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: /250 words; Text: 3,856 words

**Abstract**

The genetic and antigenic diversity of influenza A virus in swine (IAV-S) is shaped by transmission and persistence of human IAV within pig populations. The relative frequency of the phylogenetic clade H3 CIV-A that circulated for 20 years in US swine declined to 7% in 2017, but rose to 32% in 2019. To determine putative mechanisms associated with increased detection, we conducted phylogenetic and phenotypic analyses of representative strains. To visualize and the emergence, spatial spread, and genetic evolution of H3 IAV-S we developed a Nextstrain web application. These data identified two C-IVA clades that emerged in 2017 and cocirculated within multiple US states. Phylodynamic analysis of the HA gene documented low relative genetic diversity from 2017 to 2019, suggesting clonal expansion of genetically similar viruses. The clade with the majority of detections was associated with an N156H amino acid substitution, but HI assays demonstrated no significant antigenic drift associated with this mutation. Genome constellation diversity was also quantified. The minor clade was paired with N2-02B.2 clade in ancestral strains, but acquired an N2-02A.2 in 2016. An 8-fold change in NI titers between the N2 from 02B.2 and 02A.2 clades was observed, indicating antigenic drift between the N2 clades associated with the two H3 clades. The major clade HA gene was tightly linked with the nucleoprotein (NP) of the H1N1pdm09 lineage, indicating reassortment to replace the North American swine lineage NP. These data demonstrate that increased detection of the H3 Clade IV-A was not associated with increased genetic or antigenic diversity of the HA, but was associated with antigenic diversity of the NA and acquisition of the H1N1pdm09 NP. Defining factors driving spatial and temporal patterns in IAV-S diversity in swine is essential to informing vaccine strain selection and strategies to reduce the expansion and spread of potentially zoonotic swine-origin IAV.

(Word count = 301/250)

**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 identified an increase in detection frequency of a phylogenetic 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 HA 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)

**Keywords:** Influenza A virus, H3N2, Swine, Reassortment, Surveillance, Vaccines

**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 (1-3). 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 (4, 5). 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 (6, 7). 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 hemagglutinin (HA) and neuraminidase (NA) antigenic diversity. The accumulation of amino acid substitutions from polymerase mutation in the surface glycoproteins can result in changes of the antigenic phenotype of IAV (8, 9). For the H3 subtype specifically, a small number of amino acid residues have a disproportionate effect on antigenic phenotype in both humans and swine (10-12). In swine, six of these amino acid positions (145,155,156,158,159 and 189; H3 mature peptide numbering (13)) are referred to as the H3 antigenic motif (14). 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 (12). 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 (15). 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 (16, 17). 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(18-20). The HA gene from this introduction evolved into genetically distinct clades, establishing H3 cluster IV in the U.S. (20). This lineage continued to circulate in the U.S. since 2005 and genetically diversified into clades A through F (21). 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 (20, 22, 23). 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 (24) 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.

**Results**

**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; instead, the 2010.1 clade accounted 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) 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 (Supplementary Figure 1) with external branches that are shorter relative to branches on the interior of the tree.

**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 (56%) of this clade were in Iowa and Indiana. The minor clade was initially detected in the Midwest; it was rarely detected outside of these states from 2018 to present (5 detections).

**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 (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.

**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) (Figure 3). The 156 position was identified in the previously characterized antigenic motif (10, 11, 14). 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).

**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 were within 3 AU of the three test strains. The two antigens from cluster IV-B were just over and just under the 3 AU significance cutoff. The CIV-B virus MN/13 with an NYHNYK antigenic motif was closer to all three test viruses, particularly NC/19, than the CIV-B virus IA/13 with an NYNNYK antigenic motif.

**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 OK/15 and NC/19 strains had a N2-2002B.2 gene; the minor clade representative MN/18 had a N2-2002A.2 gene. The NI titer data were combined with previous experiments in ACMACS and the antigenic distances were extracted. There were differences in NI antigenic phenotype within the N2 2002 lineage, further divided into 2002.A.1, 2002.A.2, 2002.B.1 and 2002.B.2 (25) (Kaplan, submitted for publication). In particular, the minor clade MN/18 virus with the 2002.A.2 lineage had an antigenic distance of 3 AU from the 2002.B.2 lineage viruses (Figure 4b). The 2002B.2 lineage viruses were within 0.5 AU.

