



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

Wen Su¹, Rhodri Harfoot^{2,3}, Yvonne C. F. Su⁴, Jennifer DeBeauchamp², Udayan Joseph⁴, Jayanthi Jayakumar⁴, Jeri-Carol Crumpton², Trushar Jeevan², Adam Rubrum², John Franks², Philippe Noriel Q. Pascua², Christina Kackos², Yuqin Zhang¹, Mengting Zhang¹, Yue Ji¹, Huyen Trang Bui¹, Jeremy C. Jones², Lisa Kercher², Scott Krauss², Stephan Pleschka⁵, Michael C. W. Chan¹, Robert G. Webster², Chung-Yi Wu⁶, Kristien Van Reeth⁷, Malik Peiris¹, Richard J. Webby²✉, Gavin J. D. Smith⁴✉ and Hui-Ling Yen¹✉

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 influenza viruses from 1979–1983, 1984–1987 and 1988–1992 were reconstructed and characterized. Glycan-binding analyses showed stepwise changes in the haemagglutinin receptor-binding specificity of the EA swine influenza viruses—that is, from recognition of both α 2,3- and α 2,6-linked sialosides to recognition of α 2,6-linked sialosides only; however, efficient transmission in piglets was enabled by adaptive changes in the viral polymerase protein and nucleoprotein, which have been fixed since 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 jumps 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. Although 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 capability of influenza viruses crossing mammalian species barriers and becoming established in humans⁶. However, only a limited number of influenza viruses of avian origin are known to have successfully crossed the species barriers and established in a new mammalian host^{3,7}. 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 European countries^{8,11–13}; they donated the *neuraminidase* (NA) and

M gene segments to the A(H1N1)pdm09 virus⁵ and have demonstrated pandemic potential, including resistance to human MxA^{13,14}, airborne transmissibility in ferrets^{12,15} and antigenic novelty (for example, limited cross-reactivity with the antibodies that humans have 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 whether the sustained swine transmissibility of EA swine viruses was an intrinsic trait of the precursor avian influenza virus or whether 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 transmission¹⁸. We examined the receptor binding preference of 11 EA swine viruses isolated from 1979 to 2011 (Extended

¹School of Public Health, Li Ka Shing Faculty of Medicine, The University of Hong Kong, Hong Kong Special Administrative Region, China.

²Department of Infectious Diseases, St. Jude Children's Research Hospital, Memphis, TN, USA. ³Department of Microbiology and Immunology, University of Otago, Dunedin, New Zealand. ⁴Programme in Emerging Infectious Diseases, Duke–NUS Medical School, Singapore, Singapore.

⁵Institute of Medical Virology, Justus Liebig University Giessen, Giessen, Germany. ⁶Genomics Research Center, Academia Sinica, Taipei, Taiwan, Republic of China. ⁷Laboratory of Virology, Faculty of Veterinary Medicine, Department of Virology, Parasitology and Immunology, Ghent University, Merelbeke, Belgium. ✉e-mail: Richard.Webby@stjude.org; gavin.smith@duke-nus.edu.sg; hyen@hku.hk

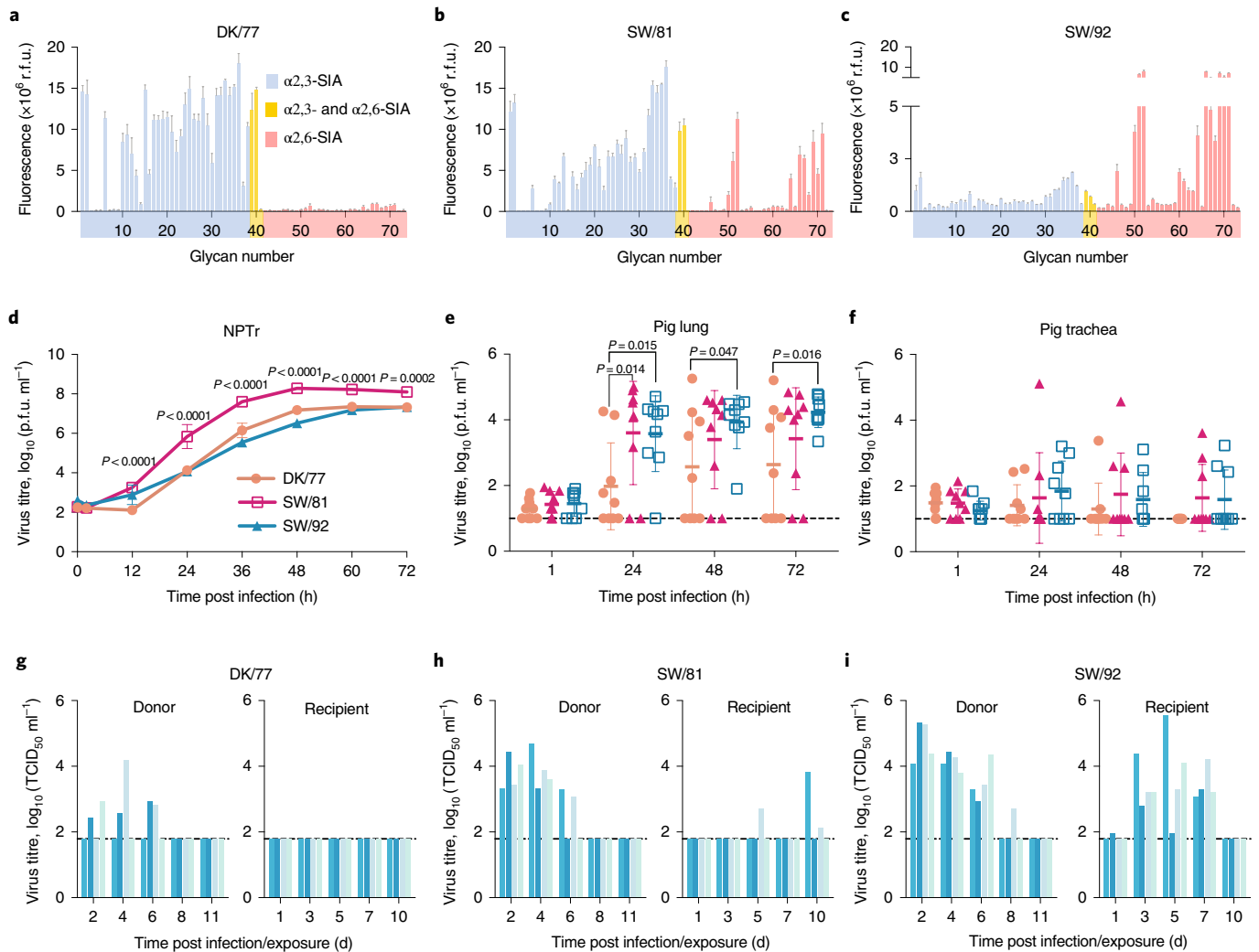


Fig. 1 | Avian and EA swine influenza viruses differed in receptor binding profiles, in vitro and ex vivo replication efficiencies and contact transmissibility in pigs. **a–c**, Glycan array analysis of the avian influenza virus DK/77 (**a**) and EA swine influenza viruses SW/81 (**b**) and SW/92 (**c**). Glycan array data are shown as the mean \pm s.d., calculated from six replicate spots of each glycan. SIA, sialoside; r.f.u., relative fluorescence units. **d**, Multicycle replication kinetics in NPTr cells. Data are the mean \pm s.d. from three repeats in one of two independently performed experiments; the significant P values ($P < 0.05$) between SW/81 and DK/77 are shown. **e, f**, Replication kinetics in ex vivo cultures of pig lungs (**e**) and trachea (**f**). Tissues from three healthy pigs were used to prepare lung and tracheal ($n = 3$ per pig for both) explants for the experiments. Each data point represents one explant sample. Data are the mean \pm s.d. from nine data points; the significant P values ($P < 0.05$) are indicated. **d–f**, Statistical differences were calculated by two-way ANOVA, followed by a Tukey's post-hoc test. The detection limit was 10 plaque forming units (p.f.u.) per ml and was shown with the dashed horizontal line. **g–i**, Contact transmissibility of DK/77 avian influenza virus (**g**) and 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 ($\log_{10}(\text{TCID}_{50} \text{ ml}^{-1}) = 1.789$; TCID $_{50}$, 50% tissue culture infectious dose) is shown by the horizontal dashed line.

Data Fig. 1). Compared with the avian H1N1 influenza virus A/duck/Bavaria/2/1977 (DK/77), which preferentially recognizes α -2,3-linked 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-linked 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 haemagglutinin (HA) fusion pH over time, with values ranging from pH 5.2 to pH 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-linked sialosides (Fig. 1a–c) were evaluated in vitro. In newborn pig trachea (NPTr) cells, SW/81 replicated to

a higher titre than the DK/77 and SW/92 viruses at all time points (two-way analysis of variance (ANOVA) and Tukey's post-hoc test, $P < 0.01$; Fig. 1d). In pig-lung explants, both SW/81 and SW/92 replicated to a higher titre than DK/77 at 24h post infection (h.p.i.; 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 the DK/77, SW/81 and SW/92 viruses were further compared in piglets (3–4 weeks old). Inoculated donor pigs were co-housed with naive contact pigs 1 d post inoculation (d.p.i.) at a 2:2 donor/contact ratio, with a total of four donors and four direct contacts in duplicated experiments. The area under the curve (AUC)

values were calculated to approximate the total viral load over the course of infection (Fig. 1g–i). In donors, the avian DK/77 virus replicated ($AUC = 1.58 \pm 1.59$, mean \pm s.d.) to a lower titre than the swine viruses SW/81 (3.81 ± 1.1) and SW/92 (5.8 ± 0.78 ; Kruskal–Wallis test, $P = 0.0031$; Dunn's post-hoc test, $P = 0.61$ and 0.013 , respectively). There was no significant difference in AUC between the 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 and contact pigs seroconverted (haemagglutination inhibition (HI) titre $< 1:20$). The SW/81 virus replicated better than DK/77, with infectious virus detected in the nasal cavity of 4/4 donors and seroconversion in 3/4 donors (HI titre of 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 the 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 the replicative capacity in pig nasal tissues, followed by developing efficient transmissibility among pigs.

Reconstructed ancestral EA swine H1N1 viruses possessed comparable phenotypes to 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 Figs. 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 Global Initiative on Sharing All Influenza Data (GISAID, <https://www.gisaid.org/>; accession number: 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 and Node 1 in Supplementary Fig. 2). RG-EA2 virus was constructed to represent the early evolutionary stage of EA swine viruses from 1979–1983 (Fig. 2a and Node 2 in Supplementary Fig. 2), whereas the RG-EA3 and RG-EA4 viruses represented EA swine viruses from 1984–1987 and 1988–1992, respectively (Fig. 2a and Nodes 3 and 4 in Supplementary Fig. 2).

