ORIGINAL ARTICLE

WILEY Evolutionary Applications Open Access

Adaptive evolution during the establishment of European avian-like H1N1 influenza A virus in swine

Udayan Joseph[1] | Dhanasekaran Vijaykrishna[1,2] | Gavin J.D. Smith[1,3] (iD) | Yvonne C.F. Su[1]

[1]Programme in Emerging Infectious Diseases, Duke-NUS Medical School, Singapore

[2]Department of Microbiology, Biomedicine Discovery Institute, Monash University, Melbourne, Vic., Australia

[3]Duke Global Health Institute, Duke University, Durham, NC, USA

Correspondence
Gavin J.D. Smith and Yvonne C.F. Su, Programme in Emerging Infectious Diseases, Duke-NUS Medical School, Singapore.
Emails: gavin.smith@duke-nus.edu.sg, yvonne.su@duke-nus.edu.sg

Abstract

An H1N1 subtype influenza A virus with all eight gene segments derived from wild birds (including mallards), ducks and chickens, caused severe disease outbreaks in swine populations in Europe beginning in 1979 and successfully adapted to form the European avian-like swine (EA-swine) influenza lineage. Genes of the EA-swine lineage that are clearly segregated from its closest avian relatives continue to circulate in swine populations globally and represent a unique opportunity to study the adaptive process of an avian-to-mammalian cross-species transmission. Here, we used a relaxed molecular clock model to test whether the EA-swine virus originated through the introduction of a single avian ancestor as an entire genome, followed by an analysis of host-specific selection pressures among different gene segments. Our data indicated independent introduction of gene segments via transmission of avian viruses into swine followed by reassortment events that occurred at least 1–4 years prior to the EA-swine outbreak. All EA-swine gene segments exhibit greater selection pressure than avian viruses, reflecting both adaptive pressures and relaxed selective constraints that are associated with host switching. Notably, we identified key amino acid mutations in the viral surface proteins (H1 and N1) that play a role in adaptation to new hosts. Following the establishment of EA-swine lineage, we observed an increased frequency of intrasubtype reassortment of segments compared to the earlier strains that has been associated with adaptive amino acid replacements, disease severity and vaccine escape. Taken together, our study provides key insights into the adaptive changes in viral genomes following the transmission of avian influenza viruses to swine and the early establishment of the EA-swine lineage.

KEYWORDS
cross-species transmission, influenza A virus, natural selection, reassortment

1 | INTRODUCTION

Influenza pandemics represent a significant threat to global public health. Four major pandemics have been recorded since the 1900s, occurring in 1918, 1957, 1968 and 2009 when influenza A viruses with genes from animal sources adapted to the human population, a process known as antigenic shift (Fraser et al., 2009; Garten et al., 2009; Kilbourne, 2006; Smith, Vijaykrishna, et al., 2009; Taubenberger, Hultin, & Morens, 2007; Webster, Bean, Gorman, Chambers, & Kawaoka, 1992). Phylogenetic analysis of viruses archived since the

1930s has shown that the H2N2/1957 and H3N2/1968 pandemic viruses resulted from reassortment of animal and circulating human viruses (Kawaoka, Krauss, & Webster, 1989; Schäffr et al., 1993; Scholtissek, Rohde, Von Hoyningen, & Rott, 1978). The H1N1/2009 pandemic virus emerged from swine that contained gene segments ultimately derived from previously circulating human and avian viruses, highlighting a key role of segmental reassortment of genes from multiple hosts for host adaption and pandemic emergence (Fraser et al., 2009; Garten et al., 2009; Smith, Vijaykrishna, et al., 2009).

The adaptation of an avian virus to mammalian hosts has been of great interest due to the pandemic potential of circulating low and highly pathogenic avian influenza (Herfst et al., 2012; Imai et al., 2012). While the origins of the 1918 virus is considered by some to have evolved directly from avian hosts (Taubenberger, Reid, Krafft, Bijwaard, & Fanning, 1997; Taubenberger et al., 2005; Worobey, Han, & Rambaut, 2014), phylogenetic studies using Bayesian molecular clock methods have shown that the origins of this virus may be staggered and that the role of the intermediate host, such as swine, may not be ruled out (Smith, Vijaykrishna, et al., 2009).

Swine influenza virus (SwIV) infection typically displays similar clinical symptoms that resemble infections in humans including high fever, nasal discharge and cough, although swine can also suffer from severe bronchitis and bronchiolitis that result in rapid breathing (tachypnea) and shortness of breath (dyspnoea) (Janke, 2013). Due to the presence of both human-like (α-2,6-linked sialic acids) and avian-like (α-2,3-linked sialic acid) virus receptors in the pig respiratory tract, swine have the ability to facilitate co-infections of avian, swine and human influenza viruses generating novel reassortant viruses (Ito et al., 1998; Kida et al., 1994; Matrosovich et al., 2000; Rogers & Paulson, 1983; Rogers, Pritchett, Lane, & Paulson, 1983). This distinguishing feature has led to swine being referred to as "intermediate hosts" or "mixing vessels" that may serve as a potential reservoir in transmitting avian influenza viruses to humans (Ma, Kahn, & Richt, 2008; Scholtissek, 1990).

The Classical swine (CS) H1N1 viruses were the earliest recorded influenza virus in swine (Dowdle & Hattwick, 1977; Janke, 2013; Shope, 1931). Evolutionary analysis (Worobey et al., 2014) indicated that the CS HA-H1 lineage directly descended from the human pandemic 1918 virus to form a stable lineage that has circulated for at least 80 years in swine. Reassortment of this lineage was first observed in 1998, when several severe outbreaks occurred in pig farms across North America (Webby et al., 2000; Zhou et al., 1999). These outbreaks were caused by the emergence of two distinctive H3N2 viruses: a double-reassortant that contained CS H1N1 and human H3N2 virus genes, and a virus that contained the triple-reassortant internal gene cassette (TRIG) that incorporated additional avian gene segments. The latter of the two viruses, which contains gene segments from the CS viruses (NS, NP and M), seasonal H3N2 viruses (HA, NA, PB1) and avian (PB2 and PA) viruses, eventually became more widespread in North America than the double-reassortant virus and continues to circulate to date (Nelson et al., 2015; Webby et al., 2000). Other SwIV viruses that commonly circulate in North American swine populations include human-like H1N1 and human-like H3N2 and H1N2 viruses (Anderson et al., 2013; Vincent et al., 2014).

In contrast, the European avian-like swine (EA-swine) H1N1 virus was first detected in January 1979 in Belgium (Pensaert, Ottis, Vandeputte, Kaplan, & Bachmann, 1981). The virus spread to neighbouring countries such as Germany and France by late 1980 and eventually replaced CS viruses in Europe (Brown et al., 1997; Zell, Scholtissek, & Ludwig, 2012). Distinct from the CS lineage, the eight gene segments of EA-swine lineage were derived from the avian influenza virus gene pool across Eurasia (Dunham et al., 2009; Lycett et al., 2012; Worobey et al., 2014; Zell et al., 2013). The EA-swine virus subsequently spread to Asia, including Hong Kong and China, as early as 1993 (Vijaykrishna et al., 2011).

The three major SIV lineages (CS, TRIG and EA-swine) differ in their ancestral origins, but more notably, they also vary in their rates of adaptive evolution (Bhatt et al., 2013): the EA-swine virus exhibit faster rates than CS, reflecting that higher substitution rates are observed during host adaptation following cross-species transmission. The EA-swine H1N1 continued to circulate for many decades in Europe and donated genetic segments to form the following major lineages: EA-swine H1N1, human-like EA-swine H3N2, human-like EA-swine H1N2 and pandemic H1N1 in humans. Furthermore, since 2009 at least 23 distinct genotypes (referred to as "A–W") were identified in European swine (Watson et al., 2015).

Of significant importance to global human health is the emergence of 2009 H1N1 pandemic virus from swine to humans, as two genes (NA and M) were derived from EA-swine viruses while the remaining segments originated from TRIG viruses (Garten et al., 2009; Smith, Vijaykrishna, et al., 2009). Furthermore, the HA of EA-swine lineages is of particular concern as they preferentially bind to the human receptor, posing a potential threat of efficient transmission of EA-swine from swine to humans (Yang et al., 2016). While recent attention has focused on understanding the evolutionary dynamics and migration patterns of SwIV, here we elucidate the evolutionary processes and adaptive changes of avian influenza viruses upon transmission to pigs. To do this, we applied a relaxed molecular clock to elucidate the molecular changes in adaptive evolution of the EA-swine lineage at a genomic level. Identifying the key changes in cross-species transmission will provide insight into the potential genetic factors that may be responsible for host adaptation and gain understanding into the patterns of adaptive evolution in establishing stable lineage descendants.

