Genotype patterns of contemporary reassorted H3N2 virus in US swine

Pravina Kitikoon,1 Martha I. Nelson,2 Mary Lea Killian,3 Tavis K. Anderson,1 Leo Koster,3 Marie R. Culhane4 and Amy L. Vincent1

1Virus and Prion Disease Research Unit, National Animal Disease Center, USDA-ARS, Ames, IA 50010, USA
2Fogarty International Center, National Institutes of Health, Bethesda, MD 20892, USA
3National Veterinary Services Laboratories, USDA-APHIS, Ames, IA 50010, USA
4Veterinary Diagnostic Laboratory, College of Veterinary Medicine, University of Minnesota, St. Paul, MN 55108, USA

To understand the evolution of swine-origin H3N2v influenza viruses that have infected 320 humans in the USA since August 2011, we performed a phylogenetic analysis at a whole genome scale of North American swine influenza viruses (n = 200). All viral isolates evolved from the prototypical North American H3 cluster 4 (c4), with evidence for further diversification into subclusters. At least ten distinct reassorted H3N2/pandemic H1N1 (rH3N2p) genotypes were identified in swine. Genotype 1 (G1) was most frequently detected in swine and all human H3N2v viruses clustered within a single G1 clade. These data suggest that the genetic requirements for transmission to humans may be restricted to a specific subset of swine viruses. Mutations at putative antigenic sites as well as reduced serological cross-reactivity among the H3 subclusters suggest antigenic drift of these contemporary viruses.

Influenza A virus (IAV) infection remains one of the most important causes of respiratory diseases in humans and animals. Humans and swine express similar IAV receptors in the respiratory tract, facilitating bidirectional transmission between the two species (Imai & Kawaoka, 2012). Pigs are capable of generating novel reassortant viruses with the potential to be highly infectious and transmissible in humans, representing a continual public health threat (Zhou et al., 1999). Consequently, understanding patterns of IAV genetic and antigenic diversity in swine is critical to assess risk in human and swine populations.

Seven antigenically distinct haemagglutinin (HA) clusters, H1α, H1β, H1γ, H1δ1, H1δ2, H1N1pdm09 and H3 cluster 4 (c4) co-circulate in the US swine population (Lorusso et al., 2012). Until 2009, all HA lineages except for the H1N1pdm09 virus contained the ‘triple reassortment internal gene’ (TRIG) constellation of swine (M, NP and NS genes), avian (PB2 and PA genes), and human (PB1) lineages identified in swine in 1998. The internal gene constellation in H1N1pdm09 is distinct from the H3N2-TRIG as the M gene is of Eurasian swine lineage (Garten et al., 2009). The level of endemic viral diversity is frequently enhanced by interspecies transmission events: within 2 years of the emergence of H1N1pdm09 in humans, at least 49 spillover events of H1N1pdm09 from humans to swine occurred (Nelson et al., 2012a), and, at least 23 additional human-to-swine transmission events of human seasonal H1 and H3 influenza viruses have been identified (Nelson et al., 2012a). These introductions have led to multiple reassortment events between H1N1pdm09 and the endemically circulating swine IAVs including H3N2-TRIG viruses (Ducatez et al., 2011). The H3N2-TRIG/H1N1pdm09 (rH3N2p) reassortant viruses in the swine population are of particular concern since, from August 2011 to November 2012, there have been 320 human cases of infection with an H3N2 virus with an M segment of H1N1pdm09 origin and the remaining genes of swine origin (H3N2v) (CDC, 2012; Lindstrom et al., 2012). These H3N2v human cases highlight the need for whole genome characterization to further understand the evolution of rH3N2p viruses and the diversity of IAV in swine and humans.

