Origin of the European avian-like swine influenza viruses

Andi Krumbholz,1* Jeannette Lange,1 Andreas Sauerbrei,1 Marco Groth,2 Matthias Platzer,2 Pumaree Kanrai,3 Stephan Pleschka,3 Christoph Scholtissek,3 Mathias Büttner,4 Ralf Dürwald5 and Roland Zell1

1Institut für Virologie und Antivirale Therapie, Universitätsklinikum Jena, Friedrich-Schiller-Universität Jena, Hans-Knöll-Str. 2, D-07745 Jena, Germany
2Genomanalyse, Leibniz-Institut für Altersforschung – Fritz-Lipmann-Institut, Beutenbergstrasse 11, D-07745 Jena, Germany
3Institut für Medizinische Virologie, Justus-Liebig-Universität, Schubertstrasse 81, D-35392 Giessen, Germany
4Bayerisches Landesamt für Gesundheit und Lebensmittelsicherheit, Veterinärstrasse 2, D-85762 Oberschleissheim, Germany
5Virale Vakzinen, Geschäftsbereich Tiergesundheit, IDT Biologika GmbH, Am Pharmapark, D-06861 Dessau-Rosslau, Germany

The avian-like swine influenza viruses emerged in 1979 in Belgium and Germany. Thereafter, they spread through many European swine-producing countries, replaced the circulating classical swine H1N1 influenza viruses, and became endemic. Serological and subsequent molecular data indicated an avian source, but details remained obscure due to a lack of relevant avian influenza virus sequence data. Here, the origin of the European avian-like swine influenza viruses was analysed using a collection of 16 European swine H1N1 influenza viruses sampled in 1979–1981 in Germany, the Netherlands, Belgium, Italy and France, as well as several contemporaneous avian influenza viruses of various serotypes. The phylogenetic trees suggested a triple reassortant with a unique genotype constellation. Time-resolved maximum clade credibility trees indicated times to the most recent common ancestors of 34–46 years (before 2008) depending on the RNA segment and the method of tree inference.

Influenza A virus is the single species of the genus Influenzavirus A in the family Orthomyxoviridae (McCauley et al., 2012). The influenza A virus (FLUAV) genome is segmented and has negative-strand polarity. The virion is composed of: (i) an envelope with three integral viral membrane proteins named haemagglutinin (HA), neuraminidase (NA) and M2 proton channel; (ii) the matrix protein (M1), which lines the inner membrane surface; and (iii) eight ribonucleoprotein complexes each comprising one RNA segment, nucleoproteins (NPs) and the RNA-dependent RNA polymerase complex (composed of PB1, PB2 and PA). Two non-structural proteins, NS1 and NS2/NEP (nuclear export protein), are expressed in all FLUAV-infected cells. Depending on the respective virus strains, other viral proteins such as PB1-F2, PB1-N40 and PA-X may also be expressed (Krumbholz et al., 2011; Jagger et al., 2012; McCauley et al., 2012).

Three factors influence the complex influenza virus ecology. First, FLUAVs circulate in a variety of hosts. The main hosts are aquatic birds, but poultry, passerine birds and mammals – notably humans, pigs and horses – may also be infected productively (Webster et al., 1992). Further influenza-like viruses have been detected recently in bats (Wu et al., 2014). Secondly, diverse types or alleles of the HA, NA and NS segments co-circulate. By serological means, 16 HA and nine NA types have been defined in aquatic birds that can recombine to form up to 144 FLUAV subtypes (McCauley et al., 2012). Thirdly, several distinct genetic lineages of each genome segment have been observed (Lu et al., 2007). Altogether, more than 180 lineages were found to recombine to form at least 110 FLUAV subtypes (HA/NA combinations) and more than 500 genotypes. Characteristic FLUAV lineages of the eastern (Eurasia) and western (America) hemispheres developed in birds in an allopatric speciation-like mechanism when geographical

*Present address: Institut für Infektionsmedizin, Christian-Albrechts-Universität zu Kiel, and Universitätsklinikum Schleswig-Holstein, Brunsbüttler Strasse 4, D-24105 Kiel, Germany.

Ten supplementary figures and two tables are available with the online Supplementary Material.
isolation of bird populations and continuous accumulation of nucleotide substitutions over time concur (Webster et al., 1992; Olsen et al., 2006). In addition, for some segments, host adaptation leads to specific lineages of geese/ducks/swans (order Anseriformes) and gulls/terns/waders (order Charadriiformes) but also in pigs, horses and humans (Webster et al., 1992). The diversity of more than 500 genotypes is generated and maintained by several mechanisms including: the (i) co-circulation of many subtypes in a host population; (ii) annual variation of circulating viruses in the bird population when birds migrate long distances from their breeding sites to non-breeding sites; and (iii) reassortments in the individual host upon double or multiple infections (Olsen et al., 2006).