**Discussion**

In this study, we investigated possible factors to explain the recent increase in detection of H3 C-IVA viruses in US swine. Though the number of C-IVA detections increased since 2018, the relative genetic diversity within the clade did not increase until mid-2020. This pattern is distinct from a prior increase in detection frequency that was paired with simultaneous diversification from 2013 to 2015. The 2018-2019 clonal expansion of C-IVA with low diversity suggests that a selective sweep occurred in the population. Sweep-related changes have been identified in human seasonal H3N2 IAV and most often detected at amino acid sites located on the HA (26, 27). In 2019, the relative genetic diversity began to rise, likely as the result of the success, spread, and subsequent diversification of the virus. We believe that the pattern of increasing detection frequency paired with low EPS may signal the need for further investigation into a potentially highly successful clade.

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 in the HA of the major clade. The 156 amino acid position was previously identified as having a disproportionate effect on the antigenic phenotype, usually in combination with substitutions in other positions on the HA (10, 14, 28). 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 (29, 30).

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, our 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. Our results support previous findings that variation at position 156 alone did not cause significant antigenic drift (14, 28) . 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 expansion. Through examination of whole genomes, it was found that the minor clade recently reassorted to obtain N2 2002A.2 genes, while the major clade remained paired with N2 2002B.2. 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 for publication). While these results showed antigenic variation within the N2 2002 lineage, the N2 2002B.2 of the major clade viruses retained close antigenic relationships to other 2002.A and 2002.B swine N2 lineages. Therefore, loss of immunity to the N2 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 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 (31). 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 (32, 33).

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. The original C-IVA internal gene constellation in 2010 was TTTTTT, with all internal genes coming from the TRIG lineage. Reassortment of H3N2-TRIG gene segments occurred with the introduction of the 2009 H1N1 pandemic virus (H1N1pdm09). The C-IVA virus that acquired only a matrix gene from the H1N1pdm09 virus, with constellation TTTTPT, was highly successful. This was the most common constellation (n=119) found in a sample of 368 H3N2 isolates collected between 2009 and 2016 (34). It was also responsible for an H3N2v outbreak in humans in 2011-2012, causing 340 cases across 13 US states (35).

The internal gene constellation found in the reassorted C-IVA major clade, TTTPPT, had been observed before; however, it was uncommon and was detected in only 7 of the 368 isolates collected between 2009 and 2016. Results from the same study showed a wildtype field strain selected from those viruses containing the TTT**P**PT constellation was more effective in viral transmission compared to wildtype strains with the TTT**T**PT constellation. This creates concern for public health, with the knowledge that a virus from the same clade was successful in causing a human outbreak in the context of reassortment. Our HI assay included a representative strain (A/swine/New York/A01104005/2011) that was genetically similar to the H3N2v from the 2011-2012 outbreak and showed that none of the contemporary C-IVA representatives had undergone significant antigenic drift. Thus the current C-IVA viruses likely have a similar antigenic phenotype to those that caused a human outbreak.

Dynamics within the H3 clade could also help explain the resurgence of the C-IVA clade. Of note is 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 (23). Many herds were vaccinated against the 2010.1 clade of viruses with custom or autogenous vaccines following its emergence and dominance over C-IVA. Waning immunity against C-IVA viruses due to focus on vaccines containing 2010.1 viruses likely allowed a competitive advantage of the C-IVA due to the lack of population immunity against this specific clade of swine H3N2. This is supported by the balanced shape of the HA phylogeny which suggests no selection on the emerged clades.

C-IVA viruses continue to make up roughly one-half of H3N2 detections into the early months of 2021. 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. 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 (35-37) 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 for publication).

