could be explored further and potentially utilized to improve prediction of clade success.

To determine whether a selective sweep occurred, we identified amino acid substitutions sustained in the major and minor clades that were circulating as detection frequency increased from 2018 to 2020. We identified an N156H substitution, an amino acid position in the HA that was previously identified as having a disproportionate, but largely additive effect with substitutions in other positions on the antigenic phenotype of the virus. This N156H substitution prompted the antigenic characterization of the major and minor clades via HI assays to assess for a potential loss in cross-reactivity. A significant loss in cross-reactivity of the contemporary 156H from prior strains with 156N would suggest a potential lack of population immunity that could explain the increased frequency of the major clade. However, these data did not demonstrate that the substitution caused significant antigenic drift. Antibodies raised against ancestral C-IVA demonstrated HI cross-reactivity against the more recent strains regardless of the substitution. It is important to note that the impact of an amino acid substitution depends on the biological properties of the specific amino acid(s) that have changed and the overall HA1 amino acid context. However, our results support previous findings that variation at position 156 alone did not cause significant antigenic drift (Abente et al. 2016; Bolton et al. 2019) . The limited change in antigenic phenotype suggests the N156H substitution may not have been the primary cause of the observed clonal expansion of the C-IVA major clade.

With no evidence of significant antigenic drift, the contemporary C-IVA major and minor clades were analyzed for evidence of other genetic signatures associated with the expanding H3N2 C-IVA clade. Through examination of whole genomes, it was found that the minor clade recently reassorted to obtain N2 2002.A genes. The antigenic effects of this reassortment event were further investigated with a panel of NI anti-sera previously used to describe antigenic variation among and between swine N2 lineages (B.S. Kaplan, submitted). While results here showed antigenic variation within the N2 2002 lineage, the N2.02B.2 of the major clade viruses retained close antigenic relationships to other 02.A and 02.B swine N2 and loss of immunity to the N2 also was not likely to explain the expansion of the major clade.

Still lacking an explanation for the increased detection of the major clade based on intrinsic properties of the IV-A H3N2, we analyzed WGS data for evidence of reassortment of the internal genes. The major clade was determined to have reassorted to acquire an NP of the H1N1pdm09 lineage. This may suggest the success of this virus was not primarily due to changes in the antigenic phenotype of the surface glycoproteins but could possibly be explained by differences between the pdm09 and TRIG genetic lineages of the NP. The influenza NP is characterized as a structural RNA-binding protein that forms the ribonucleoprotein (RNP) particle (Pons et al. 1969). However, genetic differences between lineages could also alter other putative functions of the NP, such as its role in the temporal regulation of apoptosis or import and export of vRNPs from the nucleus (Mayank et al. 2015; Portela and Digard 2002). The genotype of the internal genes was summarized as a concatenation of one-letter codes representing the genetic lineage of each gene segment (PB2, PB1, PA, NP, M, and NS) without the HA and NA segments. Prior to the NP reassortment, the internal gene constellation, TTT**T**PT, was the most common constellation found in a sample of 368 H3N2 isolates collected between 2009 and 2016 (Rajao et al. 2017). The resulting internal gene constellation of the C-IVA major clade, TTT**P**PT, was observed before; however, at that time, it was uncommon and was detected in only 7 of the 366 isolates. Results from the same study showed a wildtype field strain containing a TTT**P**PT was effective in viral transmission compared to wildtype strains with TTT**T**PT. This may be one factor contributing to the major clade expanding and spread across the U.S.

Although there was no evidence of genetic diversification or significant antigenic drift within the C-IVA viruses selected for testing to explain the increased detection of the C-IVA clade, it must be highlighted that the H3 2010.1 clade that emerged in 2012 began to outcompete the C-IVA clade in 2016. There was limited serologic cross-reactivity between 2010.1 and C-IVA swine H3N2 (Rajao? Powell?). Many herds were vaccinated against the 2010.1 clade of viruses with custom or autogenous vaccines following its emergence and dominance over C-IVA H3N2. Waning immunity against C-IVA viruses due to focus on vaccines containing 2010.1 H3N2 viruses likely allowed a competitive advantage of the C-IVA due to the lack of population immunity against this specific clade of swine H3N2. The increased detection frequency of C-IVA will re-direct autogenous and custom vaccine antigens to this clade of IAV. Continued surveillance is necessary to determine if vaccination against C-IVA will result in a decrease in detection; however, this would require additional knowledge of farm specific vaccines and vaccination strategies. Other unknown abiotic factors may have potentially played a role to influence the dynamics of the clade, such as swine transportation patterns and biosecurity protocols.