Influenza viruses have established several times independent and stable infections chains in pigs. In North America, swine influenza emerged contemporaneously to the ongoing pandemic in 1918 and gave rise to the classical H1N1 swine influenza viruses (cH1N1 swFLUAV) (Schultz-Cherry et al., 2013). This lineage circulated for decades, spread to Europe and Asia, and finally was largely replaced by swFLUAVs of various serotypes but with a characteristic triple reassortant internal gene (TRIG) cassette that first emerged in 1998 in the USA (Zhou et al., 1999). In Europe, a persistent avian-like H1N1 (avH1N1) lineage was established in 1979. The respective virus had eight segments of a Eurasian avian H1N1 virus (Scholtissek et al., 1983; Castrucci et al., 1993; Campitelli et al., 1997). Later, it recombined twice with seasonal H3N2 and H1N1 viruses to yield the so-called ‘human-like’ H3N2 and H1N2 (huH3N2 and huH1N2, respectively) swFLUAVs with the internal genes of avH1N1 (Castrucci et al., 1993; Campitelli et al., 1997; Brown et al., 1998; reviewed by Zell et al., 2013). Swine influenza activity is more complex in Asia. Despite a lack of systematic surveillance for many years, available data indicate the presence of cH1N1, North American triple reassortants and European swFLUAVs, as well as a plethora of reassortant viruses (Zhu et al., 2013). In addition to the genuine swine influenza viruses, numerous reassortant viruses of the pandemic 2009 H1N1 (pdmH1N1) and swine influenza viruses emerged in America, Asia and Europe. Most of these were unstable, but the Papenburg lineage in Germany and the H3N2v variant in the USA persisted (Nelson et al., 2012; Jhung et al., 2013; Lange et al., 2013). The latter variant exhibited a significant zoonotic potential (http://www.cdc.gov/flu/swineflu/h3n2v-case-count.htm).

It was observed previously that avian influenza viruses of a particular subtype may differ in their genetic make-up (Olsen et al., 2006). One of six different avian H1N1 genotypes that had been detected in Europe in the past 40 years was able to induce a stable infection chain in pigs. As this virus, the European avH1N1 swFLUAV, displays a unique genotype constellation, we aimed to identify its ancestors by genetic analysis of early avH1N1 isolates as well as European avian FLUAVs of the 1960s, 1970s and 1980s. The data suggested that avH1N1 is a triple reassortant that emerged in the late 1960s but remained unperceived until 1979. According to Lu et al. (2007), the avH1N1 swFLUAV has an F-G-I-1C-F-1F-F-1E genotype that is unique in a compilation of more than 500 genotypes of the FluGenome database (http://www.flugene.org, accessed 30 March 2014). This database registers more than 7135 complete influenza virus genomes. In order to exclude early reassortment events soon after the emergence of the avH1N1, 15 swFLUAV isolates from Germany, the Netherlands, Italy and France (collected in 1979–1981) were sequenced along with five avian isolates of that time and a porcine H3N6 isolate from Kazakhstan collected in 1985 (Table S1, available in the online Supplementary Material). Together with sequences retrieved from GenBank, 10 datasets were compiled that included European swFLUAVs, Eurasian avian FLUAVs and reference sequences representing all known lineages of each viral genome segment. Depending on the availability of sequence data, these datasets comprised 63–68 sequences for the internal gene segments and 32 HAH1, 35 HAH3, 42 NAN1 and 35 NAN6 sequences.

Viruses were amplified in Madin–Darby canine kidney cells or in embryonated chicken eggs. The methods for virus preparation, RNA preparation, cDNA synthesis, Sanger sequencing and Illumina sequencing (including downstream analysis of Illumina data) have been described previously (Zell et al., 2008; Lange et al., 2013, 2014). Nucleotide sequences were aligned with MEGA6 (Tamura et al., 2011). For the molecular clock analysis, the coalescent Bayesian Metropolis-coupled Markov chains (MCMC) method was applied, which is implemented in the BEAST 1.8 package. BEAST 1.8 includes the programs BEAUTi, BEAST, Tracer, LogCombiner, TreeAnnotator and FigTree (Drummond et al., 2012). At least 75 million iterations were run and every 5000 states were sampled. If the effective sample size (determined with Tracer) was >200, the calculations were rerun and the chains were combined using the LogCombiner program. We used the collection year as temporal data and either a constant population size model or the Bayesian skyline model as a tree prior. The times to the most recent common ancestors (tMRCA) were calculated using the tree prior assuming either a strict clock or an uncorrelated log-normal relaxed clock. Two consecutive runs each were performed to assure convergence of MCMC chains. The HAH3 and NAN6 trees were inferred with MrBayes 3.0 (Ronquist & Huelsenbeck, 2003). Here, four MCMC chains were calculated until convergence was reached. Optimal substitution models were selected on the basis of the BIC criterion and AICc criterion with the help of the ‘find model’ option implemented in MEGA6. Twenty-five per cent of the posterior set of trees were removed as burn-in; the remaining trees were used for the construction of maximum clade credibility (MCC) trees employing TreeAnnotator. These were displayed with FigTree.