The receptor binding profiles of the RG-EA1, RG-EA2, RG-EA3 and RG-EA4 viruses were compared. The avian-like precursor RG-EA1 showed exclusive binding to α -2,3-linked 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 RG-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 a significantly lower titre than the RG-EA2, -EA3 and -EA4 viruses at 12, 24 and 36 h.p.i. (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, determined using a minigenome assay, in human 293T cells (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. The RG-EA1, RG-EA2, RG-EA3 and RG-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 no seroconversion in the donors or contacts. RG-EA2, which genetically resembles EA swine viruses of the 1979–1983 era, was transiently detected in 2/4 donors and 1/4 contacts (Fig. 3b), with seroconversion in 1/4 donors (1:40). RG-EA3 and EA4, which resemble EA swine viruses of the post-1983 era 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). The viral load in the nasal swabs of donors inoculated with RG-EA1 ($AUC = 2.48 \pm 2.25$), RG-EA2 (0.91 ± 1.05), RG-EA3 (4.72 ± 3.23) or RG-EA4 (5.18 ± 1.39) were moderately different (Kruskal–Wallis test, $P = 0.051$). Collectively, we noted a positive correlation between the viral loads of the donor nasal swabs and viral transmissibility in pigs (Spearman's coefficient of correlation (r) = 0.90, $P = 0.014$; Fig. 3e), suggesting the importance of achieving high viral loads at the nasal epithelial cells before acquiring efficient transmissibility.

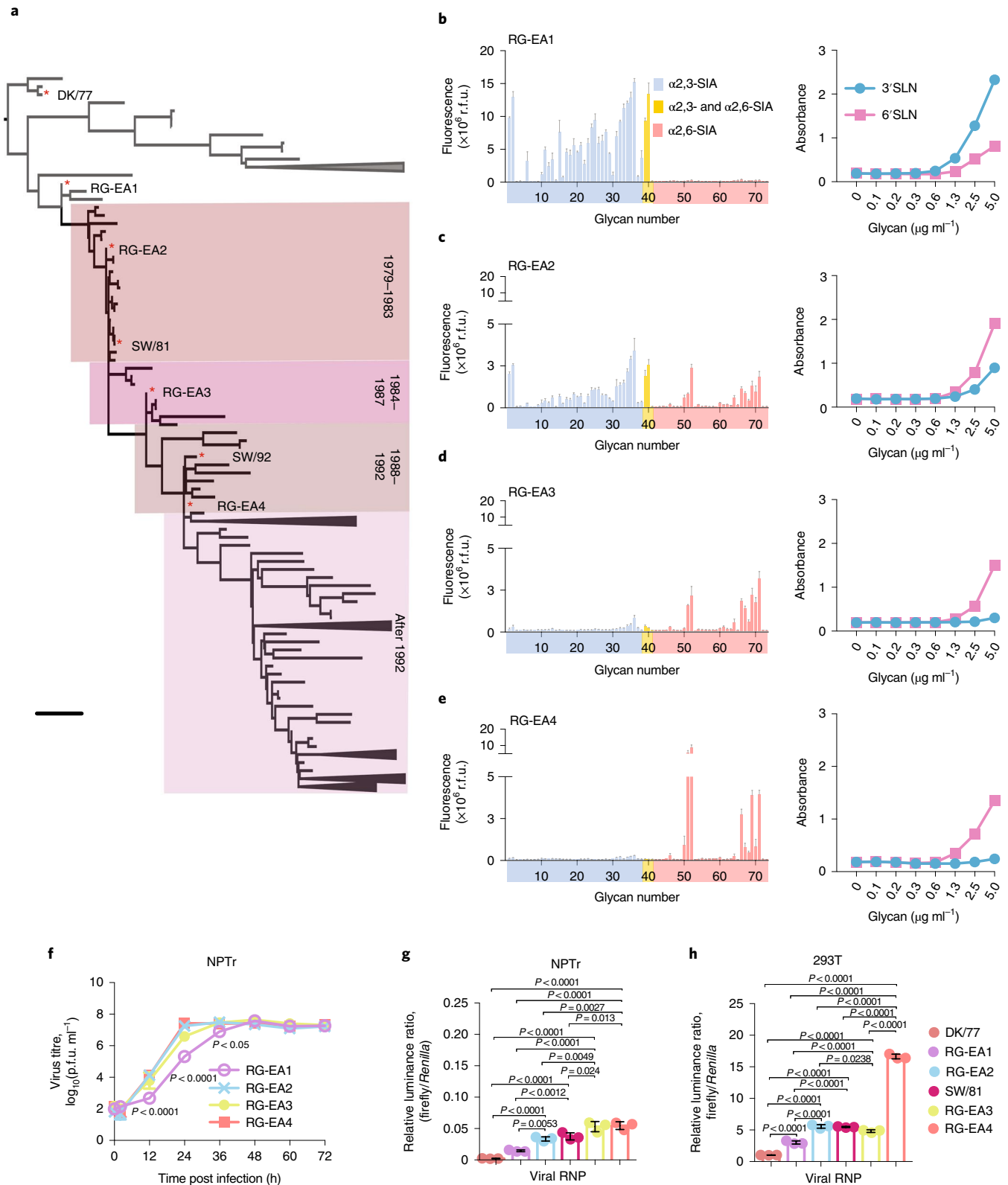
Introduction of HA and NA genes derived from RG-EA3 virus did not increase the transmissibility of RG-EA2 virus in pigs. Thirty-three amino-acid differences exist between the non-transmissible RG-EA2 and transmissible RG-EA3 viruses (Fig. 4a). Among the 12 amino-acid differences found in the HA protein of the RG-EA2 and RG-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 to 2016 (Fig. 4a). Introduction of these mutations into the HA protein of RG-EA2 reduced binding

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 DK/77. The 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 branch length representing 0.02 nucleotide substitutions per site. **b–e**, Receptor-binding profiles of the RG-EA1 (**b**), RG-EA2 (**c**), RG-EA3 (**d**) and RG-EA4 (**e**) viruses. The glycan array data are shown as the mean \pm s.d. fluorescence calculated from six replicate spots for each glycan (left) and the enzyme-linked immunosorbent assay results are shown as the mean absorbance (450 nm) from three replicates in one of two independently performed experiments (right). SIA, sialoside; 3'SLN, Neu5Ac α 2,3Gal β 1,4GlcNAc β -PAA-biotin; 6'SLN, Neu5Ac α 2,6Gal β 1,4GlcNAc β -PAA-biotin. **f**, Multicycle replication kinetics in NPTr cells. The mean from three repeats in one of two independently performed experiments is shown. Statistical differences were calculated by two-way ANOVA ($P < 0.0001$), followed by Tukey's post-hoc test, and the significant P values are shown. At 12 h post-infection, RG-EA1 was significantly different from RG-EA2 ($P < 0.0001$), RG-EA3 ($P < 0.0001$) and RA-EA4 ($P < 0.0001$); at 24 h post-infection, RG-EA1 was significantly different from RG-EA2 ($P < 0.0001$), RG-EA3 ($P < 0.0001$) and RG-EA4 ($P < 0.0001$), while RG-EA2 and RG-EA3 were also different ($P = 0.0013$); at 36 h post-infection, RG-EA1 was significantly different from RG-EA2 ($P = 0.0042$), RG-EA3 ($P = 0.0066$) and RG-EA4 ($P = 0.0182$). **g, h**, Polymerase activity in NPTr (**g**) and 293T (**h**) cells. Data show the mean \pm s.d. from three repeats (dots) in one of two independently performed experiments. Statistical differences were calculated by one-way ANOVA with Tukey's post-hoc test, and the significant P values ($P < 0.05$) are shown.

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 a Y344N mutation (N1 numbering; Fig. 4a). RG-EA1, RG-EA3 and RG-EA4 had a comparable Michaelis constant, which was higher than that of RG-EA2 (one-way ANOVA and Turkey's post-hoc test, $P=0.15$); however, the avian-like RG-EA1

had the highest velocity of an enzyme-catalysed reaction at infinite concentration of substrate ($P<0.01$; Supplementary Table 2).

To investigate whether the surface gene segments from RG-EA3 dictate the different transmission phenotypes of the RG-EA2 and RG-EA3 viruses, we generated RG-EA2 \times EA3^{SG} virus, which contained the internal genes from RG-EA2 and the surface genes



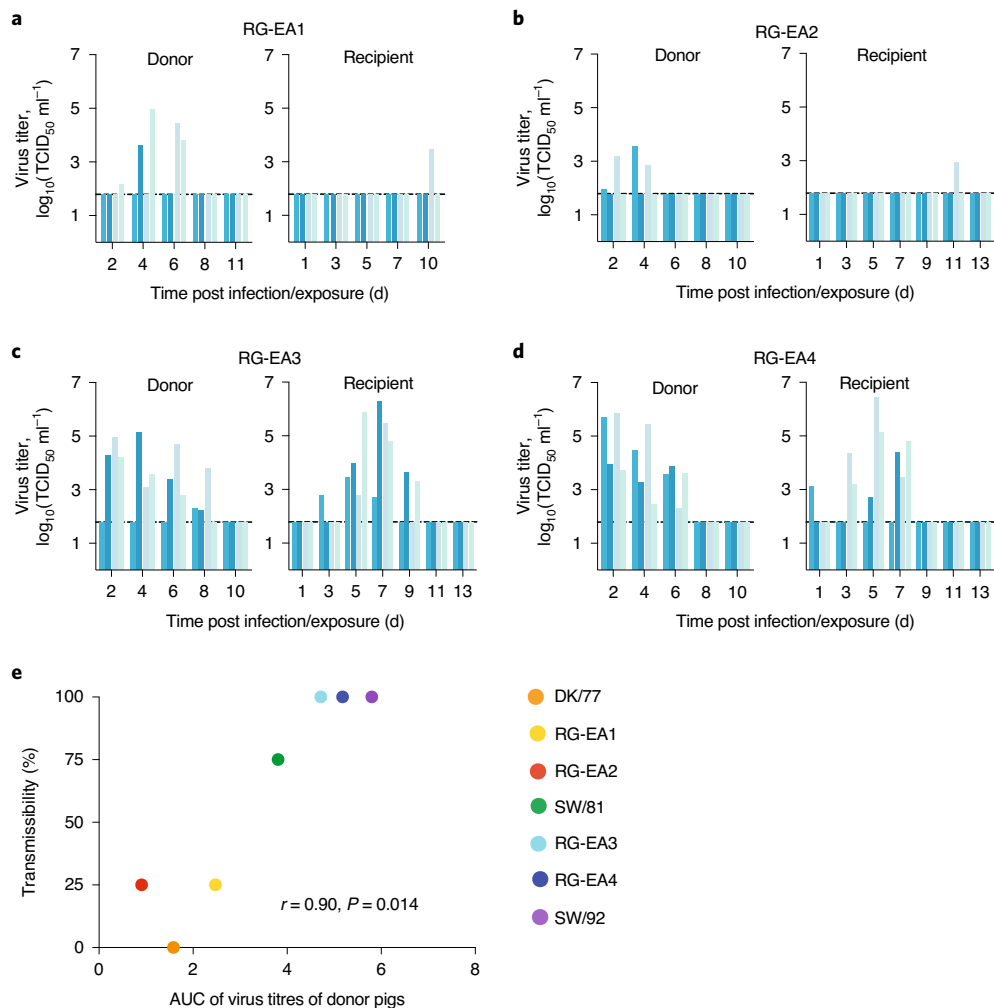


Fig. 3 | EA swine H1N1 viruses acquired efficient pig-to-pig transmissibility after 1983. **a–d**, Transmissibility of RG-EA1 (**a**), RG-EA2 (**b**), RG-EA3 (**c**) and RG-EA4 (**d**) in pigs. Contact transmission experiments were performed in duplicate, with a total of four donors and four direct contacts. Pigs housed in the same pen are denoted by the same colour. The limit of detection ($\log_{10}(\text{TCID}_{50} \text{ ml}^{-1}) = 1.789$) is shown by the horizontal dashed line. **e**, Two-sided Spearman's rank correlation coefficient analysis was used to evaluate the correlation between the viral loads in the donor nasal swabs (mean AUC from four pigs) and onward transmissibility (percentage infected out of four exposed contact pigs).

