JVI Accepted Manuscript Posted Online 22 July 2015
J. Virol. doi:10.1128/JVI.00840-15

Molecular Epidemiology and Evolution of Influenza Viruses Circulating within European Swine between 2009 and 2013

Simon J. Watson[a], Pinky Langat[a], Scott M. Reid[b], Tommy Tsan-Yuk Lam[c*], Matthew Cotten[a], Michael Kelly[b], Kristien Van Reeth[d], Yu Qiu[d], Gaëlle Simon[e], Emilie Bonin[e], Emanuela Foni[f], Chiara Chiapponi[f], Lars Larsen[g], Charlotte Hjulsager[g], Iwona Markowska-Daniel[h], Kinga Urbaniak[h], Ralf Dürrwald[i], Michael Schlegel[i], Anita Huovilainen[j], Irit Davidson[k], Ádám Dán[l], Willie Loeffen[m], Stephanie Edwards[a], Michel Bublot[n], Thais Vila[n], Jaime Maldonado[o], Laura Valls[o], ESNIP3 Consortium, Ian H. Brown[b], Oliver G. Pybus[c], Paul Kellam[a,p,#]

Wellcome Trust Sanger Institute, Hinxton, UK[a]; Animal and Plant Health Agency, Addlestone, Surrey, UK[b]; Department of Zoology, University of Oxford, Oxford, UK[c]; Laboratory of Virology, Ghent University, Merelbeke, Belgium[d]; Anses, Ploufragan-Plouzané Laboratory, Swine Virology Immunology Unit, Ploufragan, France[e]; Istituto Zooprofilattico Sperimentale della Lombardia e dell'Emilia Romagna, Parma, Italy[f]; Department of Veterinary Diagnostics and Research, Technical University of Denmark, Copenhagen, Denmark[g]; Department of Swine Diseases, Panstwowy Instytut Weterynaryjny, Pulawy, Poland[h]; IDT Biologika GmbH, Dessau-Roßlau, Germany[i]; Finnish Food Safety Authority EVIRA, Helsinki, Finland[j]; Division of Avian Diseases, Kimron Veterinary Institute, Bet Dagan, Israel[k]; National Food Chain Safety Office, Budapest, Hungary[l]; Central Veterinary Institute, Wageningen UR, Lelystad, The Netherlands[m];

Virology Department, Discovery Research, Merial, Lyon, France[n]; Veterinary Diagnostic Services DIAGNOS, Laboratorios HIPRA SA, Gerona, Spain[o]; Division of Infection & Immunity, University College London, London, UK[p]

Running title: Molecular epidemiology and evolution of European swIAV

Address correspondence to Paul Kellam, pk5@sanger.ac.uk

* Present address: State Key Laboratory of Emerging Infectious Diseases and Center of Influenza Research, School of Public Health, The University of Hong Kong, Hong Kong, SAR

Word count for abstract: 211

Word count for text *[excluding ref, footnotes, legends]*: 5215

ABSTRACT

The emergence in humans of the A(H1N1)pdm09 influenza virus, a complex reassortant virus of swine origin, highlighted the importance of worldwide influenza virus surveillance in swine. To date, large-scale surveillance studies have been reported for southern China and North America, but such data has not yet been described for Europe. We report the first large-scale genomic characterization of 290 swine influenza viruses collected from 14 European countries between 2009 and 2013. 23 distinct genotypes were identified, with the seven most common comprising 82% of the incidence. Contrasting epidemiological dynamics were observed for two of these genotypes, $H1_{hu}N2$ and H3N2, with the former showing multiple long-lived geographically-isolated lineages, whilst the latter had short-lived geographically-diffuse lineages. At least 32 human-swine transmission events have resulted in A(H1N1)pdm09 becoming established at a mean frequency of 8% across European countries. Notably, swine in the UK have largely had a replacement of the endemic Eurasian 'avian-like' genotypes with A(H1N1)pdm09-derived genotypes. The high number of reassortant genotypes observed in European swine, combined with the identification of a genotype similar to the A(H3N2)v in North America, underlines the importance of continued swine surveillance in Europe for the purposes of public health. This study further reveals that the emergence and drivers of virus evolution in swine differ at a global level.

55 **IMPORTANCE**

56 The influenza A(H1N1)pdm09 virus contains a reassortant genome with segments
57 derived from separate virus lineages that evolved in different regions of world. In
58 particular its neuraminidase and matrix segments were derived from the Eurasian
59 'avian-like' lineage that emerged in European swine in the 1970s. However, while large-
60 scale genomic characterization of swine has been reported for southern China and
61 North America, no equivalent study has yet been reported for Europe. Surveillance of
62 swine herds across Europe between 2009 and 2013 revealed that the A(H1N1)pdm09
63 virus is established in European swine, increasing the number of circulating lineages in
64 the region and increasing the possibility of the emergence of a genotype with human
65 pandemic-potential. It also has implications for veterinary health, making prevention
66 through vaccination more challenging. The identification of a genotype similar to the
67 A(H3N2)v, causing zoonoses at North American agricultural fairs, underlines the
68 importance of continued genomic characterization in European swine.

69

INTRODUCTION

Swine influenza viruses (swIAV) cause influenza in pigs, a disease that results in significant morbidity in swine herds across the world. swIAV outbreaks are due to infection with influenza A viruses (IAV), however, the genetic diversity of all circulating swIAVs can be retraced using phylogenetic methods to avian progenitor viruses, often *via* humans (1). The first documented case of swIAV was in North America during the 1918 pandemic of human H1N1 influenza A virus (IAV) (2). This H1N1 subtype is thought to have transmitted from birds to humans shortly before 1918, with the human virus likely to have transferred into swine during the pandemic (3, 4). This virus became established in swine in the United States, forming the lineage now referred to as "classical swine H1N1" (CS) (1, 2, 4). Despite the prevalence of this lineage in North America and Asia, it did not become established in Europe until 1976, when infected pigs imported from the US resulted in an outbreak in Italy that subsequently spread throughout European swine (1, 5).

European swine were infected solely by CS lineage viruses until 1979, when an avian H1N1 virus, genetically distinct from the CS lineage, was isolated from pigs in Belgium and Germany (1, 2, 6-8). This virus, now called "Eurasian 'avian-like' swine H1N1" (EA), rapidly spread throughout Europe, outcompeting the pre-existing CS viruses (1). The EA lineage continues to circulate among European swine, and since its emergence has reassorted with human seasonal-origin viruses, resulting in the co-circulation of three distinct virus subtypes in Europe: (i) Eurasian 'avian-like' $H1_{av}N1$; (ii)

92 A/swine/Gent/1/1984-like H3N2 (Gent/84); and (iii) A/swine/Scotland/410440/1994-like
93 $H1_{hu}N2$ (Scot/94) (2, 9-13).
94
95 In April 2009, a novel H1N1 IAV was isolated from humans in Mexico and the United
96 States (14). This virus rapidly spread throughout the human population, causing the first
97 global influenza pandemic of the 21st century. Studies showed that this virus, named
98 A(H1N1)pdm09, was of swine origin and arose from the reassortment of an EA $H1_{av}N1$
99 virus with a 'triple-reassortant' swine virus that has circulated in North America and Asia
100 since 1997 (15). The emergence of pandemic IAV from a swine, rather than an avian,
101 source was unexpected (16, 17). Furthermore, its complex reassortant history (involving
102 swIAVs circulating in separate regions of the world), combined with the length of time
103 that the lineage had persisted without being detected, highlighted the need for more
104 swIAV surveillance in swine worldwide (14). As a result, surveillance of swIAV has
105 increased globally since 2009 (18-20), and large-scale whole-genome sequencing studies
106 have been reported for southern China and Hong Kong (21, 22), and North America (23,
107 24). These studies discovered complex IAV diversity in swine, with high levels of
108 reassortment between the enzootic lineages. Despite the importance of European-
109 derived viruses in the genesis of the A(H1N1)pdm09 genotype, no large-scale whole-
110 genome study has been reported for Europe. Therefore, the European Surveillance
111 Network for Influenza in Pigs 3 (ESNIP3) was formed as an active swine surveillance
112 network in participating European countries, representing the largest consortium for
113 coordinated monitoring of IAV in pigs in Europe. As part of its work, the network

undertook whole-genome sequencing of isolates from consortium partners sampled since the emergence of A(H1N1)pdm09 in 2009. Here we report the genomic diversity and molecular epidemiology of swIAV in Europe. Our study characterizes a total of 290 swIAV isolates from 14 countries between 2009 and 2013, and reveals significant genotypic diversity of swIAV in European swine for the first time, as well as substantial intra-continental differences in swIAV epidemiology.