Here we analysed 200 whole genome sequences of IAV consisting of 67 swine H1 viruses, 14 human H3N2v, 119 swine H3N2 viruses including three viruses isolated from pigs in a county fair associated with human H3N2v cases in
August 2012. The objective was to illustrate the genomic patterns of rH3N2p viruses and characterize the amino acid residues at predicted antigenic sites and receptor binding sites of the HA gene. Whole genome sequences were obtained for 119 viruses (97 H3 and 22 H1 viruses) submitted through the US Department of Agriculture (USDA)–National Animal Health Laboratory Network (NAHLN) voluntary swine IAV surveillance system. H3N2 samples were subtyped and screened for the H1N1pdm09 M segment (pdm-M) (Lorusso et al., 2010) and virus isolates with pdm-M were given priority for whole genome sequencing given public health concerns, although a number of H3N2 without pdm-M were sequenced as well. Since viruses were submitted on a voluntary basis through passive surveillance, data presented here may not reflect the true prevalence of each virus genotype in the swine population. All sequence data were deposited in the Influenza Virus Resource at GenBank; H3N2 virus accession numbers are provided in Table S1, available in JGV Online (H1 virus accession numbers are available upon request). Isolates were collected from 13 states (IA, IL, IN, KS, MN, MO, NC, NE, NY, OH, OK, PA and TX) during 2009–2012. An additional 67 (22 H3 and 45 H1) North American swine influenza viruses were downloaded from GenBank, and 14 human H3N2v sequences were downloaded from GISAID (Table S2). The final dataset (n = 200) contained 119 H3N2 swine viruses, 67 H1 swine viruses and 14 human H3N2v viruses. Sequences were aligned for each individual internal gene segment (PB2, PB1, PA, NP, M and NS) and separately for the H1, H3, N1 and N2 segments using the Se-Al program (Rambaut, 2002). A maximum-likelihood (ML) tree was inferred for each of the ten alignments using the PhyML program v3.0 (Guindon & Gascuel, 2003), employing SPR (subtree pruning and regrafting) branch-swapping and a general time reversible (GTR) model of nucleotide substitution with gamma-distributed rate variation among sites. Statistical support for individual nodes was estimated by bootstrap analysis (1000 replicates) employing the same GTR + G4 model using paup* (Swofford, 2003). Genotypes were defined by the reassortment patterns of the six internal gene lineages originating either from the H3N2-TRIG and/or H1N1pdm09 viruses.

Phylogenetic analysis revealed at least ten genotypes (G1–G10) of rH3N2p currently circulating in the US swine population (Tables 1 and S1 and Fig. 1). The number of H1N1pdm09 gene segments incorporated into rH3N2p ranged from a single segment (G1) to all six internal genes (G10). Notably, genotype G1, which is associated with all 2011–2012 H3N2v human cases, was detected more frequently than any other genotype (~36 % of rH3N2p viruses). G1 was also the most spatially dispersed genotype, isolated in six US states (Fig. 2). G4 was the second most frequent genotype (5:3 reassortant with H1N1pdm09 NP, M and NS segments). Only two other genotypes exhibited sustained onward transmission in swine, as evidenced by their phylogeny: G6 (4:4 reassortant with H1N1pdm09 PA, NP, M and NS); and G9 (3:5 reassortant with H1N1pdm09 PB1, PA, NP, M and NS). Onward transmission was not observed among the remaining six genotypes, each represented by five or fewer isolates. Phylogenetic analysis of each

**Table 1. Genotype constellation identified in contemporary H3N2 viruses isolated from swine in the USA**

Origin of each gene segment is colour-coded according to the gene lineage: blue, H3 and N2 of the TRIG human seasonal lineage; green, TRIG avian lineage; pink, TRIG swine lineage; lime green, 2009 pandemic H1N1 lineage. Genotypes 1–10 of reassorted rH3N2p and H3N2-TRIG viruses are colour-coded in the first column by the pattern of gene constellation. Similar colour-coding of genotypes is used in Fig. 1. The H3 subcluster A–F and N2 gene lineage is indicated in the HA and NA (neuraminidase) columns as denoted in Figs 1 and S6, respectively.

<table>
<thead>
<tr>
<th>Genotype</th>
<th>PB2</th>
<th>PB1</th>
<th>PA</th>
<th>HA</th>
<th>NP</th>
<th>NA</th>
<th>M</th>
<th>NS</th>
<th>No. of isolates</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>A</td>
<td>E</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>23</td>
</tr>
<tr>
<td>2</td>
<td>A</td>
<td>E</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>5*</td>
</tr>
<tr>
<td>3</td>
<td>B</td>
<td>D</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>2</td>
</tr>
<tr>
<td>4</td>
<td>B</td>
<td>D</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>14*</td>
</tr>
<tr>
<td>5</td>
<td>B</td>
<td>D</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>1</td>
</tr>
<tr>
<td>6</td>
<td>D</td>
<td>F</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>3</td>
</tr>
<tr>
<td>7</td>
<td>C</td>
<td>E</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>1</td>
</tr>
<tr>
<td>8</td>
<td>C</td>
<td>D</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>7</td>
</tr>
<tr>
<td>9</td>
<td>D</td>
<td>F</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>3</td>
</tr>
<tr>
<td>10</td>
<td>C</td>
<td>D</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>55</td>
</tr>
</tbody>
</table>