2 | MATERIALS AND METHODS

2.1 | Data set curation and maximum-likelihood reconstruction

Over 2000 nucleotide sequences of all H1NX (for HA), HXN1 (for NA) and HXNX (for all internal gene segments) isolated from avian and swine hosts before the year 2000 and their associated metadata were downloaded from the NCBI Influenza Virus Resource (http://www.ncbi.nlm.nih.gov/genomes/FLU/FLU.html). Sequence lengths of <500 bp were excluded from the analysis. For each gene segment, the large data sets were aligned using MAFFT (Katoh, Misawa, Kuma,
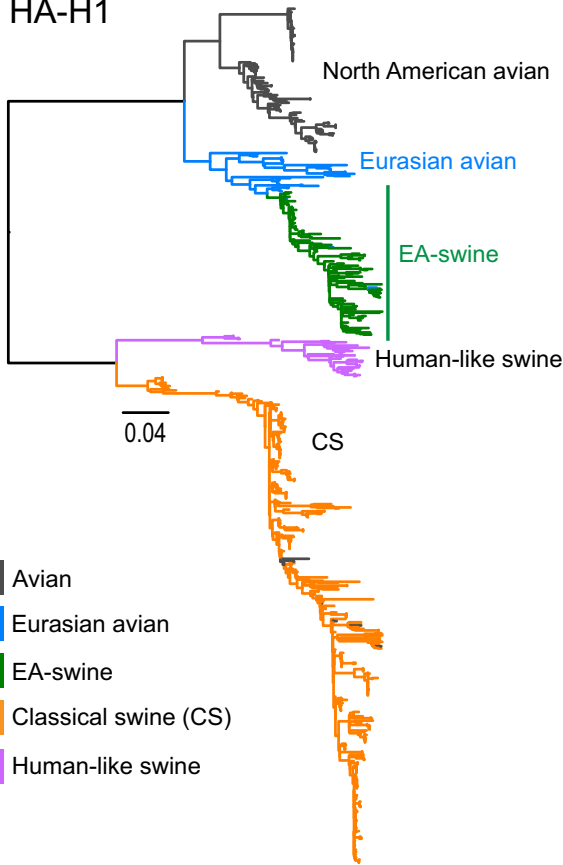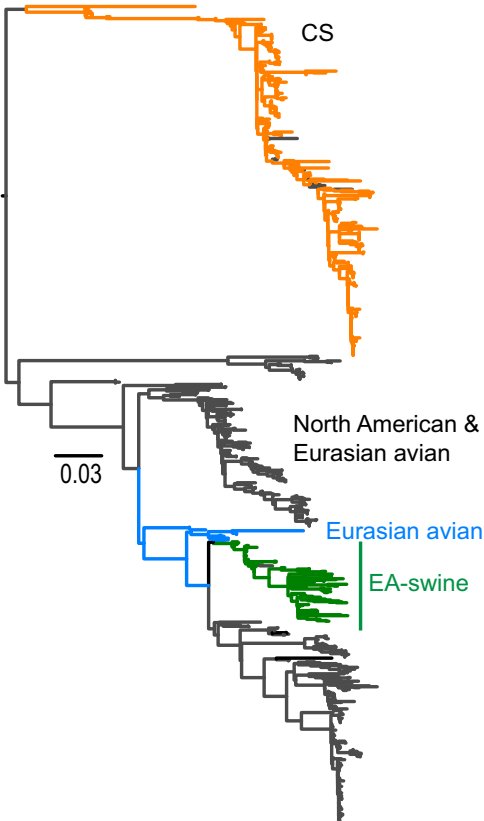
FIGURE 1 Evolution and divergence times of avian and swine HA-H1 and NA-N1 gene segments. Green branches denote the EA-swine lineage; blue branches represent the closely related avian lineage from which the new swine lineage emerged; orange branches denote classical swine (CS); grey branches denote avian strains; and purple branches represent human-like swine. The dated phylogenies were generated using 100 million MCMC generations, under a GTR+Γ substitution model, a coalescent-based GMRF tree prior and an uncorrelated relaxed clock model. Node A denotes the mean TMRCA of the combined avian and EA-swine, and node B represents the mean TMRCA of EA-swine lineage only. Horizontal blue bars represent 95% highest posterior density (HPD) intervals, and shaded grey areas represent the period of cross-species transmission. GMRF, Gaussian Markov random field; TMRCA, time to most recent common ancestor; EA-swine, European avian-like swine

& Miyata, 2002) followed by manual alignment. The sequences were trimmed to only include coding regions for subsequent analysis. Preliminary maximum-likelihood (ML) analysis was conducted for large individual gene data sets (PB2, PB1, PA, HA, NP, NA, MP and NS) using RAxML v9.0 (Stamatakis, 2014). To focus on the adaptation and evolution of EA-swine lineage, we used the resulting large ML trees (Figure 1; Figs S1–S8) to select the EA-swine lineage and its closest avian lineage (denoted by blue branches in Figure 1 and Fig. S9). The latter included all available avian sequences that are exclusively form a sister group to the entire EA-swine lineage. Therefore, distantly related avian lineages, with heterogeneous evolutionary rates that may affect phylogenetic analyses, were excluded in subsequent analyses. All duplicate viruses were manually removed from the data sets, and isolates with 100% identical residues were removed using the webserver of the program CD-HIT (Huang, Niu, Gao, Fu, & Li, 2010; Li & Godzik, 2006). The program TempEst v1.5 (http://tree.bio.ed.ac.uk/software/tempest/) was also used to plot root-to-tip divergence times to remove any outliers from the sequence data sets as a possible result of mislabelled isolation dates. To account for uncertainty of isolation date, isolates were mid-year-rooted if the exact date of sampling was unknown. For each segment, the reduced data sets (comprising the EA-swine and its closely related avian lineages) were used to reconstruct ML trees using a generalized time reversible nucleotide substitution model plus gamma distributed rates among sites (GTR+Γ) in PhyML v3.0 (Guindon et al., 2010).

2.2 | Temporal phylogenies and estimation of TMRCAs

For each gene segment, the estimates of evolutionary rates and temporal phylogenies were performed in BEAST v1.8.2 (Drummond, Suchard, Xie, & Rambaut, 2012). An uncorrelated lognormal relaxed clock model within a Bayesian Markov chain Monte Carlo (MCMC) framework was used, with a Gaussian Markov random field coalescent tree prior. At least two independent MCMC runs of 100 million steps were performed and combined to ensure adequate sampling of all parameters, with a 10%–20% "burn-in" removed in each run. The relevant statistics and values were parsed from these runs directly from the combined log files using the program Tracer v1.6 (http://tree.bio.ed.ac.uk/Tracer). Bayes factors (BF) for statistical support of differences between estimated time to most recent common ancestor (TMRCA) and nucleotide substitution rate values were calculated as described previously (Bahl, Vijaykrishna, Holmes, Smith, & Guan, 2009), where BF ≥ 150 indicates very strong support, 150 > BF ≥ 20

indicates strong support values, and 20 > BF ≥ 3 indicates supported values.

2.3 | Selection analyses and ancestral node reconstruction

The estimates of the degree of natural selection were performed as previously described (Joseph et al., 2015). Briefly, the ratio of non-synonymous to synonymous substitutions per codon (d_N/d_S ratio) was estimated for each segment data set using the single-likelihood ancestor counting (SLAC) method (Kosakovsky Pond & Frost, 2005) run through the Datamonkey webserver (Delport, Poon, Frost, & Kosakovsky Pond, 2010) with user-supplied ML trees (as above). Specific amino acid sites of selection were determined using the Tdg09 program (Tamuri, Dos Reis, Hay, & Goldstein, 2009), with statistical cut-offs of set at the false discovery rate (FDR) value of 0.20. Ancestral codon substitutions of nodes were determined using the baseml program of the PAML suite v4.7 (Yang, 2007) and transcribed onto trees generated with RAxML v8.1.6 (Stamatakis, 2014) using the treesub program (Tamuri, 2013), as described previously (Su et al., 2015).