MATERIALS AND METHODS

Surveillance and sample collection. ESNIP3 consortium partners carried out influenza surveillance between 2010 and 2013 on swine farms with outbreaks of respiratory disease. Preliminary subtyping was performed on the isolated viruses at the time of sample collection (25). Samples positive for swIAV were selected for further antigenic and genetic characterization based on their geographic location, date of collection, and viral subtype, to capture the diversity of the circulating viruses across Europe. Further details of the total numbers of swIAV detected in each country, and thus the proportion that were sent for sequencing, can be found in (25). All selected samples for each country were sent to a central virus bank at the UK Animal and Plant Health Agency, where they were cultured in embryonated fowls' eggs prior to viral RNA extraction.

PCR amplification and virus sequencing. Viral RNA amplification was performed using an 8-segment RT-PCR as previously described (26) using the modified primers described in Baillie *et al.* (27). Amplicons for each sample were pooled, then individually indexed

and processed into libraries through either the standard Roche Rapid Library Prep for the 454 sequencing platform, or as described by Quail *et al.* (2008) for the Illumina platform (28). Isolates were sequenced on either the Genome Sequencer FLX Titanium XL+ instrument (Roche/454 Life Sciences), or sequenced on the MiSeq instrument (Illumina) using the 150 bp paired-end reagent kit.

Genome assembly. Data generated by either platform were quality-controlled using QUASR version 7.01, to remove any primer sequences and trim reads by applying a median-read-quality cutoff, as previously described (29). Quality-controlled readsets were *de novo* assembled using IVA version 0.8.1 (30), and custom Python scripts were used to remove any assembled contiguous sequences ("contigs") that were either not of influenza-origin, or did not contain at least 70% of the expected open reading frame length for that segment. In addition, readsets were assembled against a reference sequence using SMALT version 0.7.4 (www.sanger.ac.uk/resources/software/smalt/), with the reference selected using custom Python scripts that performs a BLAST search on a subset of the reads and downloads the most frequent hit for each segment. Output files generated by SMALT were parsed using SAMtools version 0.1.8 (31) and QUASR to generate consensus sequences. These consensus sequences were used to fill in segments unable to be assembled by the *de novo* assembler. Python scripts are available from the authors on request.

For samples where multiple contigs could be generated for one or more segments, the lineage-of-origin for all segments was determined (*i.e.* classical swine, Eurasian 'avian-like', 2009 pandemic, human seasonal-derived, triple reassortant, or avian). Where samples had one or two segments with contigs of different lineages, the contig whose lineage matched the remaining segments was selected, with the others considered contaminants and removed. Where multiple genotypes could be constructed (*e.g.* complete pdm09 and EA genotypes in a sample), a custom Python script was used to calculate the relative abundance of each genotype in the original reads. If the genotype was present at <5% in the sample, then it was considered a contaminant and discarded. Using this stringent quality check, no mixed infections were detected in any of the samples.

Phylogenetic analysis. The influenza sequences were combined with existing sequences retrieved from the NCBI Influenza Virus Resource (32) that represented the range of genetic diversity of swIAV worldwide. All available European swine isolates were included. Each genome segment was aligned separately using the MUSCLE aligner (33) provided in MEGA version 6.06 (34). Separate alignments were made for H1, H3, N1, and N2 sequences. Alignments were then trimmed to coding regions, and sequences covering less than 50% of the coding region were removed. The resultant datasets contained between 763 (N1) and 2405 (H3) sequences. Phylogenetic trees were then inferred under a maximum-likelihood (ML) criterion using RAxML version 7.2.8 (35). Phylogenies were inferred under the general time-reversible nucleotide substitution

model, with among-site rate heterogeneity modeled as a 4-category discrete gamma-distribution (GTR+Γ_4). Tree robustness was determined through bootstrap analysis of 1,000 sequence pseudo-replicates. Trees were visualized using FigTree version 1.4.2 (http://tree.bio.ed.ac.uk/software/figtree/).

Genotype assignment. From the ML phylogenies, virus isolates were categorized into lineages circulating in swine worldwide, with a particular emphasis on Europe: (i) Eurasian 'avian-like' $H1_{av}N1$; (ii) A/swine/Gent/1/1984-like H3N2; (iii) A/swine/Scotland/410440/1994-like $H1_{hu}N2$; (iv) A/swine/Italy/4675/2003-like rH1N2; (v) North American 'triple reassortant'; (vi) 'classical' H1N1; (vii) A(H1N1)pdm09; (viii) human seasonal H3N2; and (ix) avian. Each segment for a sample was assigned to one of the nine lineages defined above to generate a complete genotype for that sample. For the purposes of genotype classification, where samples had one or more missing internal gene segments (PB2, PB1, PA, NP, MP, NS), the lineage of the missing segments was assigned to those of the sequenced internal genes. This assumption is reasonable because of the negligible rate of reassortment observed within the internal gene cassette of the other isolates in this study. Where no internal gene segments were obtained, the genotype was specified as 'Undetermined'. Furthermore, if either of the external glycoprotein segments (HA or NA) were missing, the genotype was also considered 'Undetermined'.

200 Isolates collected by the ESNIP3 project that had already been deposited in GenBank
201 were removed to avoid duplication of isolates. Furthermore, to avoid the
202 misrepresentation of genotype proportions across Europe, only a single isolate
203 representative of an outbreak was retained per country per year. The ML phylogenies
204 for all segments were scrutinized, and where viruses grouped together in a well-
205 supported clade across all eight segments, they were considered from the same
206 outbreak. A single sample was chosen to represent that outbreak per year, retaining the
207 ESNIP3 isolate where appropriate.
208
209 **Lineage-specific phylogenies.** Molecular phylogenies were estimated for the H1 and N2
210 genes for the Scot/94 lineage, and the N2 gene for the Gent/84 lineage. These were
211 inferred using MrBayes version 3.2.2 (36, 37), with three Bayesian Markov Chain Monte
212 Carlo (BMCMC) chains run for 2 million states under a GTR+Γ_4 substitution model. A 25%
213 burn-in was discarded after assessing convergence using Tracer version 1.6
214 (http://tree.bio.ed.ac.uk/software/tracer/).
215
216 **Molecular clock phylogeny.** A time-resolved phylogeny was estimated for the
217 A(H1N1)pdm09 lineage in order to investigate the transmission of this virus from
218 humans to swine. All European human influenza A(H1N1)pdm09 internal gene cassette
219 (IGC) sequences were downloaded from the NCBI Influenza Virus Resource, along with
220 their collection dates. After separately aligning and trimming each segment to its coding
221 region using MEGA version 6.06, the segments were concatenated. A Python script was

used to down-sample the sequences to remove highly-similar sequences sampled from the same geographic location and month. These were then combined with the previously-curated A(H1N1)pdm09 swine IGC sequences. Molecular clock Bayesian phylogenies were inferred using BEAST version 1.8.0 (38, 39), under a GTR+ Γ_4 substitution model. A strict molecular clock was used after assessing the clock-likeness of the data using Path-O-Gen version 1.4 (http://tree.bio.ed.ac.uk/software/pathogen/), and a Bayesian Skyride coalescent model was used to model demographic history (40). Three BMCMC chains were run for 100 million states, with samples taken every 10,000. A 10% burn-in was used, and convergence of the chains assessed using Tracer version 1.6.

Statistical analyses. The null hypothesis of no association between host species and phylogenetic ancestry of the A(H1N1)pdm09 lineage was tested using Bayesian Tip-association Significance Testing (BaTS) beta build 2 (41). All full-length human and swine A(H1N1)pdm09 PB2, HA-H1, and NA-N1 sequences were downloaded from the NCBI Influenza Virus Resource and combined with the A(H1N1)pdm09 swine sequences generated in this study. Separate alignments were made for each segment, and alignments trimmed to the coding region. Bayesian posterior sets of trees were inferred using MrBayes version 3.2.2 as above. Custom Python scripts were written to generate the BaTS input Nexus file where each taxon is labeled with its corresponding host species.

244 **Nucleotide sequence accession numbers.** The generated nucleotide sequences are
245 available in GenBank under accession numbers KR699644 to KR701609.

RESULTS

Genotypic diversity. A total of 243 viruses, provided by the ESNIP3 consortium, that circulated between 2009 and 2013 were used to generate 231 complete and 12 incomplete genomes using high-throughput sequencing platforms. These were combined with the 47 genomes present in GenBank to generate a set of 290 swIAV genomes from 14 countries across Europe: Belgium, Denmark, Finland, France, Germany, Hungary, Israel, Italy, the Netherlands, Norway, Poland, Spain, Sweden, and the United Kingdom (**Supplementary Table 1**).

