Virology 518 (2018) 45–54

Contents lists available at ScienceDirect

Virology

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

Antigenic and genetic evolution of contemporary swine H1 influenza viruses in the United States

Daniela S. Rajao[a,1], Tavis K. Anderson[a], Pravina Kitikoon[a,2], Jered Stratton[a], Nicola S. Lewis[b], Amy L. Vincent[a,*]

[a] Virus and Prion Research Unit, National Animal Disease Center, USDA-ARS, 1920 Dayton Avenue, PO Box 70, Ames, IA 50010, USA
[b] Department of Zoology, University of Cambridge, Downing St, Cambridge CB2 3EJ, UK

ARTICLE INFO

Keywords:
Antigenic diversity
H1N1
H1N2
Influenza A
Swine
Virus evolution

ABSTRACT

Several lineages of influenza A viruses (IAV) currently circulate in North American pigs. Genetic diversity is further increased by transmission of IAV between swine and humans and subsequent evolution. Here, we characterized the genetic and antigenic evolution of contemporary swine H1N1 and H1N2 viruses representing clusters H1-α (1A.1), H1-β (1A.2), H1pdm (1A.3.3.2), H1-γ (1A.3.3.3), H1-δ1 (1B.2.2), and H1-δ2 (1B.2.1) currently circulating in pigs in the United States. The δ1-viruses diversified into two new genetic clades, H1-δ1a (1B.2.2.1) and H1-δ1b (1B.2.2.2), which were also antigenically distinct from the earlier H1-δ1-viruses. Further characterization revealed that a few key amino acid changes were associated with antigenic divergence in these groups. The continued genetic and antigenic evolution of contemporary H1 viruses might lead to loss of vaccine cross-protection that could lead to significant economic impact to the swine industry, and represents a challenge to public health initiatives that attempt to minimize swine-to-human IAV transmission.

1. Importance

The hemagglutinin (HA) protein of influenza A virus (IAV) is the primary target of protective immune responses and the major component in vaccine formulation. However, repeated introductions of non-swine IAV into swine populations and the processes of antigenic shift and drift result in virus evolution and potential mismatches between vaccines and circulating strains. Further, the increasing diversity of swine IAV represents a challenge for public health initiatives, and assessment of the risk of interspecies transmission of viruses relies on assessing antigenic diversity relative to human population immunity. In this study, we found that antigenic drift of IAV in the U.S. pig population resulted in seven distinct antigenic phenotypes of the H1 subtype currently circulating. We identified changes in the HA protein associated with the observed antigenic drift, suggesting that these amino acids may be important antigenic sites. These data demonstrate that recent evolution of H1 swine IAV resulted in novel genetic clades with distinct antigenic phenotypes that are unlikely to be protected by current vaccine formulations.

2. Introduction

Influenza A virus (IAV) is endemic in pigs globally, and three different subtypes currently co-circulate in North American swine: H1N1, H1N2, and H3N2 (Anderson et al., 2013; Lorusso et al., 2013). Swine have been indicted as a potential source of reassortant viruses that can be transmitted to humans (Brown, 2000; Vijaykrishna et al., 2011). Two major events illustrate this potential: the introduction of the swine-origin 2009 pandemic H1N1 virus (H1N1pdm09) with widespread global morbidity and mortality (Smith et al., 2009b), and the occurrence of more than 300 cases of swine-origin H3N2 IAV (H3N2v) (Jhung et al., 2013). However, the transmission of human IAV into swine populations is a frequent occurrence, and the repeated introduction and subsequent evolution of these human viral lineages in swine populations has a critical role in the genetic diversity of swine IAV (Nelson et al., 2012, 2014).