**Materials and Methods**

**Data Collection**

All available U.S. swine H3 nucleotide sequences (n=3395) detected between January 2010 and March 2021 deposited into GenBank (38) were downloaded from the Influenza Research Database (IRD) (39). Duplicate strains were removed from the dataset. Sequences were then classified using octoFLU (40), 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 (22).

**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 (41). Sequences were aligned with mafft v7.450 (42) and a maximum-likelihood phylogenetic tree was inferred using the generalized time-reversible model (GTR) of nucleotide substitution in FastTree v2.1.11 (43). This tree was used in a root-to-tip regression analysis in TempEst v1.5.3 (44) 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 (45) to estimate effective population size of the C-IVA lineage over time. We applied the GMRF Bayesian Skyride coalescent model (46) 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 (47) and a maximum clade credibility (MCC) tree was generated using TreeAnnotator v1.8.4 (48).

**Deployment of Nextstrain for H3 IAV in swine**

The Nextstrain (24) 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 (49). 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/.

**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 (10, 28). 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 (8, 10, 28).

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 (50, 51). 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.

**Data Availability**

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.

**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.

**References**

1. Muramoto Y, Noda T, Kawakami E, Akkina R, Kawaoka Y.2013. Identification of novel influenza A virus proteins translated from PA mRNA. J Virol 87:2455-62.

2. Yamayoshi S, Watanabe M, Goto H, Kawaoka Y.2016. Identification of a Novel Viral Protein Expressed from the PB2 Segment of Influenza A Virus. J Virol 90:444-56.

3. Gamblin SJ, Skehel JJ.2010. Influenza hemagglutinin and neuraminidase membrane glycoproteins. J Biol Chem 285:28403-9.

4. Essere B, Yver M, Gavazzi C, Terrier O, Isel C, Fournier E, Giroux F, Textoris J, Julien T, Socratous C, Rosa-Calatrava M, Lina B, Marquet R, Moules V.2013. Critical role of segment-specific packaging signals in genetic reassortment of influenza A viruses. Proc Natl Acad Sci U S A 110:E3840-8.

5. Marshall N, Priyamvada L, Ende Z, Steel J, Lowen AC.2013. Influenza virus reassortment occurs with high frequency in the absence of segment mismatch. PLoS Pathog 9:e1003421.

6. Nicholls JM, Chan RW, Russell RJ, Air GM, Peiris JS.2008. Evolving complexities of influenza virus and its receptors. Trends Microbiol 16:149-57.

7. Ito T, Couceiro JN, Kelm S, Baum LG, Krauss S, Castrucci MR, Donatelli I, Kida H, Paulson JC, Webster RG, Kawaoka Y.1998. Molecular basis for the generation in pigs of influenza A viruses with pandemic potential. J Virol 72:7367-73.

8. Smith DJ, Lapedes AS, de Jong JC, Bestebroer TM, Rimmelzwaan GF, Osterhaus AD, Fouchier RA.2004. Mapping the antigenic and genetic evolution of influenza virus. Science 305:371-6.

9. Bedford T, Suchard MA, Lemey P, Dudas G, Gregory V, Hay AJ, McCauley JW, Russell CA, Smith DJ, Rambaut A.2014. Integrating influenza antigenic dynamics with molecular evolution. Elife 3:e01914.

10. Lewis NS, Anderson TK, Kitikoon P, Skepner E, Burke DF, Vincent AL.2014. Substitutions near the hemagglutinin receptor-binding site determine the antigenic evolution of influenza A H3N2 viruses in U.S. swine. J Virol 88:4752-63.

11. Koel BF, Burke DF, Bestebroer TM, van der Vliet S, Zondag GC, Vervaet G, Skepner E, Lewis NS, Spronken MI, Russell CA, Eropkin MY, Hurt AC, Barr IG, de Jong JC, Rimmelzwaan GF, Osterhaus AD, Fouchier RA, Smith DJ.2013. Substitutions near the receptor binding site determine major antigenic change during influenza virus evolution. Science 342:976-9.