derived from RG-EA3. The RG-EA2×EA3^{SG} virus was transiently detected in the nasal swabs of 2/4 inoculated donors, with peak titres detected 8 and 10 d.p.i. (Fig. 4c), respectively; none of the donors had seroconverted by 11 d.p.i. RG-EA2×EA3^{SG} was transiently detected in 2/4 contacts with peak titres at late time points, that is, 11 and 13 d post exposure (d.p.e.), respectively (Fig. 4c), with none of the contact pigs seroconverting. These results suggest that introduction of the HA and NA genes of the RG-EA3 virus was insufficient to confer increased transmissibility to the RG-EA2 virus.

Molecular determinants associated with efficient transmission in pigs reside in the internal genes. Among the 20 amino acids that differentiate the internal proteins of the RG-EA2 and RG-EA3 viruses, polymerase basic protein 1 (PB1)-Q621R and nucleoprotein (NP)-R351K were detected at high frequencies (>90%) among the 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 the 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 a higher titre than the RG-EA2 and RG-EA2^{NP-R351K} viruses at 24 ($P = 0.019$) and 36 h.p.i. ($P = 0.024$), respectively (Fig. 5c). Interestingly, the NP-R351K mutation facilitated the accumulation of viral RNP in the nucleus at earlier time points compared with the RG-EA2 virus ($P < 0.05$; Fig. 5d,e).

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

We focused on these two mutations and compared the transmissibility of the RG-EA2^{NP-R351K} and RG-EA2^{PB1-Q621R, NP-R351K} viruses in pigs. RG-EA2^{NP-R351K} was detected in the nasal swabs of all donors and contacts (Fig. 5g), and seroconversion was

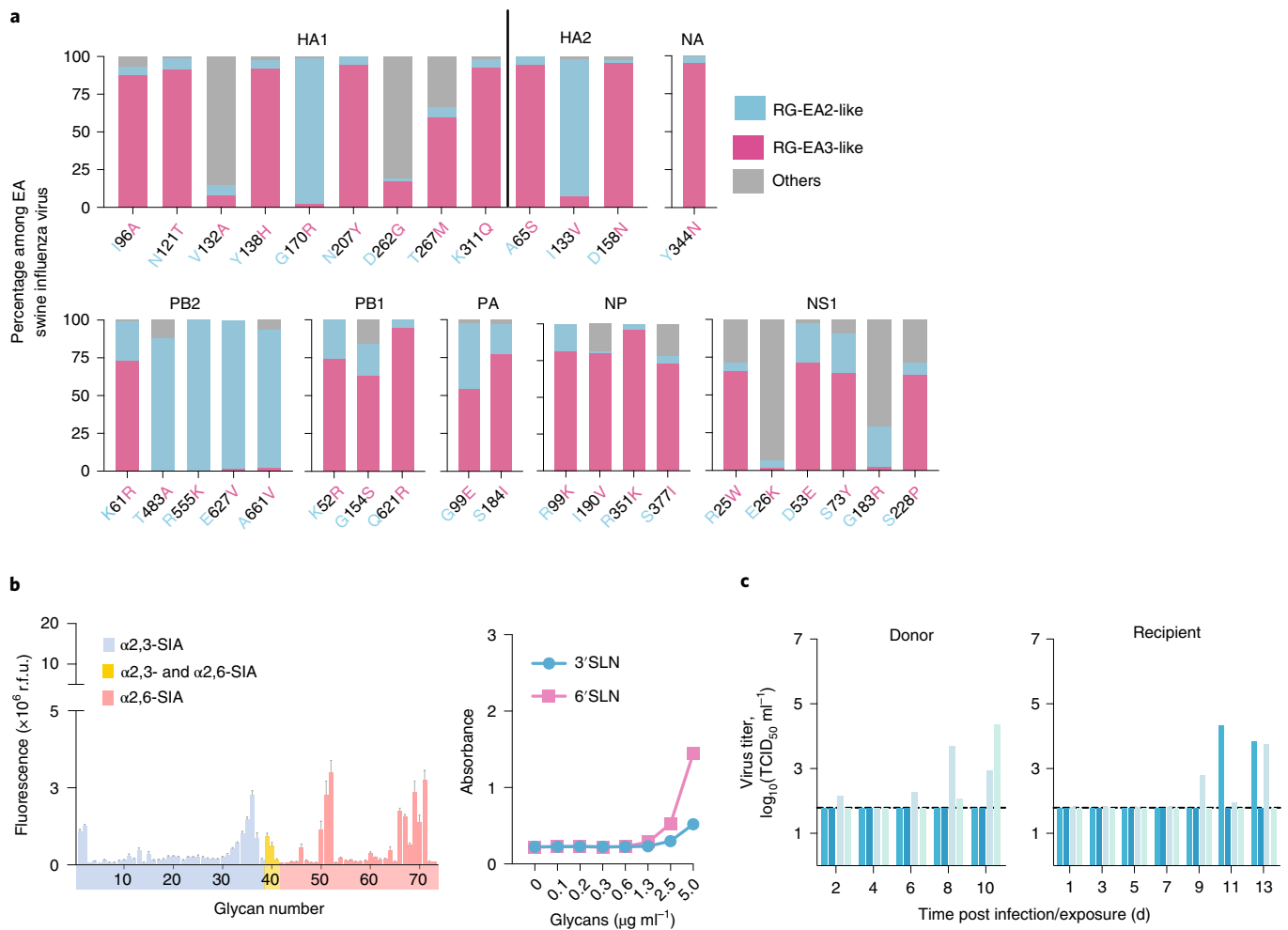


Fig. 4 | Introduction of HA and NA genes derived from the RG-EA3 virus did not enhance the efficient contact transmissibility of the RG-EA2 virus in pigs. **a**, Detection frequency of the 33 amino acids that differed between the RG-EA2 and RG-EA3 EA swine viruses among PB2 ($n = 417$), PB1 ($n = 381$), polymerase acidic protein (PA; $n = 382$), HA1 and HA2 ($n = 454$), NP ($n = 495$), NA ($n = 429$) and non-structural protein 1 (NS1; $n = 391$) of EA swine influenza A viruses isolated from 1979 to 2016. 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,HA2-D158N}. The glycan array data are shown as the mean + s.d., calculated from six replicate spots for each glycan (left), and the enzyme-linked immunosorbent assay results are shown as the mean absorbance (450 nm) from three replicates in one of two independently performed experiments (right). SIA, sialoside. **c**, Contact transmissibility of RG-EA2 \times EA3^{SG} virus containing the internal genes from RG-EA2 and the surface genes derived from RG-EA3 in pigs. The 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 ($\log_{10}(\text{TCID}_{50} \text{ ml}^{-1}) = 1.789$) is shown by the horizontal dashed line.

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 in the nasal swabs of all donors and contacts (Fig. 5h), with seroconversion in all donors (1:80 to 1:160) and contacts (1:160 to 1:640). The total amount of virus shed by the RG-EA2^{NP-R351K}-inoculated donors ($\text{AUC} = 2.74 \pm 2.63$) 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 the contact pigs, the viral load shed by RG-EA2^{NP-R351K}-infected contacts ($\text{AUC} = 3.04 \pm 1.57$) was slightly lower than that of the contacts infected with RG-EA2^{PB1-Q621R,NP-R351K} (5.4 ± 1.31 ; two-sided Mann-Whitney test, $P = 0.11$). Next-generation sequencing analyses were performed on the peak-titre 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 one pig (Supplementary Table 3). Together, these results show that introduction of the NP-R351K mutation is 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 number: EPI_ISL_1055769) and full genome of A/swine/Belgium/2/1979 (designated as Be02-lung; GISAID accession number: 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 the 33 amino acids that differed between the 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

