

The soft palate is an important site of adaptation for transmissible influenza viruses

Seema S. Lakdawala^{1†}, Akila Jayaraman², Rebecca A. Halpin³, Elaine W. Lamirande¹, Angela R. Shih¹, Timothy B. Stockwell³, Xudong Lin³, Ari Simenauer³, Christopher T. Hanson¹, Leatrice Vogel¹, Myeisha Paskel¹, Mahnaz Minai⁴, Ian Moore⁴, Marlene Orandle^{4†}, Suman R. Das³, David E. Wentworth^{3†}, Ram Sasisekharan² & Kanta Subbarao¹

Influenza A viruses pose a major public health threat by causing seasonal epidemics and sporadic pandemics. Their epidemiological success relies on airborne transmission from person to person; however, the viral properties governing airborne transmission of influenza A viruses are complex. Influenza A virus infection is mediated via binding of the viral haemagglutinin (HA) to terminally attached α 2,3 or α 2,6 sialic acids on cell surface glycoproteins. Human influenza A viruses preferentially bind α 2,6-linked sialic acids whereas avian influenza A viruses bind α 2,3-linked sialic acids on complex glycans on airway epithelial cells^{1,2}. Historically, influenza A viruses with preferential association with α 2,3-linked sialic acids have not been transmitted efficiently by the airborne route in ferrets^{3,4}. Here we observe efficient airborne transmission of a 2009 pandemic H1N1 (H1N1pdm) virus (A/California/07/2009) engineered to preferentially bind α 2,3-linked sialic acids. Airborne transmission was associated with rapid selection of virus with a change at a single HA site that conferred binding to long-chain α 2,6-linked sialic acids, without loss of α 2,3-linked sialic acid binding. The transmissible virus emerged in experimentally infected ferrets within 24 hours after infection and was remarkably enriched in the soft palate, where long-chain α 2,6-linked sialic acids predominate on the nasopharyngeal surface. Notably, presence of long-chain α 2,6-linked sialic acids is conserved in ferret, pig and human soft palate. Using a loss-of-function approach with this one virus, we demonstrate that the ferret soft palate, a tissue not normally sampled in animal models of influenza, rapidly selects for transmissible influenza A viruses with human receptor (α 2,6-linked sialic acids) preference.

Receptor-binding specificity is an important determinant of host-range restriction and transmission of influenza A viruses (refs 4, 5 and reviewed in ref. 6). The ability of zoonotic influenza A viruses to transmit via the airborne route increases their pandemic potential⁷. Recently, several investigators have attempted to identify viral determinants of airborne transmission by generating transmissible H5 and H7 avian influenza A viruses^{8–10}. We approached the question differently and used an epidemiologically successful influenza A virus in which we altered receptor preference from the human (α 2,6-linked sialic acids) to the avian receptor (α 2,3-linked sialic acids).

We previously generated H1N1pdm virus variants with highly specific binding to either α 2,6-linked or α 2,3-linked sialic acids (referred to as α 2,6 or α 2,3 H1N1pdm virus, respectively)¹¹. The α 2,3 H1N1pdm virus was generated by introducing four amino acid mutations in the receptor binding site of HA (D187E, I216A, D222G and E224A)¹¹. Unexpectedly, the α 2,6 and α 2,3 H1N1pdm viruses transmitted via the airborne route equally well in ferrets (Fig. 1 and Supplementary Table 1) and with a similar efficiency as observed previously for wild-type H1N1pdm virus^{12–15}.

A delay in peak viral shedding was noted in the airborne-contact animals (see Fig. 1 legend for details) in the α 2,3 virus group (red arrows, Fig. 1), suggesting that the virus evolves before transmission. Deep sequence analysis of viral RNA (vRNA) extracted from nasal washes of α 2,3 H1N1pdm virus-infected ferrets revealed a mixed population at amino acid position 222 (H1 numbering) with the engineered glycine (G) and wild-type aspartic acid (D), while the other three engineered changes in the HA were retained (Fig. 2a and Supplementary Table 2). Interestingly, the vRNA from the nasal washes of airborne-contact ferrets contained only the G222D HA mutation (Fig. 2a and Supplementary Table 2), suggesting that this sequence at amino acid 222 in the α 2,3 H1N1pdm virus was associated with airborne transmission. The virus inoculum did not contain a mixture at this residue (Fig. 2a), and associated changes were not observed in the neuraminidase gene (Supplementary Table 3).

A D222G change in the 2009 H1N1pdm virus HA has occurred in natural isolates, and reports suggest an association with increased virulence in humans and no effect on airborne transmission^{16–18}. Theoretical structural analysis suggests that the G222D reversion makes the receptor binding site better suited to bind α 2,6-linked sialic acids while retaining contacts with α 2,3-linked sialic acids via glutamic acid at amino acid 187 (Extended Data Fig. 1). Glycan binding data corroborated this structural prediction because the G222D mutation caused no change in α 2,3-linked sialic acid binding but substantially increased binding to long-chain α 2,6-linked sialic acids (Fig. 2b).

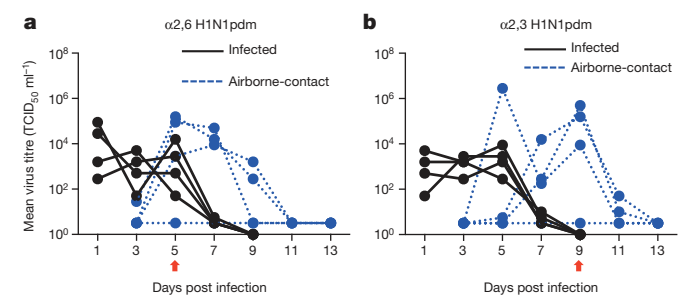


Figure 1 | Airborne transmission of receptor-specific H1N1pdm viruses. Transmission studies were performed with 4 pairs of animals (8 animals total) in double secure cages with perforated dividers¹². One ferret in each pair was infected with 10^6 50% tissue culture infectious dose (TCID₅₀) of the indicated virus; a naive ferret (referred to as airborne-contact) was introduced into the adjacent compartment 24 h later. Nasal secretions were collected every other day for 14 days. Viral titres from the nasal secretions are graphed for each infected or airborne-contact animal. Transmission of α 2,6 H1N1pdm (a) and α 2,3 H1N1pdm (b) viruses was similar. The red arrow indicates the peak day of viral shedding for airborne contact animals.