From these data, 23 different genotypes (A-W) were observed across Europe (**Figure 1**). 12 genotypes (A-L), comprising 67% of the isolates, contained an internal gene cassette (IGC) derived from the Eurasian 'avian-like' (EA) lineage, while eight genotypes (P-W), comprising 27% of the isolates, contained an IGC derived from the A(H1N1)pdm09 lineage. Three genotypes (M, N, O), each observed only once, contained a reassortant IGC with segments from both the EA and A(H1N1)pdm09 lineages. The majority of the observed genetic diversity was therefore generated through reassortment of the glycoprotein-encoding segments of the four major lineages circulating in European swine: (i) EA $H1_{av}N1$; (ii) A/swine/Scotland/410440/1994-like $H1_{hu}N2$; (iii) A/swine/Gent/1/1984-like H3N2; and (iv) A(H1N1)pdm09. Additional diversity in the neuraminidase (NA) segment was provided through the circulation of the A/swine/Italy/4675/2003 'human-like' N2 lineage, and sporadic reverse zoonoses of human H3N2 seasonal viruses.

Despite 23 distinct swIAV genotypes being identified in European swine, 82% of the samples could be attributed to the seven most frequent genotypes (A, B, C, D, P, Q, R). The remaining 18% represented rarer and geographically-constrained reassortant genotypes (**Figure 1**).

Frequency of EA-based genotypes across Europe. The proportion of each circulating swIAV genotype varied among European regions (**Figure 2**). This was most stark in the contrast between the United Kingdom and mainland Europe. Swine in mainland Europe were predominantly infected by EA-based genotypes (A-L); among countries where more than 10 isolates were analysed, the mean percentage of all these genotypes together was 83% ± 11%. UK swine, conversely, were predominantly infected by A(H1N1)pdm09-based genotypes (P-W), with the proportion of the EA-based genotypes being only 15%.

Other geographic trends within mainland Europe were less pronounced and possibly also reflected different sampling biases (**Figure 2**). EA $H1_{av}N1$ (genotype A) was found across mainland Europe at an average prevalence of 37% ± 16%; frequency was highest in Belgium (58%) and lowest in Germany (18%) (**Figure 2A**). Gent/84-like H3N2 (genotype B) was observed across mainland Europe at an average frequency of 15% ± 14%, with Spain (36%) and Hungary (33%; n=6) having highest frequency, but only a single outbreak of this genotype was isolated in France, close to the Belgian border.

This genotype was not isolated in Denmark, Poland, or the UK (**Figure 2B**). Scot/94-like $H1_{hu}N2$ genotype (genotype C) was present across mainland Europe at an average prevalence of 7% ± 10%, and at 7% in the UK. The prevalence of this genotype appears to be inversely proportional to that of genotype B, having higher relative frequency in countries where B had low frequency, and *vice versa* (**Figure 2C**). Its highest frequency was in France (30%), while Belgium, Germany, and Italy only had a single isolate, and the genotype was not isolated in Denmark. Instead, Danish swine were predominantly infected with a reassortant $rH1_{av}N2$ (genotype D). This was the most frequent genotype in Denmark (47%), but only two isolates of genotype D were identified in Sweden, and single isolates were observed in the Netherlands, Germany and Italy (**Supplementary Figure 1**).

The remaining EA-derived genotypes (E-L) were geographically-constrained reassortants found either at a significant percentage in one country with sporadic occurrences in a neighbouring country (E, G), or as transient isolated occurrences within a country (H-L). A genotype of particular note is F, containing EA-derived internal genes, and a Scot/94-derived HA, but whose NA originated from a separate human-to-swine transmission to the Gent/84 and Scot/94 transmission events. This was found circulating at 19% prevalence in Italy, where it originally arose (42).

Frequency of A(H1N1)pdm09 genotypes across Europe

Following the outbreak of the human pandemic in April 2009, A(H1N1)pdm09 influenza (genotype P) was detected in European swine as early as September 2009 (43). Since then, the pandemic virus has been isolated from swine in countries across Europe at varying frequencies, with UK and Poland each having the greatest at 27%, while mainland Europe on average had 8% ± 9% in countries with more than 10 samples (**Figure 2**). The pandemic virus was not detected in swine in Belgium or the Netherlands.

Following the introduction of A(H1N1)pdm09 into European swine, reassortants between this genotype and the established genotypes circulating within Europe soon emerged. In the UK, the pandemic virus reassorted with enzootic Scot/94-derived $H1_{hu}N2$, acquiring the external glycoproteins (genotype Q) (**Figure 2**). This genotype was first observed in 2010 and rapidly replaced genotype C; from 2010 to 2013 no genotype C was isolated in the UK (**Figure 3**), while genotype Q became the most frequent genotype, comprising 54% of the isolates (**Figure 3**). Despite its circulation in the UK, genotype Q was not observed in any other European country (**Figure 2Q**). Instead, mainland Europe saw the emergence of a reassortant between A(H1N1)pdm09 and the endemic H3N2 (genotype B), with the pandemic virus acquiring an NA-N2 segment (giving genotype R). This genotype was predominantly isolated in Germany, where it comprised 26% of the isolates, but was also found at a much lower prevalence in Italy and the Netherlands (**Figure 2R**). Aside from these three genotypes, A(H1N1)pdm09-based viruses were found in just 3% of the samples (**Figure 1**).

Phylogenetic analyses

(i) A/swine/Scotland/410440/1994-like. Phylogenetic analysis of the haemagglutinin (**Supplementary Figure 1**) and neuraminidase (**Figure 4**) segments of the Scot/94-derived $H1_{hu}N2$ lineage showed at least four long-lived clades of the virus circulating within Europe. Three of these were geographically structured, circulating exclusively within France (**Figure 4; blue box**), Spain (**Figure 4; pink box**) and the UK (**Figure 4; orange box**) respectively. The fourth clade (**Figure 4; yellow box**) has a different dynamic, containing viruses isolated from different countries.

Each of these four clades has further diversified into multiple sub-clades that are present in their geographical region. At least four sub-clades are observed within France, which, on several occasions, have reassorted with EA $H1_{av}N1$ viruses to acquire their HA-H1 segment, forming a novel reassortant genotype (genotype G). In the UK, at least three sub-clades are observed, each of which has reassorted with the A(H1N1)pdm09 virus to acquire their internal genes. A further reassortment event between a Spanish virus of genotype J and a UK Scot/94 virus resulted in the Spanish virus replacing its human seasonal-origin NA with Scot/94 NA to become a complete Scot/94 genotype. This then circulated in Spanish swine, apparently replacing the original virus (**Figure 4, Supplementary Figure 1; pink box**).

The Scot/94 haemagglutinin phylogeny shows that a distinct clade circulated in Italy. These viruses acquired an NA-N2 segment from a separate human-to-swine

355 transmission (genotype F), with the resultant virus becoming enzootic within Italy. These
356 viruses have evolved into at least two sub-clades, with one predominant between 2005
357 and 2010, while the other predominated after 2010.
358
359 **(ii) A/swine/Gent/1/1984-like.** Phylogenetic analysis of the neuraminidase segment for
360 the Gent/84 H3N2 lineage shows a markedly different dynamic to the Scot/94 $H1_{hu}N2$
361 lineage (**Figure 5**). The phylogeny does not show the same maintenance of distinct
362 lineages characteristic of the Scot/94 virus. Instead the Gent/84 H3N2 virus (genotype B)
363 has a more 'ladder-like' phylogeny, with higher lineage turnover and lower genetic
364 diversity in any year, characteristic of the human seasonal H3N2 virus. The virus is also
365 geographically heterogeneous, with the short-lived lineages isolated from multiple
366 countries concurrently.
367
368 However, a reassortment event with an EA $H1_{av}N1$ virus resulted in the replacement of
369 the HA-H3 segment with an HA-H1 (genotype D) (**Figure 5; purple box**). This genotype
370 became enzootic within Denmark, forming a separate clade within the Gent/84 NA-N2
371 phylogeny. The dynamics of the virus within this clade are more like those of the
372 Scot/94 NA-N2, with multiple sub-lineages co-circulating. This genotype shows a higher
373 propensity to reassort, with at least two sub-lineages emerging that have replaced their
374 EA segments with A(H1N1)pdm09 ones (genotype R). A second A(H1N1)pdm09
375 reassortant that replaced its EA internal genes with pdm09 ones (genotype T) was also
376 observed on at least two separate occasions.