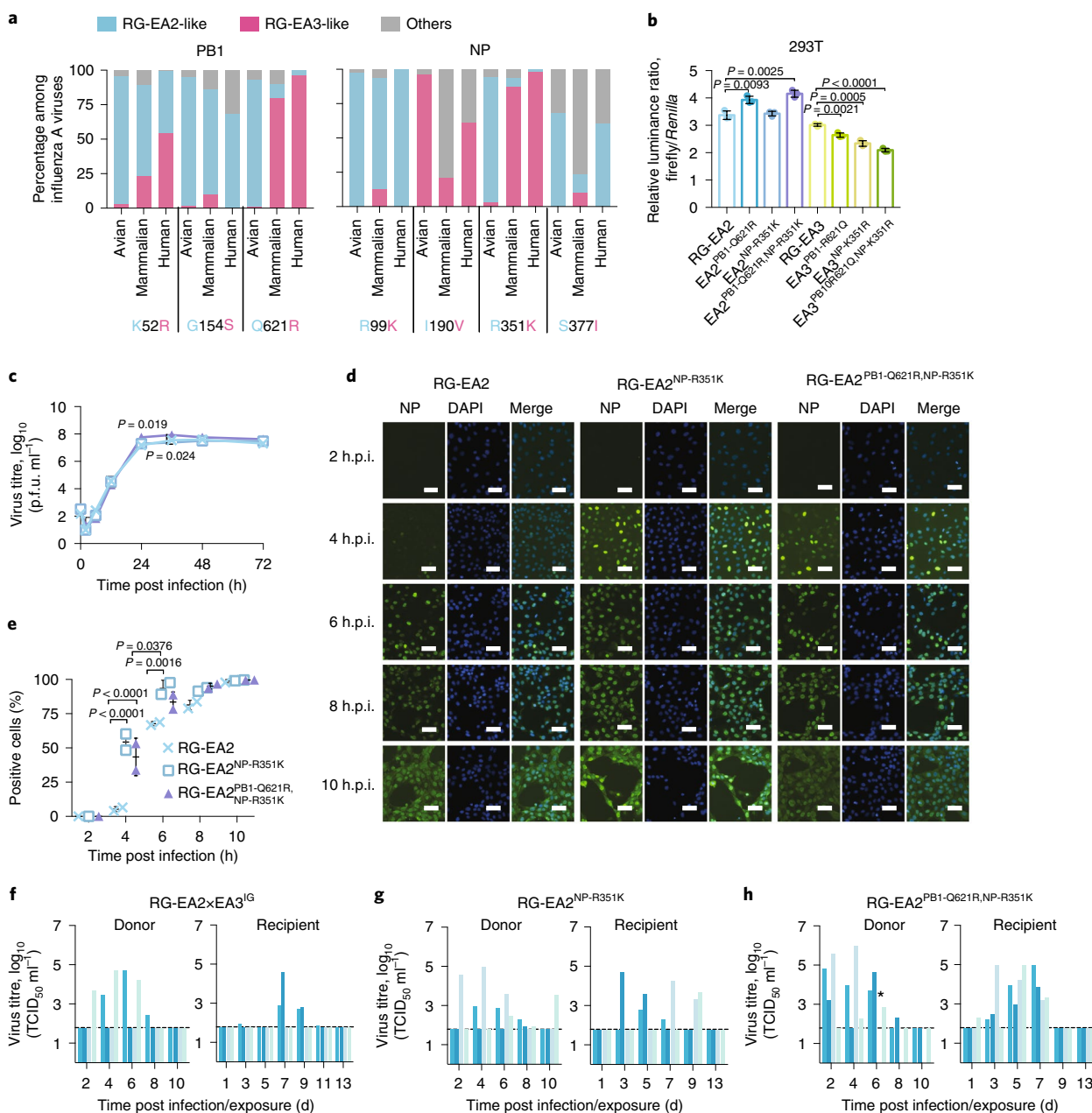


Fig. 5 | The NP-R351K mutation in EA swine influenza viruses was the minimal molecular change required to facilitate the transmission of the non-transmissible RG-EA2 virus. a, Detection frequency of Q/R at PB1 residue 621 ($n = 74,095$) and R/K at NP residue 351 ($n = 80,188$) in avian, human and other mammalian influenza A viruses. **b**, The effects of amino-acid substitutions at PB1 residue 621 and NP residue 351 on the viral polymerase activity were determined using a minigenome assay in 293T cells. Data are the mean \pm s.d. from three repeats (dots) in one of two independently performed experiments. A one-way ANOVA with Tukey's post-hoc test was performed, and the significant P values ($P < 0.05$) are shown. **c**, Multicycle replication kinetics of recombinant viruses in NPTr cells. Data are the mean from three repeats in one experiment. A two-way ANOVA with Tukey's post-hoc test was performed, and the significant P values ($P < 0.05$) are shown, including the comparison between RG-EA2 and RG-EA2^{PB1-Q621R,NP-R351K} at 24 h.p.i. and between RG-EA2^{NP-R351K} and RG-EA2^{PB1-Q621R,NP-R351K} at 36 h.p.i. **d,e**, Nucleus localization of NP protein at different time points in NPTr cells infected with RG-EA2, RG-EA2^{NP-R351K} and RG-EA2^{PB1-Q621R,NP-R351K} at a multiplicity of infection (m.o.i.) of five. **d**, Representative images from two replicates of one experiment are shown. Scale bars, 100 μ M. **e**, Mean \pm s.d. of the NP-positive rate (percentage of NP-positive cells per 4,6-diamidino-2-phenylindole (DAPI)-positive cells) from two replicates (dots) of one experiment calculated for different times following infection. Statistical significance was calculated using a two-way ANOVA with Tukey's post-hoc test and significant P values ($P < 0.05$) are shown. **f-h**, Contact transmissibility of the RG-EA2 \times EA3^G (**f**), RG-EA2^{NP-R351K} (**g**) and RG-EA2^{PB1-Q621R,NP-R351K} viruses in pigs. The 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 ($\log_{10}(\text{TCID}_{50} \text{ ml}^{-1}) = 1.789$) is shown by the horizontal dashed line; *, donor in the RG-EA2^{PB1-Q621R,NP-R351K} group that was euthanized on day 4 post infection due to respiratory distress with abdominal distention, severe lethargy and vomiting.

RG-EA2; Extended Data Fig. 2). In comparison with RG-EA2, the partial sequence of the Be01-lung sample showed four amino-acid differences at HA1 residues that are highly variable among EA swine

viruses (Extended Data Fig. 2). 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 EA swine viruses from the 1979–1983 era. Additional serial passages of recombinant RG-EA2 virus carrying the HA and NA genes derived from the Be02-lung (designated as RG-Be02-lung^{SG}×EA2^{CG}) or RG-EA1, RG-EA2, RG-EA3 and 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 before 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 recognition of both α -2,3- and α -2,6-linked sialosides (RG-EA2) to recognition of only α -2,6-linked sialosides (RG-EA3 and RG-EA4) and gradually increased polymerase activity, which probably 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, which differed in 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 illustrate 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 genes²³ and 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 the corresponding algorithms is dependent on the number of available sequences, direct sequencing of EA swine influenza viruses from archived pig-lung homogenates from 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 (1979–1983) possessed inefficient transmission potential among pigs. Transmission may be facilitated by the housing conditions of the pig farms—with close contact, contaminated surfaces and shared water—as well as 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 (for example, overlapping habitats) and viral factors that permitted replication of the avian-like H1N1 virus in the pig respiratory epithelial cells (for example, dual receptor binding specificity possessed by the early EA swine viruses). The fact that the introduction of the NP-R351K mutation was able to facilitate efficient transmission of RG-EA2, an early EA swine virus that possesses dual binding specificity, indicates that an exclusive α -2,6-linked sialoside 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 descended from the 1918 pandemic influenza virus²⁵, and the 1918 NP protein was probably of avian origin²⁶. Consistent with its putative avian origin, the 1918 virus contained NP-R351. The NP-R351K mutation was quickly fixed in the human seasonal A(H1N1) viruses from 1918 to 1957, with 39 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) of the A(H2N2)

sequences from 1957 to 1968 and in 99.93% (26,895/26,915) of the 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 virus⁶. Given that 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 continued 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, which is known to interact with importin- α proteins that mediate nucleus importation^{27,28}. In addition, 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 (ref. ³⁰) and EA swine influenza viruses¹⁴, 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 virus⁶. 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}, which 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 an early stage as avian influenza viruses adapt to mammalian hosts. Continuous surveillance that is strategically coordinated with risk assessment studies may help identify viruses with pandemic potential before they become fully adapted in mammalian species. Intervention strategies implemented before full adaptation 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 animal biosafety level 2+ 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 anticipate 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 0–4 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 probable that these viruses existed in nature but remain unsampled.

All experimental work on the reconstructed EA swine viruses were performed in biosafety level 2 and biosafety level 2-enhanced laboratories with restricted access following approved standard operating procedures. Refresher training of the study personnel is provided annually. All reagents (plasmids and viruses) were stored in locked freezers in rooms with restricted access. During the 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 US (this study was funded by a US 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 US 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 US 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 maintained in minimal essential medium (MEM) supplemented with 10% FCS, 1% penicillin-streptomycin (P/S) and 1% vitamins, and buffered with 25 mM HEPES. Human embryonic kidney 293T (293T) cells were obtained from the ATCC and maintained in Opti-MEM medium supplemented with 5% FCS and 1% P/S. The NPTr cells were obtained from Istituto Zooprofilattico Sperimentale, della Lombardia e dell'Emilia Romagna and maintained in MEM supplemented with 10% FCS, 1% P/S and 1% sodium pyruvate. African green monkey kidney (Vero) cells were obtained from the ATCC and maintained in MEM supplemented with 10% FCS and 1% P/S. The cells were cultured at 37 °C in 5% CO₂. All of the cells used in the study tested negative in routine tests for *Mycoplasma* species using real-time PCR conducted by the Faculty Core Facility, the University of Hong Kong (HKU).

Avian influenza virus DK/77 and EA swine influenza A(H1N1) viruses A/swine/Netherlands/3/1980, SW/81, A/swine/Netherlands/12/1985, A/swine/Italy/670/1987 and SW/92 were provided by R. Webster from St. Jude Children's Research Hospital, Memphis, TN, USA. The EA swine influenza virus A/swine/Arnsberg/6554/1979 (H1N1) was provided by S. Pleschka from Justus-Liebig-University Giessen, Germany. The 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 (Be01-lung and Be02-lung) were provided by K. Van Reeth from Ghent University, Mellebeke, Belgium. Viruses were cultured at an m.o.i. of 0.005 on MDCK cells cultured in MEM supplemented with 0.3% BSA (Sigma-Aldrich), 1% P/S, 1% vitamin, 25 mM HEPES and 1 µg ml⁻¹ L-1-tosylamide-2-phenylmethyl chloromethyl ketone-treated trypsin (Sigma-Aldrich). 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 their concentration (p.f.u. ml⁻¹).

Ancestral sequence reconstruction. Over 2,000 nucleotide sequences of HA-H1NX (*n* = 410), NA-HXN1 (*n* = 443) and HXNX (for all the six internal genes: *PB2* (*n* = 454), *PB1* (*n* = 427), *PA* (*n* = 397), *NP* (*n* = 417), *MP* (*n* = 495) and *NS* (*n* = 400) were downloaded from the NCBI database before 2014. Sequences that were smaller 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 analyses were performed using RAXML version 8.0 (ref. 32) in Geneious R9.0.3 (Biomatters Ltd.) to generate ten 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)³³, which utilizes baseml to reconstruct maximum-likelihood 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 visualize the evolutionary positions of the reconstructed ancestral nucleotide sequences, we included those sequences in the individual gene datasets and constructed maximum-likelihood 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, which have shared nodes for RG-EA2 and RG-EA3. We then used the treesub programme³⁴ 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 synthesised by GeneArt and Synbio Technologies. The synthesised genes were cloned into the pHW2000 vector by mega-primer PCR as described^{35,36}. Recombinant viruses were generated by transfecting eight plasmids into 293T cells. The virus titres were determined on MDCK cells in six-well plates by plaque assay. All rescued viruses were propagated twice on MDCK cells at an m.o.i. of 0.005 to prepare virus stocks. Their genomes were confirmed via Sanger sequencing.

