Ancestral sequence reconstruction pinpoints adaptations that enable avian influenza virus transmission in pigs

Wen Su[1], **Rhodri Harfoot**[2,3], **Yvonne CF Su**[4], **Jennifer DeBeauchamp**[2], **Udayan Joseph**[4], **Jayanthi Jayakumar**[4], **Jeri-Carol Crumpton**[2], **Trushar Jeevan**[2], **Adam Rubrum**[2], **John Franks**[2], **Philippe Noriel Q. Pascua**[2], **Christina Kackos**[2], **Yuqin Zhang**[1], **Mengting Zhang**[1], **Yue Ji**[1], **Huyen Trang Bui**[1], **Jeremy C Jones**[2], **Lisa Kercher**[2], **Scott Krauss**[2], **Stephan Pleschka**[5], **Michael CW Chan**[1], **Robert G Webster**[2], **Chung-Yi Wu**[6], **Kristien Van Reeth**[7], **Malik Peiris**[1], **Richard J Webby**[2,*], **Gavin JD Smith**[4,*], **Hui-Ling Yen**[1,*]

[1]School of Public Health, Li Ka Shing Faculty of Medicine, The University of Hong Kong, Hong Kong SAR, China

[2]Department of Infectious Diseases, St. Jude Children's Research Hospital, Memphis, TN, USA

[3]Department of Microbiology and Immunology, University of Otago, Dunedin, New Zealand

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

[5]Institute of Medical Virology, Justus Liebig University Giessen, Germany

[6]Genomics Research Center, Academia Sinica, Taiwan, ROC

[7]Laboratory of Virology, Faculty of Veterinary Medicine, Department of Virology, Parasitology and Immunology, Ghent University, Merelbeke, Belgium

Abstract

Understanding the evolutionary adaptations that enable avian influenza viruses to transmit in mammalian hosts could allow better detection of zoonotic viruses with pandemic potential. We applied ancestral sequence reconstruction to gain viruses representing different adaptive stages of the European avian-like (EA) H1N1 swine influenza virus as it transitioned from avian to swine hosts since 1979. Ancestral viruses representing the avian-like precursor virus and EA swine viruses from 1979–1983, 1984–1987, and 1988–1992 were reconstructed and characterized. Glycan array analyses showed stepwise changes in the hemagglutinin receptor binding specificity from recognizing both alpha2,3- and alpha2,6-sialosides to alpha2,6-sialosides; however, efficient transmission in piglets was enabled by adaptive changes in the viral polymerase

Competing interests: The authors declare no competing interests.

protein and nucleoprotein that have been fixed after 1983. PB1-Q621R and NP-R351K increased viral replication and transmission in piglets when introduced into the 1979–1983 ancestral virus that lacked efficient transmissibility. The stepwise adaptation of an avian influenza virus to a mammalian host suggests that there may be opportunities to intervene and prevent interspecies jump through strategic coordination of surveillance and risk assessment activities.

Pandemic influenza viruses must acquire an ability for sustained human-to-human transmissibility, an ability aided by antigenic novelty and a resulting absence of human population immunity. While antigenic novelty can be readily expected for influenza viruses emerging from animal reservoirs, sustained transmissibility is a rare and poorly defined trait that needs to be acquired to foster optimal virus-host interactions[1–5]. The emergence of the A(H1N1)pdm09 pandemic influenza virus from pigs is a clear reminder of the feasibility of influenza viruses to cross mammalian species barriers and become established in humans[6]. In addition to the A(H1N1)pdm09 virus, only a limited number of avian-origin influenza viruses are known to have successfully crossed the species barriers and established in a new mammalian host. These viruses offer the best opportunities to identify the molecular determinants for successful establishment of avian influenza viruses in mammals.

The European avian-like (EA) swine H1N1 lineage viruses transitioned from avian to swine hosts in the late 1970s[7,8], with all eight gene segments derived from Eurasian avian influenza viruses[9,10]. The EA swine viruses are prevalent among pigs in China and in European countries[8,11–13], they donated the NA and M gene segments to the A(H1N1)pdm09 virus[6], and they have demonstrated pandemic potential, including resistance to human MxA[13,14], airborne transmissibility in ferrets[12,15], and antigenic novelty (eg. limited cross-reactivity with the antibodies that human population has developed against circulating influenza strains)[12,13,15]. Here, we identify and characterize the cumulative molecular changes present in naturally occurring EA swine viruses that were responsible for the avian-to-swine adaptation. By generating ancestral viruses at different points in the early evolution of EA swine viruses[16,17], we specifically investigated if the sustained swine transmissibility of EA swine viruses was an intrinsic trait of the precursor avian influenza virus or if it was acquired in pigs through subsequent adaptations.

RESULTS

EA swine H1N1 viruses demonstrated stepwise changes in receptor-binding properties.

A change in receptor binding specificity is a critical step for avian-to-mammalian interspecies transmission[18], we examined the receptor binding preference of eleven EA swine viruses isolated from 1979 to 2011 (Extended Data Fig. 1). Compared to the avian H1N1 influenza virus A/duck/Bavaria/2/1977 (DK/77), which preferentially recognizes α-2,3 sialosides (Fig. 1a), early EA swine viruses represented by A/swine/Germany/2/1981 (SW/81), showed dual binding specificity to both α-2,3 and α-2,6 sialosides (Fig. 1b). The EA swine viruses isolated after 1990, represented by A/swine/Schleswig-Holstein/1/1992 (SW/92), bound exclusively to α-2,6-linked sialosides (Fig. 1c). In agreement with previous results[19,20], EA swine viruses showed no apparent changes in their hemagglutinin fusion pH over time with values ranging from pH 5.2 to 5.8 (Supplementary Table 1).

The replication efficiency of avian (DK/77) and swine (SW/81 and SW/92) influenza viruses exhibiting differential binding for α-2,3 and α-2,6 sialosides (Fig. 1a,b,c) were evaluated in vitro. In the new-born pig trachea (NPTr) cells, SW/81 replicated to higher titers than DK/77 and SW/92 viruses at all time points (Two-way ANOVA and Tukey's post hoc test, $P < 0.01$)(Fig. 1d). In pig lung explants, SW/81 and SW/92 replicated to higher titers than DK/77 at 24 hours post-infection (hpi) (Two-way ANOVA, $P < 0.01$ and Tukey's post hoc test, $P < 0.05$) (Fig. 1e). No significant differences in replication were observed in the pig trachea explants (Fig. 1f).

EA swine H1N1 viruses demonstrated stepwise changes in contact transmission potential in pigs.

The contact transmission potential of DK/77, SW/81, and SW/92 viruses were further compared in 3–4 weeks old piglets. Inoculated donor pigs were co-housed with naïve contact pigs on 1 day post-inoculation (dpi) at a 2:2 donor-contact ratio, with a total of 4 donors and 4 direct contacts in duplicated experiments. Area under the curve (AUC) were calculated to approximate total viral load over the course of infection (Fig. 1g, h, i). In donors, the avian DK/77 virus replicated (mean ± SD AUC = 1.58 ± 1.59) to lower titers than the swine viruses SW/81 (3.81 ± 1.1) and SW/92 (5.8 ± 0.78), respectively (Kruskal–Wallis test, $P = 0.0031$, Dunn's post hoc test at $P = 0.61$ and 0.013, respectively). There was no significant difference in AUC between SW/81 and SW/92 inoculated donors (Dunn's post hoc test, $P = 0.35$). Correspondingly, the avian DK/77 virus failed to transmit to any contact (Fig. 1g), and none of the donor or the contact pigs seroconverted [hemagglutination inhibition (HI) titer < 1:20]. SW/81 virus replicated better than the DK/77 virus with infectious virus detected in the nasal cavity of 4/4 donors and seroconversion in 3/4 donors (HI titers at 1:40 to 1:80). However, SW/81 transmitted inefficiently to contact pigs as infectious virus was detected transiently in 2/4 contacts at later time points post-exposure (Fig. 1h), with seroconversion detected in 2/4 contacts (1:40). In contrast, robust replication and transmission of SW/92 virus was detected in all donors and contacts (Fig. 1i), with seroconversion detected in all donors (1:160 to 1:320) and contacts (1:80 to 1:320). These results suggest that the EA swine viruses may have gone through sequential adaptation during the avian-to-pig host transition by increasing replicative capacity in pig nasal tissues followed by developing efficient transmissibility among pigs.

Reconstructed ancestral EA swine H1N1 viruses possessed comparable phenotypes as the wild-type viruses.

To comprehensively map the molecular changes associated with EA swine virus adaptation in pigs, maximum likelihood phylogenies for the eight viral gene segments were constructed using EA swine viruses isolated from 1979 to 2014 and avian viruses isolated from 1949 to 2013 (Supplementary Fig. 1 and 2). Ancestral sequence reconstruction was used to infer sequences of four major nodes that represent different evolutionary stages of EA swine influenza viruses (Fig. 2a). Four ancestral viruses were generated using gene synthesis and plasmid-based reverse genetics. Their full genomes were deposited to the public database GISAID (accession numbers: EPI_ISL_539852-5). RG-EA1 virus was constructed based on the inferred nodal sequences at the split between avian and EA swine lineages (Fig. 2a; Node 1 in Supplementary Fig. 2). RG-EA2 virus was constructed to represent the early

evolutionary stage of EA swine viruses in 1979–1983 (Fig. 2a; Node 2 in Supplementary Fig. 2), while RG-EA3 and RG-EA4 viruses represented EA swine viruses in 1984–1987 and 1988–1992, respectively (Fig. 2a; Nodes 3 and 4 in Supplementary Fig. 2).

The receptor binding profiles of the RG-EA1, EA2, EA3, EA4 viruses were compared. The avian-like precursor RG-EA1 virus showed exclusive binding to α-2,3 sialosides (Fig. 2b). The RG-EA2 virus showed dual binding for α-2,3 and α-2,6-linked sialosides (Fig. 2c) that resembled early EA swine isolates (Fig. 1b and Extended Data Fig. 1). The RG-EA3 and EA4 viruses bound predominantly to α-2,6-linked sialosides (Fig. 2d, e) and resembled late EA swine isolates (Fig. 1c and Extended Data Fig. 1). The four resurrected EA viruses showed comparable HA stability with fusion pH 5.8–5.9 (Supplementary Table 2) that resembled EA swine viruses but not the avian virus DK/77 (Supplementary Table 1). In NPTr cells, RG-EA1 replicated to significantly lower titers than RG-EA2, EA3, EA4 viruses at 12, 24 and 36 hpi (Two-way ANOVA, $P < 0.01$ and Tukey's post hoc test, $P < 0.05$; Fig. 2f). We also observed an increase in the polymerase activity from RG-EA1 to RG-EA4 in NPTr cells (Fig. 2g), with RG-EA4 showing the highest polymerase activity in human 293T cells using the minigenome assay (Fig. 2h)[21,22]. These results also support the stepwise adaptation of EA swine viruses in pigs since its introduction from the avian hosts.

EA swine H1N1 viruses sequentially acquired efficient pig-to-pig transmissibility after 1983.