377

378 **(iii) A(H1N1)pdm09.** The A(H1N1)pdm09 virus has only been observed in European
379 swine since late 2009. As a result of this recent emergence, the genetic diversity of the
380 virus is low, and the phylogenetic branching in the lineage is incompletely resolved,
381 resulting in polytomies for all of the segments. To place the swine isolates in the context
382 of the human outbreak, a molecular clock phylogeny was estimated for the internal
383 gene cassette, combining all European swine and human isolates. The phylogeny shows
384 the swine isolates interspersed throughout the tree, with an estimated minimum of 32
385 different transmissions of the virus from humans to swine (**Supplementary Figure 2**).
386 Due to the limited sampling of A(H1N1)pdm09 in swine, it is not clear from the
387 molecular clock phylogeny whether the virus circulated enzootically in European swine
388 or if it was maintained through continual short-lived introductions from humans. To
389 assess this question statistically, Bayesian phylogenies containing human and swine
390 isolates were inferred for the A(H1N1)pdm09 PB2 segment (**Supplementary Figure 3**).
391 Statistical testing of the association between host species and phylogenetic relationship
392 ancestry using BaTS showed that the swine isolates clustered more often than expected
393 by chance, suggesting that some of the introduced A(H1N1)pdm09 virus had been
394 circulating within swine in Europe (**Table 1**). The BaTS test was repeated for the HA-H1
395 (**Supplementary Figure 4**) and NA-N1 segments (**Supplementary Figure 5**), with the
396 results again showing significant clustering of the swine isolates (**Table 1**). Together,
397 these results suggest that both the internal genes and the external genes of the
398 A(H1N1)pdm09 virus are clustering more than expected by chance.

Reassortant genotype of public-health interest. A genotype of particular note is the triple reassortant isolated in Spain in 2012 containing Gent/84 external glycoproteins, and EA internal genes but an A(H1N1)pdm09 matrix segment (A/swine/Spain/28778/2012). This genetic makeup is comparable to the North American A(H3N2)v that has been associated with multiple swine-to-human zoonoses in North American swine fairs (44); both contain human-derived H3 and N2 glycoproteins that have since evolved within swine, and both contain internal gene cassettes with an acquired pdm09 matrix protein. This constellation therefore poses a potential public health risk, particularly as its external glycoproteins have been antigenically evolving within swine for over 30 years, so humans will likely be immunologically naïve against the virus. Because of the possible public health interest in this genotype, this isolate was re-extracted and re-sequenced, confirming its makeup.

DISCUSSION

The genomic characterization of 290 European swIAV, of which 243 were sequenced as part of this study, is reported here. These genomes define the diversity of swIAV across Europe between 2009 and 2013, during which time the A(H1N1)pdm09 lineage was introduced back into swine through reverse-zoonosis. The introduction of this new swIAV lineage into European swine increased the complexity of circulating genotypes, resulting in an increase in the number of reassortment possibilities. In total 23 different genotypes were observed among 278 genomes from European swine. In contrast, a

study by Liang *et al.* in southern China over a comparable time period found 29 genotypes among 387 genomes (21). This reduced reassortment diversity in European swIAV is a result of both a smaller number of genotypes and a bias towards fewer genotypes. Furthermore, reassortment involving the IGC was rare in European swIAV, with just three isolates (1%) found to contain a reassortant IGC. In contrast, reassortment of the internal segments was observed frequently in southern China, with multiple reassortant genotypes persisting in the swine population. This is likely to be due to the presence of the triple reasortant (TR) lineage in China, in addition to the EA and pdm09 lineages, because the majority of IGC reassortment events observed by Liang *et al.* involved the TR lineage (21). Only one EA/pdm09 IGC reassortant was isolated recurrently in China – an EA genotype with an A(H1N1)pdm09 matrix gene – which was also isolated in this study (genotype M). This reassortment difference extends to North America, where the TR lineage first arose and is the predominant lineage; although the frequency of inter-lineage reassortment within the internal gene segments was lower than for southern China, it was still considerably higher than in Europe (24).

Despite 23 distinct genotypes being found in European swine, only four (A, B, C, P) were found to be circulating across the whole of Europe. A further six (D, E, F, G, Q, R) were found in geographically-constrained areas, *i.e.* highly frequent in a single country with occasional outbreaks in other countries. The remaining 13 genotypes were isolated sporadically and infrequently, suggesting that these were perhaps less fit reassortants

443 that were only identified due to the extent of surveillance. It is likely that regular whole
444 genome sequencing-based surveillance of human, swine, and avian influenza will
445 provide a more compelling catalogue of the diversity and fitness landscape of IAV
446 reassortants than is possible through *in vitro* studies.
447
448 The four genotypes circulating throughout Europe include the three lineages that have
449 been enzootic and prevalent in European swine for at least 19 years: EA $H1_{av}N1$,
450 Gent/84 H3N2, and Scot/94 $H1_{hu}N2$. However these genotypes have different
451 frequencies and dynamics across mainland Europe. While $H1_{av}N1$ was found at a high
452 frequency in all countries, the prevalence of H3N2 and $H1_{hu}N2$ were inversely related.
453 These observations are consistent with the data obtained through preliminary subtyping
454 of the ESNIP3 samples (25). Differences between the lineages were also observed in
455 phylogenetic analysis of the two lineages' glycoprotein segments; the $H1_{hu}N2$ phylogeny
456 showed greater genetic diversity at any point in time through long-lived lineages that
457 circulate independently in different countries (**Figure 4**). Conversely, the H3N2 lineage
458 had less genetic diversity at any point in time due to its rapid turnover of short-lived
459 lineages that were geographically diffuse (**Figure 5**). This dynamic relationship between
460 the two lineages is similar to the relationship between the A/H1N1 and A/H3N2
461 subtypes in humans (45), or the B/Yamagata and B/Victoria lineages in humans (46).
462
463 The differing phylogenies of the $H1_{hu}N2$ and H3N2 lineages give an insight into their
464 epidemiological dynamics. The 'ladder-like' phylogeny of H3N2 suggests that the lineage

is under strong selective pressure, against which the virus fixes advantageous mutations rapidly along the trunk of the tree, with loss of the side-branches that do not contain the variation (47). The co-circulation of multiple sub-clades of the $H1_{hu}N2$ lineage within each country however, suggests that the virus is not subject to the same intense selective pressures as the H3N2 lineage. This could be due to reduced cross-reactive immunity in swine between the $H1_{hu}N2$ sub-clades. This difference in selective pressure on the two lineages may explain why the Gent/84 H3 was only found in conjunction with the Gent/84 N2, whereas the Scot/94 H1 was apparently more able to reassort with the NA of other lineages (**Figure 1**). Furthermore, this inverse relationship between H3N2 and $H1_{hu}N2$ may be influenced by evolutionary dynamics and selection – in swine populations such as France and the UK, the emergence of the $H1_{hu}N2$ correlated with the disappearance of H3N2. Possibly immunity to N2 had an influence favoring selection of the lineage with a greater diversity and opportunity for selection of fitter viruses in swine. The cause of the apparent absence of such pressures in other major swine-producing countries (Belgium, Germany, the Netherlands, and Spain), where H3N2 and $H1_{hu}N2$ co-exist, is unclear, though may be due to differences in their swine production systems compared to UK and France. The latter have a relatively low pig density (<90 heads/hectare agricultural area in 2010) compared to the former (>90 heads/hectare). The Netherlands, Belgium, Spain, and Germany have among the highest swine densities in Europe, with the Netherlands as high as 704 heads/hectare in 2010 (http://www.fao.org/docrep/017/i3138e/i3138e07.pdf). A higher swine density has been previously shown to increase the risk of seroprevalence for influenza, and as such

487 may be associated with the co-circulation of the two genotypes (48). However, a more
488 formal statistical assessment of the predictors is required, and will require examination
489 of different industry structures, production systems, and vaccination usage to better
490 understand the underlying factors influencing virus evolution.
491
492 The A(H1N1)pdm09 virus that emerged in humans in early 2009 was the third most-
493 frequent genotype found in swine across Europe. The first confirmed disease outbreak
494 in European pigs was in September 2009, but it is highly possible that the virus crossed
495 to pigs from humans earlier, after its emergence in Europe in April (43). In this study we
496 have estimated that at least 32 separate introductions of the A(H1N1)pdm09 virus from
497 humans into swine have occurred in the period to 2013, and we find phylogenetic
498 evidence that the virus circulates endemically among swine. This finding is in contrast to
499 the study in southern China, where the A(H1N1)pdm09 virus did not persist after each
500 introduction, and only its internal genes were maintained in swine through
501 reassortment with the HA and NA genes of other enzootic lineages (21). However, we
502 also observed replacement of the A(H1N1)pdm09 external glycoproteins through
503 reassortment with enzootic lineages; notably acquiring the H1 and N2 from the $H1_{hu}N2$
504 lineage (genotype C) in the UK, or acquiring the N2 from the H3N2 lineage (genotype B)
505 in Germany. The prevalence of these A(H1N1)pdm09 reassortants have previously been
506 noted (49-52), and suggest that, as in southern China, the internal genes are highly
507 compatible with glycoprotein segments from enzootic lineages, and therefore the

circulation of the A(H1N1)pdm09 in European swine increases the reassortment potential for European swIAV.