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

Glycan microarray. A synthesised glycan microarray comprising 38 glycans of α2,3-linked sialosides (glycans 1–38), three glycans containing both α2,3- and α2,6-linked sialosides (glycans 39–41) and 32 glycans of α2,6-linked sialosides (glycans 42–73; Extended Data Fig. 3) was utilized for virus-binding studies using viruses that were inactivated by 0.025% formaldehyde for 7 d at 4 °C. The glycan array slides were blocked by SuperBlock (PBS) blocking buffer (Pierce) for 1 h at room temperature. To avoid the influence of NA, the NA inhibitor zanamivir was added to the virus stocks at a final concentration of 10 µM. Formalin-inactivated viruses were diluted to 64 HA U 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 r.p.m.) for 1 h. 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). The slides were scanned by an InnoScan 710 AL microarray scanner (Innopsys) equipped with two laser sources, visible wavelength 635 nm and 532 nm. The data were analysed using the GenePix Pro 6.0 software (Molecular Devices).

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 with virus at an m.o.i. of 0.001 in 1 ml infection medium, with two repeats. The supernatants were collected at 0, 2, 12, 24, 36, 48, 60 and 72 h.p.i. and titrated on MDCK cells in a six-well plate by plaque assay.

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

Site-directed mutagenesis. Primers designed using QuickChange primer design (<https://www.agilent.com/store/primerDesignProgram.jsp>) were used to introduce specific mutations in the *HA*, *PB1* and *NP* genes. These specific primers are listed in Supplementary Table 7. The PCR reactions were performed using a QuickChange multi site-directed mutagenesis kit according to the manufacturer's instructions (Agilent).

Minigenome assay. Polymerase activity was evaluated using a minigenome assay^{21,22}. 293T or NPTr cells in six-well plates were transfected with 1 µg of the plasmids encoding the viral *PB2*, *PB1* or *PA* genes and 2 µg of the plasmid that encodes the viral *NP* gene, 1 µg reporter plasmid encoding firefly luciferase flanked by the noncoding region of the influenza *M* gene driven by either the human Polymerase I promoter³⁷ or swine Polymerase I promoter³⁸ and 0.1 µg pHRL-CMV plasmid (*Renilla* luciferase driven by the CMV promoter) using TransIT (MIRUS) reagent according to the manufacturer's recommendation. After 24 h, the luciferase activity was measured with the dual-luciferase reporter system (Promega) on a SpectraMax iD5 Multi-Mode Microplate Reader (Molecular Devices).

Transmissibility of EA viruses in pigs. Male and female Yorkshire crossbred piglets (Midwest Research Swine LLC.), 3–4 weeks old, were randomized into different groups. Before performing the experiments, the pigs were confirmed as seronegative for influenza A virus NP protein (ID.vet) and showed HI titres of ≤1:10 for the homologous EA swine influenza virus. Two donor pigs were intranasally inoculated with 1 × 10⁶ p.f.u. virus in 1 ml PBS under anaesthesia and two naive pigs were introduced to co-house with two donors 1 d.p.i. Each experiment was independently performed in duplicate with a total of four donors and four direct contacts. No statistical methods were used to predetermine the sample sizes, but our sample sizes are similar to those reported in previous publications^{39,40}. Nasal swabs were collected from both nostrils of the donors 2, 4, 6, 8 and 10 or 11 d.p.i. The donors were euthanized 11 d.p.i. for post-infection sera collection. Nasal swabs were collected from both nostrils of all contact pigs 1, 3, 5, 7, 9, 10 and 11 or 13 d.p.e. The contact pigs were euthanized 10 or 13 d.p.e. for post-exposure sera collection. All nasal swabs were placed into 1 ml of viral transport medium and stored at -80 °C. The TCID₅₀ of the nasal swabs was determined by titration in MDCK cells⁴¹.

HI assay. Antibodies to the homologous virus from donor and contact pigs were tested by HI assay following standard procedures. Pig sera were treated with receptor-destroying enzyme, RDE (Denka Seiken), for 18–20 h in a 37 °C water bath and then inactivated for 30 min at 56 °C. The final dilution of RDE-treated sera was 1:10. After preparing two-fold serial dilutions in a 96-well V-bottom microplate, the corresponding virus at the volume of 25 µl containing four HA units was added to the plates and then incubated for 30 min at room temperature. Subsequently, 50 µl of 0.5% turkey red blood cells (LAMPURE Biological Laboratories, Inc.) was added, and the plate was slightly shaken and incubated at room temperature for 30 min. The HI titres 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 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 ($n=396$) and NS ($n=391$)—of EA swine influenza viruses isolated from 1979 to 2016 were used for specific amino-acid prevalence analysis. To investigate whether mutations (that is, 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 were downloaded from GISAID. Sequences were aligned using MAFFT¹² at the CIPRES science gateway (version 3)¹³. The number of amino acids were counted using BioEdit (version 7.0).

Serial passages in vitro and in ovo. To compare mutations that may emerge after passaging in vitro and in ovo, the recombinant viruses RG-Be02-lung^{SG}×EA2^{IG}, RG-EA1, RG-EA2, RG-EA3 and RG-EA4 were serially passaged in MDCK and NPTr cells as well as ten-day-old embryonated chicken eggs using transfection supernatant as the starting material. In vitro infections were performed using an m.o.i. of 0.001–0.005 at each passage and in ovo infections were performed using 1×10^4 p.f.u. in 0.1 ml. Three independent serial passages were performed in parallel for each condition. The culture supernatants or allantoic fluids harvested from passage three were analysed by next-generation sequencing.

Next-generation sequencing. Viral RNA was extracted from samples using a QIAamp viral RNA mini kit (Qiagen). The RNA was transcribed into complementary DNA using SuperScript III reverse transcriptase (Invitrogen). The gene segments were amplified by Q5 High Fidelity DNA polymerase (NEB) using a pair of primers (454-Tag1-U12, GCCGGAGCTCTGCAGATATCAGCRAAAGCAGG; and 454-Tag1-U1, GCCGGAGCTCTGCAGATATCAGTAGAAACAAGG). Libraries were prepared based on the Nextera DNA flex library preparation standard protocol¹⁴. The PCR products (300 ng) were cleaved and tagged by bead-linked transposome at 55°C and the tagmented DNA was amplified with a pair of indexes using a five-cycle PCR programme. After purification, the libraries were finally eluted using 32 µl resuspension buffer. The libraries were quantified using an Agilent fragment analyser automated CE system with the high-sensitivity NGS fragment analysis 474 kit (Agilent). Pool libraries (100–150 pM) containing index adaptors allowing multiplex sequencing of four samples per iSeq 100 i1 cartridge were run on an iSeq 100 sequencing system (Illumina). The output data were analysed on CLC Genomic Workbench (version 20; Qiagen) using 1% as the variant-calling threshold. The nucleotide substitutions with a frequency of 5% or more were further analysed.

NA kinetics. The NA kinetics analysis is described in the Supplementary Methods.

Immunofluorescence staining. To investigate the effect of mutations in NP on the cellular distribution of NP during the virus life cycle, NPTr cells were infected with virus at an m.o.i. of five. Infections were stopped at 2, 4, 6, 8, 10 and 12 h.p.i. by fixation in 4% paraformaldehyde (Electron Microscopy Sciences). The cells were then permeabilized with 0.1% Triton X-100 in PBS for 30 min and labelled with a 1:200 dilution of mouse monoclonal IgG2a influenza A NP antibody (Santa Cruz Biotech, Inc) at 4°C overnight. The cells were subsequently incubated with a 1:200 dilution of fluorescein isothiocyanate-labelled goat anti-mouse IgG/IgM (BD Pharmingen) for 3 h at room temperature in the dark. After three washes, the nuclei were counterstained at a 1:1,000 dilution of DAPI in SlowFade gold antifade mountant (Thermo Fisher Scientific) at room temperature for 15 min. Cell imaging and fluorescence were carried out using a Nikon Eclipse Ti-S fluorescence microscope equipped with a Nikon DS-Qi2 camera and the iNIS-Elements BR imaging software (version 4.40). The positive cells were enumerated using Fiji¹⁵.

Statistical analyses. Blinding on the conditions of the experiments was not performed during data collection and analyses. 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 titres over time, followed by Tukey's multiple-comparisons post-hoc tests. The AUC was calculated from the nasal swabs of the donor and contact pigs. A two-sided Mann–Whitney *U*-test was performed to compare two groups, and Kruskal–Wallis and Dunn's multiple-comparisons post-hoc tests 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 analysed in Microsoft Excel for Mac, version 16.28, and GraphPad Prism version 8.4.1 for Windows. All statistical parameters for specific analyses are shown in the corresponding figure legends. Statistically significant *P* values ($P < 0.05$) are indicated in the figures.

Reporting Summary. Further information on research design is available in the Nature Research Reporting Summary linked to this article.

Data availability

Sequences of Sanger sequencing validated in the study were deposited to the public database GISAID under the following accession numbers: DK/77, EPI_ISL_539860; A/swine/Arnsberg/6554/1979, EPI_ISL_539869; Be01-lung, EPI_ISL_1055769; Be02-lung, EPI_ISL_1055773; A/swine/Netherlands/3/1980, EPI_ISL_539861; SW/81, EPI_ISL_539862; A/swine/Netherlands/12/1985, EPI_ISL_539863; SW/92, 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; and 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. Source data are provided with this paper.