RG-EA1, EA2, EA3, and EA4 viruses were further evaluated for their transmissibility among pigs by direct contact. The avian-like RG-EA1 virus was transiently detected in the nasal swabs of 3/4 donors and 1/4 contacts (Fig. 3a), with none of the donors or contacts seroconverted. RG-EA2, which genetically resembles 1979–1983 era EA swine viruses was transiently detected in 2/4 donors and 1/4 contacts (Fig. 3b), with 1/4 donors seroconverted (1:40). RG-EA3 and EA4, which resemble post 1983 era EA swine viruses replicated more efficiently in all donors (Fig. 3c,d) and were transmitted to all contacts with seroconversion detected in donors (1:40 to 1:320) and contacts (1:320 to 1:640). Viral loads detected in the nasal swabs of donors inoculated with RG-EA1 (mean ± SD AUC = 2.48 ± 2.25), EA2 (0.91 ± 1.05), EA3 (4.72 ± 3.23), or EA4 (5.18 ± 1.39) were moderately different (Kruskal–Wallis test, $P = 0.051$). Collectively, we noted a positive correlation between viral loads detected in donor nasal swabs and viral transmissibility in pigs (Spearman's $r = 0.90$, $P = 0.014$) (Fig. 3e), suggesting the importance of achieving high viral loads at the nasal epithelial cells prior to acquiring efficient transmissibility.

Introducing HA and NA genes derived from RG-EA3 virus did not increase transmissibility of RG-EA2 virus in pigs.

Thirty-three amino acid differences exist between the non-transmissible RG-EA2 and transmissible EA3 viruses (Fig. 4a). Among the 12 amino acid differences found in the HA protein of RG-EA2 and EA3 viruses, HA1-N121T, HA1-Y138H, HA1-N207Y, HA1-K311Q, HA2-A65S, and HA2-D158N (H1 numbering) were detected at high frequencies (>90%) among EA swine viruses isolated from 1979–2016 (Fig. 4a). Introducing these mutations into the HA protein of RG-EA2 virus reduced binding for α-2,3-linked sialosides and marginally enhanced binding for α-2,6-linked sialosides (Fig. 4b). In the NA protein,

RG-EA2 and RG-EA3 differed by the Y344N mutation (N1 numbering) (Fig. 4a). RG-EA1, EA3 and EA4 showed comparable Km that are higher than that of RG-EA2 (One-way ANOVA and Turkey's post hoc test, P = 0.15); however, the avian-like RG-EA1 showed the highest V_{max} (P < 0.01) (Supplementary Table 2).

To investigate if the surface gene segments from RG-EA3 dictate the different transmission phenotypes between RG-EA2 and RG-EA3 viruses, we generated RG-EA2xEA3^{SG} virus that contained the internal genes from RG-EA2 and the surface genes derived from RG-EA3. RG-EA2xEA3^{SG} virus was transiently detected in the nasal swabs of 2/4 inoculated donors with peak titers detected on 8 and 10 dpi (Fig. 4c), respectively; none of the donors seroconverted by 11 dpi. RG-EA2xEA3^{SG} was transiently detected in 2/4 contacts with peak titers at late time points on 11 and 13 days post-exposure (dpe), respectively (Fig. 4c), with none of the contact pigs seroconverted. These results suggest that introducing the HA and NA genes of RG-EA3 virus was not sufficient to confer increase transmissibility for RG-EA2 virus.

Molecular determinants associated with efficient transmission in pigs reside in the internal genes.

Among 20 amino acids that differentiate the internal proteins of RG-EA2 and -EA3 viruses, PB1-Q621R and NP-R351K were detected at high frequencies (> 90%) among EA swine influenza viruses isolated from 1979 to 2018 (Fig. 4a). PB1-R621 and NP-K351 were also highly enriched among classical H1N1 swine viruses and human influenza A viruses isolated from 1933 to 2019 (Fig. 5a). Introduction of the PB1-Q621R but not the NP-R351K mutation increased the polymerase activity of RG-EA2 (One-way ANOVA and Tukey's post hoc test, P < 0.01; Fig. 5b). Both PB1-R621Q and NP-K351R mutations reduced the polymerase activity of RG-EA3 (P < 0.01; Fig. 5b). In NPTr cells, RG-EA2 $^{PB1-Q621R,NP-R351K}$ replicated to higher titers than RG-EA2 and RG-EA2 $^{NP-R351K}$ viruses at 24 hpi (P = 0.019) and 36 hpi (P = 0.024), respectively (Fig. 5c). Interestingly, the NP-R351K mutation facilitated accumulation of viral RNP in the nucleus at earlier time points when compared to RG-EA2 virus (P < 0.05; Fig. 5d,e).

Next, we generated RG-EA2xEA3^{IG} virus contained surface genes from RG-EA2 and internal genes from RG-EA3. In donors, RG-EA2xEA3^{IG} virus was detected in the nasal swabs of 2/4 inoculated donors, with peak titers detected earlier on 4 and 6 dpi (Fig. 5f) than those inoculated with the RG-EA2xEA3^{SG} virus (Fig. 4c). In contact pigs, RG-EA2xEA3^{IG} virus was transiently detected in 2/4 contacts (Fig. 5f), with peak titers detected earlier on 7 dpe (Fig. 5f) than those exposed to RG-EA2xEA3^{IG} on 11 and 13 dpe (Fig. 4c); seroconversion was detected in one pig (1:160). These results suggest that RG-EA2xEA3^{IG} virus showed better replication in inoculated donors and transmitted more rapidly to contact pigs than the RG-EA2xEA3^{SG} virus.

We further focused on these two mutations and compared the transmissibility of RG-EA2$^{NP-R351K}$ and RG-EA2$^{PB1-Q621R, NP-R351K}$ viruses in pigs. RG-EA2$^{NP-R351K}$ was detected from the nasal swabs of all donors and contacts (Fig. 5g), and seroconversion was detected in 4/4 donors (1:80 to 1:160) and 3/4 contacts (1:80 to 1:160). RG-EA2$^{PB1-Q621R,NP-R351K}$ was detected from the nasal swabs of all donors and contacts

(Fig. 5h), with all donors (1:80 to 1:160) and contacts (1:160 to 1:640) seroconverted. The total amount of virus shed by the RG-EA2$^{NP-R351K}$ inoculated donors (mean ± SD AUC = 2.74 ± 2.63) was comparable to that of the RG-EA2$^{PB1-Q621R,NP-R351K}$ inoculated donors (3.78 ± 1.67; two-sided Mann-Whitney test, P = 0.69). In contact pigs, viral load shed by RG-EA2$^{NP-R351K}$ infected contacts (mean ± SD AUC = 3.04 ± 1.57) was slightly lower than that of the RG-EA2$^{PB1-Q621R,NP-R351K}$ infected contacts (5.4 ± 1.31; two-sided Mann-Whitney test, P = 0.11). Next-generation sequencing analyses were performed on the peak-titer nasal swab samples of each contact pig infected with RG-EA2$^{NP-R351K}$ or RG-EA2$^{PB1-Q621R,NP-R351K}$ and we did not observe common adaptive mutations in more than 1 pig (Supplementary Table 3). Taken together, these results showed that introducing the NP-R351K mutation was sufficient to enhance the transmissibility of the reconstructed early EA2 swine virus.

Sequence analyses of archived EA swine viruses in pig lung homogenates from 1979.

We adopted an evolution-guided approach to generate recombinant EA swine viruses based on the posterior distributions of ancestral states at their corresponding nodes of the virus phylogeny. This approach may be biased if the viral sequences contained mutations that emerge after sequential passages in embryonated chicken eggs or in cell culture. To validate the predicted ancestral sequences, direct sequencing of EA swine viruses in two archived pig lung homogenates from 1979 was performed: a partial genome of A/swine/Belgium/1/1979 (designated as Be01-lung; GISAID accession numbers: EPI_ISL_1055769) and full genome of A/swine/Belgium/2/1979 (designated as Be02-lung; GISAID accession numbers: EPI_ISL_1055773) were recovered. The Be01-lung and Be02-lung sequences shared 98.8–99.3% nucleotide homology and 98.9–99.4% amino acid homology with the RG-EA2 virus. Among 33 amino acids that differed between RG-EA2 and RG-EA3 viruses, the Be02-lung sample only differed from the RG-EA2 virus by one amino acid (PB2-483M in Be02-lung and T in RG-EA2) (Extended Data Fig. 2). In comparison with RG-EA2, the partial sequence of Be01-lung sample showed four amino acid differences at HA1 residues that are highly variable among EA swine viruses (Extended Data Fig. 2). Importantly, the PB1-621 and NP-351 residues of the two pig lung homogenates were identical to that of the RG-EA2 virus (Extended Data Fig. 2). Collectively, both archived pig lung sequences shared high homology with the ancestral sequence of RG-EA2 that genetically resembles 1979–1983 era EA swine viruses. Additional serial passages of recombinant RG-EA2 virus carrying the HA and NA genes derived from Be02-lung (designated as RG-Be02-lung^{SG}xEA2^{IG}) or RG-EA1, RG-EA2, RG-EA3, RG-EA4 viruses in embryonated chicken eggs, MDCK cells, or NPTr cells failed to identify any common amino acid changes associated with egg adaptation (Supplementary Tables 4–6).

DISCUSSION

Using ancestral sequence reconstruction, we demonstrated that sustained pig-to-pig transmissibility of the EA swine influenza viruses was not an intrinsic property possessed by the avian-like precursor virus prior to its introduction into mammals. Instead, EA swine viruses acquired efficient transmissibility after 1983 through stepwise avian-to-pig adaptations. Specifically, we observed changes in receptor binding specificity from

recognizing both α-2,3 and α-2,6-linked sialosides (RG-EA2) to recognizing only α-2,6-linked sialosides (RG-EA3 and RG-EA4) and gradually increased polymerase activity, which likely contributed to viral replication in pig nasal epithelial cells and the subsequent efficient transmissibility in pigs. Interestingly, further analyses using RG-EA2 and RG-EA3 viruses that differed in the receptor binding specificity and transmission potential in pigs, identified that the NP-R351K mutation was the minimal molecular change required to significantly enhance transmissibility in pigs. Our results illustrated the multi-step process for avian influenza viruses to sequentially adapt in mammalian hosts.

Ancestral sequence reconstruction is a powerful tool that has been used to investigate the function of unsampled genes[23] and to improve influenza vaccine designs[17,24]. Here, we show that ancestral sequence reconstruction is an excellent approach to study important adaptive processes underpinning the interspecies transmission of influenza viruses. Although the predictive power of corresponding algorithms is dependent on the number of available sequences, direct sequencing of EA swine influenza viruses from archived pig lung homogenates in 1979 has confirmed the robustness of the approach in constructing representative viruses that circulated naturally at different evolutionary nodes.

Our results suggest that the early EA swine viruses between 1979 and 1983 possessed inefficient transmission potential among pigs. Transmission may be facilitated under the housing condition of the pig farms with close contact, contaminated surfaces, and shared water, as well as the self-resolving disease signs in pigs. The first events that allowed introduction of an avian-like virus to swine hosts may have involved ecological factors that facilitated exposure to the avian-like H1N1 viruses (eg. overlapped habitats) and viral factors that permitted replication of the avian-like H1N1 virus at the pig respiratory epithelial cells (e.g. dual receptor binding specificity possessed by the early EA swine viruses). The fact that introducing the NP-R351K mutation was able to facilitate efficient transmission of RG-EA2, an early EA swine virus that possessed dual binding specificity, indicates that an exclusive α-2,6 binding specificity may not be essential for efficient influenza transmission in pigs.