¹Laboratory of Infectious Diseases, National Institute of Allergy and Infectious Diseases, National Institutes of Health, Bethesda, Maryland 20892, USA. ²Department of Biological Engineering, Koch Institute for Integrative Cancer Research, Singapore-MIT Alliance for Research and Technology, Massachusetts Institute of Technology, Cambridge, Massachusetts 02139, USA. ³J. Craig Venter Institute, Rockville, Maryland 20850, USA. ⁴Comparative Medicine Branch, National Institute of Allergy and Infectious Diseases, National Institutes of Health, Bethesda, Maryland 20892, USA. †Present addresses: Department of Microbiology and Molecular Genetics, University of Pittsburgh School of Medicine, Pittsburgh, Pennsylvania 15217, USA (S.S.L.); National Institute of Occupational Safety and Health, Centers for Disease Control and Prevention, Morgantown, West Virginia, 26505 USA (M.O.); Virology Surveillance and Diagnosis Branch, Influenza Division, Centers for Disease Control and Prevention, Atlanta, Georgia 30329, USA (D.E.W.).

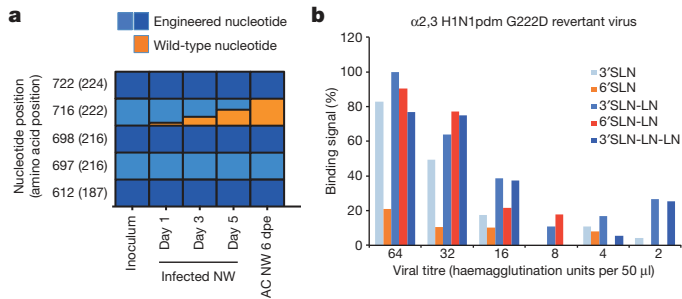


Figure 2 | Characterization of transmissible $\alpha 2,3$ H1N1pdm viruses. **a**, Deep sequencing of the $\alpha 2,3$ H1N1pdm inoculum, nasal wash (NW) from an infected ferret on 1, 3 and 5 dpi, and nasal wash from one airborne-contact (AC) animal on 6 days post-exposure (dpe) revealed a reversion at residue 222 from G to D. This data is representative of the three transmission pairs that resulted in infection of the airborne-contact animals. Graphical representation of the proportion of reads at each engineered nucleotide is shown. Blue shading represents the $\alpha 2,3$ -linked sialic acid engineered nucleotide and orange represents the wild-type nucleotide residue. All other engineered nucleotides were maintained. **b**, A G222D reversion, in the context of the other engineered mutations, affects the glycan specificity of the $\alpha 2,3$ H1N1pdm virus. The glycans are indicated in the key to the figure and are defined in the Methods; orange colours represent $\alpha 2,6$ -linked sialic acids and blue colours represent $\alpha 2,3$ -linked sialic acids. H1 numbering is used for all amino acid positions.

Previous reports have demonstrated the importance of $\alpha 2,6$ -linked sialic acid binding for transmission^{4,5,19}. We now demonstrate conclusively that airborne transmission requires gain of long-chain $\alpha 2,6$ -linked sialic acid binding and, contrary to previous suggestions⁴, loss of $\alpha 2,3$ -linked sialic acid binding is not necessary.

The presence of a distinct and identifiable HA sequence in the transmissible virus allowed us to determine whether it emerges in a specific area of the respiratory tract of experimentally infected ferrets. Tissue sections and samples from the upper and lower respiratory tract were collected several days post-infection (dpi) from groups of 3 ferrets infected with the $\alpha 2,3$ H1N1pdm virus. Virus was detected in all ferrets and all samples (Extended Data Fig. 2). Deep sequencing of vRNA from both the upper and lower respiratory tract revealed a mixed population at residue 222 (Fig. 3). Surprisingly, vRNA from the soft palate was remarkably and uniquely enriched for the G222D virus on 1 dpi and $\geq 90\%$ of the sequences encoded D222 at 3 dpi (Fig. 3c). All other engineered mutations were maintained (Extended Data Fig. 3). These data suggest that the G222D revertant virus was actively selected in the ferret soft palate.

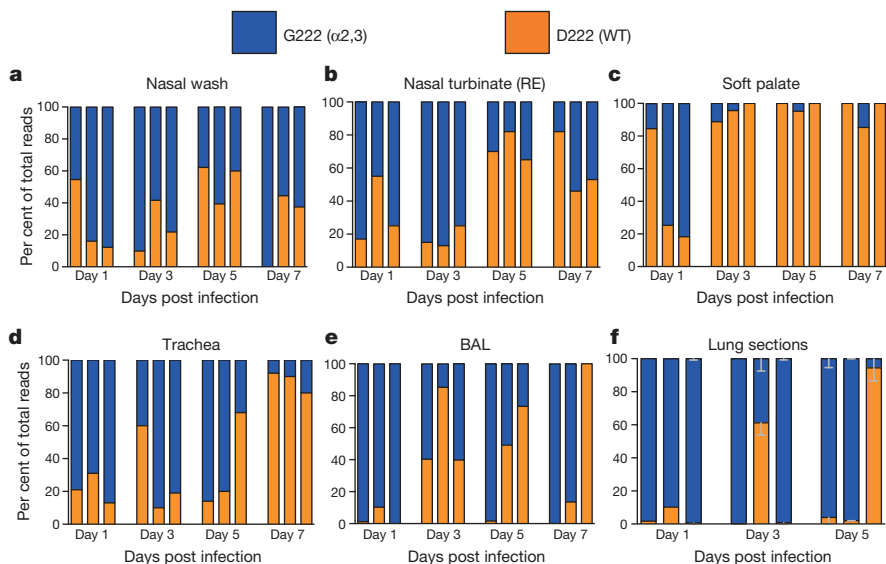


Figure 3 | Emergence of the $\alpha 2,3$ G222D H1N1pdm virus in the ferret respiratory tract. **a–f**, Different samples from the ferret respiratory tract: nasal wash (**a**), respiratory epithelium (RE) of nasal turbinates (**b**), soft palate (**c**), trachea (**d**), bronchoalveolar lavage (BAL) (**e**), and combined right, middle and left cranial lung sections (**f**) were collected from three animals each on 1, 3, 5 and 7 dpi. The respiratory epithelium region of the nasal turbinates is depicted in Extended Data Fig. 6h. The HA gene from virus populations in these samples were deep sequenced and the proportion of reads with D at position 222 is shown in orange, and G is shown in blue. Each bar represents a single animal. The standard error between the right and left lung sections is shown in **f**.

To determine whether the rapid enrichment of G222D revertant virus in the soft palate was responsible for infection of the airborne-contact animal, we performed an airborne transmission study where naive ferrets were exposed to experimentally infected donor ferrets for only 2 days. Surprisingly, even within this shortened exposure time, two airborne-contact animals shed virus and 3 out of 4 airborne-contact animals seroconverted (Extended Data Fig. 4 and Supplementary Table 1). Sequence analysis of vRNA from the two airborne-contact animals with detectable virus in the nasal washes revealed presence of the G222D revertant. These data suggest that the selection of the $\alpha 2,3$ H1N1pdm virus with the D222 sequence occurs within 3 dpi in the donor ferret and that the airborne-contact ferrets were possibly infected with virus originating in the soft palate because there was nearly complete selection of the G222D mutant by 3 dpi in this tissue.