*2 of 5 isolates, 1 of 14 isolates, and 1 of 5 isolates from G2, G4 and G6, respectively, have N2-2002 gene lineage.
gene segment indicated that these minor genotypes (G2, G3, G5, G7, G8 and G10) appear to have evolved through multiple independent reassortment events, as indicated in Figs. 1 and S1–S7, as the genes from these viruses are located on separate tree branches rather than forming monophyletic clusters. All ten of the rH3N2p genotypes contained M, eight contained NP, and seven contained the PA from H1N1pdm09. In contrast, the H1N1pdm09 PB2 and PB1 segments were the least commonly observed among the rH3N2p reassortants ($n=3$ and 2 genotypes, respectively). The first rH3N2p swine isolate was identified in November 2009 (A/swine/Minnesota/239105/2009, G5), but the first virus with G1 pattern virus was not detected until November 2010 (A/swine/Iowa/A01049034/2010), approximately 8 months before the first H3N2v virus was isolated in humans (Fig. S9). Following their emergence, G1 viruses appear to have proliferated in swine during 2011, coincident with the increase in human H3N2v cases.

Based on the HA gene, historic swine H3N2 viruses in the USA belong to one of the three clusters (c1–c3) derived from three distinct human seasonal H3N2 introductions from 1995, 1997 and 1996, respectively (Webby et al., 2000). In 2005, a separate monophyletic H3 cluster (c4) evolved from c3 viruses and was detected in Canada (Olsen et al., 2006) and subsequently in US swine populations (Kumar et al., 2011). The HA gene of all H3 viruses reported here evolved from the H3 c4 that was relatively evolutionarily stable in North American swine since 2005.
In addition to the reassortment described above, the HA gene of H3 has diversified since 2010, and the rH3N2p viruses were associated with six different clades identified on the H3 phylogeny (clades A–F, Fig. 1). The majority of G1 viruses are positioned within clade A, G4 viruses in clade F, and G9 viruses in clade D (Table S3). Three G1 isolates were found in clade E, indicating that G1 isolates are not monophyletic and the genotype has emerged through independent events. In fact, genotypes 1, 2, 3, 6, 7 and 10 were also not monophyletic in H3 phylogeny, suggesting that the majority of genotypes have evolved through more than one evolutionary pathway. In some cases, more similar genotypes cluster together, such as genotypes 1 and 2b and genotypes 9 and 10, which differ by a single segment, suggesting that in some cases similar genotypes may share common ancestry (Fig. 1). Although the ten genotypes were defined by the evolutionary origin of the six internal gene segments, two genetically distinct N2 lineages currently co-circulate in US swine, those related to the 1998 triple reassortant viruses (N2-1998) and those which were acquired circa 2002 by a subsequent reassortment event with human seasonal H3N2 viruses (N2-2002) (Nelson et al., 2012b) (Fig. S5). Including the NA lineage gives up to 14 distinct H3N2 genotypes. Whole-genome sequences from 14 H3N2v viruses isolated from humans in 2011–2012 were included in the phylogenetic analyses (Figs. 1 and S1–S7). All human H3N2v viruses clustered within clade A of the HA segment with the G1 swine viruses. The human H3N2v viruses clustered together with clade A swine viruses in the PB2, PB1, PA, H3, NP, M and NS phylogenies, but form two genetically distinct clusters within the large N2-2002 clade on the N2 phylogeny. Although the majority of 2011 H3N2v human isolates cluster within clade A on the N2 phylogeny, all five H3N2v isolates collected in 2012 and one 2011 isolate (A/West Virginia/06/2011) cluster in a different part of the tree, along with a third cluster of G1 swine viruses. This suggests that human transmission events were associated with two genetically distinct N2-2002 swine lineages of the G1 genotype. No human H3N2v viruses were located with clade E G1 swine viruses, suggesting that the requirements for transmission to humans may be more stringent than simply the G1 genotype, and may require sequence signatures specific to the HA clade.

Fig. 2. Spatial patterns of the ten rH3N2p genotypes and H3N2-TRIG viruses from the USA and Canada analysed in this study. Genotypes (Table 1) are indicated by the number and T indicates the H3N2-TRIG gene constellation.
the emergence of a new antigenic cluster in human H3N2
Given that a single amino acid substitution can result in
substitutions in only two antigenic sites (Table S4).

A of G1, the 2002-lineage NA, and possibly in other segments as well. Phylogenetic evidence of human-to-human trans-
mIssion is limited to IA (2011, three isolates) and possibly
WV (2011, two isolates), with the human H3N2v isolates clustering with swine viruses throughout clade A.