The NP-R351K amino acid change has been fixed not only in the EA swine viruses but also in other human and swine influenza viruses. The NP gene of human seasonal A(H1N1), A(H2N2), and A(H3N2) influenza viruses has descended from the 1918 pandemic influenza virus[25], and the 1918 NP protein was likely of avian origin[26]. Consistent with its putative avian origin, the 1918 virus contained NP-R351. The NP-R351K mutation has been quickly fixed among the human seasonal A(H1N1) viruses from 1918 to 1957, with 39 out of 42 (92.86%) available sequences containing K351. This adaptive mutation has been maintained in human A(H2N2) and A(H3N2) viruses, as K351 is found in 100% (129/129) A(H2N2) sequences from 1957 to 1968, and in 99.93% (26,895/26,915) A(H3N2) sequences from 1968 to 2019. The recent A(H1N1)pdm09 virus derived its NP gene from the classical swine influenza viruses that share the same origin as the 1918 pandemic virus[6]. Since the NP-R351K mutation was also fixed in 97.08% (1,993/2,053) of the classical swine influenza viruses, the A(H1N1)pdm09 virus continues to harbour the NP-R351K mutation at a high frequency (99.97%, 17,621/17,626) from 2009 to 2019. NP residue 351 is in proximity to residue 319 known to interact with importin-α proteins that mediate nucleus

importation[27,28]. Additionally, interaction with interferon-induced MxA protein may select NP residues that confer MxA resistance[29–31]. The potential interaction of the NP-R351K mutation with Mx1 and MxA has been studied in the context of A(H1N1)pdm09[30] and EA swine influenza viruses[14], highlighting that the NP-R351K mutation alone did not confer resistance to Mx1 or MxA. Further studies are needed to delineate the functional role conferred by the NP-R351K mutation.

The reconstructed ancestral influenza A viruses represent excellent models for the study of interspecies transmission and host adaptation. Influenza viruses with segmented genomes may rapidly expand genetic diversity and acquire viral fitness through genetic reassortment, as seen with the A(H1N1)pdm09 virus[6]. While this study did not address the effect of genetic reassortment, swine viruses containing genes derived from the EA swine influenza viruses have been frequently detected in recent years[13,15] that warrants further studies. RNA viruses will continue to cross species barriers and there is a need to maintain vigilance for the next pandemic virus. Our results suggest that there may be the opportunity to intervene at the early stage as avian influenza viruses adapt to mammalian hosts. Continuous surveillance that are strategically coordinated with risk assessment studies may help to identify viruses with pandemic potential before they become fully adapted in the mammalian species. Intervention strategies implemented prior to full adaption to the new host are most likely to be successful and represent the best use of limited preparedness resources.

METHODS

Ethics Statements.

Pig experiments were performed in an ABSL2+ facility at St. Jude Children Research Hospital, in compliance with the NIH and the animal Welfare and with the approval of the St. Jude Animal Care and Use Committee (Protocol 428).

Research approach and oversight.

This study adopted a retrospective study design to follow the natural evolutionary path of the EA swine influenza viruses. The ancestral viruses that we constructed based on the consensus sequences of existing low pathogenic avian influenza viruses and EA swine influenza viruses were, by design, anticipated to have characteristics of wild type, naturally occurring, virus isolates. In essence, we aimed to dissect and study viral evolution that had already occurred in nature, rather than anticipating mutations that may lead to increased mammalian transmissibility or virulence. All of the ancestral viruses that we generated were anticipated to be less mammalian adapted (loss-of-mammalian-function) than present day EA viruses circulating widely in swine in Europe and Asia. All viral protein sequences encoded by the reconstructed EA1, EA2, EA3, and EA4 viruses showed high identity (99%– 100%) with 0 to 4 amino acid differences to the avian and swine influenza virus sequences deposited in GenBank or GISAID. Given the high mutation rate of influenza viruses and limited surveillance conducted in birds and pigs in the 1970s and 1980s, it is therefore likely that these viruses existed in nature but remain unsampled.

All experimental work on the reconstructed EA swine viruses were performed in BSL2-enhanced laboratories with restricted access following approved standard operating procedures. Refresher training of study personnel is provided annually. All reagents (plasmids and viruses) were stored in locked freezers in rooms with restricted access. During animal challenge studies all staff wore full personal protective equipment including respiratory protection.

The research approach was reviewed from a dual use/gain of function perspective by institutional and funding agency committees. The U.S. (this study was funded by a U.S. federal agency) Dual Use Research of Concern policies are restricted to 15 microbial agents. As the avian and swine influenza viruses used in this study are not included in the 15 agents, these policies were not considered further. The work was assessed during the U.S. Gain-of-Function (GOF) Research Funding Pause in 2014 for activities that would enhance the pathogenicity and/or transmissibility of influenza viruses in mammals via the respiratory route. Institutional and U.S. NIAID review of the study determined that it did not meet the criteria for the GOF research funding pause. The rationale for this decision included that "the reconstruction of a precursor wild-type virus by reverse genetics that exists in nature is not gain-of-function" and that "the resultant virus is expected to be more avian-like than currently circulating swine viruses and be less pathogenic and/ or transmissible in mammals".

Cells and viruses.

Madin-Darby canine kidney (MDCK) cells were obtained from the American Type Culture Collection (ATCC) and were maintained in minimal essential medium (MEM) supplemented with 10% fetal calf serum (FCS), 1% penicillin-streptomycin (P/S), and 1% vitamins, and buffered with 25 mM HEPES. Human embryonic kidney 293T (293T) cells were obtained from ATCC and were maintained in Opti-MEM supplemented with 5% FCS and 1% P/S. New-born pig trachea (NPTr) cells were obtained from Istituto Zooprofilattico Sperimentale, della Lombardia e dell'Emilia Romagna and were maintained in MEM supplemented with 10% FCS, 1% P/S, and 1% sodium pyruvate. African green monkey kidney (Vero) cells were obtained from ATCC and were maintained in MEM supplemented with 10% FCS and 1% P/S. Cells were cultured at 37°C in 5% CO_2. All cells used in the study were routinely tested negative for *Mycoplasma sp.* using real-time PCR conducted by the Faculty Core Facility, the University of Hong Kong (HKU).

Avian influenza virus A/duck/Bavaria/2/1977 (H1N1) and Eurasian avian-like (EA) swine influenza A(H1N1) viruses A/swine/Netherlands/3/1980, A/swine/Germany/2/1981, A/swine/Netherlands/12/1985, A/swine/Italy/670/1987, and A/swine/Schleswig-Holstein/1/1992 were kindly provided by Dr. Robert Webster from St. Jude Children's Research Hospital, Memphis, TN, USA. EA swine influenza viruses A/swine/Arnsberg/6554/1979 (H1N1) was kindly provided by Prof. Stephan Pleschka from Justus-Liebig-University Giessen, Germany. EA swine influenza viruses, A/swine/Hong Kong/8512/2001, A/swine/Hong Kong/72/2007, A/swine/Hong Kong/1559/2008, A/swine/Hong Kong/NS29/2009, and A/swine/Hong Kong/NS4848/2011 were isolated and stored at the University of Hong Kong. Archived pig lung homogenates of EA swine viruses (A/swine/

Belgium/1/1979 and A/swine/Belgium/2/1979) were kindly provided by Prof. Kristien Van Reeth from Ghent University, Merelbeke, Belgium. Viruses were grown at a multiplicity of infection (MOI) of 0.005 on MDCK cells cultured in MEM supplemented with 0.3% bovine serum albumin (BSA, Sigma-Aldrich, Missouri, USA), 1% P/S, and 1% vitamin, 25 mM HEPES and 1μg/mL L-1-tosylamide-2-phenylmethyl chloromethyl ketone (TPCK)-treated trypsin (Sigma-Aldrich, Missouri, USA). The viruses were passaged twice on MDCK cells and their genomes were confirmed by Sanger sequencing (Centre for PanorOmic Sciences, HKU). Stock viruses were stored at −80 °C. The stock viruses were titrated in MDCK cells to determine the plaque forming units per milliliter (pfu/mL).

Ancestral sequence reconstruction.

Over 2,000 nucleotide sequences of HA-H1NX (N = 410), NA-HXN1 (N = 443) and HXNX (for all the 6 internal genes: PB2 (N = 454), PB1 (N = 427), PA (N = 397), NP (N = 417), MP (N = 495) and NS (N = 400) were downloaded from NCBI database before the year 2014. Sequences less than 500 bp were removed from the datasets. Identical sequences and outliers were also removed from the datasets. For each gene segment, independent maximum likelihood (ML) analyses were performed with RAxML ver 8.0[33] in Geneious R9.0.3 (Biomatters Ltd.) to generate 10 input trees for ancestral state reconstruction. For each gene segment, internal nodes were assigned to indicate the transmission events of EA swine virus from avian to swine and to also represent different evolutionary stages of EA swine viruses in pigs. We used the Lazarus software package (version 2.0)[34] that utilises baseml to reconstruct ML ancestral states for ancestral nodes using a general time reversible nucleotide substitution model with four gamma-distributed discrete categories of among-site rate variation.

To visualise the evolutionary positions of the reconstructed ancestral nucleotide sequences, we included those sequences in the individual gene datasets and constructed ML phylogenies using RAxML. The reconstructed ancestral nodes for RG-EA1 to RG-EA4 are located in the same position in individual gene trees, with the exception of the highly conserved M1 and M2 genes that have shared nodes for RG-EA2 and RG-EA3. We then used the treesub programme[35] to infer amino acid substitutions at the tree nodes, which were mapped onto the gene phylogenies to show fixation of amino acids within each gene.

Generation of recombinant viruses.

The nodal sequences derived from the phylogenetic analyses of eight gene segments of avian and EA swine influenza viruses were synthesized by GeneArt and Synbio Techonogies (Monmouth Junction, New Jersey, USA). The synthesized genes were cloned into pHW2000 vector by mega-primers PCR as described[36,37]. Recombinant viruses were generated by transfecting eight plasmids into 293T cells. Virus titers were determined in 6-well plates on MDCK cells by plaque assay. All rescued viruses were propagated twice on MDCK cells at a MOI of 0.005 to prepare virus stocks. Their genomes were confirmed via Sanger sequencing.

Solid-phase binding assay.

Solid-phase binding assay is described in the Supplementary Methods.

Glycan microarray.