2.4 | Measurement of uracil content of avian and EA-swine lineage

Differences in uracil content between EA-swine and closely related avian lineages were performed similar to as previously described (Joseph et al., 2015) as it has been postulated that the uracil content of influenza A virus genomes tends to increase steadily in mammalian hosts over time (Rabadan, Levine, & Robins, 2006; Worobey et al., 2014). Briefly, uracil content values of the avian and EA-swine gene segments were measured using PAUP* 4.0b10 (Swofford, 2003). Uracil content values were visualized and plotted using the R statistical software package (R Core Team 2013).

2.5 | Detection of reassortment within the EA-swine lineage

To identify the intralineage reassortment events in the EA-swine virus lineage, we used the software Dendroscope v.3.0 (Huson & Scornavacca, 2012) to generate tanglegrams using the ML phylogenies. The ML phylogenies for individual gene segments of 38 EA-swine isolates were reconstructed using IQ-Tree v1.3.0 (Nguyen, Schmidt, von Haeseler, & Minh, 2015). The runs were performed using 10,000 ultrafast bootstrap replicates and automatic selection of
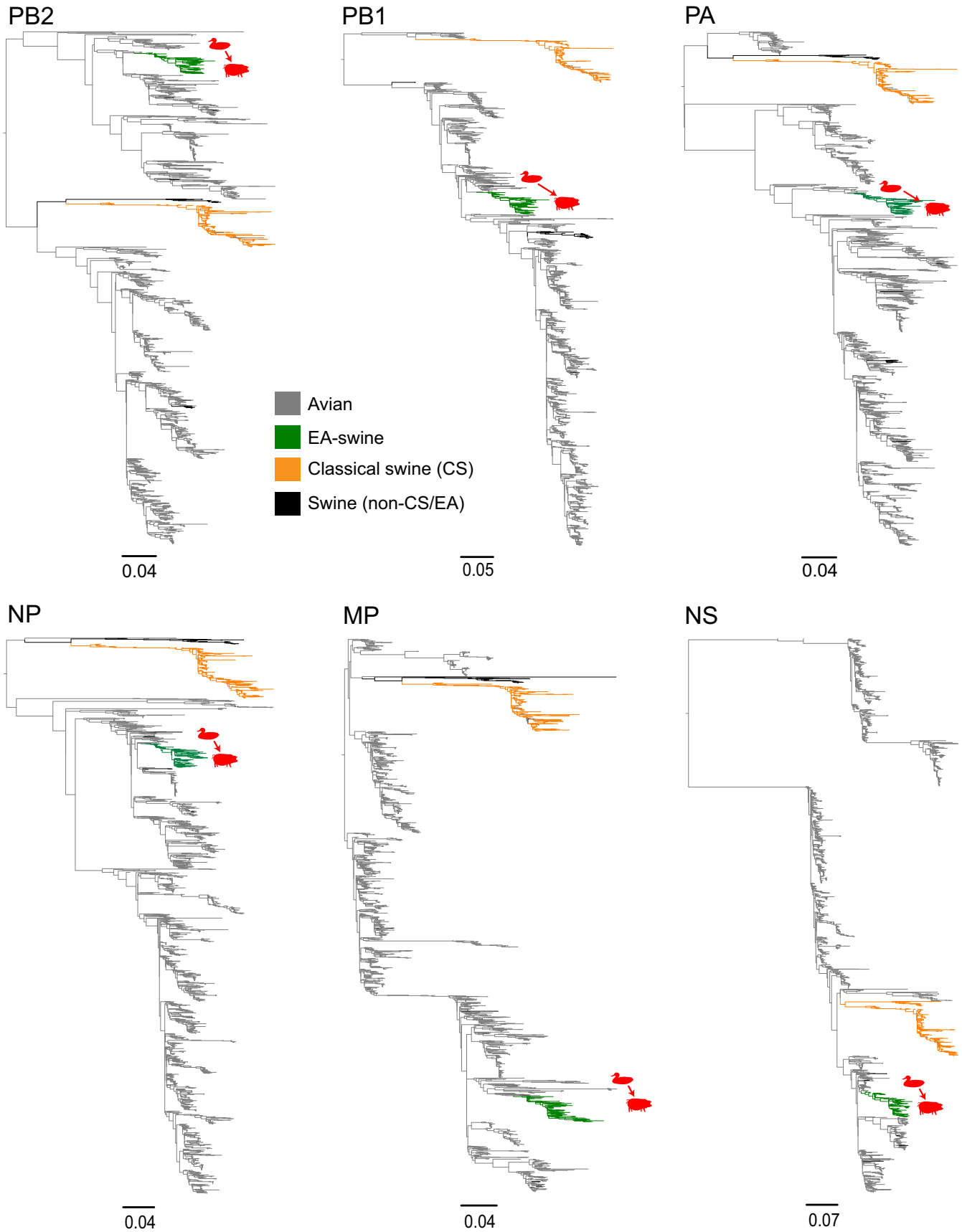
FIGURE 2 Evolution of avian and swine internal gene segments. Green branches denote the EA-swine lineage; orange branches denote classical swine (CS); and grey branches denote avian strains. Red arrows indicate the avian-to-swine cross-species transmission events for each gene. EA-swine, European avian-like swine

best-fit substitution model. The resulting ML phylogenies were rooted using the earliest EA-swine strain (A/swine/Arnsberg/1979). To infer and visualize the reassortment events, auxiliary lines were then drawn between same set of virus isolates in the phylogenies of two gene segments (i.e., between HA and non-HA phylogenies).

3 | RESULTS

3.1 | Large-scale phylogenetic analysis of avian and swine H1N1 lineages in 1924–2000

Maximum-likelihood phylogenies of all eight individual gene segments (Figures 1 and 2; Figs S1–S8) show that the EA-swine lineages (denoted by green branches) were monophyletic and clearly distinct to the CS swine lineage (denoted by orange branches). In all eight gene segments, the EA-swine lineages were derived from various Eurasian avian lineages. The most closely related avian gene segments were from different subtypes and were predominantly found in aquatic birds across Europe. For instance, the EA-swine PB2 was most closely related to genes from duck H2N3 (Germany) and turkey H5N2 (United Kingdom), while the NS was most closely related to genes from duck H1N1 (France) and H2N3, H7N7 (Germany). A few avian H1 strains, isolated from domestic turkey in Europe, nested within the EA-swine lineage in all eight genes although without apparent evidence of sustained transmission, suggesting incidental infection arising due to close proximity of high-density turkey and swine populations in Europe during that period (Bonfante et al., 2016). Taken together, these results indicate that the EA-swine lineage virus is unlikely to have resulted from a single avian-to-swine transmission.

3.2 | Divergence times and natural selection of the EA-swine lineage and its closest avian relatives

The availability of genome sequences from the EA-swine lineage since 1979 and data from parallel surveillance in aquatic bird surveillance in Europe that contained the closest avian relatives (Figures 1 and 2) enabled us to estimate the time of divergence of the EA-swine lineage. Dated phylogenies of all eight genes (Figure 1: HA and NA; Fig. S9: six internal genes) showed that the TMRCAs of the EA-swine and the closest avian lineages (combined avian and EA-swine lineage in Table S1) ranged from 1962 for the MP gene to 1974 for the NS gene, indicating an unsampled diversity of 4 to 17 years of these gene segments during which cross-species transmission may have occurred. However, it is not conclusive if these avian gene segments were introduced into swine as early as 17 years before first EA-swine detection from 1979 onwards as influenza surveillance in wild bird populations was not widespread during this period.