Infections in swine of the A(H1N1)pdm09 virus varied across Europe; mainland Europe was found to have an average frequency of 8%, which is in agreement with the 9% proportion found from the preliminary subtyping of European swine (25). Swine in the UK, however, have seen a near-complete replacement of their enzootic $H1_{av}N1$ and $H1_{hu}N2$ viruses with A(H1N1)pdm09 and pdm09-$H1_{hu}N2$ viruses. Phylogenetic analysis of the Scot/94 H1 and N2 segments showed that at least four separate sub-lineages of EA-$H1_{hu}N2$ each replaced their EA IGC with the pdm09 one, consistent with the idea that the pdm09 IGC is fitter in swine. The relative proportions of the $H1_{av}N1$ and $H1_{hu}N2$ subtypes have remained the same since the replacement of the EA lineage by the pdm09 one; a serological study of UK swine between 2008 and 2009 showed that $H1_{hu}N2$ was the predominant subtype, detected in 45% of all farms, with the $H1_{av}N1$ subtype found in approximately 21% of farms (53). Consistent with this, we showed here that since the introduction of the pdm09 virus in UK swine, the frequency of the pdm09-$H1_{hu}N2$ is approximately 54% while the pdm09 virus is 27%. The reasons for the difference in prevalence between mainland Europe and the UK are unclear and warrant further investigation.

A triple reassortant genotype containing an EA IGC with a pdm09 matrix gene and Gent/84 H3 and N2 segments (genotype N) was observed in a single Spanish pig

(A/swine/Spain/28778/2012). Acquisition of the pdm09 matrix protein by enzootic swIAVs has been previously noted in China (21), and also in the United States, where the resultant A(H3N2)v genotypes were able to infect humans (54, 55). Genotype N has a similar genetic makeup to the A(H3N2)v, and importantly its external glycoproteins are derived from the Gent/84 H3N2 lineage that has been evolving in swine since the early-to-mid 1970s (1). As such, humans are likely to be immunologically-naïve against this virus, and as such it poses a potential public health risk. Given the level of reassortment observed in European swine, further surveillance efforts should be sought to track the emergence and potential spread of such genotypes with human-pandemic potential. Whole-genome sequencing of swIAV isolates is an important aspect of this surveillance effort, without which the dynamics of the circulating lineages cannot be determined. Furthermore, only through whole-genome sequencing can the rare, but potentially important, reassortants involving the IGC be observed. However, the limited IGC reassortment indicates that preliminary subtyping of the HA and NA segments is still suitable for routine surveillance of European swIAV.

This study reveals that the emergence and drivers of virus evolution in pigs differ at the global level. The factors favoring virus emergence and selection are complex, but we show that establishment of new genotypes and lineages are complex and less frequent at the population level. Whole system analyses, both at the virus host level, together with the influence of natural or vaccine-derived immunity, require further investigation.

ACKNOWLEDGEMENTS

The authors thank the veterinarians and swine holders who submitted samples to the partner laboratories. They also acknowledge all scientists and technicians who contributed towards the provision of these data. ESNIP3 consortium: Karen Van der Meulen, Ghent University, Belgium; Montserrat Agüero Garcia, Laboratorio Central de Veterinaria, Spain; Charalambos Billinis and Constantinos Kyriakis, University of Thessaly, Greece; Nicola S. Lewis, University of Cambridge, United Kingdom; Timm Harder, Friedrich-Loeffler-Institut, Germany; Giovanni Cattoli, Instituto Zooprofilattico Sperimentale delle Venezie, Italy; Richard Webby, St. Jude Children's Research Hospital, USA; Hualan Chen, Harbin Veterinary Research Institute, China; Amy L. Vincent, US Department of Agriculture, USA. This work was funded by the European Union FP7-INFLUENZA-2010 ESNIP3 programme (grant #259949).

REFERENCES

1. **Brown I**. 2000. The epidemiology and evolution of influenza viruses in pigs. Veterinary Microbiology **74**:29–46.
2. **Zell R**, **Scholtissek C**, **Ludwig S**. 2012. Genetics, Evolution, and the Zoonotic Capacity of European Swine Influenza Viruses, pp. 29–55. *In* Richt, JA, Webby, RJ (eds.), Swine Influenza. Springer Berlin Heidelberg, Berlin, Heidelberg.
3. **Taubenberger JK**, **Reid AH**, **Lourens RM**, **Wang R**, **Jin G**, **Fanning TG**. 2006. Molecular virology: Was the 1918 pandemic caused by a bird flu? Was the 1918 flu avian in origin? (Reply). Nature **440**:E9–E10.
4. **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 201324197.
5. **Lange J**, **Groth M**, **Kanrai P**, **Pleschka S**, **Scholtissek C**, **Dürrwald R**, **Platzer M**, **Sauerbrei A**, **Zell R**. 2014. Circulation of classical swine influenza virus in Europe between the wars? Arch Virol **159**:1467–1473.
6. **Dunham EJ**, **Dugan VG**, **Kaser EK**, **Perkins SE**, **Brown I**, **Holmes EC**, **Taubenberger JK**. 2009. Different evolutionary trajectories of European avian-like and classical swine H1N1 influenza A viruses. J. Virol. **83**:5485–5494.
7. **Bhatt S**, **Lam TT**, **Lycett SJ**, **Leigh Brown AJ**, **Bowden TA**, **Holmes EC**, **Guan Y**, **Wood JLN**, **Brown I**, **Kellam P**, **Combating Swine Influenza Consortium**, **Pybus O**. 2013. The evolutionary dynamics of influenza A virus adaptation to mammalian hosts. Philos. Trans. R. Soc. Lond., B, Biol. Sci. **368**:20120382.
8. **Pensaert M**, **Ottis K**, **Vandeputte J**, **Kaplan MM**, **Bachmann PA**. 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**:75–78.
9. **Castrucci MR**, **Donatelli I**, **Sidoli L**, **Barigazzi G**, **Kawaoka Y**, **Webster RG**. 1993. Genetic Reassortment between Avian and Human Influenza A Viruses in Italian Pigs. Virology **193**:503–506.
10. **Schrader C**, **Süss J**. 2004. Molecular epidemiology of porcine H3N2 influenza A viruses isolated in Germany between 1982 and 2001. Intervirology **47**:72–77.
11. **Van Reeth K**, **Brown I**, **Dürrwald R**, **Foni E**, **Labarque G**, **Lenihan P**, **Maldonado J**, **Markowska Daniel I**, **Pensaert M**, **Pospisil Z**, **Koch G**. 2008. Seroprevalence of H1N1, H3N2 and H1N2 influenza viruses in pigs in seven European countries in 2002–2003. Influenza Other Respi Viruses **2**:99–105.
12. **Brown I**, **Harris PA**, **McCauley JW**, **Alexander DJ**. 1998. Multiple genetic reassortment of avian and human influenza A viruses in European pigs, resulting in the emergence of an H1N2 virus of novel genotype. J. Gen. Virol. **79 (Pt 12)**:2947–2955.
13. **Lam TT**, **Hon C-C**, **Wang Z**, **Hui RK-H**, **Zeng F**, **Leung FC-C**. 2008. Evolutionary analyses of European H1N2 swine influenza A virus by placing timestamps on the multiple reassortment events. Virus Research **131**:271–278.