A synthesized glycan microarray comprised of 38 glycans of α2,3-linked sialosides (glycan 1–38), 3 glycans containing both α2,3 & α2,6-linked sialosides (glycan 39–41) and 32 glycans of α2,6-linked sialosides (glycan 42–73) (Extended Data Fig. 3) was utilized for virus binding study using viruses that were inactivated by 0.025% formaldehyde for 7-day at 4 °C. Glycan array slides were blocked by SuperBlock™ (PBS) Blocking Buffer (Pierce, Rockford, USA) for 1 hour at room temperature. To avoid the influence of neuraminidase (NA), NA inhibitor zanamivir was added to virus stocks at a final concentration of 10 μM. Formalin-inactivated viruses were diluted to 64 HAU per 50 μL and 100 μL of the inactivated viruses were added to the wells of the glycan array. The array slides were incubated at room temperature with slow shaking (13 rpm) for 1 hour. The bound viruses were detected using a 1:6.7 dilution of HA2-targeting human monoclonal antibody (MED18852) followed by a 1:40 dilution of goat anti-human IgG (H+L) secondary antibody labelled with Alexa Fluor 647 (Invitrogen TM, Thermo Fisher, USA). The slides were scanned by InnoScan 710 AL Microarray Scanner (Innopsys, Chicago, USA) equipped with two laser sources, visible wavelength 635nm and 532nm. The data were analyzed in GenePix Pro 6.0 software (Molecular Devices, San Jose, USA).

Syncytium formation assay.

The syncytium formation assay is described in the Supplementary Methods.

Viral replication kinetics *in vitro*.

Confluent MDCK and NPTr cells in 12-well plates were inoculated at an MOI of 0.001 in 1 mL infection medium, with two repeats. Supernatants were collected at 0, 2, 12, 24, 36, 48, 60 and 72 hpi and were titrated on MDCK cells in 6-well plate by plaque assay.

Preparation and infection of swine trachea and lung explants.

Preparation and infection of swine trachea and lung explants are described in the Supplementary Methods.

Site-directed mutagenesis.

Primers designed via QuickChange Primer Design (https://www.agilent.com/store/primerDesignProgram.jsp) were used to introduce specific mutations to HA, PB1, NP genes. These specific primers were listed in Supplementary Table 7. The PCR reactions were performed using QuickChange Multi Site-Directed Mutagenesis Kit in accordance with the manufacturer's instructions (Agilent, Santa Clara, USA).

Minigenome assay.

The polymerase activity was evaluated by minigenome assay[21,22]. 1 μg of each polymerase gene plasmid, 2 μg of NP plasmid, 1 μg reporter plasmid encoding the codon firefly luciferase flanked by the non-coding region of influenza M gene either driven by human Polymerase I promoter[38] or swine Polymerase I promoter[39] and 0.1 μg phRL-CMV plasmid (renilla luciferase driven by CMV promotor) were transfected respectively into 293T cells or NPTr cells in the 6-well plates by using TransIT (MIRUS, Madison, USA) reagent

according to the manufacturer's recommendation protocol. After 24 hours, the luciferase activity was measured with the dual-luciferase reporter system (Promega, Madison, USA) on a SpectraMax iD5 Multi-Mode Microplate Reader (Molecular Devices, San Jose, USA).

Transmissibility of EA viruses in pigs.

Male and female Yorkshire Crossbred piglets (Midwest Research Swine LLC., Glencoe, MN USA) at 3 to 4 weeks old were randomized into different groups. Prior to the experiments, pigs were confirmed as seronegative for influenza A virus NP protein (ID.vet, Grabels, France) and showed HI titer 1:10 for the homologous EA swine influenza virus. Two donor pigs were inoculated intranasally with 10^6 pfu in 1 mL PBS under anesthesia and two naïve pigs were introduced to co-house with two donors on 1 day post-inoculation (dpi). Each experiment was independently performed in duplicate with a total of 4 donors and 4 direct contacts. No statistical methods were used to pre-determine sample sizes but our sample sizes are similar to those reported in previous publications[40,41]. Nasal swabs from donors were collected from both nostrils on 2, 4, 6, 8 and 10 or 11 dpi. Donors were euthanized on 11 dpi for post-infection sera collection. Nasal swabs from all contact pigs were collected from both nostrils on 1, 3, 5, 7, 9, 10 and 11 or 13 dpe. Contact pigs were euthanized on 10 or 13 dpe for post-exposure sera collection. All nasal swabs were placed into 1 mL of viral transport medium (VTM) and stored at −80 °C. Infectious viral titers in the nasal swabs were titrated in MDCK cells to determine the 50% Tissue Culture Infectious Dose $(TCID_{50})$[42].

Hemagglutination inhibition (HI) assay.

Antibodies against the homologous virus from donor or contact pigs were tested by HI assay following standard procedures. Pig sera were treated with receptor-destroying enzyme, RDE (Denka Seiken Tokyo, Japan), for 18–20 hours in a 37°C water-bath and then inactivated for 30 min at 56°C. The final dilution of RDE-treated sera is 1:10. After preparing two-fold serial dilution in 96-well V-bottom microplate, the corresponding virus at the volume of 25 μL containing 4 HA Units was added into the plates and then incubated for 30 min at room temperature. Subsequently, 50μL of 0.5% Turkey red blood cells (LAMPIRE Biological Laboratories, Inc., Pipersville, PA, USA) was added, and the plate was slightly shaken and incubated at room temperature for 30 min. HI titers were expressed as the highest dilution of serum that completely inhibited agglutination of virus and erythrocytes. The detection limit was 1:10. Seroconversion was applied using HI titre 1:20 as the cut-off value.

Sequence analysis.

H1N1 subtype swine influenza viruses and avian influenza viruses were downloaded from NCBI and Global Initiative on Sharing All Influenza Data (GISAID, https://www.gisaid.org/). The sequences that clustered with the gene segments of A/duck/Bavaria/1/1977 (H1N1) were extracted from these maximum likelihood phylogenetic trees. Finally, eight gene segments, PB2 (N = 417), PB1 (N = 381), PA (N = 382), HA (N = 454), NP (N = 495), NA (N = 429), M (N = 396), NS (N = 391) of EA swine influenza viruses isolated from 1979 to 2016 were used for specific amino acids prevalence analysis. To investigate if mutations (i.e. PB1-Q621R and NP-R351K) are conserved among influenza A viruses, the PB1 of 74,095 influenza A viruses and NP of 80,188 influenza A viruses

isolated from 1933 to 2019, respectively, were downloaded from GISAID. Sequences were aligned using MAFFT[43] at the CIPRES science gateway (Version 3)[44]. The number of amino acids were counted using BioEdit (Version 7.0).

Serial passages *in vitro* and *in ovo*.

To compare for mutations that may emerge after passaging in vitro and in ovo, recombinant viruses RG-Be02-lung[SG]xEA2[IG], RG-EA1, RG-EA2, RG-EA3, RG-EA4 were serial passaged in MDCK cells, NPTr cells, and in 10-day old embryonated chicken eggs using transfection supernatant as starting materials. Infections *in vitro* were performed using MOI of 0.001–0.005 at each passage and infections *in ovo* were performed using 10^4 pfu in 0.1 mL. Three independent serial passages were performed in parallel for each condition. Culture supernatant or allantoic fluid harvested from passage three (P3) were analysed by next generation sequencing.

Next-Generation Sequencing.

Viral RNA was extracted from samples using QIAamp viral RNA mini kit (Qiagen, Hilden, Germany). RNA was transcribed into cDNA with SuperScript™ III Reverse Transcriptase (Invitrogen, California, USA). The gene segments were amplified by Q5 High Fidelity DNA polymerase (NEB, Massachusetts, USA) using a pair of primers (454-Tag1-U12:GCC GGA GCT CTG CAG ATA TCA GCR AAA GCA GG; 454-Tag1-U13:GCC GGA GCT CTG CAG ATA TCA GTA GAA ACA AGG). Libraries were prepared based on the Nextera DNA Flex Library Preparation standard protocol[45]. 300ng PCR products were cleaved and tagged by bead-linked transposome at 55°C, then the tagmented DNA were amplified with a pair of indexes using a 5-cycle PCR program. After purification, the libraries were finally eluted using 32 μL resuspension buffer. The libraries were quantified with Agilent fragment analyzer automated CE system with the high sensitivity NGS fragment analysis 474 kit (Agilent, California, USA). 100–150 pM pool libraries containing index adapters allowing multiplex sequencing of 4 samples per iSeq 100 i1 Cartridge was run on iseq100 sequencing system (Illumina, California, USA). Output data were analysed on CLC Genomic workbench (version 20) (Qiagen, Hilden, Germany) using 1% as variant calling threshold. The nucleotide substitutions with a frequency of 5% or more were further analysed.

NA kinetics.

NA kinetics analysis is described in the Supplementary Methods.

Immunofluorescence staining.

To investigate the effect of mutation in NP on cellular distribution of NP during the virus life cycle, NPTr cells were infected with virus at an MOI of 5. Infections were stopped at 2, 4, 6, 8, 10 and 12 hpi by fixation in 4% paraformaldehyde (Electron Microscopy Sciences, PA, USA). The cells were then permeabilized with 0.1% Triton™ X-100 in PBS for 30 min and labelled with 1:200 dilution of mouse monoclonal IgG2a Influenza A NP antibody (Santa Cruz Biotech, Inc, Santa Cruz, USA) at 4 °C overnight. The cells were subsequently incubated with a 1:200 dilution of FITC Goat Anti-Mouse IgG/IgM (BD Pharmingen™,

San Jose, USA) for 3 hours at room temperature in the dark. After three washes, the nuclei were counterstained at a 1:1000 dilution of DAPI (4′,6′-diamidino-2-phenylindole) in SlowFade™ Gold Antifade Mountant (Thermo Fisher Scientific, Eugene, USA) at room temperature for 15 min. Cell imaging and fluorescence were carried out using the Nikon Eclipse Ti-S fluorescence microscope (Nikon, Tokyo, Japan) equipped with a Nikon DS-Qi2 camera and the iNIS-Elements BR imaging software (Version 4.40). The positive cells were numerated using FIJI[46].

Data availability.

Sequences of Sanger sequencing validated in the study were deposited to the public database GISAID under the following accession number. A/duck/Bavaria/2/1977: EPI_ISL_539860; A/swine/Arnsberg/6554/1979: EPI_ISL_539869; A/swine/Belgium/1/1979: EPI_ISL_1055769; A/swine/Belgium/2/1979: EPI_ISL_1055773; A/swine/Netherlands/3/1980: EPI_ISL_539861; A/swine/Germany/2/1981: EPI_ISL_539862; A/swine/Netherlands/12/1985: EPI_ISL_539863; A/swine/Schleswig-Holstein/1/1992: EPI_ISL_539864; A/swine/Hong Kong/8512/2001: EPI_ISL_539865; A/swine/Hong Kong/72/2007 EPI_ISL_539866; A/swine/Hong Kong/1559/2008: EPI_ISL_539867; A/swine/Hong Kong/NS4848/2011: EPI_ISL_539868; RG-EA1: EPI_ISL_539852; RG-EA2: EPI_ISL_539853; RG-EA3: EPI_ISL_539854; RG-EA4: EPI_ISL_539855. Samples of next generation sequencing have been deposited in the Sequence Read Archive of the NCBI, under project accession number PRJNA725802. All source data shown in figures 1, 2, 3, 4, 5 and Extended Data figure 1 are provided.

Statistical Analyses.