To determine the TMRCAs of the EA-swine viruses, we performed independent analyses of the EA-swine lineage. The mean nucleotide substitution rates of EA-swine lineage varied in different gene segments (Table 1), ranging from $2.31–4.20 \times 10^{-3}$ substitutions per site per year (subs site^{-1} year^{-1}). The HA-H1 and NA-N1 segments exhibited higher substitution rates (4.20×10^{-3} and 3.50×10^{-3} subs site^{-1} year^{-1}, respectively) than the internal segments. For each gene segment, the mean TMRCAs of the EA-swine lineages were estimated as between September 1976 and August 1978 (Figure 3b, Table 1, Fig. S10). The PB1 and NS genes estimated TMRCA were older than other genes, at 1976.86 (95% HPD: 1975.45–1978.02) and 1976.68 (95% HPD: 1974.96–1978.32), respectively. This suggests that the avian PB1 and NS genes may have been introduced into the

TABLE 1 Times to most common recent ancestor (TMRCA) and nucleotide substitution rates of all eight segments of the EA-swine lineage viruses (Fig. S10)

Gene segment	TMRCA of EA-swine (year)			Nucleotide substitution rate of EA-swine (subs site^{-1} year^{-1})		
	Mean	Lower 95% HPD	Upper 95% HPD	Mean	Lower 95% HPD	Upper 95% HPD
PB2	1977.88 (19 November 1977)	1977.09	1978.66	3.06E-03	2.73E-03	3.40E-03
PB1	1976.86 (12 November 1976)	1975.45	1978.02	2.65E-03	2.24E-03	3.06E-03
PA	1977.08 (30 January 1977)	1976.03	1978.11	2.89E-03	2.55E-03	3.23E-03
HA	1978.38 (20 May 1978)	1977.78	1978.97	4.20E-03	3.67E-03	4.72E-03
NP	1977.45 (14 June 1977)	1976.34	1978.50	2.47E-03	2.11E-03	2.85E-03
NA	1978.63 (19 August 1978)	1977.83	1979.35	3.50E-03	3.01E-03	4.02E-03
MP	1977.30 (21 April 1977)	1975.91	1978.52	2.31E-03	1.83E-03	2.82E-03
NS	1976.68 (06 September 1976)	1974.96	1978.32	2.41E-03	2.91E-03	3.44E-03

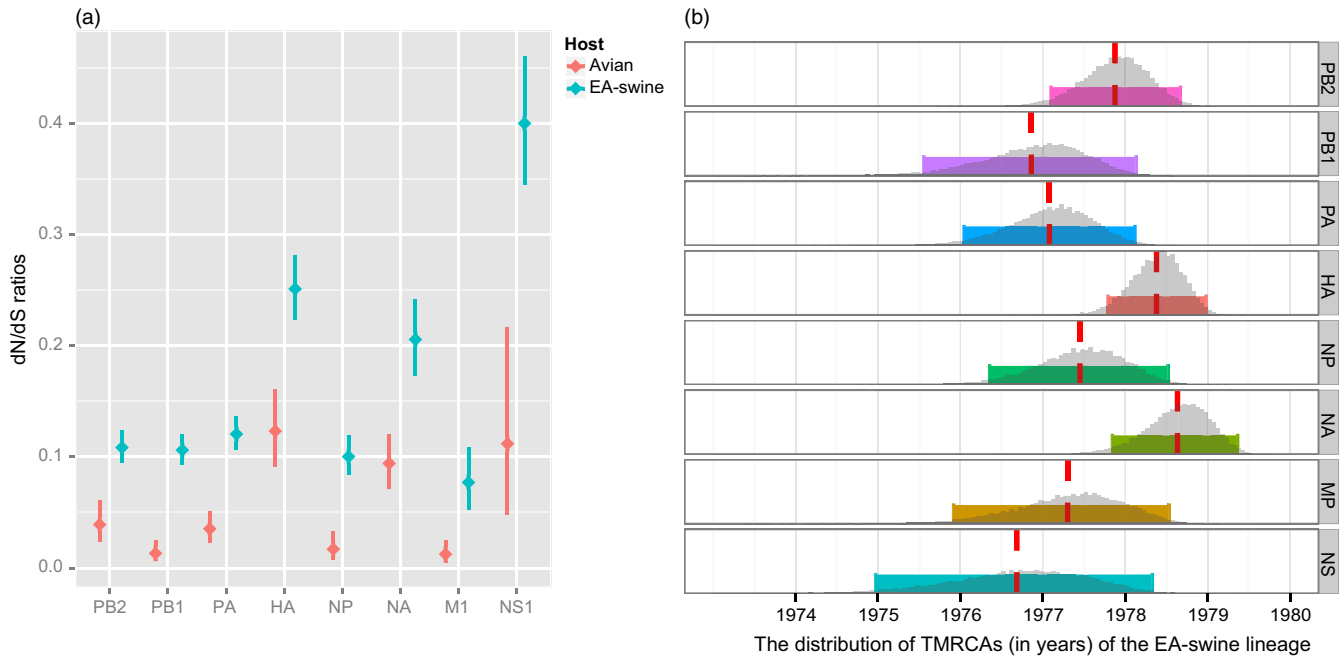EA-swine, European avian-like swine; HPD, highest posterior density.

FIGURE 3 Selection pressure and estimated time of the most recent common ancestors (TMRCAs) of the EA-swine lineage. (a) Average d_N/d_S ratios (with 95% confidence intervals) of proteins compared between avian (pink) and swine (blue) lineages. Only ratios of M1 and NS1 proteins depicted as representation of the in-frame translation of the MP and NS segments, respectively. (b) Dotted red lines represent the mean TMRCAs, solid coloured bars denote the 95% HPD intervals, and the grey histograms represent the TMRCA distribution across 18,000 sampled Bayesian MCMC trees. EA-swine, European avian-like swine; MCMC, Markov chain Monte Carlo

new swine host populations as early as December 1974. In contrast, the HA and NA genes were introduced into the swine populations later, at 1978.38 and 1978.63, respectively. More notably, our BF test demonstrated the differences in TMRCA between segments were statistically significant (Table S2): all internal segments were older (i.e., circulating longer in swine) than the HA and NA segments. Among the internal gene segments, the NS was significantly older than PB2, PB1 and NP, whereas the PB2 was younger than the PB1 and PA, and the NP as significantly younger than the PB1. It is therefore possible that these gene segments were independently introduced from birds to pigs through a series of reassortment events and circulated undetected during this period.

We further estimated and compared the average d_N/d_S ratios between the avian and EA-swine lineages across all eight gene segments (Figure 3a, Table 2). The SLAC analysis revealed that the EA-swine lineage exhibited a markedly greater level of selection pressures (i.e., higher d_N/d_S) in most genes (with exception to alternatively spliced M2 and NS2 genes) than the avian lineages. More specifically, the HA, NA and NS1 genes displayed more adaptive pressures than the other segments. Greater d_N/d_S in the NS1 gene could be indicative of antagonizing host type I interferon innate immune response (Wei, Chen, Lin, & Pan, 2014). However, it should also be noted that the mean d_N/d_S value (0.40; 95% HPD 0.35–0.46) of the NS1 in our analysis is comparable with other a previous study of EA viruses from swine in Hong Kong (0.30; 95% HPD 0.23–0.39) that only included viruses sampled from 2001 to 2010 (Vijaykrishna et al., 2011). These results are therefore consistent with previous observations.

In addition, we reconstructed the ancestral nonsynonymous mutations occurring along the branches that were involved in the host switch from avian to swine. Our results identified at least 30 amino acid mutations across all eight gene segments, with a larger proportion of mutations occurring in the surface HA and NA genes (10 and eight nonsynonymous substitutions, respectively) (Table 2). Greater HA and NA mutations may also be correlated with higher mean substitutions rates that are observed in both genes (see above).

We also applied the Tdg09 method and designated the host-specific lineages to estimate the site-specific selection of amino acids that are likely responsible for the host switching. The results indicated that 19 of the 30 nonsynonymous mutations that occurred across all eight segments with a statistical cut-off of FDR ≤ 0.20 (Table 2), suggesting that the remaining residue changes are likely founder mutations or a by-product of relaxed selection occurring in the establishment of a new host species (Su et al., 2015; Wertheim, Murrell, Smith, Kosakovsky Pond, & Scheffler, 2015). The surface HA and NA proteins showed a greater number of significant amino acid mutations compared to the internal gene segments. Of these statistically significant residues, mutations in the HA genes may have major implications in the cross-species transmission of the avian H1 segment into swine host (Figure 4): amino acid positions at 169, 204 and 209 (155, 190 and 225 in H3 numbering) situated on the HA globular head were deemed to be responsible for mediating host receptor switching between the use of α-2,3 and α-2,6 sialic acid residues (Glaser et al., 2005; Lin et al., 2012; Stevens et al., 2006), and may be involved in facilitating the adaptation of avian influenza virus to the new swine