The soft palate, with mucosal surfaces facing the oral cavity and nasopharynx, is not usually examined in animal models of influenza. To understand what drives the enrichment of the long-chain $\alpha 2,6$ -linked-sialic acid-binding G222D revertant virus at this site, we stained the soft palate with lectins specific for $\alpha 2,6$ or $\alpha 2,3$ sialic acids (Extended Data Fig. 5). The ciliated respiratory epithelium and mucus secreting goblet cells in the respiratory epithelium and submucosal glands (SMG) contained $\alpha 2,6$ -linked sialic acids (SNA staining) (Extended Data Fig. 5). Expression of $\alpha 2,3$ -linked sialic acid (MAL II staining) was present in the connective tissue underlying the respiratory epithelium and in the serous cells of the SMG. Using a purified HA protein (SC18) that selectively binds long-chain $\alpha 2,6$ -linked sialic acids²⁰, we found high expression of long-chain $\alpha 2,6$ -linked sialic acids in the soft palate compared to the trachea and lungs of ferrets (Fig. 4 and Extended Data Fig. 6). A recent report detailing the glycan profile of the ferret respiratory tract confirms that the soft palate abundantly expresses $\alpha 2,6$ sialylated LacNAc structures²¹, similar to the long-chain $\alpha 2,6$ -linked sialic acids recognized by SC18 HA. Interestingly, both the respiratory epithelium and olfactory epithelium from the nasal turbinates of ferrets expressed high levels of long-chain $\alpha 2,6$ -linked sialic acid, but the respiratory epithelium of the nasal turbinates was not enriched for the G222D mutant (Fig. 3b and Extended Data Fig. 6). These data suggest that the soft palate is unusual in driving selection for the G222D virus.

To determine the relevance for humans, we evaluated the expression of long-chain $\alpha 2,6$ -linked sialic acids in the soft palate of humans and pigs. Interestingly, expression of long-chain $\alpha 2,6$ -linked sialic acids was conserved on the respiratory epithelium and goblet cells of the soft palate of both species (Fig. 4). In addition, staining with plant lectins specific for $\alpha 2,6$ -linked or $\alpha 2,3$ -linked sialic acids (Extended Data Fig. 7) revealed

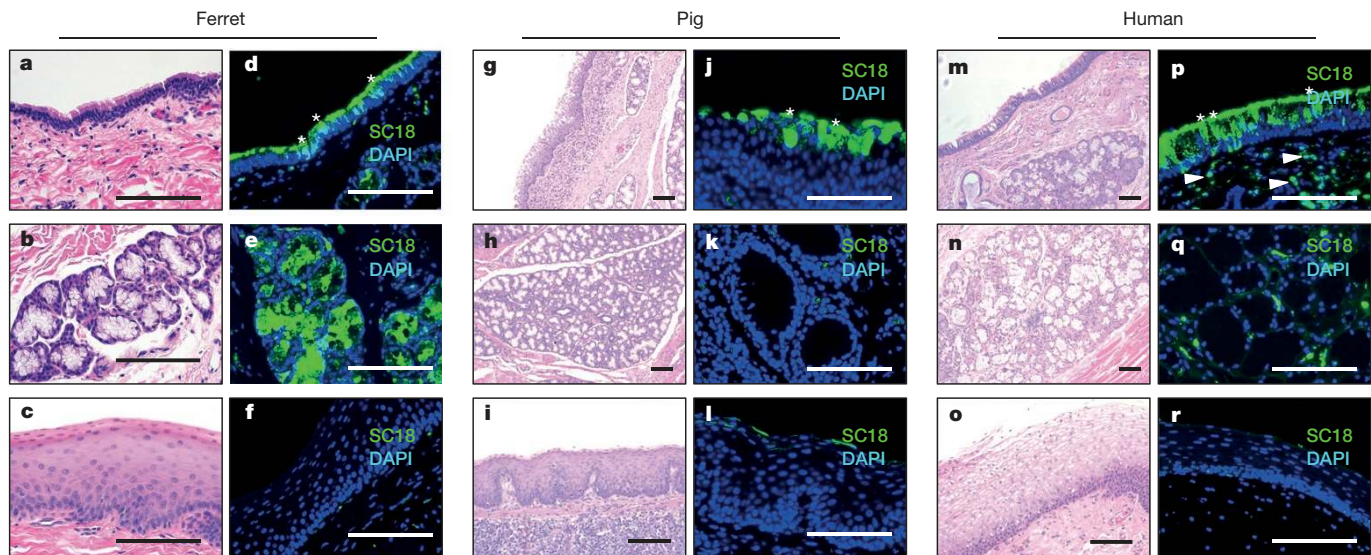


Figure 4 | Comparison of long-chain α 2,6-linked sialic acid expression in the soft palate of ferrets, pigs and humans. a–c, g–i, m–o, Haematoxylin and eosin staining of the soft palate from an uninfected ferret (a–c), pig (g–i) and human (m–o) highlights the nasopharyngeal, SMG and oral surfaces. d–f, j–l, p–r, Purified SC18 HA was used to define long-chain α 2,6-linked sialic acids in these sections from an uninfected ferret (d–f), pig (j–l) and human (p–r). Staining of the nasopharyngeal surface is depicted for each species across

the first row, SMG in the second row and oral surface on the last row. At least two independent tissue samples were stained and analysed for each species. A sialidase-A-treated control was run for each sample to ensure specificity of SC18 HA (not shown). Scale bars, 100 μ m in all images. Asterisks highlight SC18-positive goblet cells and white arrowheads indicate SC18-positive plasma cells in human soft palate.

that α 2,6-linked sialic acids were present on the nasopharyngeal surface and SMG of both pigs and humans. Expression of α 2,3-linked sialic acids was detected in the basal cells of the oral surface and on the nasopharyngeal surface of the human soft palate; these findings are consistent with reports describing the sialic acid distribution in the human nasopharynx²². Other investigators have also reported replication of seasonal and pandemic influenza A viruses in tissue sections obtained from the human nasopharynx²³. Taken together, these data highlight the importance of the nasopharynx, of which the soft palate forms the floor, as a site for host adaptation of influenza A viruses.

Influenza A virus infection of the soft palate may contribute to airborne transmission by providing a mucin-rich microenvironment for generation of airborne virus during coughing, sneezing or breathing. Infection with α 2,3 H1N1pdm virus resulted in severe inflammation and necrosis of the respiratory epithelial cells and SMG in the soft palate (Extended Data Fig. 8). Since the soft palate is innervated by the trigeminal nerve, inflammation of this tissue could stimulate sneezing. Alternatively, the soft palate may be the site where infection is initiated during airborne transmission; therefore binding to this tissue would provide a fitness advantage.