14. **Smith GJD**, **Vijaykrishna D**, **Bahl J**, **Lycett SJ**, **Worobey M**, **Pybus O**, **Ma SK**, **Cheung CL**, **Raghwani J**, **Bhatt S**, **Peiris JSM**, **Guan Y**, **Rambaut A**. 2009. Origins and evolutionary genomics of the 2009 swine-origin H1N1 influenza A epidemic. Nature **459**:1122–1125.
15. **Neumann G**, **Noda T**, **Kawaoka Y**. 2009. Emergence and pandemic potential of swine-origin H1N1 influenza virus. Nature **459**:931–939.
16. **Guan Y**, **Poon LLM**, **Cheung CY**, **Ellis TM**, **Lim W**, **Lipatov AS**, **Chan KH**, **Sturm-Ramirez KM**, **Cheung CL**, **Leung YHC**, **Yuen KY**, **Webster RG**, **Peiris JSM**. 2004. H5N1 influenza: a protean pandemic threat. PNAS **101**:8156–8161.
17. **Breiman RF**, **Nasidi A**, **Katz MA**, **Njenga MK**, **Vertefeuille J**. 2007. Preparedness for Highly Pathogenic Avian Influenza Pandemic in Africa. Emerging Infect. Dis. **13**:1453–1458.
18. **Deng Y-M**, **Iannello P**, **Smith I**, **Watson J**, **Barr IG**, **Daniels P**, **Komadina N**, **Harrower B**, **Wong FYK**. 2012. Transmission of influenza A(H1N1) 2009 pandemic viruses in Australian swine. Influenza Other Respi Viruses **6**:e42–e47.
19. **Nelson MI**, **Gramer MR**, **Vincent AL**, **Holmes EC**. 2012. Global transmission of influenza viruses from humans to swine. Journal of General Virology **93**:2195–2203.
20. **Nelson MI**, **Wentworth DE**, **Culhane MR**, **Vincent AL**, **Viboud C**, **LaPointe MP**, **Lin X**, **Holmes EC**, **Detmer SE**. 2014. Introductions and evolution of human-origin seasonal influenza a viruses in multinational Swine populations. J. Virol. **88**:10110–10119.
21. **Liang H**, **Lam TT**, **Fan X**, **Chen X**, **Zeng Y**, **Zhou J**, **Duan L**, **Tse M**, **Chan C-H**, **Li L**, **Leung T-Y**, **Yip C-H**, **Cheung CL**, **Zhou B**, **Smith DK**, **Poon LL-M**, **Peiris M**, **Guan Y**, **Zhu H**. 2014. Expansion of genotypic diversity and establishment of 2009 H1N1 pandemic-origin internal genes in pigs in China. J. Virol. JVI.01327–14.
22. **Vijaykrishna D**, **Smith GJD**, **Pybus O**, **Zhu H**, **Bhatt S**, **Poon LLM**, **Riley S**, **Bahl J**, **Ma SK**, **Cheung CL**, **Perera RAPM**, **Chen H**, **Shortridge KF**, **Webby RJ**, **Webster RG**, **Guan Y**, **Peiris JSM**. 2011. Long-term evolution and transmission dynamics of swine influenza A virus. Nature **473**:519–522.
23. **Anderson TK**, **Nelson MI**, **Kitikoon P**, **Swenson SL**, **Korslund JA**, **Vincent AL**. 2013. Population dynamics of cocirculating swine influenza A viruses in the United States from 2009 to 2012. Influenza Other Respi Viruses **7 Suppl 4**:42–51.
24. **Nelson MI**, **Detmer SE**, **Wentworth DE**, **Tan Y**, **Schwartzbard A**, **Halpin RA**, **Stockwell T**, **Lin X**, **Vincent A**, **Gramer MR**, **Holmes EC**. 2012. Genomic reassortment of influenza A virus in North American swine, 1998-2011. Journal of General Virology.
25. **Simon G**, **Larsen LE**, **Dürrwald R**, **Foni E**, **Harder TC**, **Van Reeth K**, **Markowska Daniel I**, **Reid SM**, **Dan A**, **Maldonado J**, **Huovilainen A**, **Billinis C**, **Davidson I**, **Agüero M**, **Vila T**, **Hervé S**, **Breum SØ**, **Chiapponi C**, **Urbaniak K**, **Kyriakis CS**, **ESNIP3 consortium**, **Brown I**, **Loeffen W**. 2014. European surveillance network for influenza in pigs: surveillance programs, diagnostic tools and Swine influenza virus subtypes identified in 14 European countries from 2010 to 2013. PLoS ONE **9**:e115815.

26. **Bin Zhou**, **Donnelly ME**, **Scholes DT**, **George KS**, **Hatta M**, **Kawaoka Y**, **Wentworth DE**. 2009. Single-Reaction Genomic Amplification Accelerates Sequencing and Vaccine Production for Classical and Swine Origin Human Influenza A Viruses.
27. **Baillie GJ**, **Galiano M**, **Agapow PM**, **Myers R**, **Chiam R**, **Gall A**, **Palser AL**, **Watson SJ**, **Hedge J**, **Underwood A**. 2012. Evolutionary dynamics of local pandemic H1N1/2009 influenza virus lineages revealed by whole-genome analysis. J. Virol. **86**:11–18.
28. **Quail MA**, **Kozarewa I**, **Smith F**, **Scally A**, **Stephens PJ**, **Durbin R**, **Swerdlow H**, **Turner DJ**. 2008. A large genome center's improvements to the Illumina sequencing system. Nat Meth **5**:1005–1010.
29. **Watson SJ**, **Welkers MRA**, **Depledge DP**, **Coulter E**, **Breuer JM**, **de Jong MD**, **Kellam P**. 2013. Viral population analysis and minority-variant detection using short read next-generation sequencing. Philos. Trans. R. Soc. Lond., B, Biol. Sci. **368**:20120205.
30. **Hunt M**, **Gall A**, **Ong SH**, **Brener J**, **Ferns B**, **Goulder P**, **Nastouli E**, **Keane JA**, **Kellam P**, **Otto TD**. 2015. IVA: accurate de novo assembly of RNA virus genomes. Bioinformatics btv120.
31. **Li H**, **Handsaker B**, **Wysoker A**, **Fennell T**, **Ruan J**, **Homer N**, **Marth G**, **Abecasis G**, **Durbin R**. 2009. The Sequence Alignment/Map format and SAMtools. Bioinformatics **25**:2078–2079.
32. **Bao Y**, **Bolotov P**, **Dernovoy D**, **Kiryutin B**, **Zaslavsky L**, **Tatusova T**, **Ostell J**, **Lipman D**. 2008. The Influenza Virus Resource at the National Center for Biotechnology Information.
33. **Edgar RC**. 2004. MUSCLE: multiple sequence alignment with high accuracy and high throughput. Nucleic Acids Res. **32**:1792–1797.
34. **Tamura K**, **Stecher G**, **Peterson D**, **Filipski A**, **Kumar S**. 2013. MEGA6: Molecular Evolutionary Genetics Analysis version 6.0. Molecular Biology and Evolution **30**:2725–2729.
35. **Stamatakis A**, **Ludwig T**, **Meier H**. 2005. RAxML-III: a fast program for maximum likelihood-based inference of large phylogenetic trees. Bioinformatics **21**:456–463.
36. **Huelsenbeck JP**, **Ronquist F**. 2001. MRBAYES: Bayesian inference of phylogenetic trees. Bioinformatics **17**:754–755.
37. **Ronquist F**, **Huelsenbeck JP**. 2003. MrBayes 3: Bayesian phylogenetic inference under mixed models. Bioinformatics **19**:1572–1574.
38. **Drummond A**, **Rambaut A**. 2007. BEAST: Bayesian evolutionary analysis by sampling trees. BMC Evol. Biol.
39. **Drummond AJ**, **Suchard MA**, **Xie D**, **Rambaut A**. 2012. Bayesian Phylogenetics with BEAUti and the BEAST 1.7. Molecular Biology and Evolution **29**:1969–1973.
40. **Minin VN**, **Bloomquist EW**, **Suchard MA**. 2008. Smooth Skyride through a Rough Skyline: Bayesian Coalescent-Based Inference of Population Dynamics. Molecular Biology and Evolution **25**:1459–1471.
41. **Parker J**, **Rambaut A**, **Pybus O**. 2008. Correlating viral phenotypes with phylogeny: Accounting for phylogenetic uncertainty. Infection, Genetics and

Evolution **8**:239–246.