12. Santos JJS, Abente EJ, Obadan AO, Thompson AJ, Ferreri L, Geiger G, Gonzalez-Reiche AS, Lewis NS, Burke DF, Rajao DS, Paulson JC, Vincent AL, Perez DR.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.

13. Burke DF, Smith DJ.2014. A recommended numbering scheme for influenza A HA subtypes. PLoS One 9:e112302.

14. Abente EJ, Santos J, Lewis NS, Gauger PC, Stratton J, Skepner E, Anderson TK, Rajao DS, Perez DR, Vincent AL.2016. The Molecular Determinants of Antibody Recognition and Antigenic Drift in the H3 Hemagglutinin of Swine Influenza A Virus. J Virol 90:8266-80.

15. Vincent AL, Perez DR, Rajao D, Anderson TK, Abente EJ, Walia RR, Lewis NS.2017. Influenza A virus vaccines for swine. Vet Microbiol 206:35-44.

16. Loving CL, Lager KM, Vincent AL, Brockmeier SL, Gauger PC, Anderson TK, Kitikoon P, Perez DR, Kehrli ME, Jr.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 87:9895-903.

17. Kitikoon P, Nilubol D, Erickson BJ, Janke BH, Hoover TC, Sornsen SA, Thacker EL.2006. The immune response and maternal antibody interference to a heterologous H1N1 swine influenza virus infection following vaccination. Vet Immunol Immunopathol 112:117-28.

18. Webby RJ, Swenson SL, Krauss SL, Gerrish PJ, Goyal SM, Webster RG.2000. Evolution of swine H3N2 influenza viruses in the United States. J Virol 74:8243-51.

19. Zhou NN, Senne DA, Landgraf JS, Swenson SL, Erickson G, Rossow K, Liu L, Yoon K, Krauss S, Webster RG.1999. Genetic reassortment of avian, swine, and human influenza A viruses in American pigs. J Virol 73:8851-6.

20. Anderson TK, Chang J, Arendsee ZW, Venkatesh D, Souza CK, Kimble JB, Lewis NS, Davis CT, Vincent AL.2020. Swine Influenza A Viruses and the Tangled Relationship with Humans. Cold Spring Harb Perspect Med doi:10.1101/cshperspect.a038737.

21. Kitikoon P, Nelson MI, Killian ML, Anderson TK, Koster L, Culhane MR, Vincent AL.2013. Genotype patterns of contemporary reassorted H3N2 virus in US swine. J Gen Virol 94:1236-41.

22. Zeller MA, Anderson TK, Walia RW, Vincent AL, Gauger PC.2018. ISU FLUture: a veterinary diagnostic laboratory web-based platform to monitor the temporal genetic patterns of Influenza A virus in swine. BMC Bioinformatics 19:397.

23. Rajao DS, Gauger PC, Anderson TK, Lewis NS, Abente EJ, Killian ML, Perez DR, Sutton TC, Zhang J, Vincent 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 89:11213-22.

24. Hadfield J, Megill C, Bell SM, Huddleston J, Potter B, Callender C, Sagulenko P, Bedford T, Neher RA.2018. Nextstrain: real-time tracking of pathogen evolution. Bioinformatics 34:4121-4123.

25. Zeller MA, Chang J, Vincent AL, Gauger PC, Anderson TK.2020. Coordinated evolution between N2 neuraminidase and H1 and H3 hemagglutinin genes increased influenza A virus genetic diversity in swine. bioRxiv.

26. Klingen TR, Reimering S, Loers J, Mooren K, Klawonn F, Krey T, Gabriel G, McHardy AC.2018. Sweep Dynamics (SD) plots: Computational identification of selective sweeps to monitor the adaptation of influenza A viruses. Sci Rep 8:373.

27. Rambaut A, Pybus OG, Nelson MI, Viboud C, Taubenberger JK, Holmes EC.2008. The genomic and epidemiological dynamics of human influenza A virus. Nature 453:615-9.

28. Bolton MJ, Abente EJ, Venkatesh D, Stratton JA, Zeller M, Anderson TK, Lewis NS, Vincent AL.2019. Antigenic evolution of H3N2 influenza A viruses in swine in the United States from 2012 to 2016. Influenza Other Respir Viruses 13:83-90.