Received: 18 April 2021; Accepted: 7 September 2021;

Published online: 26 October 2021

References

- Kuiken, T. et al. Host species barriers to influenza virus infections. *Science* **312**, 394–397 (2006).
- Parrish, C. R. et al. Cross-species virus transmission and the emergence of new epidemic diseases. *Microbiol. Mol. Biol. Rev.* **72**, 457–470 (2008).
- Taubenberger, J. K. & Kash, J. C. Influenza virus evolution, host adaptation, and pandemic formation. *Cell Host Microbe* **7**, 440–451 (2010).
- Lipsitch, M. et al. Viral factors in influenza pandemic risk assessment. *Elife* **5**, e18491 (2016).
- Long, J. S., Mistry, B., Haslam, S. M. & Barclay, W. S. Host and viral determinants of influenza A virus species specificity. *Nat. Rev. Microbiol.* **17**, 67–81 (2019).
- Smith, G. J. et al. Origins and evolutionary genomics of the 2009 swine-origin H1N1 influenza A epidemic. *Nature* **459**, 1122–1125 (2009).
- Pensaert, M., Ottis, K., Vandeputte, J., Kaplan, M. M. & Bachmann, P. A. 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).
- Brown, I. H. History and epidemiology of swine influenza in Europe. *Curr. Top. Microbiol. Immunol.* **370**, 133–146 (2013).
- Krumbholz, A. et al. Origin of the European avian-like swine influenza viruses. *J. Gen. Virol.* **95**, 2372–2376 (2014).
- Joseph, U., Vijaykrishna, D., Smith, G. J. D. & Su, Y. C. F. Adaptive evolution during the establishment of European avian-like H1N1 influenza A virus in swine. *Evol. Appl.* **11**, 534–546 (2018).
- Vijaykrishna, D. et al. Long-term evolution and transmission dynamics of swine influenza A virus. *Nature* **473**, 519–522 (2011).
- 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).
- 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).
- Dornfeld, D., Petric, P. P., Hassan, E., Zell, R. & Schwemmler, M. Eurasian avian-like swine influenza A viruses escape human MxA restriction through distinct mutations in their nucleoprotein. *J. Virol.* **93**, e00997-18 (2019).
- 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).
- Thornton, J. W., Need, E. & Crews, D. Resurrecting the ancestral steroid receptor: ancient origin of estrogen signaling. *Science* **301**, 1714–1717 (2003).
- Ducatez, M. F. et al. Feasibility of reconstructed ancestral H5N1 influenza viruses for cross-clade protective vaccine development. *Proc. Natl Acad. Sci. USA* **108**, 349–354 (2011).
- Rogers, G. N. & Paulson, J. C. 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).
- Baumann, J., Kouassi, N. M., Foni, E., Klenk, H. D. & Matrosovich, M. H1N1 swine influenza viruses differ from avian precursors by a higher pH optimum of membrane fusion. *J. Virol.* **90**, 1569–1577 (2016).
- 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).
- Pleschka, S. et al. A plasmid-based reverse genetics system for influenza A virus. *J. Virol.* **70**, 4188–4192 (1996).
- 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).
- Thornton, J. W. Resurrecting ancient genes: experimental analysis of extinct molecules. *Nat. Rev. Genet.* **5**, 366–375 (2004).

24. Giles, B. M. & Ross, T. M. A computationally optimized broadly reactive antigen (COBRA) based H5N1 VLP vaccine elicits broadly reactive antibodies in mice and ferrets. *Vaccine* **29**, 3043–3054 (2011).
25. Taubenberger, J. K. & Morens, D. M. The 1918 influenza pandemic and its legacy. *Cold Spring Harb. Perspect. Med.* **10**, a038695 (2020).
26. Reid, A. H., Fanning, T. G., Janczewski, T. A., Lourens, R. M. & Taubenberger, J. K. Novel origin of the 1918 pandemic influenza virus nucleoprotein gene. *J. Virol.* **78**, 12462–12470 (2004).
27. Gabriel, G., Herwig, A. & Klenk, H. D. Interaction of polymerase subunit PB2 and NP with importin α 1 is a determinant of host range of influenza A virus. *PLoS Pathog.* **4**, e11 (2008).
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).
29. Zimmermann, P., Mänz, B., Haller, O., Schwemmler, M. & Kochs, G. The viral nucleoprotein determines Mx sensitivity of influenza A viruses. *J. Virol.* **85**, 8133–8140 (2011).
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).
31. Ashenberg, O., Padmakumar, J., Doud, M. B. & Bloom, J. D. Deep mutational scanning identifies sites in influenza nucleoprotein that affect viral inhibition by MxA. *PLoS Pathog.* **13**, e1006288 (2017).
32. Stamatakis, A. RAxML version 8: a tool for phylogenetic analysis and post-analysis of large phylogenies. *Bioinformatics* **30**, 1312–1313 (2014).
33. Finnigan, G. C., Hanson-Smith, V., Stevens, T. H. & Thornton, J. W. Evolution of increased complexity in a molecular machine. *Nature* **481**, 360–364 (2012).
34. Tamuri, A. U. Treesub: annotating ancestral substitution on a tree. *GitHub* <https://github.com/tamuri/treesub> (2013).
35. 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).
36. Hoffmann, E., Neumann, G., Kawaoka, Y., Hobom, G. & Webster, R. G. A DNA transfection system for generation of influenza A virus from eight plasmids. *Proc. Natl Acad. Sci. USA* **97**, 6108–6113 (2000).
37. Hoffmann, E., Neumann, G., Hobom, G., Webster, R. G. & Kawaoka, Y. ‘Ambisense’ approach for the generation of influenza A virus: vRNA and mRNA synthesis from one template. *Virology* **267**, 310–317 (2000).
38. Moncorge, O. et al. Investigation of influenza virus polymerase activity in pig cells. *J. Virol.* **87**, 384–394 (2013).
39. Zhu, H. et al. Infectivity, transmission, and pathology of human-isolated H7N9 influenza virus in ferrets and pigs. *Science* **341**, 183–186 (2013).
40. Nishiura, H., Yen, H. L. & Cowling, B. J. Sample size considerations for one-to-one animal transmission studies of the influenza A viruses. *PLoS ONE* **8**, e55358 (2013).
41. Reed, L. J. & Muench, H. A simple method of estimating fifty percent endpoints. *Am. J. Hyg.* **27**, 493–497 (1938).
42. 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).
43. Miller, M. A., Pfeiffer, W. & Schwartz, T. Creating the CIPRES Science Gateway for inference of large phylogenetic trees. In *2010 Gateway Computing Environments Workshop (GCE)* 1–8 (IEEE, 2010).
44. Bruinsma, S. et al. Bead-linked transposomes enable a normalization-free workflow for NGS library preparation. *BMC Genomics* **19**, 722 (2018).
45. Schindelin, J. et al. Fiji: an open-source platform for biological-image analysis. *Nat. Methods* **9**, 676–682 (2012).

Acknowledgements

We thank the St. Jude core facilities, including the Animal Resources Centre and the Veterinary Pathology Core for technical support. This study was supported by contract HHSN272201400006C from NIAID, NIH, USA, and the Theme-Based Research Scheme (grant nos T11-705/14N and T11-712/19-N) from the Research Grants Council, Hong Kong SAR, China. Y.C.F.S., U.J., J.J. and G.J.D.S. are supported by the Duke–NUS Signature Research Programme funded by the Ministry of Health, Singapore.

Author contributions

H.-L.Y., G.J.D.S. and R.J.W. conceived and designed the experiments. W.S., Y.C.F.S., U.J., J.J. and G.J.D.S. performed the sequence analyses. W.S., Y.Z., M.Z. and Y.J. performed in vitro experiments. W.S. and H.T.B. performed ex vivo experiments. W.S., R.H., J.D., J.-C.C., T.J., A.R., J.F., P.N.Q.P., C.K., J.C.J. and L.K. performed in vivo animal experiments. S.K., S.P., M.C.W.C., R.G.W., C.-Y.W., K.V.R. and M.P. provided novel reagents to support the study. W.S., Y.C.F.S., G.J.D.S., R.J.W., M.P. and H.-L.Y. wrote the draft manuscript, and all authors reviewed the paper.

Competing interests

The authors declare no competing interests.

Additional information

Extended data is available for this paper at <https://doi.org/10.1038/s41564-021-00976-y>.

Supplementary information The online version contains supplementary material available at <https://doi.org/10.1038/s41564-021-00976-y>.

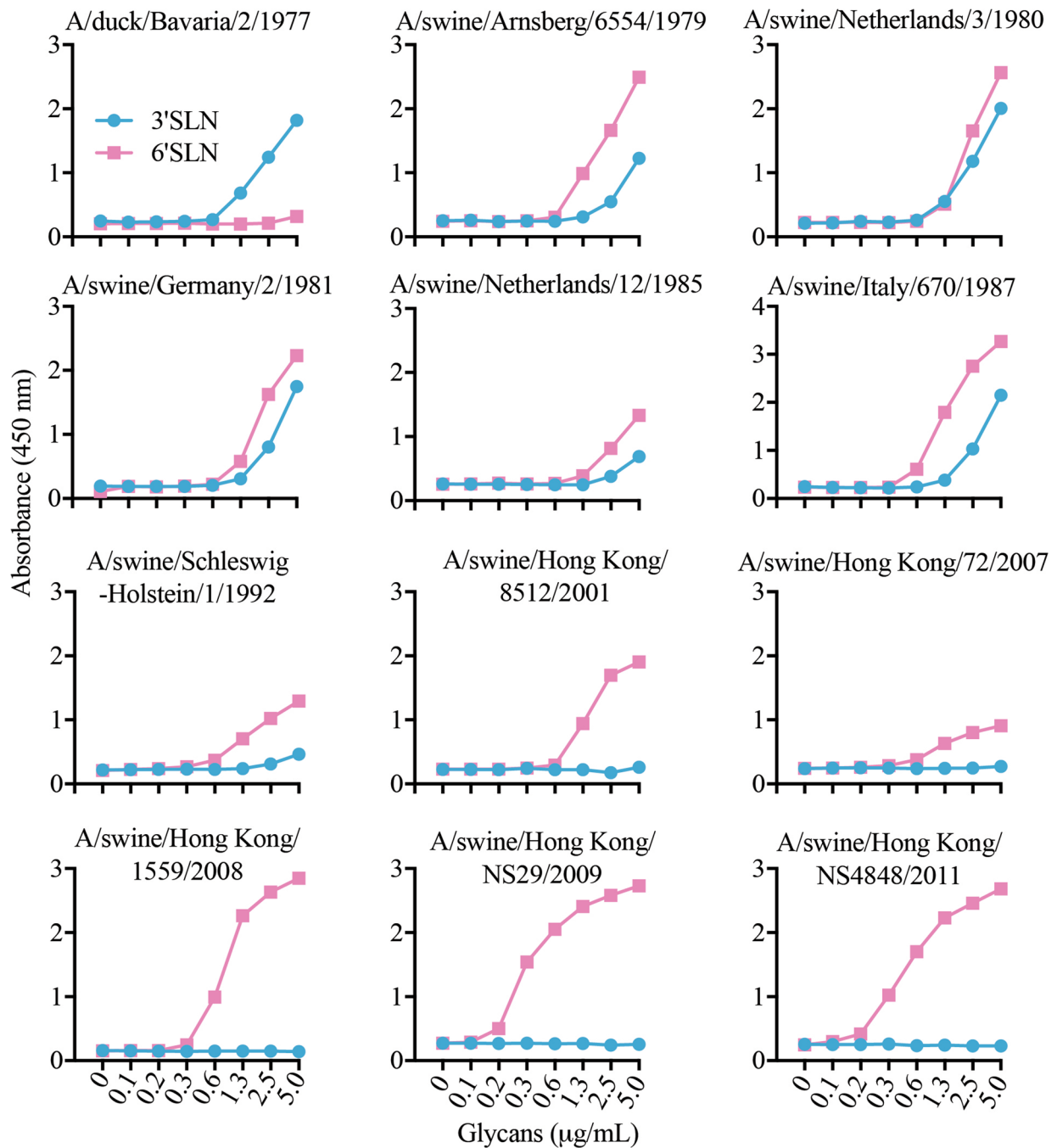
Correspondence and requests for materials should be addressed to Richard J. Webby, Gavin J. D. Smith or Hui-Ling Yen.