42. **Moreno A**, **Gabanelli E**, **Sozzi E**, **Lelli D**, **Chiapponi C**, **Ciccozzi M**, **Zehender G**, **Cordioli P**. 2013. Different evolutionary trends of swine H1N2 influenza viruses in Italy compared to European viruses. Vet. Res. **44**:112.
43. **Welsh MD**, **Baird PM**, **Guelbenzu-Gonzalo MP**, **Hanna A**, **Reid SM**, **Essen S**, **Russell C**, **Thomas S**, **Barrass L**, **McNeilly F**, **McKillen J**, **Todd D**, **Harkin V**, **McDowell S**, **CHOUDHURY B**, **Irvine RM**, **Borobia J**, **Grant J**, **Brown I**. 2010. Initial incursion of pandemic (H1N1) 2009 influenza A virus into European pigs. Veterinary Record **166**:642–645.
44. **Wong KK**, **Greenbaum A**, **Moll ME**, **Lando J**, **Moore EL**, **Ganatra R**, **Biggerstaff M**, **Lam E**, **Smith EE**, **Storms AD**, **Miller JR**, **Dato V**, **Nalluswami K**, **Nambiar A**, **Silvestri SA**, **Lute JR**, **Ostroff S**, **Hancock K**, **Branch A**, **Trock SC**, **Klimov A**, **Shu B**, **Brammer L**, **Epperson S**, **Finelli L**, **Jhung MA**. 2012. Outbreak of influenza A (H3N2) variant virus infection among attendees of an agricultural fair, Pennsylvania, USA, 2011. Emerging Infect. Dis. **18**:1937–1944.
45. **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.
46. **Vijaykrishna D**, **Holmes EC**, **Joseph U**, **Fourment M**, **Su YCF**, **Halpin R**, **Lee RTC**, **Deng Y-M**, **Gunalan V**, **Lin X**, **Stockwell TB**, **Fedorova NB**, **Bin Zhou**, **Spirason N**, **Kühnert D**, **Boskova V**, **Stadler T**, **Costa A-M**, **Dwyer DE**, **Huang QS**, **Jennings LC**, **Rawlinson W**, **Sullivan SG**, **Hurt AC**, **Maurer-Stroh S**, **Wentworth DE**, **Smith GJD**, **Barr I**, **Neher RA**. 2015. The contrasting phylodynamics of human influenza B viruses. Elife **4**:e05055.
47. **Bedford T**, **Cobey S**, **Pascual M**. 2011. Strength and tempo of selection revealed in viral gene genealogies. BMC Evol. Biol. **11**:220.
48. **Maes D**, **Deluyker H**, **Verdonck M**, **Castryck F**, **Miry C**, **Vrijens B**, **de Kruif A**. 2000. Herd factors associated with the seroprevalences of four major respiratory pathogens in slaughter pigs from farrow-to-finish pig herds. Vet. Res. **31**:313–327.
49. **Starick E**, **Lange E**, **Fereidouni S**, **Bunzenthal C**, **Höveler R**, **Kuczka A**, **Grosse Beilage E**, **Hamann H-P**, **Klingelhöfer I**, **Steinhauer D**, **Vahlenkamp T**, **Beer M**, **Harder TC**. 2011. Reassorted pandemic (H1N1) 2009 influenza A virus discovered from pigs in Germany. Journal of General Virology **92**:1184–1188.
50. **Lange J**, **Groth M**, **Schlegel M**, **Krumbholz A**, **Wieczorek K**, **Ulrich R**, **Köppen S**, **Schulz K**, **Appl D**, **Selbitz H-J**, **Sauerbrei A**, **Platzer M**, **Zell R**, **Dürrwald R**. 2013. Reassortants of the pandemic (H1N1) 2009 virus and establishment of a novel porcine H1N2 influenza virus, lineage in Germany. Veterinary Microbiology **167**:345–356.
51. **Howard WA**, **Essen SC**, **Strugnell BW**, **Russell C**, **Barrass L**, **Reid SM**, **Brown I**. 2011. Reassortant pandemic (H1N1) 2009 virus in pigs, United Kingdom. Emerging infectious ….
52. **Harder TC**, **Grosse Beilage E**, **Lange E**, **Meiners C**, **Döhring S**, **Pesch S**, **Noé T**, **Grund C**, **Beer M**, **Starick E**. 2013. Expanded cocirculation of stable subtypes, emerging lineages, and new sporadic reassortants of porcine influenza viruses in

swine populations in Northwest Germany. J. Virol. **87**:10460–10476.

53. **Mastin A**, **Alarcon P**, **Pfeiffer D**, **Wood J**, **Williamson S**, **Brown I**, **COSI Consortium**, **Wieland B**. 2011. Prevalence and risk factors for swine influenza virus infection in the English pig population. PLoS Curr **3**:RRN1209.

54. **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. Journal of General Virology **94**:1236–1241.

55. **Lindstrom S**, **Garten R**, **Balish A**, **Shu B**, **Emery S**, **Berman L**, **Barnes N**, **Sleeman K**, **Gubareva L**, **Villanueva J**, **Klimov A**. 2012. Human Infections with Novel Reassortant Influenza A(H3N2)v Viruses, United States, 2011. Emerg. Infect. Dis. **18**:834–837.

TABLES

Table 1. Statistical support for the association of host species with ancestry for the HA-H1, NA-N1, and PB2 segments of the A(H1N1)pdm09 lineage

Segment	Statistic	Observed			Null			Significance
		Mean	Lower 95% CI	Upper 95% CI	Mean	Lower 95% CI	Upper 95% CI	
HA-H1	AI	6.369	5.287	7.432	9.022	8.019	10.084	0.039
	PS	35.842	34	37	40.909	39.477	41.750	<0.01
NA-N1	AI	4.605	3.561	5.714	8.134	7.264	9.017	<0.01
	PS	32.886	31	35	45.541	44.012	46.573	<0.01
PB2	AI	4.883	3.987	5.804	6.925	5.951	7.792	<0.01
	PS	28.680	28	30	38.070	36.591	38.858	<0.01

AI: association index

PS: parsimony score

CI: Bayesian credible interval

FIGURE LEGENDS

Figure 1. swIAV genotypes isolated from European swine between 2009 and 2013. The 23 distinct genetic constellations are labeled A to W, with the lineage-of-origin for each segment indicated by a colored block.

Figure 2. Frequency of the swIAV genotypes across Europe. **Main panel** Pie charts indicate the proportion of samples isolated in each country that are either EA H1N1 (dark green), A(H1N1)pdm09 (dark red), or contain either an EA IGC (light green) or a pdm09 IGC (light red). Isolates whose IGC contains both EA and pdm09 segments are colored grey. Size of the pie chart reflects the number of samples used for analysis from that country, namely Belgium=24, Denmark=17, Finland=2, France=47, Germany=38, Hungary=6, Italy=42, the Netherlands=30, Norway=1, Poland=11, Spain=28, Sweden=2, UK=41. For clarity, Israel (n=1; pdm09) is not shown. IGC = internal gene cassette. **Panels A-R** Relative frequency of the six most prevalent swIAV genotypes in countries across Europe. Each genotype is specified as given in Figure 1. The intensity of the color in each panel reflects the relative frequency of the genotype in that country. Countries that provided no samples are not outlined. Numbers in paratheses indicate the number of countries the genotype was isolated from.

Figure 3. Comparison of swIAV genotypes isolated in **A)** the UK and **B)** mainland Europe. Bar charts are colored according to their genotype, with genotypes specified as given in

778 Figure 1. The six most prevalent genotypes are shown separately, with the remaining
779 genotypes clustered according to the lineage of their internal gene cassette.
780
781 **Figure 4.** Bayesian-inferred phylogeny of the Scot/94 lineage N2 gene. Taxa sequenced
782 through the ESNIP3 consortium are highlighted in red, while those in black were
783 obtained from the Influenza Virus Resource. Colored squares to the right of each taxa
784 indicate its genotype, with coloring and segment order as in Figure 1. White squares
785 indicate that no sequence was available for that segment. Posterior probabilities are
786 given at selected nodes. Colored highlights indicate well-supported circulating clades.
787 The scale bar is given in substitutions per site.
788
789 **Figure 5.** Bayesian-inferred phylogeny of the Gent/84 lineage N2 gene. Posterior
790 probabilities are given at selected nodes, and the scale bar is given in substitutions per
791 site. See Figure 4 legend for other details.

	Internal segments							External segments		Isolates analysed	
	PB2	PB1	PA	NP	MP	NS		HA	NA	Count	Percentage
A										85	29
B										38	13
C										26	9
D										13	5
E										11	4
F										8	3
G										5	2
H										3	1
I										2	<1
J										2	<1
K										2	<1
L										1	<1
M										1	<1
N										1	<1
O										1	<1
P										35	12
Q										22	8
R										13	5
S										3	1
T										2	<1
U										2	<1
V										1	<1
W										1	<1
Undetermined										12	4
Total samples										290	100