Blinding on the conditions of the experiments was not performed during data collection and analysis. No animal or data point was excluded from the analyses. One-way ANOVA was used to compare multiple groups and two-way ANOVA was used to compare virus titers over time, followed by Tukey's multiple comparisons post-hoc tests. Area under the curve was calculated from the nasal swabs of the donor and contact pigs. Two-sided Mann-Whitney test was performed to compare two groups, and Kruskal-Wallis test and Dunn's multiple comparisons post hoc test were used to compare multiple groups. Spearman's rank correlation coefficient analysis was performed for the monotonic relationship between virus replication efficiency and viral transmissibility. Data were analyzed in Microsoft Excel for Mac, version 16.28 and GraphPad Prism version 8.4.1 for Windows (GraphPad Software, La Jolla, USA). All statistical parameters for specific analyses are shown in the corresponding figure legends. Statistically significant P values ($P < 0.05$) are indicated in the corresponding figures.

Extended Data
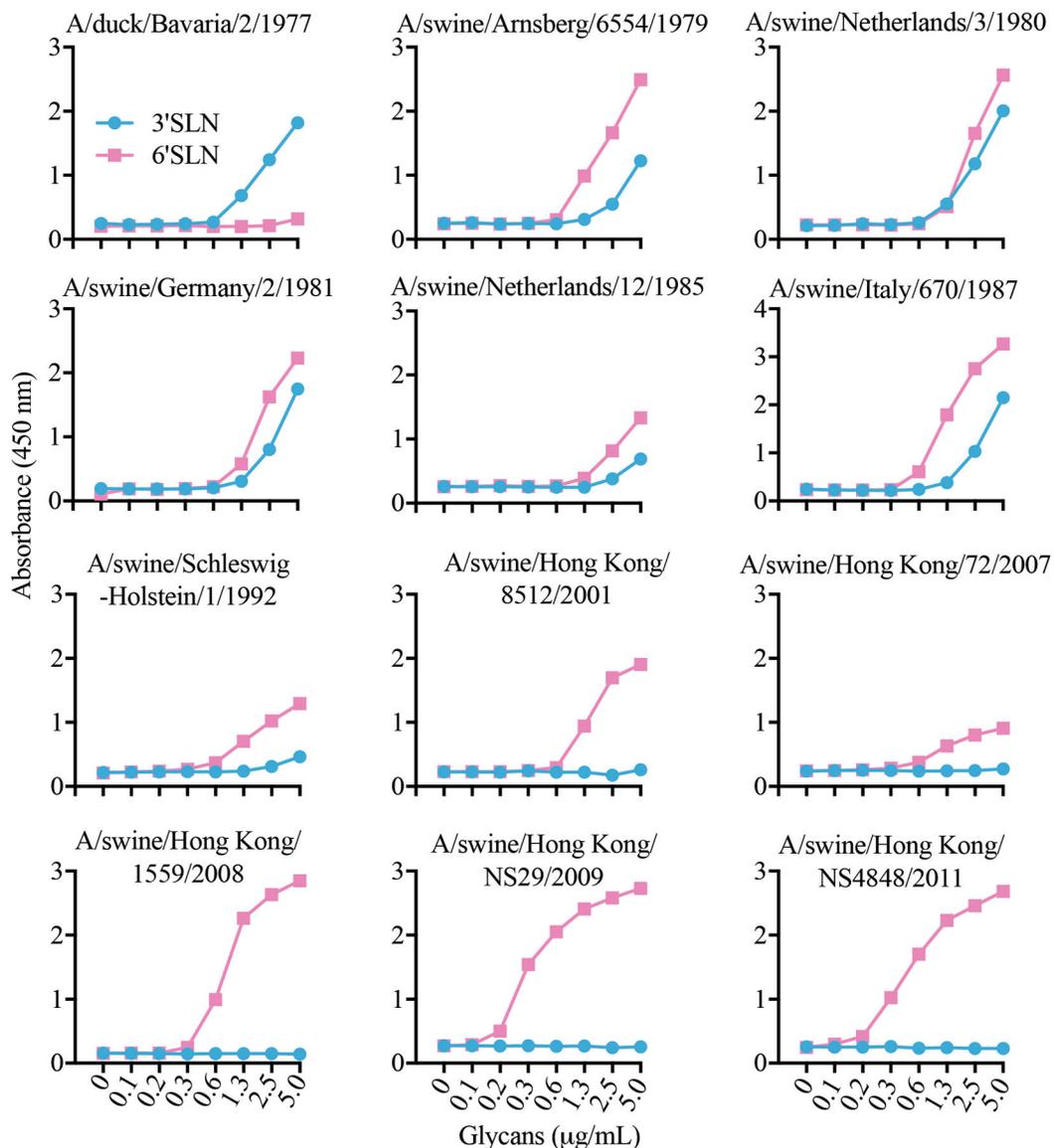
Extended Data Fig. 1. Receptor binding profile of avian and EA swine influenza A viruses.
An ELISA-based assay was employed to investigate receptor binding preference of
avian and EA swine H1N1 influenza viruses. 3'SLN (shown in cyan circles) and
6'SLN (shown in pink squares) were coated in 96 well plates at concentrations of 5
to 0.1 µg/mL. Infectious viruses (diluted to 64 HAU/50µL) were added to each well
and co-incubated with 3'SLN and 6'SLN. The plates were washed with PBS and the
detection of bound virus was done using a rabbit polyclonal antibody against HA from A/
Duck/NZL/160/1976(H1N3) (Sino Biological Inc., Beijing, China) followed by polyclonal
Goat Anti-Rabbit Immunoglobulins/HRP (Dako, Denmark). The absorbance at 450 nm was
plotted against glycan concentration. Results are shown with the mean and SD absorbance
from 3 replicates in one of two independently performed experiment.

Extended Data Fig. 2. EA swine virus sequence from archived pig lung homogenates in 1979 shared high homology with the ancestral sequence of RG-EA2 that genetically resembles 1979–1983 era EA swine viruses.

Full genome of A/swine/Belgium/2/1979 (Be02-lung) and a partial genome of A/swine/Belgium/1/1979 (Be01-lung) EA swine influenza viruses were sequenced from two archived pig lung homogenates from 1979. Four recombinant viruses were resurrected using ancestral sequence reconstruction, with RG-EA1 representing the precursor virus of EA swine viruses, RG-EA2 representing the early evolutionary stage of EA swine viruses in 1979–1983, and RG-EA3 and RG-EA4 representing EA swine viruses in 1984–1987 and 1988–1992, respectively. Among the 33 amino acids that differentiate the non-transmissible RG-EA2 virus and transmissible RG-EA3 virus, the sequences of Be01 and Be02 showed high homology to RG-EA2. /, unsuccessfully determined.
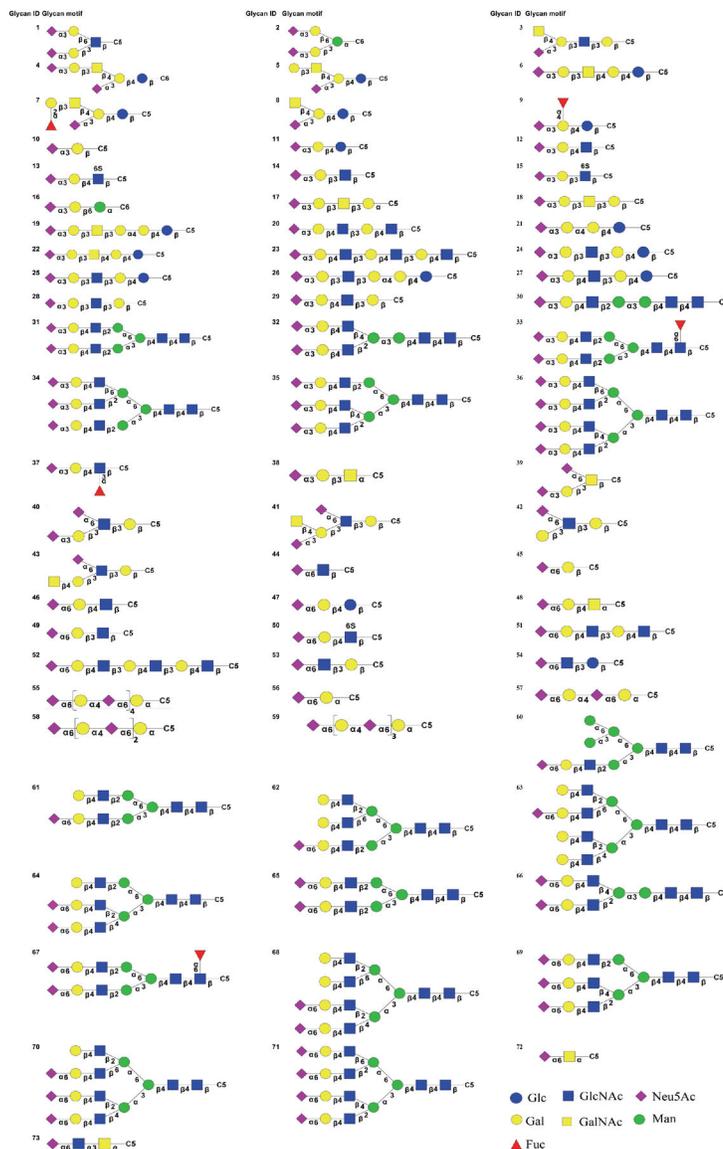
Extended Data Fig. 3. Structure of glycans used for glycan array.
A synthesized glycan microarray comprised of 38 glycans of α2,3-linked sialosides (glycan 1–38), 3 glycans containing both α2,3 & α2,6-linked sialosides (glycan 39–41) and 32 glycans of α2,6-linked sialosides (glycan 42–73) was utilized for glycan array assay.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgements

References

1. Kuiken T et al. Host species barriers to influenza virus infections. Science 312, 394–397 (2006). [PubMed: 16627737]

2. Parrish CR et al. Cross-species virus transmission and the emergence of new epidemic diseases. Microbiol. Mol. Biol. Rev 72, 457–470 (2008). [PubMed: 18772285]

3. Taubenberger JK & Kash JC Influenza virus evolution, host adaptation, and pandemic formation. Cell Host Microbe 7, 440–451 (2010). [PubMed: 20542248]

4. Lipsitch M et al. Viral factors in influenza pandemic risk assessment. Elife 5, e18491 (2016). [PubMed: 27834632]

5. Long JS, Mistry B, Haslam SM & Barclay WS Host and viral determinants of influenza A virus species specificity. Nat. Rev. Microbiol 17, 67–81 (2019). [PubMed: 30487536]

6. Smith GJ et al. Origins and evolutionary genomics of the 2009 swine-origin H1N1 influenza A epidemic. Nature 459,1122–1125 (2009). [PubMed: 19516283]

7. Pensaert M, Ottis K, Vandeputte J, Kaplan MM & Bachmann PA Evidence for the natural transmission of influenza A virus from wild ducts to swine and its potential importance for man. Bull. World Health Organ 59, 75–78 (1981). [PubMed: 6973418]

8. Brown IH History and epidemiology of Swine influenza in Europe. Curr Top Microbiol. Immunol 370, 133–146 (2013). [PubMed: 22234411]

9. Krumbholz A et al. Origin of the European avian-like swine influenza viruses. J. Gen. Virol 95, 2372–2376 (2014). [PubMed: 25073465]

10. Joseph U, Vijaykrishna D, Smith GJD & Su YC F. Adaptive evolution during the establishment of European avian-like H1N1 influenza A virus in swine. Evol. Appl 11, 534–546 (2018). [PubMed: 29636804]