C-IVA viruses continue to make up roughly one-half of H3N2 detections into the early months of 2021 (Zeller et al. 2018). The lack of evidence for increased genetic diversity or antigenic drift suggests that the C-IVA resurgence could be caused by interplay between multiple factors, such as reassortment attributed to NP or competitive advantages due to vaccine mismatch with the 2010.1 H3N2. It is necessary to continue monitoring the evolutionary dynamics of the C-IVA genetic clade and its relative abundance within the H3N2 subtype. Better understanding of the factors that contribute to IAV clade expansion is necessary to inform and improve prediction methods for more successful control measures and reduced economic loss. These control measures are also important for public health, as dominant swine H3N2 clades have caused numerous zoonotic events through human-swine agricultural interfaces (citations) and these resurging contemporary C-IVA swine strains may be antigenically drifted from the pandemic preparedness candidate vaccine virus A/Minnesota/11/2010 (C.K. Souza, submitted).

**Acknowledgements**

We gratefully acknowledge pork producers, swine veterinarians, and laboratories for participating in the USDA Influenza A Virus in Swine Surveillance System and publicly sharing sequences. This work was supported by the Iowa State University Presidential Interdisciplinary Research Initiative; the Iowa State University Veterinary Diagnostic Laboratory; the U.S. Department of Agriculture (USDA) Agricultural Research Service [ARS project number 5030-32000-120-00-D]; an National Institute of Allergy and Infectious Diseases (NIAID) at the National Institutes of Health interagency agreement associated with the Center of Research in Influenza Pathogenesis, an NIAID funded Center of Excellence in Influenza Research and Surveillance [grant HHSN272201400008C to A.L.V.]; the USDA Agricultural Research Service Research Participation Program of the Oak Ridge Institute for Science and Education (ORISE) through an interagency agreement between the U.S. Department of Energy (DOE) and USDA Agricultural Research Service [contract number DE-AC05- 06OR23100 to B.K. and C.S.]; and the SCINet project of the USDA Agricultural Research Service [ARS project number 0500-00093-001-00-D]. The funders had no role in study design, data collection and interpretation, or the decision to submit the work for publication. Mention of trade names or commercial products in this article is solely for the purpose of providing specific information and does not imply recommendation or endorsement by the USDA, DOE, or ORISE. USDA is an equal opportunity provider and employer.

**Supplementary Data**

The code and data associated with this manuscript is provide at https://github.com/flu-crew. Supplemental figures are available online at XXX. The Nextstrain for swine IAV is hosted at https://flu-crew.org

**References**



Figure 1. H3 Clade detection frequency and C-IV-A relative genetic diversity from 2011 to 2021. (A) Proportional yearly detection frequency of H3 genetic clades from public data. “Other” includes cluster I, cluster IV C-F and one-off human-to-swine transmission events. (B) Effective population size (EPS) and average detection frequency per month of C-IV-A viruses. EPS is used to estimate relative genetic diversity within the HA genes of the C-IV-A clade. Blue shading is the 95% highest posterior density (HPD) interval.