The first documented occurrence of bidirectional transmission between swine and humans dates to the 1918 H1N1 pandemic (Smith et al., 2009a), with the resulting swine strain referred to as classical H1N1 (cH1N1). The cH1N1 remained antigenically and genetically stable in the U.S. for nearly a century. In the late 1990s, a novel triple-

* Corresponding author.
E-mail address: amy.vincent@ars.usda.gov (A.L. Vincent).
[1] Present address: Department of Population Health, The University of Georgia, Athens, GA 30602, USA.
[2] Present address: Merck Animal Health, DeSoto, KS 66018, USA.

https://doi.org/10.1016/j.virol.2018.02.006
Received 14 November 2017; Received in revised form 3 February 2018; Accepted 5 February 2018

reassortant H3N2 virus was detected in U.S. swine herds (Zhou et al., 1999) and reassorted with endemic cH1N1, leading to a period of rapid diversification of the cH1N1 hemagglutinin (HA) into three distinct HA clades: H1-α (1A.1), H1-β (1A.2), and H1-γ (1A.3.3.3) (Lorusso et al., 2011; Vincent et al., 2008b). Viruses containing HA and/or neuraminidase (NA) derived from human-seasonal IAV were introduced to U.S. pigs in the early 2000s (Karasin et al., 2006; Vincent et al., 2009), giving rise to two HA phylogenetic clades, H1-δ1 (1B.2.2) and H1-δ2 (1B.2.1) (Anderson et al., 2013; Lorusso et al., 2013). The 2009 H1N1 pandemic viruses were reintroduced from humans into U.S. pigs, resulting in reassortment with endemic swine viruses (Ducatez et al., 2011) and the establishment of a sixth swine H1 genetic clade (H1N1pdm09: 1A.3.3.2). This dynamic evolution is not restricted to the HA; N2 genes were derived from two human seasonal-lineages introduced in 1998 and 2002 or N1 genes from the classical or H1N1pdm09 lineages (Anderson et al., 2013; Lorusso et al., 2013). Moreover, the current diversity of lineages within IAV circulating in swine populations is complicated by antigenic drift of the HA (de Jong et al., 1999), and this gradual evolution may alter antigenic properties of IAV that result in vaccine failure (Both et al., 1983; Luoh et al., 1992). However, antigenic diversity is not always linearly correlated with genetic diversity, and few amino acid mutations can lead to biologically significant antigenic differences that may facilitate immune escape (Both et al., 1983; Luoh et al., 1992).

Vaccination against IAV in swine is an important tool to limit clinical disease, but it may not prevent infection and/or transmission. Products currently available in the U.S. provide good protection against homologous strains, however their efficacy against antigenically distinct viruses may diminish (Vincent et al., 2010, 2008a). Vaccine strains should be updated when there is loss in cross-reactivity against contemporary strains to improve the match between vaccine antigens and circulating antigenic diversity. Although genetic analysis of the HA gene has utility in predicting vaccine efficacy (Neher et al., 2016), antigenic cross-reactivity between field-sourced viruses and vaccine strains remains the gold standard for vaccine strain candidate selection. Thus, understanding the molecular epidemiology of IAV circulating in swine populations in association with antigenic properties are crucial to improve vaccine composition to better control IAV in pigs.

Several studies have mapped the antigenic and genetic evolution of H3N2 viruses in swine (Abente et al., 2016; de Jong et al., 2007; Lewis et al., 2014), and recently a comprehensive study characterized the high level global antigenic diversity of H1 and H3 viruses circulating in human and swine (Lewis et al., 2016). However, an extensive description of the links between genetic diversity, amino acid identity, and antigenic phenotype remains unclear for H1 viruses in the U.S. Here, we performed a comprehensive characterization of the genetic and antigenic evolution of swine H1N1 and H1N2 viruses, representing all genetic clades currently circulating in pigs in the U.S. We then conducted an in-depth antigenic characterization on contemporary strains from the H1-δ (1B.2.1 and 1B.2.2) clades due to their predominance in surveillance detections and apparent genetic expansion. Further, we examined the underlying genetic basis for significant antigenic differences among these circulating H1-δ viruses.