29. Zeller MA, Gauger PC, Arendsee ZW, Souza CK, Vincent AL, Anderson TK.2021. Machine Learning Prediction and Experimental Validation of Antigenic Drift in H3 Influenza A Viruses in Swine. mSphere 6.

30. Sun H, Yang J, Zhang T, Long LP, Jia K, Yang G, Webby RJ, Wan XF.2013. Using sequence data to infer the antigenicity of influenza virus. mBio 4.

31. Pons MW, Schulze IT, Hirst GK, Hauser R.1969. Isolation and characterization of the ribonucleoprotein of influenza virus. Virology 39:250-9.

32. Portela A, Digard P.2002. The influenza virus nucleoprotein: a multifunctional RNA-binding protein pivotal to virus replication. J Gen Virol 83:723-734.

33. Mayank AK, Sharma S, Nailwal H, Lal SK.2015. Nucleoprotein of influenza A virus negatively impacts antiapoptotic protein API5 to enhance E2F1-dependent apoptosis and virus replication. Cell Death Dis 6:e2018.

34. Rajao DS, Walia RR, Campbell B, Gauger PC, Janas-Martindale A, Killian ML, Vincent 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.

35. Control CfD, Prevention.2012. Influenza A (H3N2) variant virus-related hospitalizations: Ohio, 2012. MMWR Morbidity and mortality weekly report 61:764-767.

36. Kitikoon P, Gauger PC, Anderson TK, Culhane MR, Swenson S, Loving CL, Perez DR, Vincent AL.2013. Swine influenza virus vaccine serologic cross-reactivity to contemporary US swine H3N2 and efficacy in pigs infected with an H3N2 similar to 2011-2012 H3N2v. Influenza Other Respir Viruses 7 Suppl 4:32-41.

37. Shu B, Garten R, Emery S, Balish A, Cooper L, Sessions W, Deyde V, Smith C, Berman L, Klimov A, Lindstrom S, Xu X.2012. Genetic analysis and antigenic characterization of swine origin influenza viruses isolated from humans in the United States, 1990-2010. Virology 422:151-60.

38. Benson DA, Cavanaugh M, Clark K, Karsch-Mizrachi I, Lipman DJ, Ostell J, Sayers EW.2013. GenBank. Nucleic Acids Res 41:D36-42.

39. Zhang Y, Aevermann BD, Anderson TK, Burke DF, Dauphin G, Gu Z, He S, Kumar S, Larsen CN, Lee AJ, Li X, Macken C, Mahaffey C, Pickett BE, Reardon B, Smith T, Stewart L, Suloway C, Sun G, Tong L, Vincent AL, Walters B, Zaremba S, Zhao H, Zhou L, Zmasek C, Klem EB, Scheuermann RH.2017. Influenza Research Database: An integrated bioinformatics resource for influenza virus research. Nucleic Acids Res 45:D466-D474.

40. Chang J, Anderson TK, Zeller MA, Gauger PC, Vincent AL.2019. octoFLU: Automated Classification for the Evolutionary Origin of Influenza A Virus Gene Sequences Detected in U.S. Swine. Microbiol Resour Announc 8.

41. Arendsee ZW, Chang J, Leung E. 2018. incertae-sedis/smof: first release (version 2.13.1) doi:10.5281/zenodo.1434656.

42. Katoh K, Standley DM.2013. MAFFT multiple sequence alignment software version 7: improvements in performance and usability. Mol Biol Evol 30:772-80.

43. Price MN, Dehal PS, Arkin AP.2010. FastTree 2--approximately maximum-likelihood trees for large alignments. PLoS One 5:e9490.

44. Rambaut A, Lam TT, Max Carvalho L, Pybus OG.2016. Exploring the temporal structure of heterochronous sequences using TempEst (formerly Path-O-Gen). Virus Evol 2:vew007.