These results, albeit with one virus, enhance our understanding of the properties necessary for airborne transmission of influenza A viruses in the ferret model. Loss of α 2,3-linked sialic acid specificity is not necessary but gain of long-chain α 2,6-linked sialic acid binding is critical for efficient airborne transmission of influenza A viruses. H7N9 viruses from China show dual receptor binding but variable airborne transmission efficiency in ferrets^{24,25}. Interestingly, the 1918 H1N1 virus (A/New York/1/18), which has a similar sialic acid binding preference as the α 2,3 H1N1pdm virus, did not transmit via the airborne route or adapt within the ferret host⁴, suggesting that the 2009 H1N1pdm virus may be unusual for this rapid adaptation. However, the detection of a mutation that enhanced α 2,6-linked sialic acid binding in nasal washes of ferrets infected with avian H2 viruses was recently reported²⁶, demonstrating that rapid adaptation of influenza A viruses to gain human receptor preference occurs in other influenza A virus subtypes as well.

Studies with transmissible H5 viruses suggest that increased pH and thermal stability of the HA enhance airborne transmission^{8,9,27}.

Although we did not observe adaptive mutations in the HA stalk of the α 2,3 H1N1pdm virus, perhaps because H1N1pdm HA is already adapted to humans, a mixed population was observed at four lysine residues around the receptor binding site (Extended Data Fig. 9 and Supplementary Table 2). Some are known to be egg adaptive mutations²⁸ or are components of the proposed positively charged 'lysine fence' around the base of the receptor binding site, positioned to anchor the *N*-acetylneuraminic acid and galactose sugar of α 2,3-linked and α 2,6-linked sialic acid glycans²⁹. Interestingly, the lysine residues were restored in the vRNA isolated from nasal washes of airborne-contact ferrets and the soft palate of experimentally infected ferrets (Extended Data Figs 9 and 10).

Taken together with our previously published data, long-chain α 2,6-linked sialic acid binding and a highly active neuraminidase contribute to the airborne transmission of the H1N1pdm virus^{12,30}. Importantly, we have identified the previously overlooked soft palate as an important site of isolation of transmissible virus and perhaps the initial site of infection. Analysis of the replicative fitness of influenza A viruses in this tissue may be warranted in assessment of their pandemic potential.

Online Content Methods, along with any additional Extended Data display items and Source Data, are available in the online version of the paper; references unique to these sections appear only in the online paper.

Received 15 October 2014; accepted 5 August 2015.

Published online 23 September 2015.

- Shinya, K. *et al.* Avian flu: influenza virus receptors in the human airway. *Nature* **440**, 435–436 (2006).
- van Riel, D. *et al.* H5N1 virus attachment to lower respiratory tract. *Science* **312**, 399 (2006).
- Maines, T. R. *et al.* Lack of transmission of H5N1 avian-human reassortant influenza viruses in a ferret model. *Proc. Natl Acad. Sci. USA* **103**, 12121–12126 (2006).
- Tumpey, T. M. *et al.* A two-amino acid change in the hemagglutinin of the 1918 influenza virus abolishes transmission. *Science* **315**, 655–659 (2007).
- Pappas, C. *et al.* Receptor specificity and transmission of H2N2 subtype viruses isolated from the pandemic of 1957. *PLoS ONE* **5**, e11158 (2010).
- Cauldwell, A. V., Long, J. S., Moncorge, O. & Barclay, W. S. Viral determinants of influenza A virus host range. *J. Gen. Virol.* **95**, 1193–1210 (2014).
- Lakdawala, S. S. & Subbarao, K. The ongoing battle against Influenza: The challenge of flu transmission. *Nature Med.* **18**, 1468–1470 (2012).

8. Herfst, S. *et al.* Airborne transmission of influenza A/H5N1 virus between ferrets. *Science* **336**, 1534–1541 (2012).
9. Imai, M. *et al.* Experimental adaptation of an influenza H5 HA confers respiratory droplet transmission to a reassortant H5 HA/H1N1 virus in ferrets. *Nature* **486**, 420–428 (2012).
10. Sutton, T. C. *et al.* Airborne transmission of highly pathogenic H7N1 influenza virus in ferrets. *J. Virol.* **88**, 6623–6635 (2014).
11. Lakdawala, S. S. *et al.* Receptor specificity does not affect replication or virulence of the 2009 pandemic H1N1 influenza virus in mice and ferrets. *Virology* **446**, 349–356 (2013).
12. Lakdawala, S. S. *et al.* Eurasian-origin gene segments contribute to the transmissibility, aerosol release, and morphology of the 2009 pandemic H1N1 influenza virus. *PLoS Pathog.* **7**, e1002443 (2011).
13. Itoh, Y. *et al.* *In vitro* and *in vivo* characterization of new swine-origin H1N1 influenza viruses. *Nature* **460**, 1021–1025 (2009).
14. Maines, T. R. *et al.* Transmission and pathogenesis of swine-origin 2009 A(H1N1) influenza viruses in ferrets and mice. *Science* **325**, 484–487 (2009).
15. Munster, V. J. *et al.* Pathogenesis and transmission of swine-origin 2009 A(H1N1) influenza virus in ferrets. *Science* **325**, 481–483 (2009).
16. Liu, Y. *et al.* Altered receptor specificity and cell tropism of D222G hemagglutinin mutants isolated from fatal cases of pandemic A(H1N1) 2009 influenza virus. *J. Virol.* **84**, 12069–12074 (2010).
17. Mak, G. C. *et al.* Association of D222G substitution in haemagglutinin of 2009 pandemic influenza A(H1N1) with severe disease. *Euro Surveill.* **15**, 19534 (2010).
18. Belsler, J. A. *et al.* Effect of D222G mutation in the hemagglutinin protein on receptor binding, pathogenesis and transmissibility of the 2009 pandemic H1N1 influenza virus. *PLoS ONE* **6**, e25091 (2011).
19. Zhang, Y. *et al.* Key molecular factors in hemagglutinin and PB2 contribute to efficient transmission of the 2009 H1N1 pandemic influenza virus. *J. Virol.* **86**, 9666–9674 (2012).
20. Srinivasan, A. *et al.* Quantitative biochemical rationale for differences in transmissibility of 1918 pandemic influenza A viruses. *Proc. Natl Acad. Sci. USA* **105**, 2800–2805 (2008).
21. Jia, N. *et al.* Glycomic characterisation of respiratory tract tissues of ferrets: implications for its use in influenza virus infection studies. *J. Biol. Chem.* **289**, 28489–28504 (2014).
22. Nicholls, J. M., Bourne, A. J., Chen, H., Guan, Y. & Peiris, J. S. Sialic acid receptor detection in the human respiratory tract: evidence for widespread distribution of potential binding sites for human and avian influenza viruses. *Respir. Res.* **8**, 73 (2007).
23. Chan, R. W., Chan, M. C., Nicholls, J. M. & Malik Peiris, J. S. Use of *ex vivo* and *in vitro* cultures of the human respiratory tract to study the tropism and host responses of highly pathogenic avian influenza A (H5N1) and other influenza viruses. *Virus Res.* **178**, 133–145 (2013).
24. Belsler, J. A. *et al.* Pathogenesis and transmission of avian influenza A (H7N9) virus in ferrets and mice. *Nature* **501**, 556–559 (2013).
25. Richard, M. *et al.* Limited airborne transmission of H7N9 influenza A virus between ferrets. *Nature* **501**, 560–563 (2013).
26. Pappas, C. *et al.* Assessment of transmission, pathogenesis and adaptation of H2 subtype influenza viruses in ferrets. *Virology* **477**, 61–71 (2015).
27. Linster, M. *et al.* Identification, characterization, and natural selection of mutations driving airborne transmission of A/H5N1 virus. *Cell* **157**, 329–339 (2014).
28. Chen, Z. *et al.* Generation of live attenuated novel influenza virus A/California/7/09 (H1N1) vaccines with high yield in embryonated chicken eggs. *J. Virol.* **84**, 44–51 (2010).
29. Soundararajan, V. *et al.* Extrapolating from sequence—the 2009 H1N1 ‘swine’ influenza virus. *Nature Biotechnol.* **27**, 510–513 (2009).
30. 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).