3. Materials and methods

3.1. Genetic evolution and amino acid sequence analysis

All available swine IAV HA and NA sequences from H1N1 and H1N2 viruses collected in the U.S. were downloaded from the Influenza Research Database (IRD) (Squires et al., 2012; Zhang et al., 2016) on December 5, 2016. To restrict our analyses to relevant field viruses, we excluded sequences with “lab” or “laboratory” host. From these data, alignments for the HA and the N1 and N2 NA genes were generated using MAFFT v7.294 (Katoh et al., 2002; Katoh and Standley, 2013). In addition, sequences with 100% identity were deleted using mothur v1.36.0 (Schloss et al., 2009), and poor quality data was removed using two criteria: a sequence was removed if > 50% of the gene was missing or if it had more than 5 nucleotide base ambiguities. This process resulted in a set of 3211 non-identical H1 HA, 1606 N1 and 1406 N2 NA swine IAV sequences that represent the full extent of the published swine H1N1 and H1N2 genetic diversity in the U.S. (Table S1). For each alignment (i.e., the HA, the NA-N1 and the NA-N2) we inferred the best-known maximum likelihood tree using RAxML (v8.2.4; (Stamatakis, 2014)) employing the rapid bootstrap algorithm, a general time-reversible (GTR) model of nucleotide substitution, and Γ-distributed rate variation among sites. The statistical support for individual branches was estimated by bootstrap analysis with the number of replicates determined automatically using an extended majority-rule consensus tree criterion (Pattengale et al., 2010). These analyses used the computational resources of the USDA-ARS computational cluster Ceres on ARS SCINet.

To determine the temporal evolution and relative genetic diversity of U.S. swine H1 HA, we implemented a time-scaled Bayesian approach on a second dataset. We downloaded complete HA H1N1 and H1N2 swine IAV genes collected from 2000 to present from the IRD on December 5, 2016. Given deep evolutionary divergence between the two major H1 HA lineages, we separated these data into the classical lineage H1 HA including H1-α (1A.1), H1-β (1A.2), H1-γ (1A.3.3.3), and H1N1pdm09 (1A.3.3.2) (n = 2208), and human-seasonal H1-δ lineage H1-δ1 (1B.2.2) and H1-δ2 (1B.2.1) (n = 1740). From each of these datasets, we randomly sampled to create smaller datasets of 750 HA genes to overcome computational limitations (i.e., we randomly sampled 707 H1-δ lineage viruses and added 43 reference antigens; similarly, we randomly sampled 713 H1-classical lineage viruses, then added the 37 reference antigens). These data were then aligned using MAFFT v7.294 (Katoh et al., 2002; Katoh and Standley, 2013), and HA genes that were duplicated due to the addition of reference antigen data were removed. The resultant data were then screened using root-to-tip regression in TempEst v.1.5 (Rambaut et al., 2016) and sequences with incongruent genetic divergence and sampling date were removed (Hicks and Duffy, 2012) resulting in a dataset of 730 and 726 HA genes for the H1-δ and H1-classical lineages respectively. The remaining data were analyzed using an uncorrelated relaxed lognormal molecular clock (Drummond et al., 2006), the SRD06 codon position model (Shapiro et al., 2006) that partitions codon positions (1 + 2 positions and 3 position) with an HKY85 + Γ substitution model applied to each partition. To reconstruct population dynamics we used the coalescent-based Gaussian Markov random field (GMRF) method with time-aware smoothing (Minin et al., 2008). The precision function in BEAUTi was used to sample uniformly within a one-year or one-month window for those viruses for which an exact date of collection was not available. All analyses were implemented in BEAST v1.8.4 (Drummond et al., 2012) with the BEAGLE library (Ayres et al., 2012) with two independent analyses of 100 million generations with sampling every 10,000 generations. Convergence of runs was checked in Tracer v1.6.0, runs were combined with LogCombiner v1.8.4, and evolutionary history was summarized and visualized using an annotated maximum clade credibility tree using TreeAnnotator v1.8.4 and FigTree v1.4.2. These analyses used the computational resources of the USDA-ARS computational cluster Ceres on ARS SCINet.

Hemagglutinin sequences used for antigenic characterization were translated to amino acids, trimmed to the HA1 domain, and aligned using MAFFT v7.294 (Katoh et al., 2002; Katoh and Standley, 2013). The HA1 amino acid alignments were used to identify amino acid substitutions likely to be clade-defining or to have resulted in antigenic differences between or within a genetic clade. We followed a criterion in which a substitution was considered pertinent as clade-defining if all (or all but one) viruses in one clade had the same amino acid at a specific position in comparison to a different amino acid in the same position of all (or all but one) viruses in another clade (Lewis et al., 2014, 2011; Smith et al., 2004). Antigenic outliers within the H1-δ