TABLE 2 Node reconstruction and selection pressures of EA-swine lineage viruses

Gene segment	Ancestral node substitutions	# Significant residues (Site number)[a]	Data set	SLAC Mean	95% CI Lower	95% CI Upper
PB2	R251K, F446L	7 (21, 251*, 338, 400, 624, 649, 701)	Avian	0.0391	0.0233	0.0608
			Swine	0.1084	0.0945	0.1236
PB1	R584H, A587T	0	Avian	0.0132	0.0061	0.0247
			Swine	0.1061	0.0928	0.1206
PA	I62V, E258D, V387I, E399K	7 (262, 387*, 399*, 400, 684, 688, 712)	Avian	0.0353	0.0232	0.0511
			Swine	0.1205	0.1059	0.1363
HA	K130R, S138N, T169I, A202T, E204D, G239E, T393S, N416D, R419K, S457F	33 (3, 4, 36, 60, 113, 130*, 154, 155, 159, 169*, 173, 204*, 224, 239*, 277, 279, 284, 293, 306, 312, 328, 338, 387, 393*, 405, 409, 416*, 419*, 444, 457*, 465, 490, 502)	Avian	0.1232	0.0917	0.1611
			Swine	0.2511	0.2232	0.2814
NP	V190I	8 (48, 84, 98, 99, 190*, 284, 351, 384)	Avian	0.017	0.0073	0.0329
			Swine	0.1002	0.0837	0.1188
NA	V16I, V17I, T40I, I53V, I106V, A166V, V321I, G354D	31 (16*, 17*, 34, 40*, 46, 73, 75, 81, 82, 84, 94, 105, 106*, 111, 149, 166*, 234, 260, 262, 265, 285, 313, 314, 344, 354*, 374, 385, 386, 389, 395, 454)	Avian	0.0941	0.0717	0.1208
			Swine	0.2056	0.1732	0.2418
M1	None	4 (15, 116, 214, 248)	Avian	0.0124	0.0049	0.0251
			Swine	0.077	0.0522	0.1085
M2[b]	None	1 (18)	Avian	0.6904	0.418	1.0621
			Swine	0.7086	0.5227	0.9347
NS1	Q25R, E66K, T197A, L214F	9 (4, 129, 152, 158, 200, 208, 213, 214*, 227)	Avian	0.1118	0.048	0.2163
			Swine	0.4	0.3451	0.4605
NS2[b]	None	8 (4, 26, 49, 50, 52, 56, 60, 70)	Avian	0.1418	0.044	0.3298
			Swine	0.2774	0.2073	0.3617

EA-swine, European avian-like swine.

[a]False discovery rate (FDR). Sites with asterisks indicate residues undergoing mutation in the branch of interest (i.e., avian to EA-swine lineage).

[b]Selection calculations for alternatively spliced proteins must be interpreted with caution, as selection pressure acts on the main segment/protein reading frame.
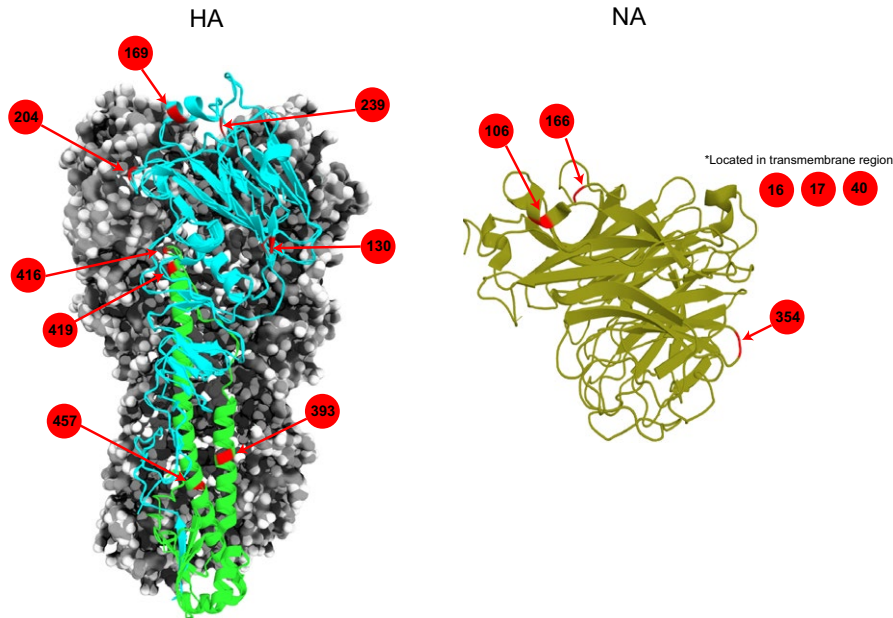
FIGURE 4 Mapping of the significant amino acid sites on the 3D structures of HA and NA molecules. Structural images were generated in MacPyMOL (The PyMOL Molecular Graphics System, version 1.8.0.3, Schrödinger, LLC): HA trimer [PDB accession code 4F3Z], and NA monomer crystal structures [PDB accession codes 4B7N]. Red arrows with numbers denote the amino acid positions that are predicted to be significant in adaptation to swine hosts

host. Other significant mutations were found in the conserved stalk region of the HA2 protein that occurred at positions 416, 419 and 457 (HA2–72, 75 and 113 in H3 numbering). The mutations at these sites in the swine have proven to play an important role in altering conformational HA stability and increasing optimal membrane fusion pH of the H1N1 HA protein, and thus are important adaptive mutations in avian-to-swine cross-species transmission (Baumann, Mounogou Kouassi, Foni, Klenk, & Matrosovich, 2015). To a lesser extent, other amino acid mutations at 130 and 393 (HA–120 and HA2–49 in H3 numbering) have been reported to have implications in mediating antibody binding (Ekiert et al., 2009; Zhu et al., 2013).

For the host jump from avian NA-N1 to swine, we identified amino acid positions 16 and 17 that are located in the transmembrane region of the NA protein (Figure 4). Their functions were reported to anchor the protein to membrane as well as be part of the signal peptide that permits transport of the protein across the endoplasmic membrane (Air, 2012). Therefore, adaptive mutations in this region may have altered the stability of the NA protein attachment to the membrane for better stability in the new swine hosts. In contrast, fewer amino acid mutations are present in the internal gene segments of EA-swine virus. Of note, mutation at the PB2-701 has been reported to be correlated with replication efficiency and these play a role in host adaptation of avian influenza viruses to mammalian hosts (Czudai-Matwich, Otte, Matrosovich, Gabriel, & Klenk, 2014; Dunham et al., 2009; Hayashi, Wills, Bussey, & Takimoto, 2015).

3.3 | Patterns of uracil content variation in avian and EA-swine lineages

To understand the potential role of changes in uracil content in the adaptive evolution of the EA-swine lineage, we evaluated and compared uracil content for all EA-swine segments in comparison with avian viruses. Within the avian lineage, no observable changes in

uracil contents for most genes, except for a slight decrease in NS gene, were observed, while the reverse was found in the PA gene (Figure 5). In contrast, we observed differential uracil patterns in the swine segments from 1980 to 2000: the HA, PA, NS and NP segments underwent a gradual increase in uracil content, suggesting greater uracil content in mammals compared to avian species may be required for the adaptation of new hosts. However, the swine PB2, PB1 and MP displayed a slight decrease in uracil content, whereas the swine NA did not undergo any apparent change uracil content.

3.4 | Reassortment within the EA-swine lineage

The extent of reassortment among the EA-swine lineage was assessed by comparing topological congruence between swine gene segments (Figs S11–S19). The HA tanglegram was rather incongruent with that of the internal genes, particularly in the PB1, NP, MP and NS genes. In addition, the more recent EA-swine strains have undergone more intense intrasubtype reassortment compared with the earlier strains, and coincided with the virus diversification that began in the late 1980s.

4 | DISCUSSION

The EA-swine virus successfully replaced the CS virus in Europe, as the infection in swine has been shown to enhance greater viral replication and induce prolonged virus shedding with higher viral loads in comparison with the CS swine virus, and the pigs were found to have low level of cross-reactive antibodies to EA-swine virus (Vijaykrishna et al., 2011). Furthermore, faster adaptive evolution was clearly observed in the EA-swine lineage compared to the CS lineage, which correlated with the establishment in the new swine host (Bhatt et al., 2013).