Supplementary Information is available in the online version of the paper.

Acknowledgements This research was supported in part by the Intramural Research Program of NIAID, NIH and with federal funds from NIAID, NIH, DHHS under contract number HHSN272200900007C and NIAID/NIH Genomic Centers for Infectious Diseases (GCID) program (U19-AI-110819). This manuscript was reviewed by the NIH’s Intramural Research Program’s Committee on Dual Use Research of Concern (DURC), who concluded that the methods and results do not meet DURC criteria. We thank the NIAID Comparative Medical Branch for technical assistance, Subbarao laboratory members for critical input, N. B. Fedorova from JCVI for technical help, X. J. Meng (Virginia Tech College of Veterinary Medicine) and P. Pineyro (Iowa State University) for pig soft palate tissues, and the Consortium for Functional Glycomics for providing glycans for the glycan array analysis. The data for this manuscript and its preparation were generated while D.E.W. was employed at JCVI. The opinions expressed in this article are the authors’ own and do not reflect the views of the Centers for Disease Control, the Department of Health and Human Services, or the United States government. A.J. and R.S. are supported in part by NIH Merit Award (R37 GM057073-13), National Research Foundation supported Interdisciplinary Research group in Infectious Diseases of SMART (Singapore MIT alliance for Research and Technology) and the Skolkovo Foundation supported Infectious Diseases Center at MIT.

Author Contributions S.S.L., A.J., R.S., D.E.W. and K.S. designed the study. S.S.L., A.J., E.W.L., A.R.S., X. L., A.S., C.T.H., L.V., M.P. and M.M. performed the experiments. S.S.L., R.A.H., T.B.S., I.M., M.O. and S.R.D. analysed the data. S.S.L., K.S., A.J. and R.S. wrote the paper.

Author Information The sequences detailed in this manuscript can be found in GenBank under accession numbers CY184674–CY185309. Reprints and permissions information is available at www.nature.com/reprints. The authors declare no competing financial interests. Readers are welcome to comment on the online version of the paper. Correspondence and requests for materials should be addressed to R.S. (rams@mit.edu) and K.S. (ksubbarao@niaid.nih.gov).

METHODS

Ethics statement and animal studies. This study was carried out in strict accordance with the recommendations in the Guide for the Care and Use of Laboratory Animals of the National Institutes of Health. The National Institutes of Health Animal Care and Use Committee (ACUC) approved the animal experiments that were conducted. All studies were conducted under ABSL2 conditions and all efforts were made to minimize suffering. No statistical methods were used to predetermine sample size. In our animal study protocol, we state that the number of animals in each experimental group varies, and is based on our prior experience. We use the minimum number of animals per group that will provide meaningful results. Randomization was not used to allocate animals to experimental groups and the animal studies were not blinded.

Virus rescue. The 2009 H1N1pdm virus used in this study is influenza A/California/07/2009. Generation and characterization of the α 2,3 H1N1pdm and α 2,6 H1N1pdm viruses have been described previously¹¹. Genomic sequencing and dose-dependent glycan binding assays confirmed the identity and receptor specificity of viruses generated by reverse genetics. All experiments were performed using viruses passaged no more than three times in MDCK cells (ATCC) or 10-day old embryonated chicken eggs.

Ferret transmission study. All ferrets (*Mustela putorius furo*) were obtained from Triple F Farms (Sayre, PA) and screened by haemagglutination inhibition (HAI) assay before infection to ensure that they were naive to seasonal influenza A and B viruses and the viruses used in this study. The transmission studies were conducted in adult ferrets as previously described¹², male and female ferrets were used in a 3:1 ratio and sample size was based on the capacity of the transmission cages. Ferrets reaching 15–20% weight loss were provided with enriched diet and monitored closely by veterinary staff for altered behaviour.

Environmental conditions inside the laboratory were monitored daily and were consistently $19 \pm 1^\circ\text{C}$ and $56 \pm 2\%$ relative humidity. The transmission experiments were conducted in the same room, to minimize any effects of caging and airflow differences on aerobiology. On day 0 four animals were infected intranasally with 10^6 TCID₅₀ of either α 2,3 H1N1pdm or α 2,6 H1N1pdm virus and placed into the transmission cage. Twenty-four hours post-infection, a naive animal (airborne-contact) was placed into the transmission cage on the other side of a perforated stainless steel barrier. The airborne-contact ferrets were always handled before the infected ferrets. Nasal washes were collected and clinical signs were recorded on alternate days from days 0 to 14. Great care was taken during nasal wash collections and husbandry to ensure that direct contact did not occur between the ferrets. On 14 days post-infection (dpi), blood was collected from each animal for serology. The shortened exposure time study was done similarly except 48 h after the naive recipient animal (airborne contact) was placed into the transmission cage the ferrets were separated into micro-isolator cages. Infected ferrets in the shortened exposure time study were killed on 7 dpi and the airborne contact animals were killed on 21 dpi. The airborne-contact animals were always handled before infected ferrets and all husbandry tools were decontaminated three times between handling of each airborne-contact animal.