To identify a biological basis for the observed diversity in the HA gene we constructed amino acid alignment
sequences including swine and human seasonal reference sequences (Table S3). Positively selected sites in the HA gene were predicted using the fixed effect likelihood (FEL) method in
Datamonkey version 2.10b (Kosakovsky Pond & Frost, 2005; Pond et al., 2005) and validated using the mixed effects model of evolution (MEME) implemented on the
Datamonkey server (Delport et al., 2010; Murrell et al., 2012). Thirteen sites were found to be under positive selection, 12 of which are within putative antigenic sites
(Table S3). The substitution rate (dN/dS) accelerated in swine H3 HA genes from viruses isolated during the year following the emergence of the H1N1pdm09 virus compared with isolates collected after July 2010, indicating higher viral adaptation possibly due to reassortment events
(Tables S6–S8). Pairwise amino acid identity of swine rH3N2p HA and human seasonal H3N2 was 79.0–91.3 %, and identity between swine rH3N2p and a swine H3 c4-
prototype from 2005 virus ranged from 93.7 to 97.8 %. Compared with a swine c4 prototype virus, H3 clades A, C, D, E and F viruses had 1–6 amino acid changes in at least four of five antigenic sites. The H3 clade B viruses and the majority of TRIG genotype viruses had amino acid substitutions in only two antigenic sites (Table S4).

Given that a single amino acid substitution can result in the emergence of a new antigenic cluster in human H3N2 and equine H3 viruses (Lewis et al., 2011; Smith et al., 2004), we conducted cross-HA inhibition assays. Specific antisera of swine H3N2 isolates from different H3 subclusters and human H3N2 vaccine strains were generated using previously described methods (Lorusso et al., 2011). In general, swine H3N2 antisera had low cross-reactivity with the human H3N2 strains and vice versa (Table S9). Antisera produced against contemporary swine H3 isolates had low cross-reactivity to c1 virus (A/
sw/TX/4199-2/1998), older c4 isolates (A/sw/MN/01146/2006 and A/sw/IL/02970/2011) and inconsistent cross-reactivity among the new subclusters. Our data concur with previous studies that demonstrate increased viral evolution after an introduction of a virus into a new host or geographical locality (Lam et al., 2012), and the role of reassortment in antigenic change in swine influenza viruses (Vijaykrishna et al., 2011). Whether the documented diversity will result in frequent swine-to-human transmis-
sion events is unknown.

In conclusion, we identified ten distinct genomic patterns of H3N2 viruses isolated from swine from 2009 to 2012, representing a substantial increase in the diversity of circulating swine H3N2 genotypes since the introduction of H1N1pdm09 viruses in 2009. Only viruses from the G1 genotype and HA clade A have been identified in the ~300 human cases of H3N2v in 2011–2012. The frequency of H3N2 submissions to the USDA Swine Influenza Virus (SIV) Surveillance system increased in the last quarter of 2012 compared with previous months (J. A. Korslund, unpublished data). This may indicate an increase in cases of respiratory disease due to rH3N2p and correlates to the
HA gene antigenic drift demonstrated here and potential loss of population immunity against newly emerging clades. The G1 constellation and HA clade A of the
H3N2v viruses may enhance the capacity for human infection that other genotypes and HA clades of swine H3N2 do not possess. It is not possible to evaluate at this time whether the clustering of human H3N2v viruses on the PB2, PB1, PA, NP, M and NS phylogenies represents requirements on the genetic composition of these other segments for swine-to-human transmission, or whether this is simply an artefact of linkage with the clade A H3 segment. Further whole-genome sequencing of human H3N2v viruses will determine whether these genetic restrictions on pig-to-human transmission continue. Overall, these findings highlight the value of whole-genome sequencing for understanding the evolution of genetic diversity among H3N2 IAVs, with important implications for both public and animal health. Swine vaccine manufacturers should closely monitor these contemporary H3N2 viruses for establishment of the clades in the context of population immunity for vaccine strain purposes.

Acknowledgements

We are grateful to the pork producers and swine veterinarians for participating in the USDA–National Animal Health Laboratory Network (NAHLN) Influenza Virus Surveillance System for swine. We acknowledge Dr Sabrina Swenson (NVSL–USDA–APHIS) and Dr John Korslund (CEAH–USDA–APHIS) for kindly sharing swine influenza genetic and surveillance data from the NAHLN Influenza Virus Surveillance System for swine. Thank you to Alicia Janas-Martindale (NVSL–USDA–APHIS) for assistance in virus isolation whole-genome sequencing. Thank you to Dr Marcus Kehri (NADC–USDA–ARS) for critical reading of the manuscript. Funding was provided by USDA–ARS and USDA–APHIS–VS. 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 US Department of Agriculture. USDA is an equal opportunity provider and employer.

References