FIGURE 5 Uracil content of each gene segments in avian and EA-swine viruses. Different colour dots represent different gene segments. The avian isolates are denoted by the solid square symbol, whereas the swine isolates are indicated by the solid diamond symbol. Coloured lines are the best-fit regression lines, and the shaded grey areas represent the 95% confidence intervals. EA-swine, European avian-like swine
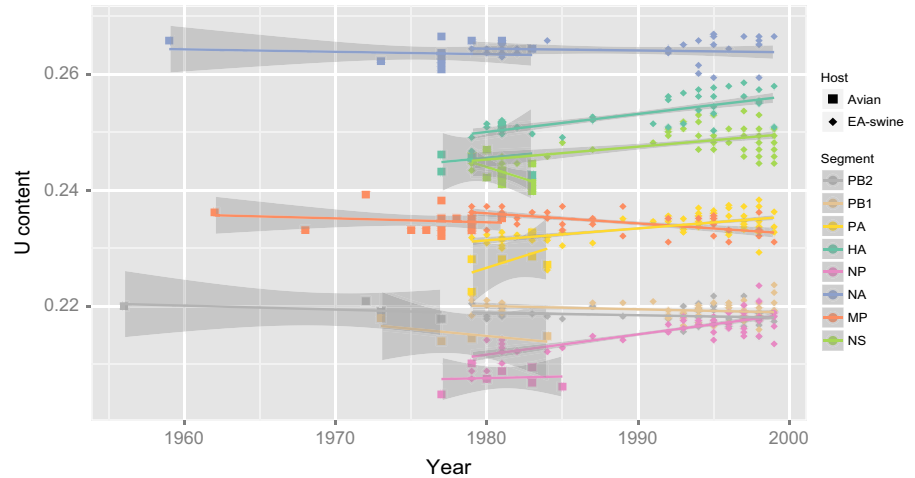
In the present study, we reconstructed large phylogenies of avian and swine viruses from Eurasia and North America. All individual gene phylogenies indicated the EA-swine virus formed a distinct lineage that is segregated from its closest avian relatives as well as CS. We further evaluated the selection pressure of this monophyletic EA-swine lineage. Our data demonstrated marked variation in the selection pressure between the avian and EA-swine viruses: the EA-swine virus experienced greater d_N/d_S ratios (0.10–0.28) than the avian virus (generally < 0.5) in most segments, especially in the swine HA, NA and NS segments. Notably, the M2 of both avian (d_N/d_S of 0.69) and EA-swine (d_N/d_S of 0.71) viruses had exceptionally greater relaxed selection pressure in comparison with other segments. The differences in selection pressures in the M1 and M2 could be associated with host immune response and tropism (Furuse, Suzuki, Kamigaki, & Oshitani, 2009).

Independent estimation of ancestral nonsynonymous mutations of EA-swine lineage revealed that adaptive mutations were likely involved with the establishment of the EA-swine lineage. Strikingly, estimates of site-specific selection indicated 30 nonsynonymous mutations across most segments, probably because of relaxed selection or founder effects. The EA-swine HA and NA segments have undergone greater amino acid mutations than other segments, which also correspond with faster rates of nucleotide substitutions in the surface proteins. Out of 30 sites, 19 sites are significant mutations that may play a crucial role in host switching of influenza A viruses from avian to swine hosts. However, the adaptive roles of most amino acid changes are still largely unknown, and further experimental characterization of these identified sites of substitution is needed to understand the adaptation of avian influenza A viruses into mammalian hosts. It is noteworthy that the swine MP gene did not seem to undergo ancestral nonsynonymous mutations. This suggests that the gene may not necessarily require significant adaptive change to achieve cross-species transfer of this gene.

The mean TMRCAs of all gene segments of the monophyletic avian and EA-swine lineage ranged from 1962 to 1974, which is consistent with previous age estimations (Dunham et al., 2009; Krumbholz et al., 2014). This also represents 4–17 years of unsampled diversity prior

to the outbreak during which cross-species transmission may have occurred. The mean TMRCAs of EA-swine lineage gene segments, when calculated alone, ranged from 1976 to 1978 with the HA and NA segment estimates significantly younger than the internal segments, at 1978.38 (May 1978) and 1978.63 (August 1978), respectively. Among the internal segments, the BF test also demonstrated the NS was markedly older than PB2, PB1 and NP segments, whereas the PB2 was younger than PB1, PA and NP segments. The previous hypothesis that the EA-swine virus was derived from an avian source in the absence of reassortment (Pensaert et al., 1981; Schultz, Fitch, Ludwig, Mandler, & Scholtissek, 1991; Vincent et al., 2014) is not supported by our age estimates, which differ significantly between gene segments. If the EA-swine virus had resulted from a single transmission of an entire virus from birds to pigs, then there would not be a statistical difference in the ages of the EA genes in pigs (Guan et al., 2010; Smith, Bahl, et al., 2009). Moreover, a different group of virus subtypes from wild and domestic birds were ancestral to each gene segment of the EA-swine lineage viruses, reflecting the acquisition of gene segments in swine from multiple sources. Taken together, our results highlight that the establishment of the EA-swine H1N1 virus was achieved through the independent introduction of avian gene segments via transmission of viruses into swine followed by reassortment events that occurred at least 1–4 years prior to the EA-swine outbreak.

Following the establishment of EA-swine lineage, we observed intrasubtype reassortment of segments occurred more frequently in subsequent EA-swine strains than the earlier strains. Frequent intrasubtype reassortment as in the case of H3N2 has been shown to be associated with adaptive amino acid replacements, disease severity and vaccine escape (Berry et al., 2016; Neverov, Lezhnina, Kondrashov, & Bazykin, 2014). We speculate this could be due to the close geographical proximity of farmyards within the Eurasian region and increased reassortment could facilitate sustained swine-to-swine transmission in Eurasian swine. In addition, EA-swine virus has been reported to experience greater rate of intersubtype reassortment with H1N2 virus within swine herds, with an average of one reassortment every 2–3 years (Lycett et al., 2012).

In summary, some notable adaptation phenomena could be indicative of cross-species transmission of influenza A viruses: internal viral proteins began to alter and appear to be better adapted to new hosts, followed by the introduction of the HA and NA surface proteins with greater levels of selective pressure. New reassortant viruses are then capable of efficient replication and sustained transmission in the new host; this is similarly observed in the H2N2 pandemic (Joseph et al., 2015) and H1N1 pandemic (Su et al., 2015). There is little doubt that influenza reassortment and the establishment of novel gene constellations in new host species have been repeatedly shown to cause panzootics and pandemics that could threaten global public health and cause significant economic losses. More crucially, the HA of EA-swine virus is of pandemic risk as it is antigenically different to currently circulating human H1 viruses in humans, and therefore, there would be little population immunity in humans against viruses carrying this HA. While the enhanced efforts to limit the mixing of different livestock have been implemented by increasing biosecurity measures on farms, continued global surveillance of swine and human populations is paramount in identifying and preventing zoonotic infections of humans.

ACKNOWLEDGEMENTS

DATA ARCHIVING STATEMENT

Genetic sequences used in the analyses are available from GenBank and accession numbers are provided in Figs S1–S8.

ORCID

Gavin J.D. Smith ![ORCID] http://orcid.org/0000-0001-5031-468X

REFERENCES

Air, G. M. (2012). Influenza neuraminidase. *Influenza and Other Respiratory Viruses*, 6(4), 245–256.

Anderson, T. K., Nelson, M. I., Kitikoon, P., Swenson, S. L., Korslund, J. A., & Vincent, A. L. (2013). Population dynamics of cocirculating swine influenza A viruses in the United States from 2009 to 2012. *Influenza and Other Respiratory Viruses*, 7(Suppl 4), 42–51.

Bahl, J., Vijaykrishna, D., Holmes, E. C., Smith, G. J., & Guan, Y. (2009). Gene flow and competitive exclusion of avian influenza A virus in natural reservoir hosts. *Virology*, 390(2), 289–297.

Baumann, J., Mounogou Kouassi, N., Foni, E., Klenk, H. D., & Matrosovich, M. (2015). H1N1 swine influenza viruses differ from avian precursors by a higher pH optimum of membrane fusion. *Journal of Virology*, 90(3), 1569–1577.

Berry, I. M., Melendrez, M. C., Li, T., Hawksworth, A. W., Brice, G. T., Blair, P. J., ... Yoon, I.-K. (2016). Frequency of influenza H3N2 intra-subtype reassortment: Attributes and implications of reassortant spread. *BMC Biology*, 14(1), 117.

Bhatt, S., Lam, T., Lycett, S., Brown, A. L., Bowden, T., Holmes, E., ... Kellam, P. (2013). The evolutionary dynamics of influenza A virus adaptation to mammalian hosts. *Philosophical Transactions of the Royal Society of London B: Biological Sciences*, 368(1614), 20120382.

Bonfante, F., Fusaro, A., Tassoni, L., Patrono, L. V., Milani, A., Maniero, S., ... Terregino, C. (2016). Spillback transmission of European H1N1 avian-like swine influenza viruses to turkeys: A strain-dependent possibility? *Veterinary Microbiology*, 186, 102–110.

Brown, I. H., Ludwig, S., Olsen, C. W., Hannoun, C., Scholtissek, C., Hinshaw, V. S., ... Alexander, D. J. (1997). Antigenic and genetic analyses of H1N1 influenza A viruses from European pigs. *Journal of General Virology*, 78(Pt 3), 553–562.