Dose-dependent direct binding of influenza viruses. To determine the receptor specificity of the G222D α 2,3 H1N1pdm virus, virus from the nasal wash of a single airborne-contact animal on day 6 post-exposure was propagated once in MDCK cells. This virus stock was inactivated with betapropiolactone and the haemagglutination titre was determined. For the glycan binding assay, 50 μl of 2.4 μM biotinylated glycans were added to wells of streptavidin-coated high binding capacity 384-well plates (Pierce) and incubated overnight at 4°C . The glycans included were 3'SLN, 3'SLN-LN, 3'SLN-LN-LN, 6'SLN and 6'SLN-LN (LN corresponds to lactosamine (Gal β 1-4GlcNAc) and 3'SLN and 6'SLN respectively correspond to Neu5Ac α 2-3 and Neu5Ac α 2-6 linked to LN) that were obtained from the Consortium of Functional Glycomics (<http://www.functionalglycomics.org>). The inactivated G222D virus was diluted to 250 μl with $1 \times$ PBS + 1% BSA. 50 μl of diluted virus was added to each of the glycan-coated wells and incubated overnight at 4°C . This was followed by three washes with $1 \times$ PBST (1X PBS + 0.1% Tween-20) and three washes with $1 \times$ PBS. The wells were blocked with $1 \times$ PBS + 1% BSA for 2 h at 4°C followed by incubation with primary antibody (ferret anti-CA07/09 antisera; 1:200 diluted in $1 \times$ PBS + 1% BSA) for 5 h at 4°C . This was followed by three washes with $1 \times$ PBST and three washes with $1 \times$ PBS. Finally, the wells were incubated with the secondary antibody (goat anti-ferret HRP conjugated antibody from Rockland; 1:200 diluted in 1X PBS + 1% BSA). The wells were washed with $1 \times$ PBST and $1 \times$ PBS as before. The binding signals were determined based on the HRP activity using the Amplex Red Peroxidase Assay (Invitrogen) according to the manufacturer's instructions. Negative controls were uncoated wells (without any glycans) to which just the virus, the antisera and the antibody were added and glycan-coated wells to which only the antisera and the antibody were added.

Ferret replication. We evaluated the replication kinetics of the α 2,3 H1N1pdm virus in the respiratory tract of 6–8-month-old male ferrets as previously

described¹¹. Briefly, all ferrets were screened before infection by HAI assay to ensure that they were naive to seasonal influenza A and B viruses. Animals were infected intranasally with 10^6 TCID₅₀ of α 2,3 H1N1pdm virus in 500 μl . Tissues were harvested to assess viral titres. Tissues were weighed and homogenized in Leibovitz's L-15 (L-15, Invitrogen) at 5% (nasal turbinates and trachea) or 10% (lung) weight per volume (W/V). The soft palate was homogenized in 1 ml of L-15. Clarified supernatant was aliquoted and titred on MDCK cells. The 50% tissue culture infectious dose (TCID₅₀) per gram of tissue was calculated by the Reed and Muench method³¹.

Influenza A virus full genome sequencing. The influenza A genomic RNA segments were simultaneously amplified from 3 μl of purified RNA (from homogenized ferret tissue) using a multi-segment RT-PCR strategy (M-RTPCR)³². In a separate reaction, each HA segment was amplified using HA-specific primers (swH1ps-1A-F: 5'-AGCAAAGCAGGGGAAAAACAAAAGCAAC-3'; swH1ps-1777A-R: 5'-AGTAGAAACAAGGGTGTTCCTCATGC-3'). Analysis of influenza viral RNA from ferret trachea and region of nasal turbinates enriched for respiratory epithelium, between the canine and second premolar teeth, was collected from tissue stored in RNAlater (Ambion) and total RNA was extracted using the RNeasy kit (Qiagen). For these samples, nested HA-specific small amplicons were generated using HA-specific PCR primers (outer primer pair H1-399F: 5'-AGCTCAGTGCATCA TTTGAAAG-3' and H1-961R: 5'-TGAATGGGAGGCTGGTGTT-3'; and inner primer pair H1-468 F: 5'-AACAAAGGTGTAACGGCAGC-3' and H1-884R: 5'-AATGATAATACCAGATCCAGCAT-3'). Illumina libraries were prepared from M-RTPCR products and from HA-specific RT-PCR products using the Nextera DNA Sample Preparation Kit (Illumina, Inc.) with half-reaction volumes.

After PCR amplification, 10 μl of each library derived from M-RTPCR products was pooled into a 1.5 ml tube; separately, 10 μl of each library derived from HA-specific amplicons was pooled into a 1.5 ml tube. Each pool was cleaned two times with Ampure XP Reagent (Beckman Coulter) to remove all leftover primers and small DNA fragments. The first and second cleanings used $1.2 \times$ and $0.6 \times$ volumes of Ampure XP Reagent, respectively. The cleaned pool derived from M-RTPCR products was sequenced on the Illumina HiSeq 2000 instrument (Illumina, Inc.) with 100-bp paired-end reads, while the cleaned pool derived from HA-specific amplicons was sequenced on the Illumina MiSeq v2 instrument with 300-bp paired-end reads. For additional sequencing coverage, and the HA-specific small amplicons, samples were re-sequenced using the Ion Torrent platform. M-RTPCR products were sheared for 7 min, and Ion-Torrent-compatible bar-coded adapters were ligated to the sheared DNA using the Ion Xpress Plus Fragment Library Kit (Thermo Fisher Scientific, Waltham, MA, USA) to create 400-bp libraries. Libraries were pooled in equal volumes and cleaned with the Ampure XP Reagent. Quantitative PCR was performed on the pooled, barcoded libraries to assess the quality of the pool and to determine the template dilution factor for emulsion PCR. The pool was diluted appropriately and amplified on ion sphere particles (ISPs) during emulsion PCR on the Ion One Touch 2 instrument (Thermo Fisher Scientific). The emulsion was broken, and the pool was cleaned and enriched for template-positive ISPs on the Ion One Touch ES instrument (Thermo Fisher Scientific). Sequencing was performed on the Ion Torrent PGM using a 318v2 chip (Thermo Fisher Scientific).

Deep sequencing analysis. Deep sequencing preparation, collection and analysis were conducted by investigators who were blinded to the experimental groups. For virus sequence assembly, all sequence reads were sorted by barcode, trimmed, and *de novo* assembled using CLC Bio's *clc_novo_assemble* program (Qiagen, Hilden, Germany). The resulting contigs were searched against custom full-length influenza segment nucleotide databases to find the closest reference sequence for each segment. All sequence reads were then mapped to the selected reference influenza A virus segments using CLC Bio's *clc_ref_assemble_long* program.