11. Vijaykrishna D et al. Long-term evolution and transmission dynamics of swine influenza A virus. Nature 473, 519–522 (2011). [PubMed: 21614079]

12. Yang H et al. Prevalence, genetics, and transmissibility in ferrets of Eurasian avian-like H1N1 swine influenza viruses. Proc. Natl. Acad. Sci. USA 113, 392–397 (2016). [PubMed: 26711995]

13. Henritzi D et al. Surveillance of European Domestic Pig Populations Identifies an Emerging Reservoir of Potentially Zoonotic Swine Influenza A Viruses. Cell Host Microbe 28, 614–627 (2020). [PubMed: 32721380]

14. Dornfeld D, Petric PP, Hassan E, Zell R & Schwemmle M Eurasian Avian-Like Swine Influenza A Viruses Escape Human MxA Restriction through Distinct Mutations in Their Nucleoprotein. J. Virol 93, e00997–18 (2019). [PubMed: 30355693]

15. Sun H et al. Prevalent Eurasian avian-like H1N1 swine influenza virus with 2009 pandemic viral genes facilitating human infection. Proc. Natl. Acad. Sci. USA 117, 17204–17210 (2020). [PubMed: 32601207]

16. Thornton JW, Need E & Crews D Resurrecting the ancestral steroid receptor: ancient origin of estrogen signaling. Science 301, 1714–1717 (2003). [PubMed: 14500980]

17. Ducatez MF et al. Feasibility of reconstructed ancestral H5N1 influenza viruses for cross-clade protective vaccine development. Proc. Natl. Acad. Sci. USA 108, 349–354 (2011). [PubMed: 21173241]

18. Rogers GN & Paulson JC Receptor determinants of human and animal influenza virus isolates: differences in receptor specificity of the H3 hemagglutinin based on species of origin. Virology 127, 361–373 (1983). [PubMed: 6868370]

19. Baumann J, Kouassi NM, Foni E, Klenk HD & Matrosovich M H1N1 Swine Influenza Viruses Differ from Avian Precursors by a Higher pH Optimum of Membrane Fusion. J. Virol 90, 1569–1577 (2016). [PubMed: 26608319]

20. Russier M et al. H1N1 influenza viruses varying widely in hemagglutinin stability transmit efficiently from swine to swine and to ferrets. PLoS Pathog. 13, e1006276 (2017). [PubMed: 28282440]

21. Pleschka S et al. A plasmid-based reverse genetics system for influenza A virus. J. Virol 70, 4188–4192 (1996). [PubMed: 8648766]

22. Salomon R et al. The polymerase complex genes contribute to the high virulence of the human H5N1 influenza virus isolate A/Vietnam/1203/04. J. Exp. Med 203, 689–697 (2006). [PubMed: 16533883]

23. Thornton JW Resurrecting ancient genes: experimental analysis of extinct molecules. Nat. Rev. Genet 5, 366–375 (2004). [PubMed: 15143319]

24. Giles BM & Ross TM A computationally optimized broadly reactive antigen (COBRA) based H5N1 VLP vaccine elicits broadly reactive antibodies in mice and ferrets. Vaccine 29, 3043–3054 (2011). [PubMed: 21320540]

25. Taubenberger JK & Morens DM The 1918 Influenza Pandemic and Its Legacy. Cold Spring Harb. Perspect. Med 10, a038695 (2020). [PubMed: 31871232]

26. Reid AH, Fanning TG, Janczewski TA, Lourens RM & Taubenberger JK Novel origin of the 1918 pandemic influenza virus nucleoprotein gene. J. Virol 78, 12462–12470 (2004). [PubMed: 15507633]

27. Gabriel G, Herwig A, Klenk HD. Interaction of polymerase subunit PB2 and NP with importin alpha1 is a determinant of host range of influenza A virus. PLoS Pathog 4, e11 (2008). [PubMed: 18248089]

28. Gabriel G et al. Differential use of importin-α isoforms governs cell tropism and host adaptation of influenza virus. Nat. Commun 2, 156 (2011). [PubMed: 21245837]

29. Zimmermann P, Mänz B, Haller O, Schwemmle M & Kochs G The viral nucleoprotein determines Mx sensitivity of influenza A viruses. J. Virol 85, 8133–8140 (2011). [PubMed: 21680506]

30. Mänz B et al. Pandemic influenza A viruses escape from restriction by human MxA through adaptive mutations in the nucleoprotein. PLoS Pathog 9, e1003279 (2013). [PubMed: 23555271]

31. Ashenberg O, Padmakumar J, Doud MB & Bloom JD Deep mutational scanning identifies sites in influenza nucleoprotein that affect viral inhibition by MxA. PLoS Pathog 13, e1006288 (2017). [PubMed: 28346537]

32. Yen H-L et al. Hemagglutinin–neuraminidase balance confers respiratory-droplet transmissibility of the pandemic H1N1 influenza virus in ferrets. Proc. Natl. Acad. Sci. USA 108, 14264–14269 (2011). [PubMed: 21825167]

33. Stamatakis A RAxML version 8: a tool for phylogenetic analysis and post-analysis of large phylogenies. Bioinformatics 30, 1312–1313 (2014). [PubMed: 24451623]

34. Finnigan GC, Hanson-Smith V, Stevens TH & Thornton JW Evolution of increased complexity in a molecular machine. Nature 481, 360–364 (2012). [PubMed: 22230956]

35. Tamuri AU Treesub: annotating ancestral substitution on a tree Available at: https://github.com/tamuri/treesub (2013) (Accessed June 12, 2018).

36. Stech J et al. Rapid and reliable universal cloning of influenza A virus genes by target-primed plasmid amplification. Nucleic Acids Res 36, e139 (2008). [PubMed: 18832366]

37. Hoffmann E, Neumann G, Kawaoka Y, Hobom G & Webster RG A DNA transfection system for generation of influenza A virus from eight plasmids. Proc. Natl. Acad. Sci. USA 97, 6108–6113 (2000). [PubMed: 10801978]

38. Hoffmann E, Neumann G, Hobom G, Webster RG & Kawaoka Y "Ambisense" approach for the generation of influenza A virus: vRNA and mRNA synthesis from one template. Virology 267, 310–317 (2000). [PubMed: 10662626]

39. Moncorge O et al. Investigation of influenza virus polymerase activity in pig cells. J. Virol 87, 384–394 (2013). [PubMed: 23077313]

40. Zhu H et al. Infectivity, transmission, and pathology of human-isolated H7N9 influenza virus in ferrets and pigs. Science 341, 183–186 (2013). [PubMed: 23704376]

41. Nishiura H, Yen HL, & Cowling BJ Sample size considerations for one-to-one animal transmission studies of the influenza A viruses. PloS one, 8, e55358 (2013). [PubMed: 23383167]

42. Reed LJ & Muench H A simple method of estimating fifty percent endpoints. Am. J. Hyg 27, 493–497 (1938).

43. Katoh K, Misawa K, Kuma K & Miyata T MAFFT: a novel method for rapid multiple sequence alignment based on fast Fourier transform. Nucleic Acids Res 30, 3059–3066 (2002). [PubMed: 12136088]

44. Miller MA, W. P., Schwartz T Creating the CIPRES Science Gateway for inference of large phylogenetic trees. 2010 Gateway Computing Environments Workshop (GCE) 23, 1–8 (2010).

45. Bruinsma S et al. Bead-linked transposomes enable a normalization-free workflow for NGS library preparation. BMC Genomics 19, 722 (2018). [PubMed: 30285621]

46. Schindelin J et al. Fiji: an open-source platform for biological-image analysis. Nat. Methods 9, 676–682 (2012). [PubMed: 22743772]
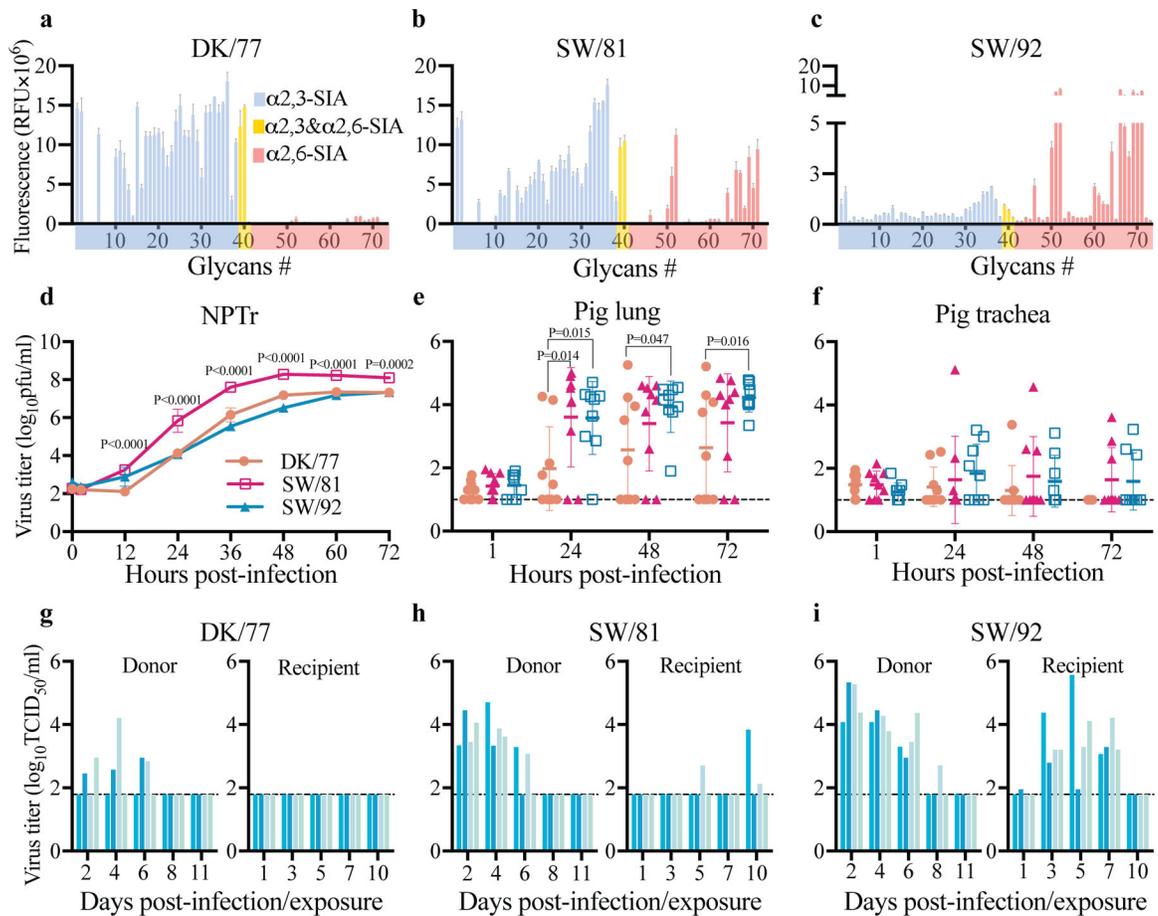
Fig. 1. Avian and EA swine influenza viruses differed in receptor binding profiles, replication efficiency *in vitro* and *ex vivo*, and contact transmissibility in pigs.