Peer review information *Nature Microbiology* thanks Eric Gaucher, Larisa Rudenko and the other, anonymous, reviewer(s) for their contribution to the peer review of this work. Peer reviewer reports are available.

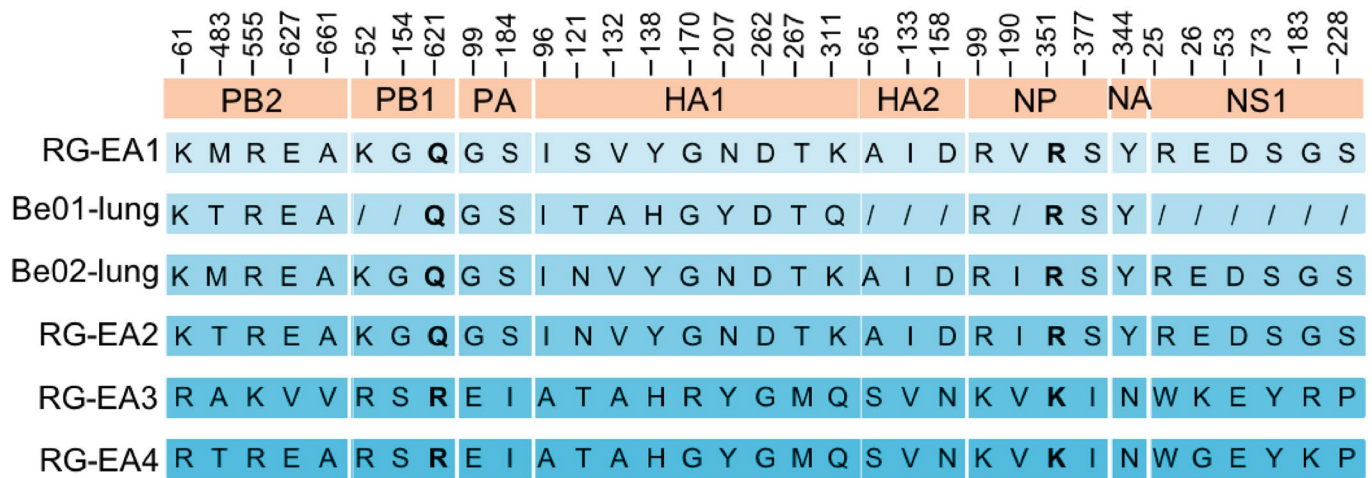
Reprints and permissions information is available at www.nature.com/reprints.

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

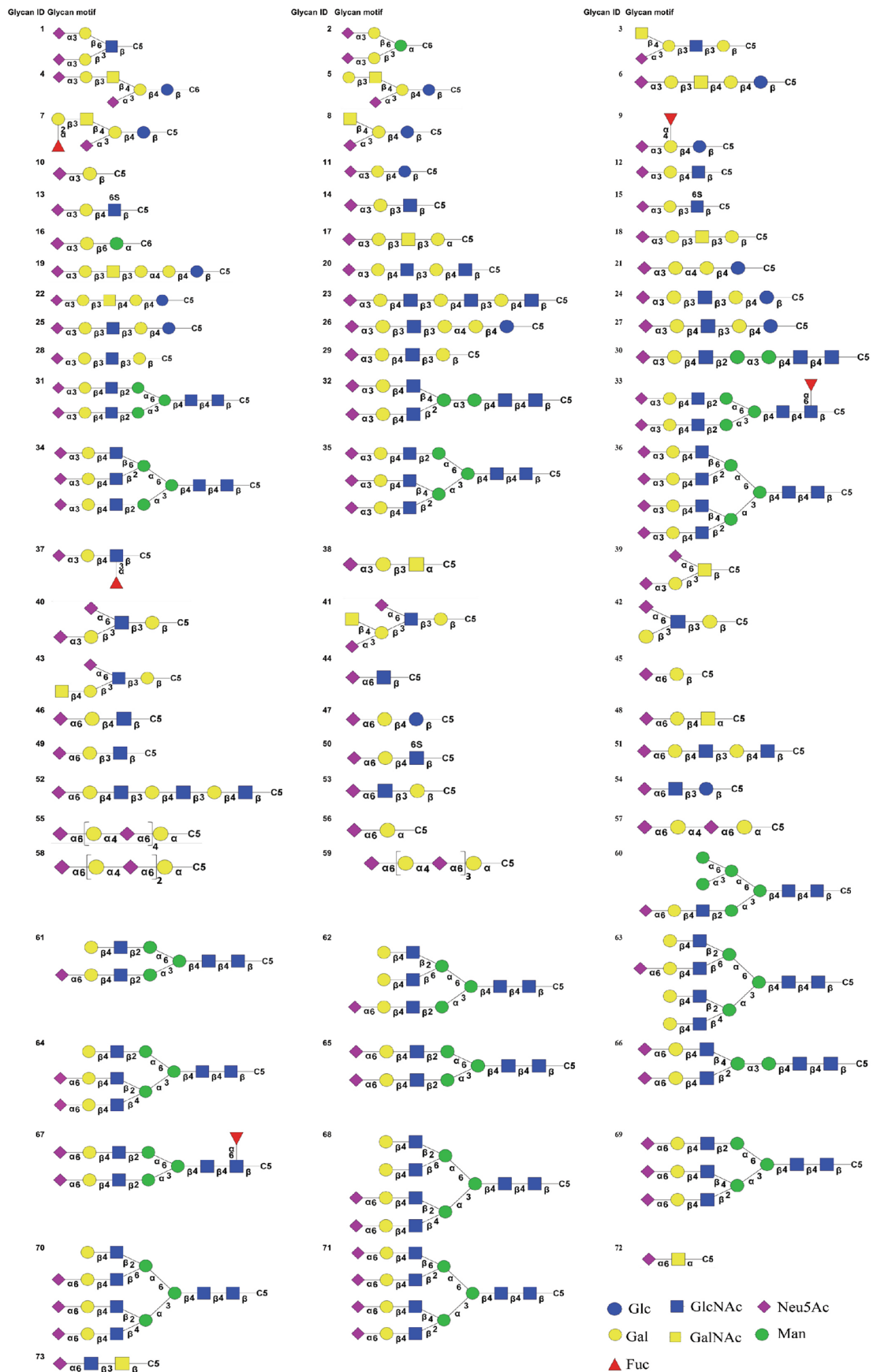
© The Author(s), under exclusive licence to Springer Nature Limited 2021



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 $\mu\text{g}/\text{ml}$. Infectious viruses (diluted to 64 HAU/50 μL) 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 s.d. 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 synthesised glycan microarray comprised of 38 glycans of α 2,3-linked sialosides (glycan 1–38), 3 glycans containing both α 2,3 & α 2,6-linked sialosides (glycan 39–41) and 32 glycans of α 2,6-linked sialosides (glycan 42–73) was utilized for glycan array assay.

Reporting Summary

Nature Research wishes to improve the reproducibility of the work that we publish. This form provides structure for consistency and transparency in reporting. For further information on Nature Research policies, see our [Editorial Policies](#) and the [Editorial Policy Checklist](#).

Statistics

For all statistical analyses, confirm that the following items are present in the figure legend, table legend, main text, or Methods section.

n/a Confirmed

- The exact sample size (n) for each experimental group/condition, given as a discrete number and unit of measurement
- A statement on whether measurements were taken from distinct samples or whether the same sample was measured repeatedly
- The statistical test(s) used AND whether they are one- or two-sided
Only common tests should be described solely by name; describe more complex techniques in the Methods section.
- A description of all covariates tested
- A description of any assumptions or corrections, such as tests of normality and adjustment for multiple comparisons
- A full description of the statistical parameters including central tendency (e.g. means) or other basic estimates (e.g. regression coefficient) AND variation (e.g. standard deviation) or associated estimates of uncertainty (e.g. confidence intervals)
- For null hypothesis testing, the test statistic (e.g. F , t , r) with confidence intervals, effect sizes, degrees of freedom and P value noted
Give P values as exact values whenever suitable.
- For Bayesian analysis, information on the choice of priors and Markov chain Monte Carlo settings
- For hierarchical and complex designs, identification of the appropriate level for tests and full reporting of outcomes
- Estimates of effect sizes (e.g. Cohen's d , Pearson's r), indicating how they were calculated

Our web collection on [statistics for biologists](#) contains articles on many of the points above.

Software and code

Policy information about [availability of computer code](#)

Data collection

Fluorescence data were obtained by InnoScan 710 AL Microarray Scanner (Innopsys, Chicago, USA); The absorbance measurement data was acquired by a FLUOstar OPTIMA microplate reader (BMG Labtech, Ortenberg, Germany); data in the immuno-fluorescence staining assay were obtained using NIS-Elements BR imaging software (Version 4.40) and Nikon DS-Qi2 cameras in a fluorescence microscope Nikon Eclipse Ti-S and were numerated using FIJI (<https://imagej.net/Fiji>); Data in syncytium formation assay were acquired using NIS-Elements BR imaging software (Version 4.40).

Data analysis

Sequences were aligned using MAFFT at the CIPRES science gateway (Version 3) (<https://www.phylo.org/>); Sequences were viewed using MEGA7. The frequency of specific amino acid changes were numerated using BioEdit (Version 7.0); Maximum likelihood analyses were performed using RAXML ver 8.0 in Geneious R9.0.3; Ancestral nucleotide sequences were inferred at internal nodes using the phylogenetic trees and the baseml program as implemented in the Lazarus software (version 2.0); Amino acid substitutions were mapped to maximum likelihood phylogenies using the treesub programme (<https://github.com/tamuri/treesub>). Fluorescence data were analyzed in in GenePix Pro 6.0 software. All NGS data were analysed on CLC Genomic workbench (version 20) (Qiagen, Hilden, Germany). Statistical analyses were performed in Microsoft Excel for Mac, version 16.28 and GraphPad Prism version 8.4.1 for Windows. No additional computer codes were used for data analysis.

For ancestral sequence reconstruction, independent maximum likelihood (ML) analyses were performed with RAXML ver 8.037 in Geneious R9.0.3 (Biomatters Ltd.) to generate 10 input trees for each gene segment. The Lazarus software package (version 2.0) that utilises baseml was used 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. The treesub programme was used 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.