Minor allele variants were identified using FindStatisticallySignificantVariants (FSSV) software (<http://sourceforge.net/projects/elvira/>). The FSSV software applies statistical tests to minimize false-positive SNP calls generated by Illumina sequence-specific errors (SSEs) described in ref. 33. SSEs usually result in false SNP calls if sequences are read in one sequencing direction. The FSSV analysis tool requires observing the same SNP at a statistically significant level in both sequencing directions. Once a minimum minor allele frequency threshold and significance level are established, the number of minor allele observations and major allele observations in each direction and the minimum minor allele frequency threshold are used to calculate *P*-values based on the binomial distribution cumulative probability. If the *P*-values calculated in both sequencing directions are less than the Bonferroni-corrected significance level, then the SNP calls are accepted. A significance level of 0.05 (Bonferroni-corrected for tests in each direction to 0.025) and a minimum minor allele frequency threshold of 3% were applied for this analysis. Differences in the consensus sequence compared to the reference sequence were identified using CLC Bio's *find_variations* software. The identified consensus and minor allele variations were analysed by assessing

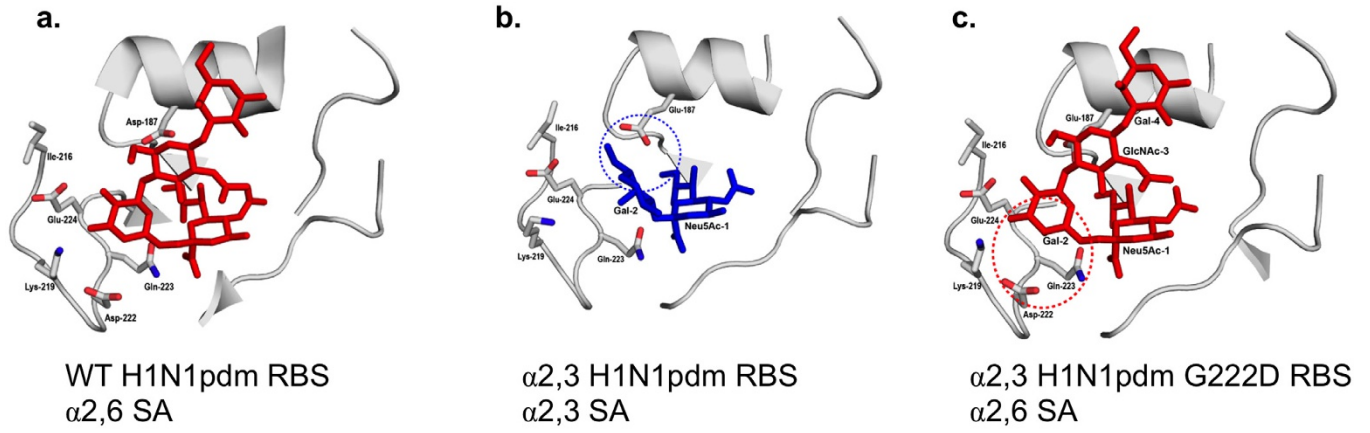
the functional impact on coding sequences or other regions based on overlap with identified features of the genome. For each sample, the reference sequence was annotated using VIGOR software³⁴, and then the variant data and genome annotation were combined using VariantClassifier software³⁵ to produce records describing the impacts of the identified variations.

Lectin and immunohistochemistry. Lectin histochemistry was performed as described previously for plant lectins³⁶ and purified HA protein³⁷. For plant lectin staining, the soft palate was subjected to microwave-based antigen retrieval using a citrate buffer and was then incubated with FITC-conjugated *Sambucus nigra* agglutinin (SNA) and biotinylated *Maackia amurensis* agglutinin (MAL II) lectins (Vector Laboratories), followed by a streptavidin-Alexa-Fluor594 conjugate (Invitrogen). For SC18 staining, the tissue sections were incubated with pre-complexed purified His-tagged SC18 HA protein, mouse anti-His antibody (Abcam), and goat anti-mouse IgG secondary antibody conjugated to Alexa-Fluor 488 (Molecular Probes) at a 4:2:1 ratio. Nuclei were counter stained with DAPI (Vector Laboratories) and sections were mounted with either ProLong Gold anti-fade reagent (Invitrogen) or Fluoromount-G (Southern Biotech). Images were captured either on an Olympus BX51 microscope with an Olympus DP80 camera or a Leica SP5 confocal microscope.

Ferret nasal turbinate biopsy samples were obtained from an uninfected 8 month old ferret as follows: the head was dissected sagittally to expose two halves of the ferret nasal turbinates; biopsy of turbinates between the canine and second

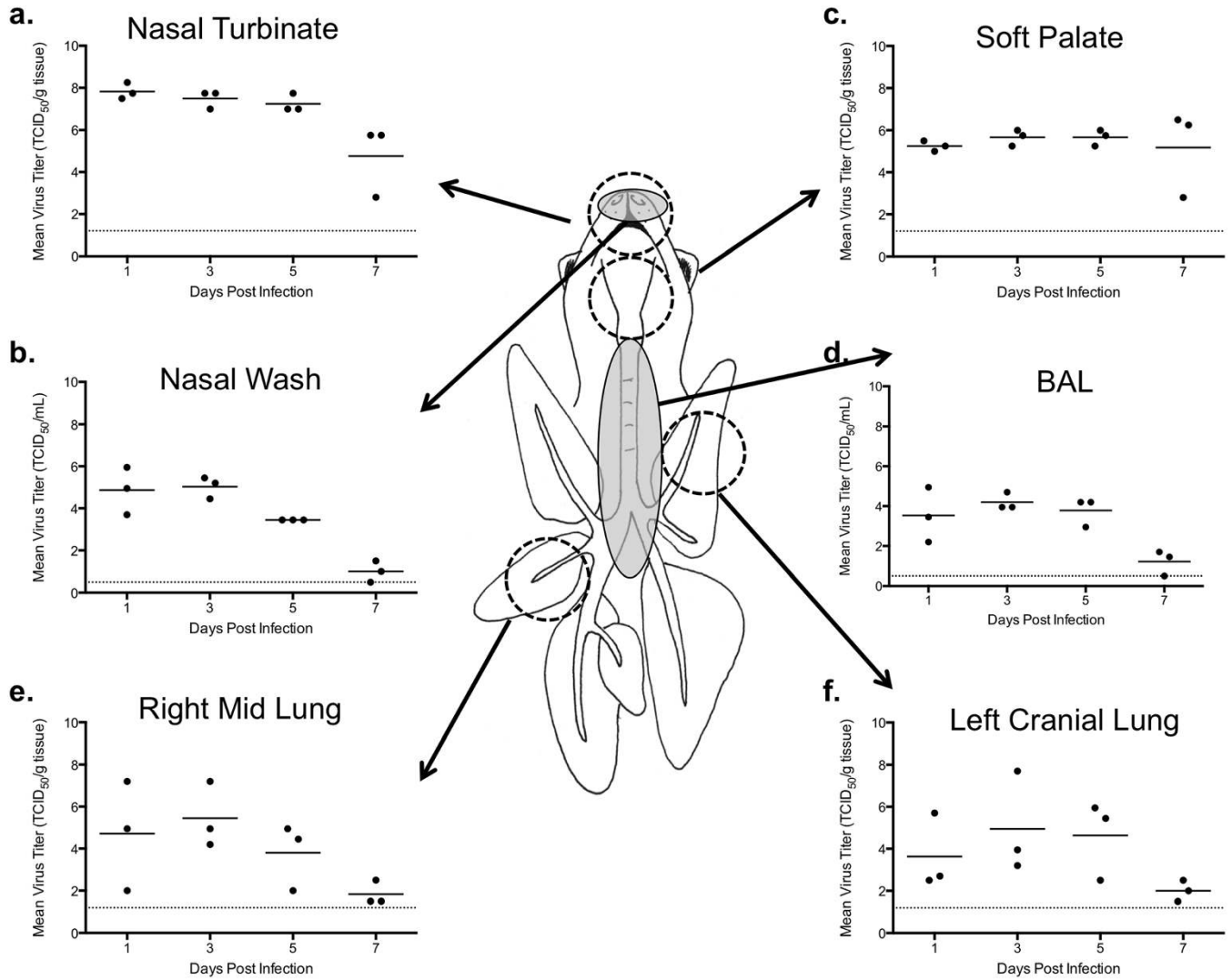
premolar represented respiratory epithelium and biopsy of turbinates at the molar tooth represented olfactory epithelium. A schematic depicting these two areas is shown in Extended Data Fig. 6h. Pig soft palate tissue sections were a gift from X. J. Meng and P. Pineyro. Pig soft palate tissues were collected from four 56-day-old mixed-breed commercial swine and fixed in 10% formalin. Soft palate tissues from four adult cadavers were obtained from the Maryland State Anatomy Board, Department of Health and Mental Hygiene in Baltimore, Maryland.