45. Drummond AJ, Rambaut A.2007. BEAST: Bayesian evolutionary analysis by sampling trees. BMC Evol Biol 7:214.

46. Minin VN, Bloomquist EW, Suchard MA.2008. Smooth skyride through a rough skyline: Bayesian coalescent-based inference of population dynamics. Mol Biol Evol 25:1459-71.

47. Rambaut A, Drummond AJ, Xie D, Baele G, Suchard MA.2018. Posterior Summarization in Bayesian Phylogenetics Using Tracer 1.7. Syst Biol 67:901-904.

48. Rambaut A, Drummond AJ.2015. TreeAnnotator v1. 8.4. MCMC Output analysis.

49. Huddleston J, Hadfield J, Sibley TR, Lee J, Fay K, Ilcisin M, Harkins E, Bedford T, Neher RA, Hodcroft EB.2021. Augur: a bioinformatics toolkit for phylogenetic analyses of human pathogens. Journal of Open Source Software 6:2906.

50. Gao J, Couzens L, Eichelberger MC.2016. Measuring Influenza Neuraminidase Inhibition Antibody Titers by Enzyme-linked Lectin Assay. J Vis Exp doi:10.3791/54573.

51. Kaplan BS, Vincent AL.2020. Detection and Titration of Influenza A Virus Neuraminidase Inhibiting (NAI) Antibodies Using an Enzyme-Linked Lectin Assay (ELLA). Methods Mol Biol 2123:335-344.

**Figure Legends**

**Figure 1.** Relative H3 clade detection frequency and relative C-IVA 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-IVA viruses. EPS estimates relative genetic diversity within the HA genes of the C-IVA clade. Blue shading is 95% highest posterior density (HPD) interval.

**Figure 2.** Tanglegram of corresponding C-IVA HA and NA gene segment time-scaled trees with sequences from 2010 to March 2021. The major and minor contemporary C-IVA clades are labeled on the HA tree. N2 genetic clades are labeled on the NA tree. Lines between the two trees indicate HA-NA pairing. Tree branches and leaves colored on blue to red gradient by yearly progression of time in accordance with the x-axis. Blue, dotted line indicates pairing of the minor clade with N2 2002B.2 prior to reassortment and orange, dashed line indicates reassortment of the minor clade with N2 2002A.2. The Nextstrain platform can be used to visualize the tanglegram in finer detail at flu-crew.org/nextstrain.

**Figure 3.** C-IVA HA time-scaled tree with genome constellation. The corresponding NA and internal gene (PB2, PB1, PA, NP, M, NS) lineage of each tip are indicated by the color-coded rectangles on the right side of figure. The major contemporary C-IVA clade is boxed in dark grey on the tree and the minor contemporary C-IVA clade is boxed in light grey on the tree. The two substitutions defining the initial expansion of the major clade (N156H and K368E) are annotated onto the tree. Dark Orange=N2 2002A; Light Orange = N2 2002A.2; Pink = N2 2002B.2; 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 three contemporary C-IVA test antigens and relevant H3 reference antigens. Distances are computed by merging the raw HI results from the assay described in this experiment with results from previous HI assays in ACMACS. Points are colored by H3 antigenic motif and their shape corresponds to their H3 clade classification. The three test antigens on the x-axis are colored by H3 antigenic motif. Significant HA antigenic drift is defined as antigenic distance of at least 3 AU and is denoted by the black dashed line. **(B)** NA antigenic distance between three contemporary C-IVA test antigens and antigens from four N2 2002 lineages. Distances are computed by merging the raw NI results from the assay described in this experiment with results from previous NI assays in ACMACS. Points are colored by NA lineage. The three test antigens on the x-axis are also colored by NA lineage. The test antigens, OK/15, NC/10 and MN/18, are labeled with black text in both panels.

**Figure S1.** H3 C-IVA maximum likelihood tree. Branches are colored by the progression of time on a blue-to-red gradient beginning in 2010 and ending in March 2021.

**Table S1.** Results from a pair of hemagglutination inhibition (HI) assays performed on November 13, 2020 and January 22nd, 2021.

**Table S2.** Results from Neuraminidase Inhibition (NI) Assay.