Czudai-Matwich, V., Otte, A., Matrosovich, M., Gabriel, G., & Klenk, H.-D. (2014). PB2 mutations D701N and S714R promote adaptation of an influenza H5N1 virus to a mammalian host. *Journal of Virology*, 88(16), 8735–8742.

Delport, W., Poon, A. F., Frost, S. D., & Kosakovsky Pond, S. L. (2010). Datamonkey 2010: A suite of phylogenetic analysis tools for evolutionary biology. *Bioinformatics*, 26(19), 2455–2457.

Dowdle, W. R., & Hattwick, M. A. (1977). Swine influenza virus infections in humans. *Journal of Infectious Diseases*, 136(Suppl 3), S386–S389.

Drummond, A. J., Suchard, M. A., Xie, D., & Rambaut, A. (2012). Bayesian phylogenetics with BEAUti and the BEAST 1.7. *Molecular Biology and Evolution*, 29(8), 1969–1973.

Dunham, E. J., Dugan, V. G., Kaser, E. K., Perkins, S. E., Brown, I. H., Holmes, E. C., & Taubenberger, J. K. (2009). Different evolutionary trajectories of European avian-like and classical swine H1N1 influenza A viruses. *Journal of Virology*, 83(11), 5485–5494.

Ekiert, D. C., Bhabha, G., Elsliger, M. A., Friesen, R. H., Jongeneelen, M., Throsby, M., ... Wilson, I. A. (2009). Antibody recognition of a highly conserved influenza virus epitope. *Science*, 324(5924), 246–251.

Fraser, C., Donnelly, C. A., Cauchemez, S., Hanage, W. P., Van Kerkhove, M. D., Hollingsworth, T. D., ... Lyons, E. J. (2009). Pandemic potential of a strain of influenza A (H1N1): Early findings. *Science*, 324(5934), 1557–1561.

Furuse, Y., Suzuki, A., Kamigaki, T., & Oshitani, H. (2009). Evolution of the M gene of the influenza A virus in different host species: Large-scale sequence analysis. *Virology Journal*, 6(1), 67.

Garten, R. J., Davis, C. T., Russell, C. A., Shu, B., Lindstrom, S., Balish, A., ... Deyde, V. (2009). Antigenic and genetic characteristics of swine-origin 2009 A (H1N1) influenza viruses circulating in humans. *Science*, 325(5937), 197–201.

Glaser, L., Stevens, J., Zamarin, D., Wilson, I. A., García-Sastre, A., Tumpey, T. M., ... Palese, P. (2005). A single amino acid substitution in 1918 influenza virus hemagglutinin changes receptor binding specificity. *Journal of Virology*, 79(17), 11533–11536.

Guan, Y., Vijaykrishna, D., Bahl, J., Zhu, H., Wang, J., & Smith, G. J. (2010). The emergence of pandemic influenza viruses. *Protein & Cell*, 1(1), 9–13.

Guindon, S., Dufayard, J. F., Lefort, V., Anisimova, M., Hordijk, W., & Gascuel, O. (2010). New algorithms and methods to estimate maximum-likelihood phylogenies: Assessing the performance of PhyML 3.0. *Systematic Biology*, 59(3), 307–321.

Hayashi, T., Wills, S., Bussey, K. A., & Takimoto, T. (2015). Identification of influenza A virus PB2 residues involved in enhanced polymerase activity and virus growth in mammalian cells at low temperatures. *Journal of Virology*, 89(15), 8042–8049.

Herfst, S., Schrauwen, E. J., Linster, M., Chutinimitkul, S., de Wit, E., Munster, V. J., ... Smith, D. J. (2012). Airborne transmission of influenza A/H5N1 virus between ferrets. *Science*, 336(6088), 1534–1541.

Huang, Y., Niu, B., Gao, Y., Fu, L., & Li, W. (2010). CD-HIT Suite: A web server for clustering and comparing biological sequences. *Bioinformatics*, 26(5), 680–682.

Huson, D. H., & Scornavacca, C. (2012). Dendroscope 3: An interactive tool for rooted phylogenetic trees and networks. *Systematic Biology*, 61(6), 1061–1067.

Imai, M., Watanabe, T., Hatta, M., Das, S. C., Ozawa, M., Shinya, K., … Watanabe, S. (2012). Experimental adaptation of an influenza H5 HA confers respiratory droplet transmission to a reassortant H5 HA/H1N1 virus in ferrets. *Nature*, *486*(7403), 420–428.

Ito, T., Couceiro, J. N. S., Kelm, S., Baum, L. G., Krauss, S., Castrucci, M. R., … Webster, R. G. (1998). Molecular basis for the generation in pigs of influenza A viruses with pandemic potential. *Journal of Virology*, *72*(9), 7367–7373.

Janke, B. (2013). Clinicopathological features of swine influenza. *Current Tropics in Microbiology and Immunology*, *370*, 69–83.

Joseph, U., Linster, M., Suzuki, Y., Krauss, S., Halpin, R. A., Vijaykrishna, D., … Webby, R. J. (2015). Adaptation of pandemic H2N2 influenza A viruses in humans. *Journal of Virology*, *89*(4), 2442–2447.

Katoh, K., Misawa, K., Kuma, K., & Miyata, T. (2002). MAFFT: A novel method for rapid multiple sequence alignment based on fast Fourier transform. *Nucleic Acids Research*, *30*(14), 3059–3066.

Kawaoka, Y., Krauss, S., & Webster, R. G. (1989). Avian-to-human transmission of the PB1 gene of influenza A viruses in the 1957 and 1968 pandemics. *Journal of Virology*, *63*(11), 4603–4608.

Kida, H., Ito, T., Yasuda, J., Shimizu, Y., Itakura, C., Shortridge, K. F., … Webster, R. G. (1994). Potential for transmission of avian influenza viruses to pigs. *Journal of General Virology*, *75*(9), 2183–2188.

Kilbourne, E. D. (2006). Influenza pandemics of the 20th century. *Emerging Infectious Diseases*, *12*(1), 9–14.

Kosakovsky Pond, S. L., & Frost, S. D. (2005). Not so different after all: A comparison of methods for detecting amino acid sites under selection. *Molecular Biology and Evolution*, *22*(5), 1208–1222.

Krumbholz, A., Lange, J., Sauerbrei, A., Groth, M., Platzer, M., Kanrai, P., … Dürrwald, R. (2014). Origin of the European avian-like swine influenza viruses. *Journal of General Virology*, *95*(11), 2372–2376.

Li, W., & Godzik, A. (2006). Cd-hit: A fast program for clustering and comparing large sets of protein or nucleotide sequences. *Bioinformatics*, *22*(13), 1658–1659.

Lin, Y. P., Xiong, X., Wharton, S. A., Martin, S. R., Coombs, P. J., Vachieri, S. G., … McCauley, J. W. (2012). Evolution of the receptor binding properties of the influenza A(H3N2) hemagglutinin. *Proceedings of the National Academy of Sciences of the United States of America*, *109*(52), 21474–21479.

Lycett, S., Baillie, G., Coulter, E., Bhatt, S., Kellam, P., McCauley, J., … Brown, A. L. (2012). Estimating reassortment rates in co-circulating Eurasian swine influenza viruses. *Journal of General Virology*, *93*(11), 2326–2336.

Ma, W., Kahn, R. E., & Richt, J. A. (2008). The pig as a mixing vessel for influenza viruses: Human and veterinary implications. *Journal of Molecular and Genetic Medicine*, *3*(1), 158–166.

Matrosovich, M., Tuzikov, A., Bovin, N., Gambaryan, A., Klimov, A., Castrucci, M. R., … Kawaoka, Y. (2000). Early alterations of the receptor-binding properties of H1, H2, and H3 avian influenza virus hemagglutinins after their introduction into mammals. *Journal of Virology*, *74*(18), 8502–8512.

Nelson, M. I., Viboud, C., Vincent, A. L., Culhane, M. R., Detmer, S. E., Wentworth, D. E., … Lemey, P. (2015). Global migration of influenza A viruses in swine. *Nature Communications*, *6*, 6696.

Neverov, A. D., Lezhnina, K. V., Kondrashov, A. S., & Bazykin, G. A. (2014). Intrasubtype reassortments cause adaptive amino acid replacements in H3N2 influenza genes. *PLoS Genetics*, *10*(1), e1004037.