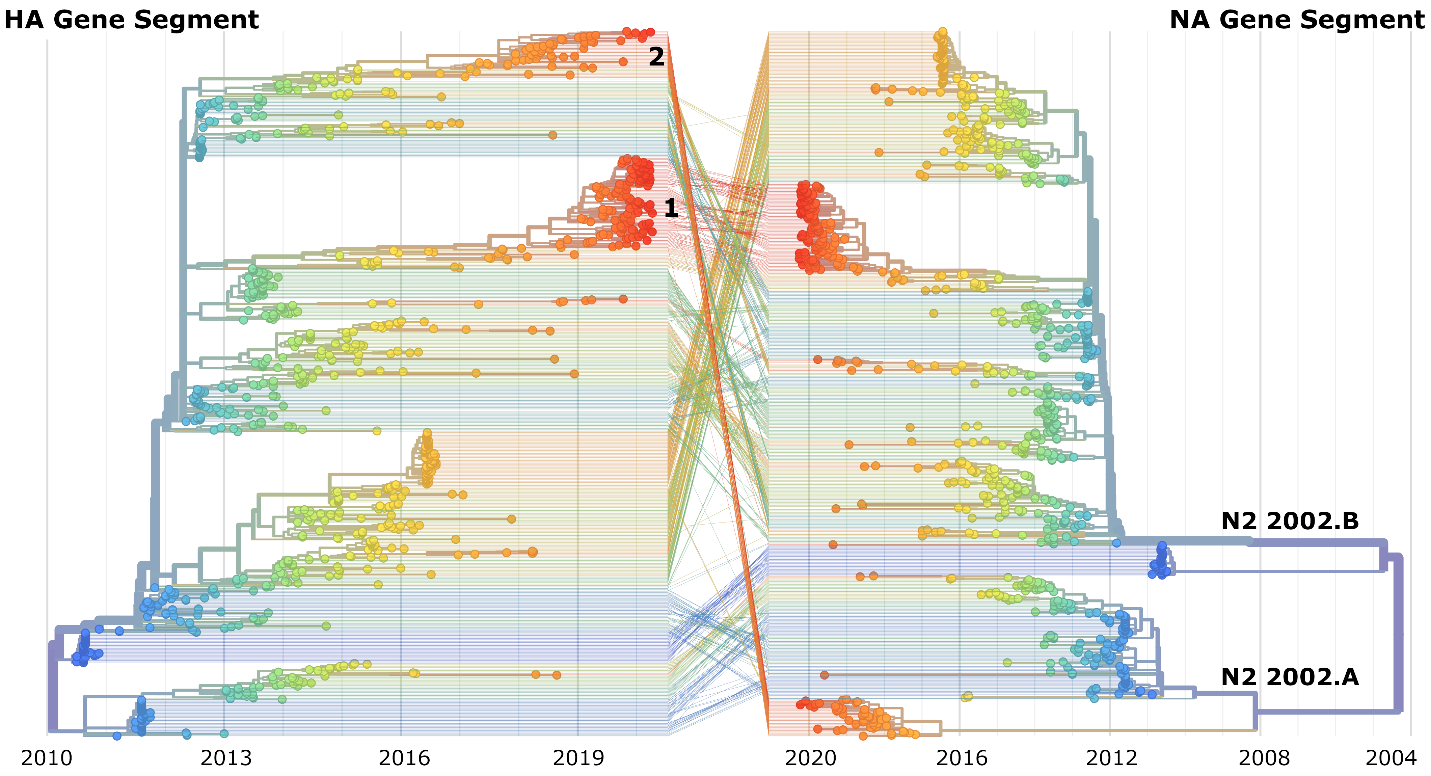


Figure 2. Tanglegram of time-scaled HA and NA gene segment trees with sequences from 2010 to April 2020. Colored by progression of time in accordance with the x-axis.



Figure 3. Time-scaled C-IVA HA tree annotated with internal gene lineages on the horizontal axis in the columns to the right. Yellow=LAIV; Red=H1N1pdm09 lineage; green=swine triple reassortant internal gene (TRIG) lineage.



Figure 4. HA and NA antigenic distance. (A) HA antigenic distance between relevant reference antigens. Distances are the result of merging raw HI results with previously completed HI assay data in ACMACS (B) NA antigenic distance between three contemporary C-IVA reference antigens and antisera from each of the 4 N2 2002 lineages. Distances are the result of merging raw NI results with previously completed HI assay data in ACMACS.

Abente, E. J., et al. (2016), 'The Molecular Determinants of Antibody Recognition and Antigenic Drift in the H3 Hemagglutinin of Swine Influenza A Virus', *J Virol,* 90 (18), 8266-80.

Anderson, T. K., et al. (2020), 'Swine Influenza A Viruses and the Tangled Relationship with Humans', *Cold Spring Harb Perspect Med*.