For manuscripts utilizing custom algorithms or software that are central to the research but not yet described in published literature, software must be made available to editors and reviewers. We strongly encourage code deposition in a community repository (e.g. GitHub). See the Nature Research [guidelines for submitting code & software](#) for further information.

Data

Policy information about [availability of data](#)

All manuscripts must include a [data availability statement](#). This statement should provide the following information, where applicable:

- Accession codes, unique identifiers, or web links for publicly available datasets
- A list of figures that have associated raw data
- A description of any restrictions on data availability

Sequences of Sanger sequencing have been 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/Italy/670/1987: EPI_ISL_68983; 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/NS29/2009: EPI_ISL_29774; 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; A/swine/Belgium/1/1979: EPI_ISL_1055769; A/swine/Belgium/2/1979: EPI_ISL_1055773. 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 supplementary figure 1 are provided with this paper.

Field-specific reporting

Please select the one below that is the best fit for your research. If you are not sure, read the appropriate sections before making your selection.

- Life sciences Behavioural & social sciences Ecological, evolutionary & environmental sciences

For a reference copy of the document with all sections, see [nature.com/documents/nr-reporting-summary-flat.pdf](https://www.nature.com/documents/nr-reporting-summary-flat.pdf)

Life sciences study design

All studies must disclose on these points even when the disclosure is negative.

Sample size	No statistical methods were used to pre-determine sample sizes for the animal experiment, but our sample sizes are similar to those reported in previous publications (PMID: 23704376 and PMID: 23383167).
Data exclusions	No data was excluded from the analyses.
Replication	Experiments were independently repeated 2-3 times. Each pig experiment was independently performed in duplicate with a total of 4 donors and 4 direct contacts. The number of technical and biological replicates are indicated throughout.
Randomization	Randomization was performed while assigning the animals into experimental groups.
Blinding	Investigators were not blinded to group allocation during data collection and or analysis.

Reporting for specific materials, systems and methods

We require information from authors about some types of materials, experimental systems and methods used in many studies. Here, indicate whether each material, system or method listed is relevant to your study. If you are not sure if a list item applies to your research, read the appropriate section before selecting a response.

Materials & experimental systems

- | n/a | Involved in the study |
|-------------------------------------|--|
| <input type="checkbox"/> | <input checked="" type="checkbox"/> Antibodies |
| <input type="checkbox"/> | <input checked="" type="checkbox"/> Eukaryotic cell lines |
| <input checked="" type="checkbox"/> | <input type="checkbox"/> Palaeontology and archaeology |
| <input type="checkbox"/> | <input checked="" type="checkbox"/> Animals and other organisms |
| <input checked="" type="checkbox"/> | <input type="checkbox"/> Human research participants |
| <input checked="" type="checkbox"/> | <input type="checkbox"/> Clinical data |
| <input type="checkbox"/> | <input checked="" type="checkbox"/> Dual use research of concern |

Methods

- | n/a | Involved in the study |
|-------------------------------------|---|
| <input checked="" type="checkbox"/> | <input type="checkbox"/> ChIP-seq |
| <input checked="" type="checkbox"/> | <input type="checkbox"/> Flow cytometry |
| <input checked="" type="checkbox"/> | <input type="checkbox"/> MRI-based neuroimaging |

Antibodies

Antibodies used

Rabbit polyclonal antibody against HA of A/duck/NZL/160/1976 (H1N3) (Sino Biological Inc., Beijing, China); Polyclonal Goat Anti-Rabbit Immunoglobulins/HRP (Dako, Glostrup, Denmark); Human monoclonal antibody (MED18852); Goat anti-human IgG (H+L) secondary antibody labelled with Alexa Fluor 647 (Invitrogen TM, Thermo Fisher, USA); Mouse monoclonal IgG2a Influenza A NP

antibody (Santa Cruz Biotech, Inc, Santa Cruz, USA); FITC Goat Anti-Mouse IgG/IgM (BD Pharmingen™, San Jose, USA); DAPI (4',6'-diamidino-2-phenylindole) in SlowFade™ Gold Antifade Mountant (Thermo Fisher Scientific, Eugene, USA)

Validation

All commercial antibodies were validated by the manufacturer as demonstrated by western blotting on the manufacturer's website. The details are as following.

1. Rabbit polyclonal antibody against HA of A/duck/NZL/160/1976 (H1N3) (Sino Biological Inc., Beijing, China). The HA of this virus was a member of European avian-like (EA) swine H1N1 lineage, as described previously in Obenauer et al. (PMID: 16439620). This antibody was produced from a hybridoma resulting from the fusion of a mouse myeloma with B cells obtained from a mouse immunized with purified, recombinant Influenza A virus H1N3 Hemagglutinin. The IgG fraction of the cell culture supernatant was purified by Protein A affinity chromatography. (<https://www.sinobiological.com/>)
2. Polyclonal Goat Anti-Rabbit Immunoglobulins/HRP (Dako, Glostrup, Denmark);
The goat antibody used for horseradish peroxidase-conjugation reacts with rabbit immunoglobulins of all classes. The cross-reaction with human, mouse and rat immunoglobulins, and fetal calf serum proteins is negligible. (<https://www.agilent.com/>)
3. Human monoclonal antibody (MED18852) ;
It is a broad-spectrum human monoclonal antibody, effectively reacting with all influenza A hemagglutinin (HA) subtypes. The antibody has been previously used to detect all influenza A subtypes in Kallewaard et al. (PMID: 27453466)
4. Goat anti-human IgG (H+L) secondary antibody labelled with Alexa Fluor 647 (Invitrogen™, Thermo Fisher, USA);
This antibody binds to heavy chains on human IgG and light chains on all human immunoglobulins. This antibody does not bind non-immunoglobulin human serum proteins or serum proteins/IgG from bovine, mouse, and rabbit. (<https://www.thermofisher.com/hk/en/home.html>)
5. Mouse monoclonal IgG2a Influenza A NP antibody (Santa Cruz Biotech, Inc, Santa Cruz, USA) ;
Influenza A NP (9G8) is a mouse monoclonal antibody raised against recombinant Influenza A NP. (<https://www.scbt.com/p/influenza-a-np-antibody-9g8>)
6. FITC Goat Anti-Mouse IgG/IgM (BD Pharmingen™, San Jose, USA) ;This antibody reacts with the heavy chains IgG and IgM and with light chains common to most mouse immunoglobulins. (<https://wwwbdbiosciences.com/en-us>)
7. DAPI (4',6'-diamidino-2-phenylindole) in SlowFade™ Gold Antifade Mountant (Thermo Fisher Scientific, Eugene, USA);
SlowFade Gold Antifade Mountant is a liquid mountant applied directly to fluorescently labeled cell or tissue samples on microscope slides. It contains chemical components designed to protect fluorescent dyes from fading (photobleaching) during fluorescence microscopy experiments. (<https://www.thermofisher.com/hk/en/home.html>)

Eukaryotic cell lines

Policy information about [cell lines](#)

Cell line source(s)	Madin-Darby Canine Kidney (MDCK) cells, Human embryonic kidney 293T (293T) and Vero cells were obtained from the American Type Culture Collection (ATCC); New-born pig trachea (NPTr) cells were obtained from Istituto Zooprofilattico Sperimentale, della Lombardia e dell'Emilia Romagna.
Authentication	All cell lines were regularly checked for cellular morphology and growth rates. But, none of the cell lines used were authenticated.
Mycoplasma contamination	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.
Commonly misidentified lines (See ICLAC register)	No commonly mis-identified lines were used.

Animals and other organisms

Policy information about [studies involving animals](#); [ARRIVE guidelines](#) recommended for reporting animal research

Laboratory animals	Female or male 3-4 weeks old Yorkshire Crossbred piglets (Midwest Research Swine LLC., Glencoe, MN USA) that were serologically negative for influenza A virus NP protein (ID.Vet, Grabels, France) were randomly used in pig transmission experiments; 10-day old embryonated chicken eggs were used for serial passages.
Wild animals	The study did not involve wild animals.
Field-collected samples	No field collected samples were used.
Ethics oversight	All pig transmission experiment 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).

Note that full information on the approval of the study protocol must also be provided in the manuscript.

Dual use research of concern

Policy information about [dual use research of concern](#)

Hazards

Could the accidental, deliberate or reckless misuse of agents or technologies generated in the work, or the application of information presented in the manuscript, pose a threat to:

No	Yes
<input checked="" type="checkbox"/>	<input type="checkbox"/> Public health
<input checked="" type="checkbox"/>	<input type="checkbox"/> National security
<input checked="" type="checkbox"/>	<input type="checkbox"/> Crops and/or livestock
<input checked="" type="checkbox"/>	<input type="checkbox"/> Ecosystems
<input checked="" type="checkbox"/>	<input type="checkbox"/> Any other significant area

Experiments of concern

Does the work involve any of these experiments of concern:

No	Yes
<input checked="" type="checkbox"/>	<input type="checkbox"/> Demonstrate how to render a vaccine ineffective
<input checked="" type="checkbox"/>	<input type="checkbox"/> Confer resistance to therapeutically useful antibiotics or antiviral agents
<input checked="" type="checkbox"/>	<input type="checkbox"/> Enhance the virulence of a pathogen or render a nonpathogen virulent
<input type="checkbox"/>	<input checked="" type="checkbox"/> Increase transmissibility of a pathogen
<input checked="" type="checkbox"/>	<input type="checkbox"/> Alter the host range of a pathogen
<input checked="" type="checkbox"/>	<input type="checkbox"/> Enable evasion of diagnostic/detection modalities
<input checked="" type="checkbox"/>	<input type="checkbox"/> Enable the weaponization of a biological agent or toxin
<input checked="" type="checkbox"/>	<input type="checkbox"/> Any other potentially harmful combination of experiments and agents

Precautions and benefits

Biosecurity precautions	All experimental work on European avian-like H1N1 swine viruses including the reconstructed EA swine viruses were performed in BSL2 laboratories with restricted access following approved standard operating procedures. Refresher training course is provided annually. All reagents (plasmids and viruses) were stored in locked freezers in rooms with restricted access.
Biosecurity oversight	The research approach and the potential dual-use research of concern of the study has been reviewed by institutional and funding agency committees.
Benefits	Understanding the evolutionary adaptations that enable avian influenza viruses to transmit in mammalian hosts could allow better detection of zoonotic viruses with pandemic potential. In this study, we used ancestral sequence reconstruction to identify the key adaptive changes of the European avian-like H1N1 (EA) swine influenza viruses as it transitioned from avian to swine hosts since 1979. 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 retrospectively dissect and study viral evolution that has already occurred in nature, rather than anticipating mutations that may lead to increased mammalian transmissibility or virulence. This retrospective approach allowed us to follow the natural evolutionary path of EA swine viruses without forward-selecting avian influenza variants that adapts to mammalian hosts under experimental conditions.
Communication benefits	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.