Nguyen, L.-T., Schmidt, H. A., von Haeseler, A., & Minh, B. Q. (2015). IQ-TREE: A fast and effective stochastic algorithm for estimating maximum-likelihood phylogenies. *Molecular Biology and Evolution*, *32*(1), 268–274.

Pensaert, M., Ottis, K., Vandeputte, J., Kaplan, M. M., & Bachmann, P. (1981). Evidence for the natural transmission of influenza A virus from wild ducks to swine and its potential importance for man. *Bulletin of the World Health Organization*, *59*(1), 75.

R Core Team. (2013). *R: A language and environment for statistical computing*, Version 3.0.2. Vienna, Austria: R Foundation for Statistical Computing.

Rabadan, R., Levine, A. J., & Robins, H. (2006). Comparison of avian and human influenza A viruses reveals a mutational bias on the viral genomes. *Journal of Virology*, *80*(23), 11887–11891.

Rogers, G. N., & Paulson, J. C. (1983). Receptor determinants of human and animal influenza virus isolates: Differences in receptor specificity of the H3 hemagglutinin based on species of origin. *Virology*, *127*(2), 361–373.

Rogers, G. N., Pritchett, T. J., Lane, J. L., & Paulson, J. C. (1983). Differential sensitivity of human, avian, and equine influenza A viruses to a glycoprotein inhibitor of infection: Selection of receptor specific variants. *Virology*, *131*(2), 394–408.

Schäffr, J. R., Kawaoka, Y., Bean, W. J., Süss, J., Senne, D., & Webster, R. G. (1993). Origin of the pandemic 1957 H2 influenza A virus and the persistence of its possible progenitors in the avian reservoir. *Virology*, *194*(2), 781–788.

Scholtissek, C. (1990). Pigs as 'mixing vessels' for the creation of new pandemic influenza A viruses. *Medical Principles and Practice*, *2*(2), 65–71.

Scholtissek, C., Rohde, W., Von Hoyningen, V., & Rott, R. (1978). On the origin of the human influenza virus subtypes H2N2 and H3N2. *Virology*, *87*(1), 13–20.

Schultz, U., Fitch, W. M., Ludwig, S., Mandler, J., & Scholtissek, C. (1991). Evolution of pig influenza viruses. *Virology*, *183*(1), 61–73.

Shope, R. E. (1931). Swine influenza: I. Experimental transmission and pathology. *The Journal of Experimental Medicine*, *54*(3), 349.

Smith, G. J., Bahl, J., Vijaykrishna, D., Zhang, J., Poon, L. L., Chen, H., … Guan, Y. (2009). Dating the emergence of pandemic influenza viruses. *Proceedings of the National Academy of Sciences of the United States of America*, *106*(28), 11709–11712.

Smith, G. J., Vijaykrishna, D., Bahl, J., Lycett, S. J., Worobey, M., Pybus, O. G., … Bhatt, S. (2009). Origins and evolutionary genomics of the 2009 swine-origin H1N1 influenza A epidemic. *Nature*, *459*(7250), 1122–1125.

Stamatakis, A. (2014). RAxML version 8: A tool for phylogenetic analysis and post-analysis of large phylogenies. *Bioinformatics*, *30*(9), 1312–1313.

Stevens, J., Blixt, O., Tumpey, T. M., Taubenberger, J. K., Paulson, J. C., & Wilson, I. A. (2006). Structure and receptor specificity of the hemagglutinin from an H5N1 influenza virus. *Science*, *312*(5772), 404–410.

Su, Y. C. F., Bahl, J., Joseph, U., Butt, K. M., Peck, H. A., Koay, E. S. C., … Smith, G. J. D. (2015). Phylodynamics of H1N1/2009 influenza reveals the transition from host adaptation to immune-driven selection. *Nature Communications*, *6*, 7952.

Swofford, D. L.. (2003). *PAUP*. Phylogenetic analysis using parsimony (*and other methods)*, Version 4. Sunderland, MA: Sinauer Associates.

Tamuri, A. U.. (2013). Treesub: annotating ancestral substitution on a tree. Retrieved from https://github.com/tamuri/treesub.

Tamuri, A. U., Dos Reis, M., Hay, A. J., & Goldstein, R. A. (2009). Identifying changes in selective constraints: Host shifts in influenza. *PLoS Computational Biology*, *5*(11), e1000564.

Taubenberger, J. K., Hultin, J. V., & Morens, D. M. (2007). Discovery and characterization of the 1918 pandemic influenza virus in historical context. *Antiviral Therapy*, *12*(4 Pt B), 581.

Taubenberger, J. K., Reid, A. H., Krafft, A. E., Bijwaard, K. E., & Fanning, T. G. (1997). Initial genetic characterization of the 1918 "Spanish" influenza virus. *Science*, *275*(5307), 1793–1796.

Taubenberger, J. K., Reid, A. H., Lourens, R. M., Wang, R., Jin, G., & Fanning, T. G. (2005). Characterization of the 1918 influenza virus polymerase genes. *Nature*, *437*(7060), 889–893.

Vijaykrishna, D., Smith, G. J., Pybus, O. G., Zhu, H., Bhatt, S., Poon, L. L., … Cheung, C. L. (2011). Long-term evolution and transmission dynamics of swine influenza A virus. *Nature*, *473*(7348), 519–522.

Vincent, A., Awada, L., Brown, I., Chen, H., Claes, F., Dauphin, G., … Lewis, N., et al. (2014). Review of influenza A virus in swine worldwide: A call

for increased surveillance and research. *Zoonoses Public Health*, *61*(1), 4–17.

Watson, S. J., Langat, P., Reid, S. M., Lam, T. T., Cotten, M., Kelly, M., ... Bonin, E. (2015). Molecular epidemiology and evolution of influenza viruses circulating within European swine between 2009 and 2013. *Journal of Virology*, *89*(19), 9920–9931.

Webby, R. J., Swenson, S. L., Krauss, S. L., Gerrish, P. J., Goyal, S. M., & Webster, R. G. (2000). Evolution of swine H3N2 influenza viruses in the United States. *Journal of Virology*, *74*(18), 8243–8251.

Webster, R. G., Bean, W. J., Gorman, O. T., Chambers, T. M., & Kawaoka, Y. (1992). Evolution and ecology of influenza A viruses. *Microbiological Reviews*, *56*(1), 152–179.

Wei, K., Chen, Y., Lin, Y., & Pan, Y. (2014). Genetic dynamic analysis of the influenza A H5N1 NS1 gene in China. *PLoS One*, *9*(7), e101384.

Wertheim, J. O., Murrell, B., Smith, M. D., Kosakovsky Pond, S. L., & Scheffler, K. (2015). RELAX: Detecting relaxed selection in a phylogenetic framework. *Molecular Biology and Evolution*, *32*(3), 820–832.

Worobey, M., Han, G.-Z., & Rambaut, A. (2014). Genesis and pathogenesis of the 1918 pandemic H1N1 influenza A virus. *Proceedings of the National Academy of Sciences of the United States of America*, *111*(22), 8107–8112.

Yang, Z. (2007). PAML 4: Phylogenetic analysis by maximum likelihood. *Molecular Biology and Evolution*, *24*(8), 1586–1591.

Yang, H., Chen, Y., Qiao, C., He, X., Zhou, H., Sun, Y., ... Zhang, Q. (2016). Prevalence, genetics, and transmissibility in ferrets of Eurasian avian-like H1N1 swine influenza viruses. *Proceedings of the National Academy of Sciences of the United States of America*, *113*(2), 392–397.

Zell, R., Scholtissek, C., & Ludwig, S. (2013). Genetics, evolution, and the zoonotic capacity of European Swine influenza viruses. *Current Tropics in Microbiology and Immunology*, *370*, 29–55.

Zhou, N. N., Senne, D. A., Landgraf, J. S., Swenson, S. L., Erickson, G., Rossow, K., ... Webster, R. G. (1999). Genetic reassortment of avian, swine, and human influenza A viruses in American pigs. *Journal of Virology*, *73*(10), 8851–8856.

Zhu, X., Guo, Y. H., Jiang, T., Wang, Y. D., Chan, K. H., Li, X. F., ... Yuen, K. Y. (2013). A unique and conserved neutralization epitope in H5N1 influenza viruses identified by an antibody against the A/Goose/Guangdong/1/96 hemagglutinin. *Journal of Virology*, *87*(23), 12619–12635.

SUPPORTING INFORMATION

Additional Supporting Information may be found online in the supporting information tab for this article.