Arendsee, Z. W., Chang, J., and Leung, E. (2018), ' incertae-sedis/smof: first release (version 2.13.1)'.

Bedford, T., et al. (2014), 'Integrating influenza antigenic dynamics with molecular evolution', *Elife,* 3, e01914.

Benson, D. A., et al. (2013), 'GenBank', *Nucleic Acids Res,* 41 (Database issue), D36-42.

Bolton, M. J., et al. (2019), 'Antigenic evolution of H3N2 influenza A viruses in swine in the United States from 2012 to 2016', *Influenza Other Respir Viruses,* 13 (1), 83-90.

Burke, D. F. and Smith, D. J. (2014), 'A recommended numbering scheme for influenza A HA subtypes', *PLoS One,* 9 (11), e112302.

Chang, J., et al. (2019), 'octoFLU: Automated Classification for the Evolutionary Origin of Influenza A Virus Gene Sequences Detected in U.S. Swine', *Microbiol Resour Announc,* 8 (32).

Drummond, A. J. and Rambaut, A. (2007), 'BEAST: Bayesian evolutionary analysis by sampling trees', *BMC Evol Biol,* 7, 214.

Essere, B., et al. (2013), 'Critical role of segment-specific packaging signals in genetic reassortment of influenza A viruses', *Proc Natl Acad Sci U S A,* 110 (40), E3840-8.

Gamblin, S. J. and Skehel, J. J. (2010), 'Influenza hemagglutinin and neuraminidase membrane glycoproteins', *J Biol Chem,* 285 (37), 28403-9.

Gao, J., Couzens, L., and Eichelberger, M. C. (2016), 'Measuring Influenza Neuraminidase Inhibition Antibody Titers by Enzyme-linked Lectin Assay', *J Vis Exp,* (115).

Hadfield, J., et al. (2018), 'Nextstrain: real-time tracking of pathogen evolution', *Bioinformatics,* 34 (23), 4121-23.

Huddleston, John, et al. (2021), 'Augur: a bioinformatics toolkit for phylogenetic analyses of human pathogens', *Journal of Open Source Software,* 6 (57), 2906.

Ito, T., et al. (1998), 'Molecular basis for the generation in pigs of influenza A viruses with pandemic potential', *J Virol,* 72 (9), 7367-73.

Kaplan, B. S. and Vincent, A. L. (2020), 'Detection and Titration of Influenza A Virus Neuraminidase Inhibiting (NAI) Antibodies Using an Enzyme-Linked Lectin Assay (ELLA)', *Methods Mol Biol,* 2123, 335-44.

Katoh, K. and Standley, D. M. (2013), 'MAFFT multiple sequence alignment software version 7: improvements in performance and usability', *Mol Biol Evol,* 30 (4), 772-80.

Kitikoon, P., et al. (2006), 'The immune response and maternal antibody interference to a heterologous H1N1 swine influenza virus infection following vaccination', *Vet Immunol Immunopathol,* 112 (3-4), 117-28.

Kitikoon, P., et al. (2013), 'Genotype patterns of contemporary reassorted H3N2 virus in US swine', *J Gen Virol,* 94 (Pt 6), 1236-41.

Klingen, T. R., et al. (2018), 'Sweep Dynamics (SD) plots: Computational identification of selective sweeps to monitor the adaptation of influenza A viruses', *Sci Rep,* 8 (1), 373.

Koel, B. F., et al. (2013), 'Substitutions near the receptor binding site determine major antigenic change during influenza virus evolution', *Science,* 342 (6161), 976-9.

Lewis, N. S., et al. (2014), 'Substitutions near the hemagglutinin receptor-binding site determine the antigenic evolution of influenza A H3N2 viruses in U.S. swine', *J Virol,* 88 (9), 4752-63.

Loving, C. L., et al. (2013), 'Efficacy in pigs of inactivated and live attenuated influenza virus vaccines against infection and transmission of an emerging H3N2 similar to the 2011-2012 H3N2v', *J Virol,* 87 (17), 9895-903.

Marshall, N., et al. (2013), 'Influenza virus reassortment occurs with high frequency in the absence of segment mismatch', *PLoS Pathog,* 9 (6), e1003421.