a,b,c, Glycan array analysis of avian influenza virus DK/77 (**a**) and EA swine influenza viruses SW/81 (**b**) and SW/92 (**c**). Glycan array data are shown with the mean and SD fluorescence (RFU) calculated from 6 replicate spots of each glycan. **d**, Multicycle replication kinetics in NPTr cells. Mean and SD from three repeats in one of two independently performed experiment are shown. Statistical differences were calculated by two-way ANOVA followed by a Tukey's post-hoc test and P-values < 0.05 between SW/81 and DK/77 are shown. **e,f**, Replication kinetics in pig lung (**e**) and pig trachea (**f**) ex vivo cultures. Tissues from three healthy pigs were used to prepare lung (N=3 per pig) and tracheal (N=3 per pig) explants for the experiments. Each data point represents one explant sample. Mean and SD from 9 data points are shown. Statistical differences were calculated by two-way ANOVA followed by a Tukey's post-hoc test and P-values < 0.05 are shown. **g,h,i**, Contact transmissibility of DK/77 avian influenza virus (**g**), SW/81 (**h**), and SW/92 (**i**) EA swine influenza viruses in pigs. Each virus was tested in duplicate with a total of four donors and four direct contacts. Pigs housed in the same pen are denoted by the same colour. The limit of detection (1.789 $\log_{10}$TCID$_{50}$/mL) is shown with the horizontal dashed line.
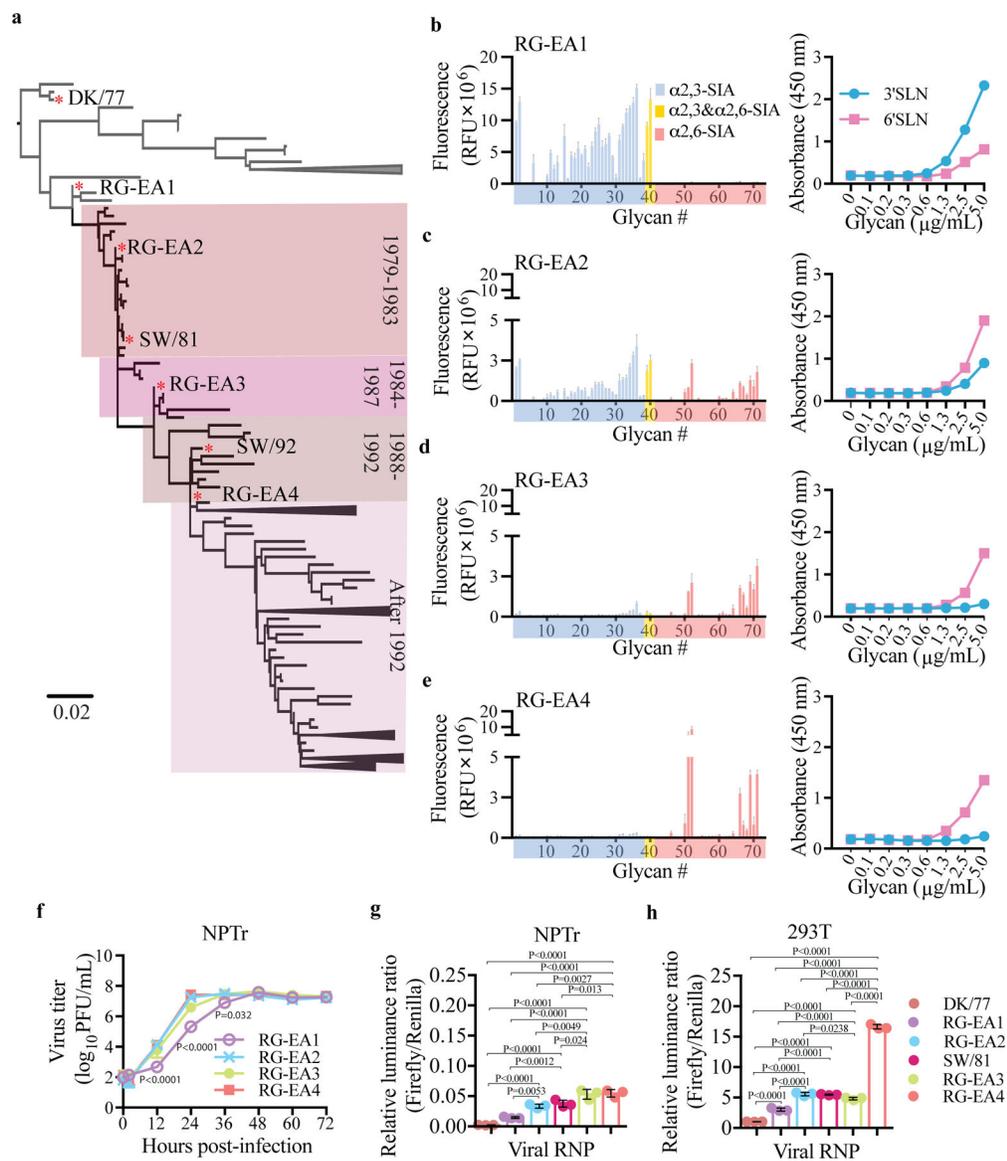
Fig. 2. Characterization of reconstructed EA influenza viruses representing different evolutionary stages of EA swine influenza viruses.

a, The maximum likelihood phylogeny of HA gene sequences was constructed using avian (N = 69) and EA swine viruses (N = 344) isolated from 1977 to 2014. The phylogenetic tree was rooted to the branch of A/duck/Bavaria/2/1977 (H1N1) (DK/77). Red asterisks indicate the phylogenetic position of the reconstructed sequences of RG-EA1 to RG-EA4 and wild-type viruses used in the experiments. The scale bar indicates the length of branch that represents 0.02 nucleotide substitutions per site. **b,c,d,e**, Receptor binding profiles of RG-EA1 (**b**), RG-EA2 (**c**), RG-EA3 (**d**), RG-EA4 (**e**) viruses. Glycan array data (left panel) are shown with the mean and SD fluorescence (RFU) calculated from 6 replicate spots of each glycan (left panel), and the enzyme-linked immunosorbent assay (right panel) results are shown with the mean and SD absorbance (450 nm) from 3 replicates in one of two independently performed experiment. **f**, Multicycle replication kinetics in NPTr cells. Mean and SD from three repeats in one of two independently performed experiment is shown.

Statistical differences were calculated by two-way ANOVA followed by a Tukey's post-hoc test and P-values < 0.05 are shown. **g**,**h**, Polymerase activity in NPTr cells (g) and 293T cells (h). The mean and SD ratio (bar) of firefly/ renilla luciferase from three repeats (dots) in one of two independently performed experiment is shown. One-way ANOVA with Tukey's post hoc test was performed and P-values < 0.05 are shown.
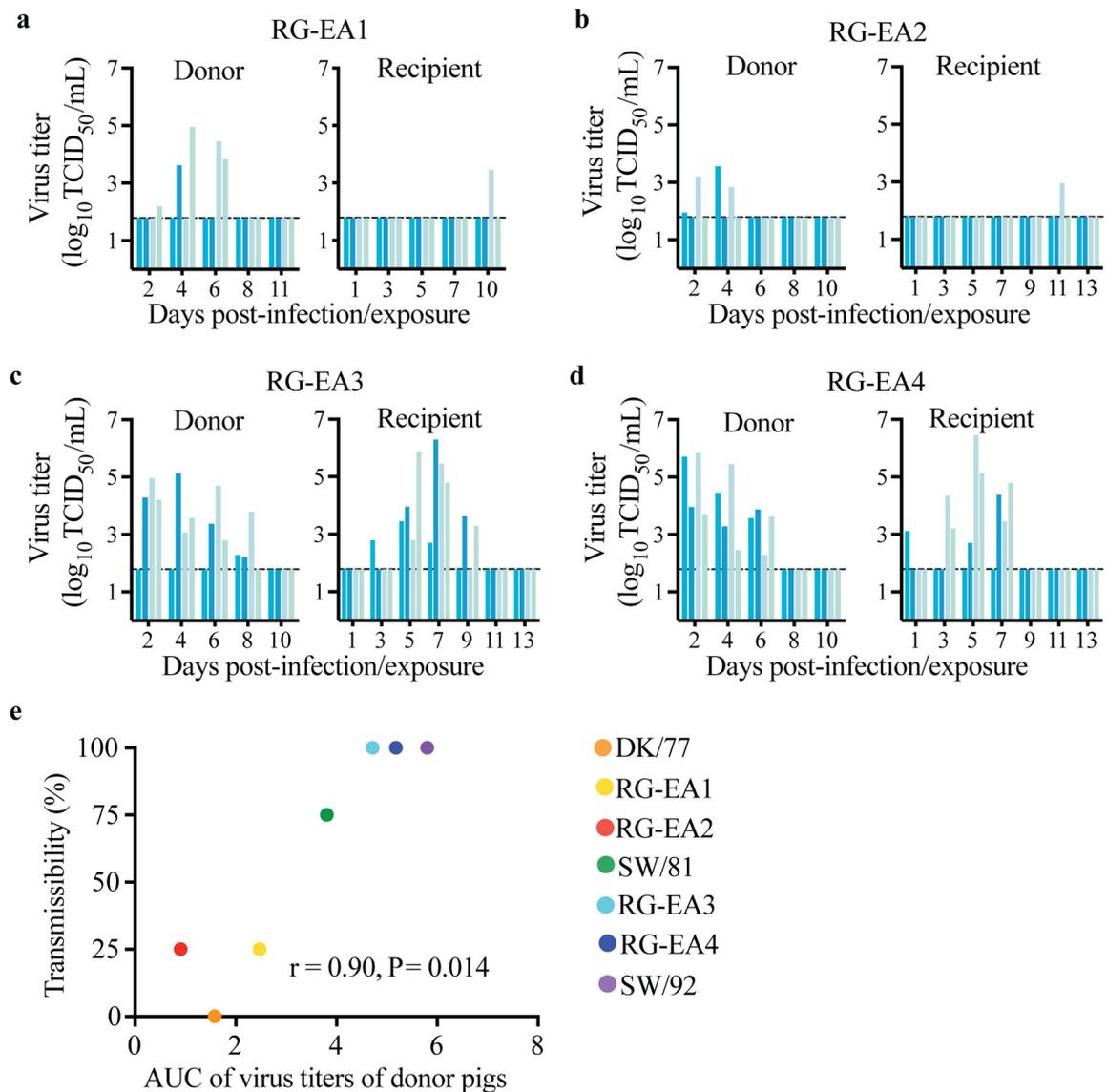
Fig. 3. EA swine H1N1 viruses acquired efficient pig-to-pig transmissibility after 1983.
a, **b, c**, **d**, The transmissibility of RG-EA1 (**a**), EA2 (**b**), EA3 (**c**), and EA4 (**d**) in pigs. Contact transmission experiments were performed in duplicate with a total of four donors and four direct contacts. Pigs housed in the same pen are denoted by the same colour. The limit of detection (1.789 $\log_{10}$TCID$_{50}$/mL) is shown with the horizontal dashed line. **e**, Two-sided Spearman's rank correlation coefficient analysis was used to evaluate the correlation between viral loads detected in donor nasal swabs (mean AUC from 4 pigs, X axis) and onward transmissibility (% of infected out of 4 exposed contact pigs, Y axis).
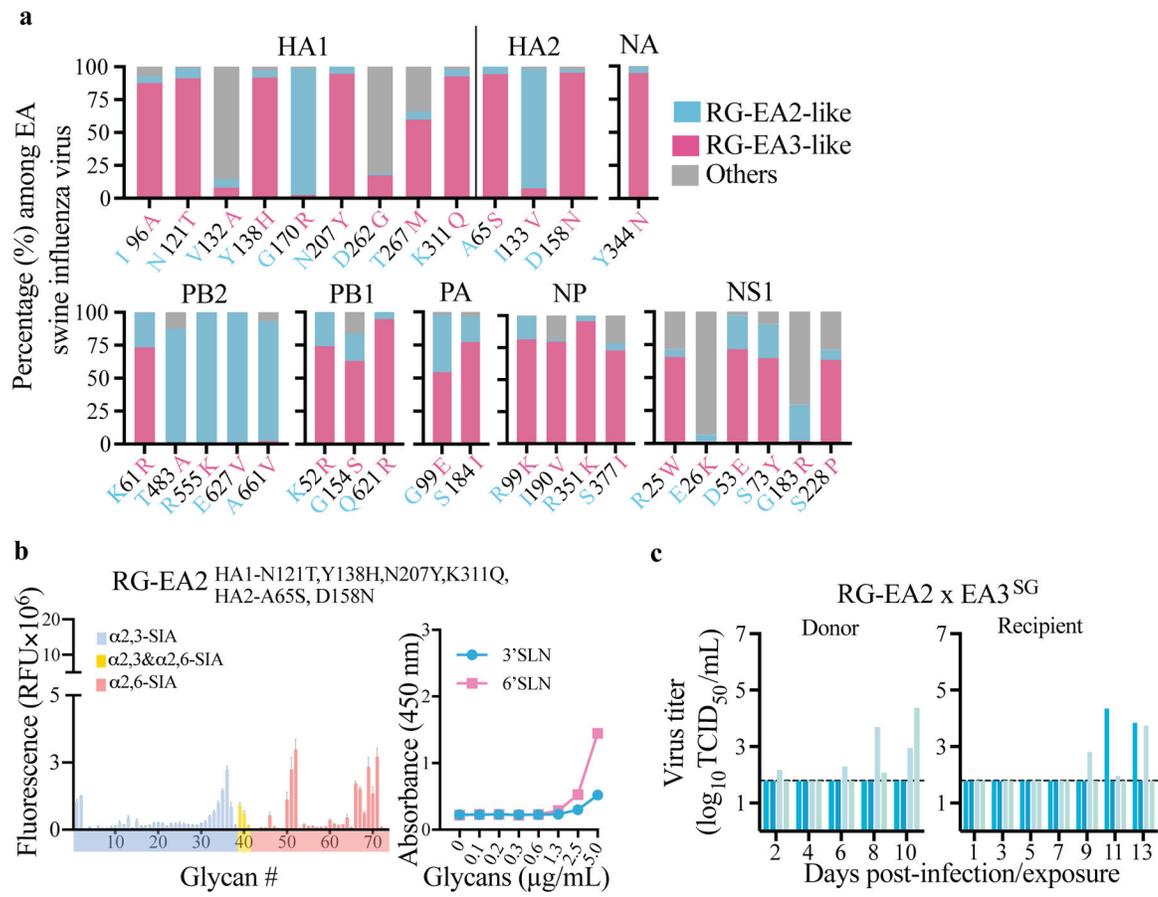
Fig. 4. Introducing HA and NA genes derived from RG-EA3 virus did not enhance efficient contact transmissibility of RG-EA2 virus in pigs.