Time-resolved phylogenetic trees inferred with the constant population size model and uncorrelated log-normal relaxed clock are presented in Figs S1–S8. All trees demonstrated inconsistent clustering of avH1N1 swFLUAV with various avian FLUAV strains of that time. For the
polymerase subunits PB2 and PB1, A/duck/Germany/1215/1973 (H2N3) was the closest relative, whereas the PA and NS segments clustered with A/duck/Rugen/78-6/1981 (H2N3). The remaining gene segments were closely related to avian H1 isolates (HA, NA, M: A/duck/Schleswig/21/1979; NP: A/mallard/Stralsund/40-6/1981). The topologies of the other trees (strict clock/constant population size model; uncorrelated log-normal relaxed clock/Bayesian skyline; strict clock/Bayesian skyline) were essentially identical (data not shown). The data suggested that avH1N1 swFLUAVs are triple reassortants comprising segments 1 and 2 of an A/duck/Germany/1215/1973-like ancestor, segments 3 and 8 of an A/duck/Rugen/78-6/1981-like ancestor and segments 4–7 of an A/duck/Schleswig/21/1979-like ancestor (Fig. 1).

Other parental viruses cannot be excluded on the basis of available sequence data, but each alternative scenario required at least three parental viruses and two reassortment events. As the phylogenetic trees also demonstrated monophyly of all avH1N1 swFLUAV strains, further reassortment events could be excluded once the avH1N1 genotype was established. Using BEAST 1.8, $t_{	ext{MRCA}}$ was estimated as 34–46 years (before 2008) and corresponded to the years 1962–1974 (Table S2). This indicated that the avH1N1 swFLUAVs may have circulated many years prior to their emergence in pigs in 1979.

It was striking that most of the ancestral viruses of the avH1N1 swFLUAVs were of German provenance. However, as many anatids are migratory birds, one cannot exclude the possibility that the crucial reassortments occurred elsewhere. Hence, this observation simply reflects the brisk virus collection activity in Germany at that time and is in line with the first reports on swH1N1-induced influenza of pigs: Pensaert et al. (1981) described outbreaks of swine influenza that started in January 1979 in Belgium, whereas the first known German outbreaks were recorded in the winter of 1979/1980 with the first German avH1N1 isolate in December 1979 (Witte et al., 1981). Nevertheless, the first European H1N1 isolates of birds were collected in Bavaria, Germany, in 1977 (Ottis & Bachmann, 1980), and our time-resolved phylogenetic trees indicated a long circulation time.

![Fig. 1. Genetic composition of avian-like H1N1 swine influenza viruses (represented by A/swine/Arnsberg/1/1979 (H1N1) and nine avian influenza viruses collected in Europe in the 1970s and early 1980s. FLUAV genome segments are indicated by boxes. The genetic lineages according to Lu et al. (2007) are given. Asterisks denote genome segments with closest similarity to swFLUAVs. The assumed origin of the gene segments of A/swine/Arnsberg/1/1979 is indicated by hatched and grey and black shaded boxes.](image-url)
of the novel genotype. It was generally believed that the avian H1N1 viruses had the same genetic composition as the avH1N1 swFLUAVs isolated from 1979 onwards (reviewed by Zell et al., 2013). The sequence data presented here clearly indicate that this is not the case. Neither the lineages of the PA, NA and NS segments of A/duck/Bavaria/2/1977, nor those of the polymerase subunits of A/duck/Schleswig/21/1979, A/mallard/Stralsund/40-6/1981 or A/mallard/Marquenterre/Z237/1983 matched the avH1N1 swFLUAV genotype, despite their H1N1 subtype (Fig. 1). Therefore, two consecutive reassortment events in birds could be a parsimonious explanation for this unique genotype constellation. Due to the abundance of avian FLUAV genotypes of that time, it is conceivable that these reassortments occurred in birds, but direct proof is lacking for two reasons: (i) currently, additional sequence data for avian FLUAVs are not available; and (ii) it is unknown if all relevant virus isolates that could solve this particular question were collected at that time or survived so many years in freezers.