Mayank, A. K., et al. (2015), 'Nucleoprotein of influenza A virus negatively impacts antiapoptotic protein API5 to enhance E2F1-dependent apoptosis and virus replication', *Cell Death Dis,* 6, e2018.

Minin, V. N., Bloomquist, E. W., and Suchard, M. A. (2008), 'Smooth skyride through a rough skyline: Bayesian coalescent-based inference of population dynamics', *Mol Biol Evol,* 25 (7), 1459-71.

Muramoto, Y., et al. (2013), 'Identification of novel influenza A virus proteins translated from PA mRNA', *J Virol,* 87 (5), 2455-62.

Nicholls, J. M., et al. (2008), 'Evolving complexities of influenza virus and its receptors', *Trends Microbiol,* 16 (4), 149-57.

Pons, M. W., et al. (1969), 'Isolation and characterization of the ribonucleoprotein of influenza virus', *Virology,* 39 (2), 250-9.

Portela, A. and Digard, P. (2002), 'The influenza virus nucleoprotein: a multifunctional RNA-binding protein pivotal to virus replication', *J Gen Virol,* 83 (Pt 4), 723-34.

Price, M. N., Dehal, P. S., and Arkin, A. P. (2010), 'FastTree 2--approximately maximum-likelihood trees for large alignments', *PLoS One,* 5 (3), e9490.

Rajao, D. S., et al. (2017), 'Reassortment between Swine H3N2 and 2009 Pandemic H1N1 in the United States Resulted in Influenza A Viruses with Diverse Genetic Constellations with Variable Virulence in Pigs', *J Virol,* 91 (4).

Rajao, D. S., et al. (2015), 'Novel Reassortant Human-Like H3N2 and H3N1 Influenza A Viruses Detected in Pigs Are Virulent and Antigenically Distinct from Swine Viruses Endemic to the United States', *J Virol,* 89 (22), 11213-22.

Rambaut, A. and Drummond, A. J. (2015), 'TreeAnnotator v1. 8.4', *MCMC Output analysis*.

Rambaut, A., et al. (2016), 'Exploring the temporal structure of heterochronous sequences using TempEst (formerly Path-O-Gen)', *Virus Evol,* 2 (1), vew007.

Rambaut, A., et al. (2018), 'Posterior Summarization in Bayesian Phylogenetics Using Tracer 1.7', *Syst Biol,* 67 (5), 901-04.

Rambaut, A., et al. (2008), 'The genomic and epidemiological dynamics of human influenza A virus', *Nature,* 453 (7195), 615-9.

Santos, J. J. S., et al. (2019), 'Plasticity of Amino Acid Residue 145 Near the Receptor Binding Site of H3 Swine Influenza A Viruses and Its Impact on Receptor Binding and Antibody Recognition', *J Virol,* 93 (2).

Smith, D. J., et al. (2004), 'Mapping the antigenic and genetic evolution of influenza virus', *Science,* 305 (5682), 371-6.

Vincent, A. L., et al. (2017), 'Influenza A virus vaccines for swine', *Vet Microbiol,* 206, 35-44.

Webby, R. J., et al. (2000), 'Evolution of swine H3N2 influenza viruses in the United States', *J Virol,* 74 (18), 8243-51.

Yamayoshi, S., et al. (2016), 'Identification of a Novel Viral Protein Expressed from the PB2 Segment of Influenza A Virus', *J Virol,* 90 (1), 444-56.

Zeller, M. A., et al. (2018), 'ISU FLUture: a veterinary diagnostic laboratory web-based platform to monitor the temporal genetic patterns of Influenza A virus in swine', *BMC Bioinformatics,* 19 (1), 397.

Zeller, M. A., et al. (2020), 'Coordinated evolution between N2 neuraminidase and H1 and H3 hemagglutinin genes increased influenza A virus genetic diversity in swine', *bioRxiv*.

Zhang, Y., et al. (2017), 'Influenza Research Database: An integrated bioinformatics resource for influenza virus research', *Nucleic Acids Res,* 45 (D1), D466-D74.

Zhou, N. N., et al. (1999), 'Genetic reassortment of avian, swine, and human influenza A viruses in American pigs', *J Virol,* 73 (10), 8851-6.