a, The detection frequency of the 33 amino acids that differed between RG-EA2 and RG-EA3 EA swine viruses among PB2 (N = 417), PB1 (N = 381), PA (N = 382), HA1/HA2 (N = 454), NP (N = 495), NA (N = 429), NS1 (N = 391) of EA swine influenza A viruses isolated in 1979–2016 were analyzed. Amino acid residues that resembled RG-EA2, RG-EA3, or other amino acids are colored in cyan, magenta, and gray, respectively. HA and NA are numbered according to H1 and N1 numbering, respectively. **b**, Receptor binding profile of RG-EA2[HA1-N121T, HA1-Y138H, HA1-N207Y, HA1-K311Q, HA2-A65S, and HA2-D158N]. Glycan array data (left panel) are shown with the mean and SD fluorescence (RFU) calculated from 6 replicate spots of each glycan (left panel), and the enzyme-linked immunosorbent assay (right panel) results are shown with the mean and SD absorbance (450 nm) from 3 replicates in one of two independently performed experiment. **c**, The contact transmissibility of RG-EA2xEA3[SG] virus containing the internal genes from RG-EA2 and the surface genes derived from RG-EA3 in pigs. Contact transmission experiments were performed in duplicate with a total of four donors and four direct contacts. Pigs housed in the same pen are denoted by the same colour. The limit of detection (1.789 $\log_{10}TCID_{50}$/mL) is shown with the horizontal dashed line.
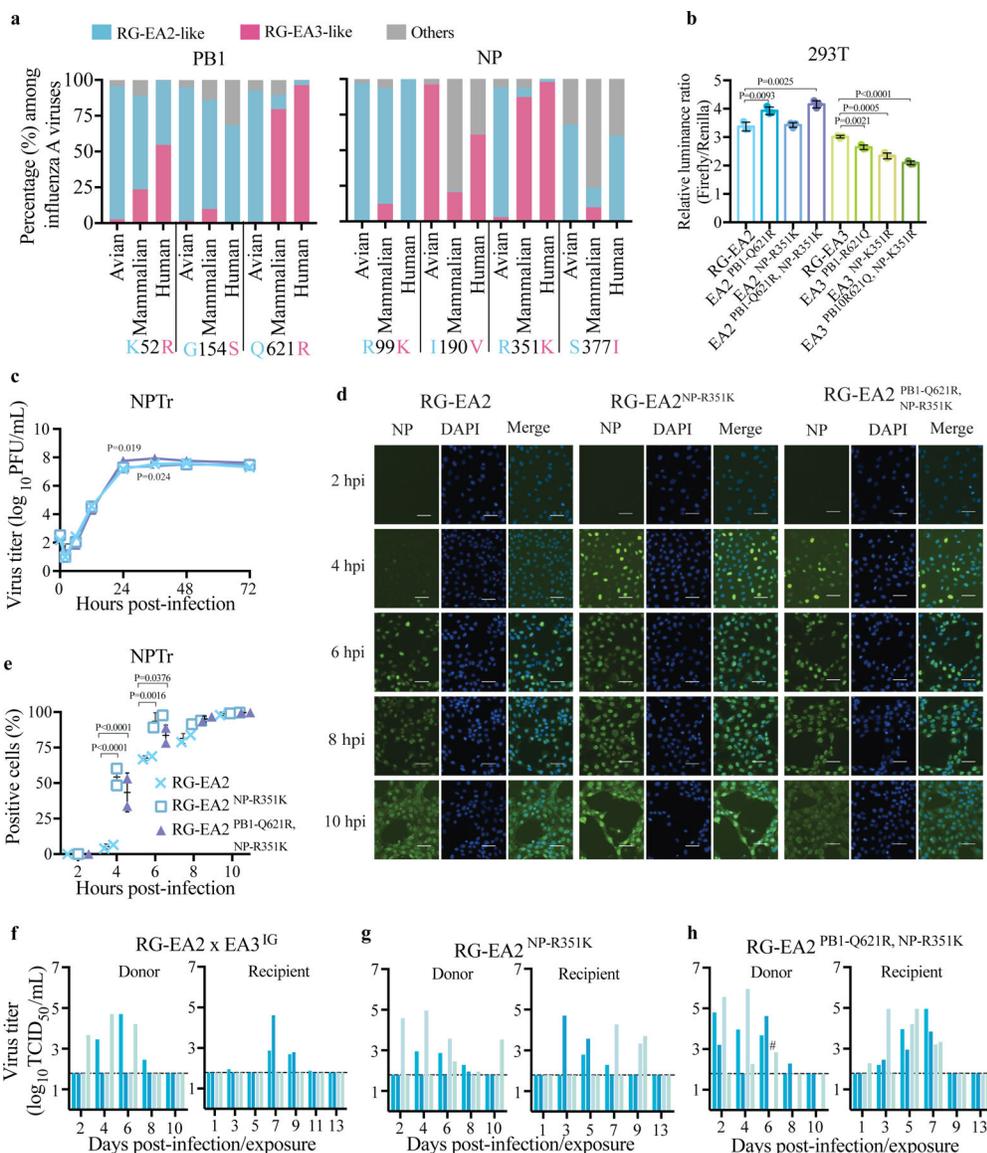
Fig. 5. The NP-R351K mutation in EA swine influenza viruses was the minimal molecular change required to facilitate transmission of the non-transmissible RG-EA2 virus.

a, Detection frequency of Q/R at PB1 residue 621 (N = 74,095) and R/K at NP residue 351 (N = 80,188) among mammalian and human influenza A viruses. **b**, The effects of amino acid substitutions at PB1 residue 621 and NP residue 351 on viral polymerase activity were determined using minigenome assay in 293T cells. The mean and SD ratio (bar) of firefly/ renilla luciferase from three repeats (dots) in one of two independently performed experiment is shown. One-way ANOVA with Tukey's post hoc test was performed and P-values < 0.05 are shown. (**c**) Multicycle replication kinetics of recombinant viruses in NPTr cells. Mean and SD from three repeats in one experiment is shown. Two-way ANOVA with Tukey's post hoc test was performed and the P-values < 0.05 are shown, including the comparison between RG-EA2 and RG-EA2 $^{PB1-Q621R, NP-R351K}$ at 24 hpi and between RG-EA2$^{NP-R351K}$ and RG-EA2$^{PB1-Q621R, NP-R351K}$ at 36 hpi. **d,e**, Nucleus localization of NP protein in NPTr cells infected with RG-EA2, RG-EA2$^{NP-R351K}$, and

RG-EA2$^{\text{PB1-Q621R, NP-R351R}}$ at MOI of 5 at different time points. Representative images from two replicates of one experiment are shown (Scale bar: 100 μM) (**d**). Mean and SD of NP positive rate (NP positive cells per DAPI positive cells, %) from two replicates (dots) of one experiment are calculated at different hours post-infection (**e**). Two-way ANOVA with Tukey's post hoc test were performed and P-values < 0.05 are shown. **f,g,h,** The contact transmissibility of RG-EA2xEA3$^{\text{IG}}$ virus (**f**), RG-EA2 $^{\text{NP-R351K}}$ (**g**), and RG-EA2 $^{\text{PB 1-Q621R, NP-R351K}}$ in pigs. Contact transmission experiments were performed in duplicate with a total of four donors and four direct contacts. Pigs housed in the same pen are denoted by the same colour. The limit of detection (1.789 $\log_{10}\text{TCID}_{50}$/mL) is shown with the horizontal dashed line. # indicates a donor in the RG-EA2 $^{\text{PB 1-Q621R,NP-R351K}}$ group euthanized on day 4 post-infection due to respiratory distress with abdominal distention, severe lethargy, and vomiting.