A/swine/Gent/1/1984-like H3N2

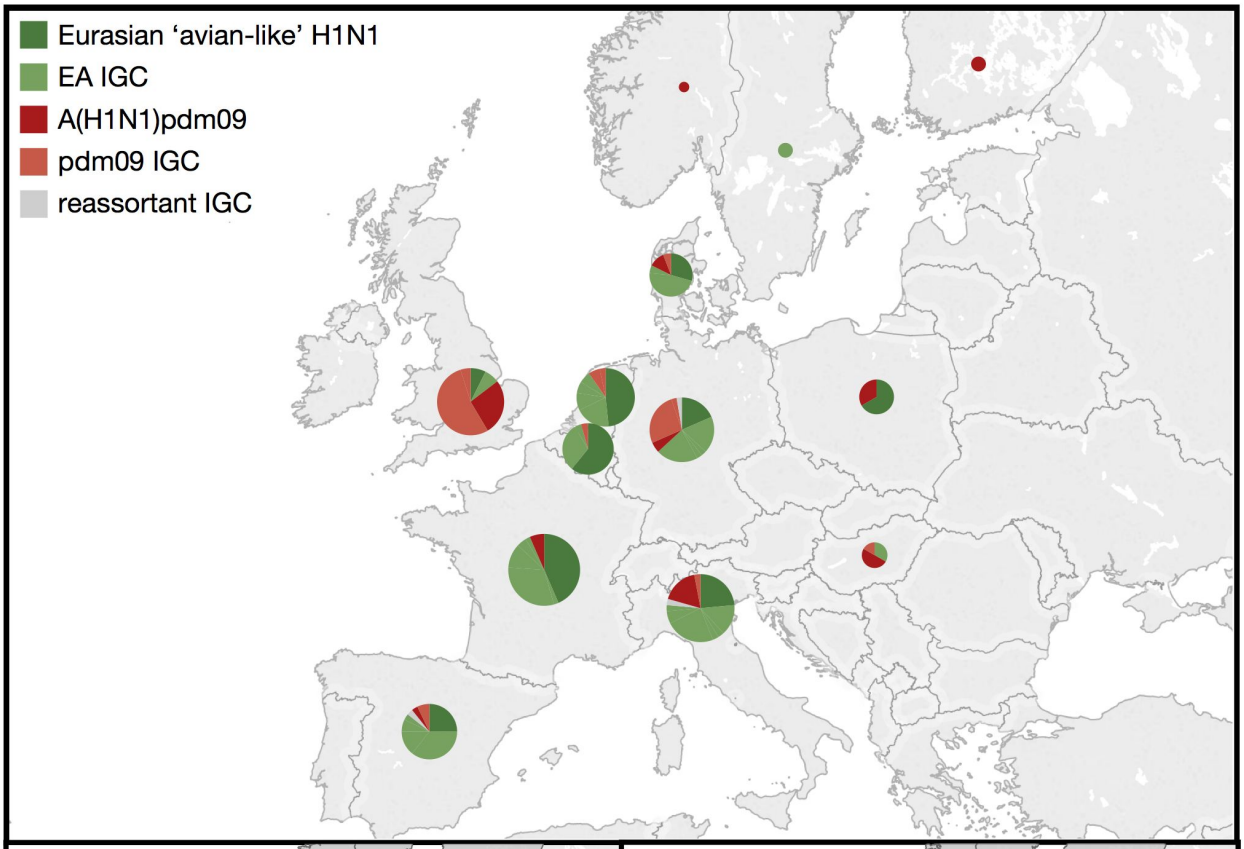
A/swine/Scotland/410440/1994-like $H1_{hu}N2$

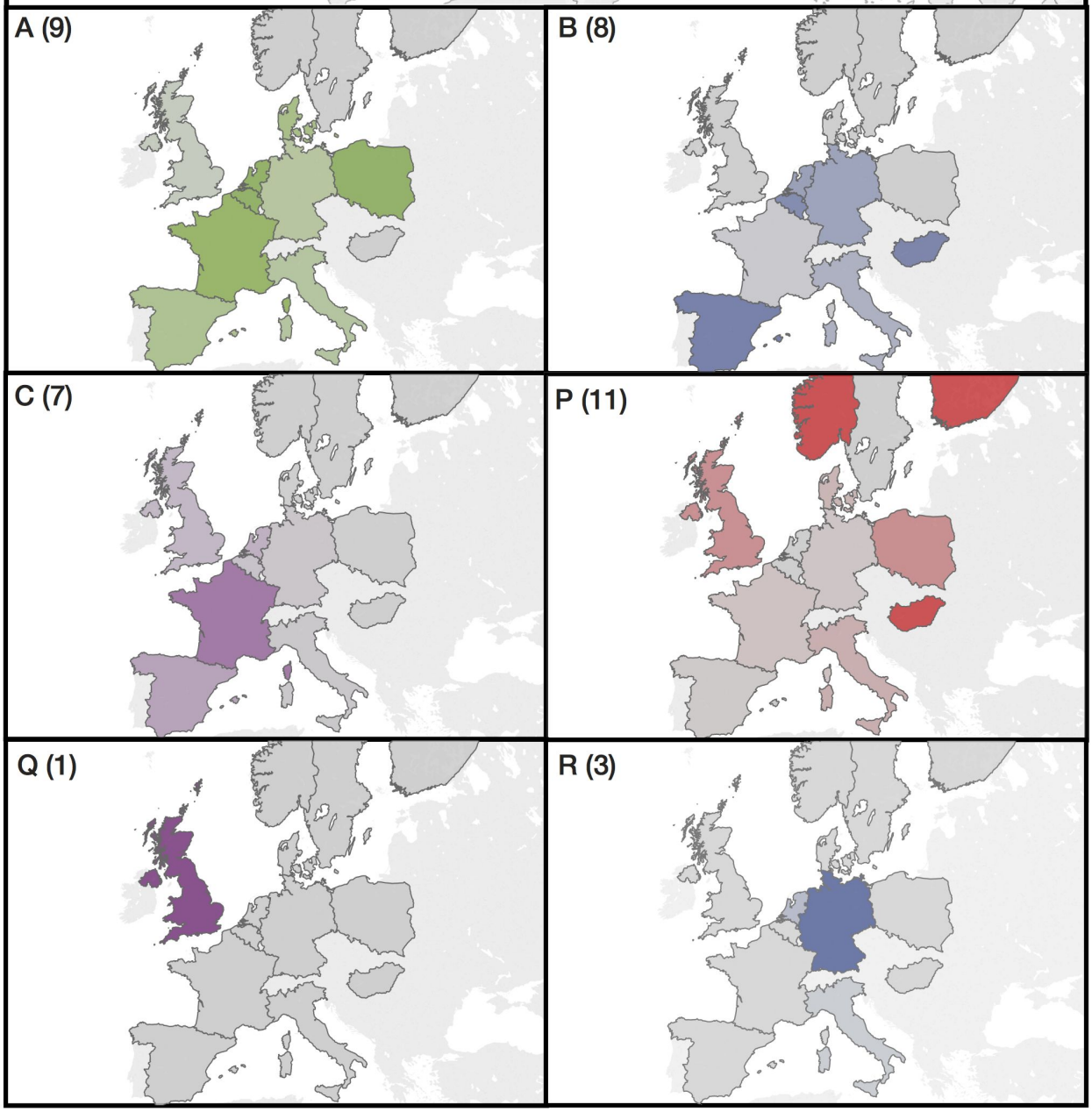
Eurasian avian-like $H1_{av}N1$

A(H1N1)pdm09

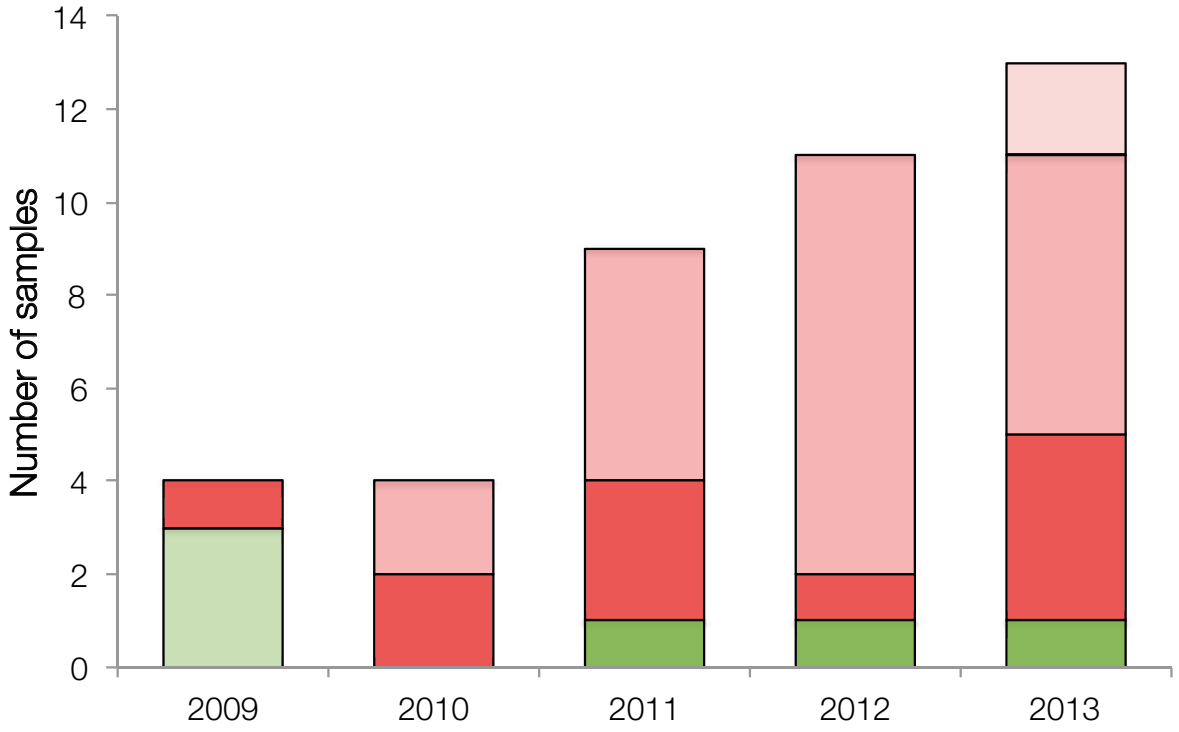
A/swine/Italy/4675/2003-like N2

Human seasonal-like N2





A) United Kingdom



B) Mainland Europe