The isolate A/swine/Kazakhstan/106/1985 (H3N6) exhibited another unique genetic setting (genotype: G-G-E-3B-F-6C-F-2A). The lineages of the internal segments PB2, PB1, PA, NP, M and NS corresponded to the European avian viruses of the late 1970s and early 1980s including H1N1 (Figs S1–S3, S5, S7 and S8). However, the distribution of the HA 3B and NA 6C genes in Europe is not well documented for the 1980s (Figs S9 and S10). There are only two older HA sequences available (A/duck/Ukraine/1963 and A/turkey/England/1969) and one contemporary NA sequence (A/duck/Potsdam/2216-4/1984). Hence, this unusual genetic make-up indicates a spillover infection of a pig with a rare avian influenza virus. Details on clinical signs of the affected pigs and the place of sampling are no longer available.

The phylogenetic trees also affirmed the mechanism of perpetuation of the high avian FLUAV diversity. The trees indicated co-circulation of each two PB2, PA, NANA and NS lineages in Europe in the 1970s and 1980s. Whereas PB2 lineages G and K have continued to co-circulate for 60 years (Fig. S1), the co-circulation of the PA lineages E and I lasted 20 years only. Thus, lineage I may have ceased in the bird population in the late 1980s as it was detected thereafter in swFLUAVs only. Contemporaneous to the disappearance of lineage I in birds, the PA lineage D emerged in Germany (Fig. S3). The genetic diversity of the NA gene is even more complex. In the 1970s and early 1980s, lineages 1E and 1F were present in Europe. Both lineages then disappeared in European birds, and three lineages, 1H, 1I and 1K, surged. Other NA lineages emerged in Asia (1G, 1J and 1L). However, the 1F lineage survived in the European swFLUAVs and in pdmH1N1, whereas the 1E lineage persisted in North American birds (Fig. S6). Finally, co-circulation of the NS lineages 1E and 2A has now been observed in Europe for more than 30 years.

The PB2 F lineage of the avH1N1 swFLUAVs evolved from the K lineage (Fig. S1). We hypothesize that modern practices of animal husbandry may have contributed to the evolution of a novel FLUAV lineage: it is conceivable that the initial cross-species infection occurred in a single pig farm followed by virus spread from herd to herd. The isolated lodging of farm animals prevents them from freely exchanging influenza viruses with the main virus population in aquatic birds. Therefore, evolution of the PB2 F lineage resembles peripatric speciation, which is defined by the isolation of a small population in a new environment. As the isolated population was numerically much smaller than the parental virus population, replication bottlenecks (when virus is transmitted from one pig herd to another) and the founder effect (Mayr, 1954; Templeton, 1980) may have allowed the establishment of rapid genetic changes. These led to the assignment of a new lineage, which is defined by a cut-off of 10% nucleotide difference by p-distance (http://www.flugenome.org).

Humans and horses usually host few FLUAV serotypes with rather stable genotypes that circulate for years and decades, whereas avian FLUAVs exhibit a great genetic diversity with numerous but ephemeral genotypes. The concept of pigs as the ideal mixing vessel for reassortment of avian and human influenza viruses (Scholtissek et al., 1985) received some recent limitations with regard to the proposed mechanism. Several studies indicate that sialic acids with α(2,3)- and α(2,6)-linked galactose residues that serve as receptors for human and avian influenza viruses, respectively, exhibit similar distribution in the respiratory tracts of pigs and humans (Nelli et al., 2010; Van Poucke et al., 2010; Trebbien et al., 2011). Hence, one has to assume additional factors rather than receptor distribution to explain the great number of genetic reassortments observed in pigs. Pig-farming practices and the contact of pigs with avian faecal material that is contaminated with influenza viruses provide possible means of cross-species infections, e.g. A/swine/Kazakhstan/106/1985, but direct proof of the ‘mixing vessel’ hypothesis is still lacking. Therefore, it is as yet unclear whether the avian-like swine H1N1 originated from reassortment events in pigs or birds. There was a dynamic influenza activity in the European avifauna, which was only rudimentally explored in the 1970s and early 1980s: at least 29 FLUAV subtypes (HA/NA combinations) co-circulated at that time. Sequence data for many of these are still either incomplete or not available. Despite this lack of sequence information, FLUAV genome sequencing activities of recent years have unravelled at least 36 European genotypes of that time (data not shown). Although available data allow only a sketchy picture of the genetic diversity of the European avian FLUAVs, a view arose that underlines the greater importance of reassortments in the evolution of avian and porcine FLUAVs, whereas genetic drift is the more dominant driver in the evolution of human FLUAVs.

Acknowledgements

We thank Drs R. Heckler (Niedersächsisches Landesgesundheitsamt Hannover), C. Linne-Jonas (Staatliches Veterinäruntersuchungsamt Arnsberg) and J. Süss (Bundesamt für Gesundheitlichen Verbraucherschutz Berlin) for the gift of virus strains and Martina
Müller and Ivonne Görlch for technical assistance. This study was supported by grants from the German Ministry for Education and Research (01KI1006L, 01KI1006E and 01KI1006J) awarded to R.D., S.P. and R.Z.

References