31. Reed, L. J. & Muench, H. A simple method of estimating fifty percent endpoints. *Am. J. Hyg.* **27**, 493–497 (1938).
32. Zhou, B. *et al.* Single-reaction genomic amplification accelerates sequencing and vaccine production for classical and Swine origin human influenza A viruses. *J. Virol.* **83**, 10309–10313 (2009).
33. Nakamura, K. *et al.* Sequence-specific error profile of Illumina sequencers. *Nucleic Acids Res.* **39**, e90 (2011).
34. Wang, S., Sundaram, J. P. & Stockwell, T. B. VIGOR extended to annotate genomes for additional 12 different viruses. *Nucleic Acids Res.* **40**, W186–W192 (2012).
35. Li, K. & Stockwell, T. B. VariantClassifier: A hierarchical variant classifier for annotated genomes. *BMC Res. Notes* **3**, 191 (2010).
36. Matsuoka, Y. *et al.* African green monkeys recapitulate the clinical experience with replication of live attenuated pandemic influenza virus vaccine candidates. *J. Virol.* **88**, 8139–8152 (2014).
37. Jayaraman, A. *et al.* Decoding the distribution of glycan receptors for human-adapted influenza A viruses in ferret respiratory tract. *PLoS ONE* **7**, e27517 (2012).



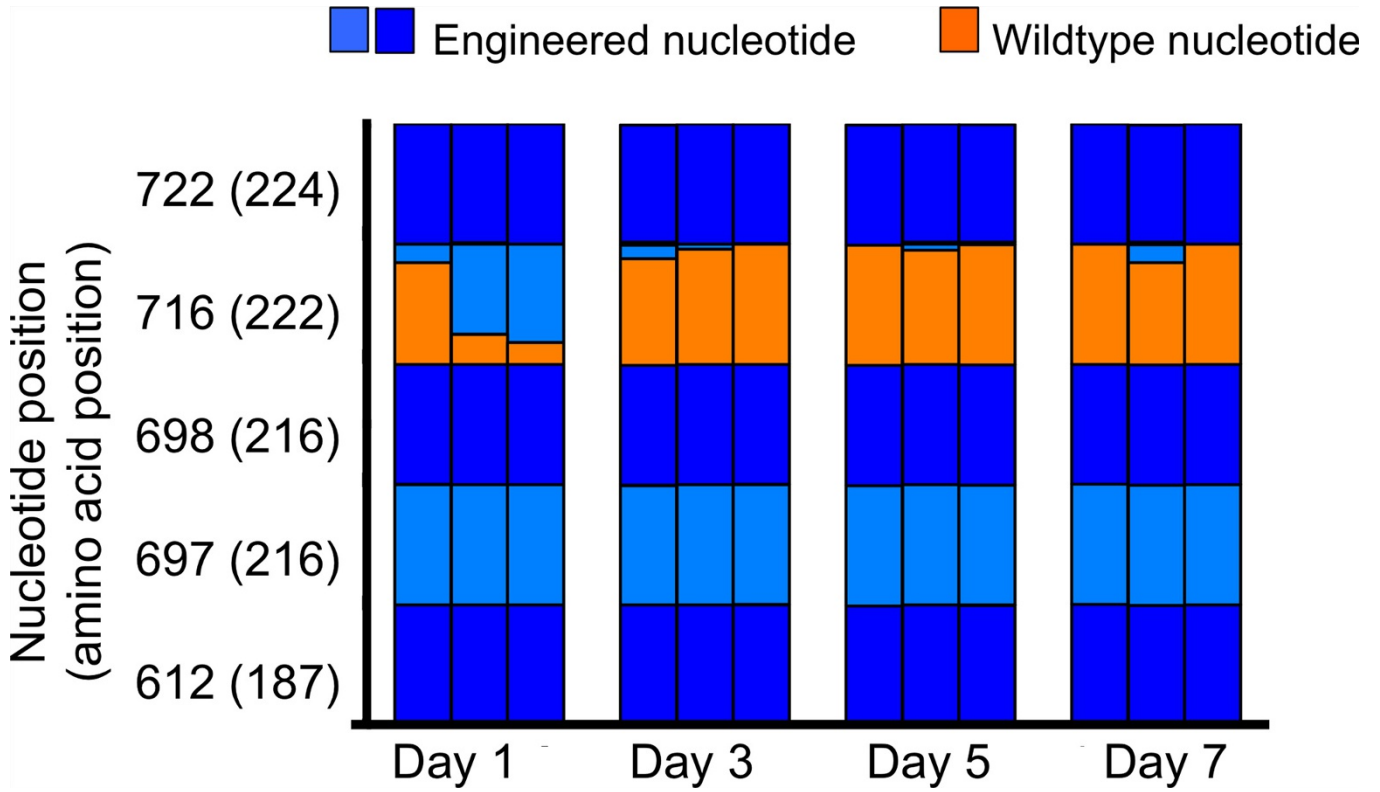
Extended Data Figure 1 | Amino acids in the receptor binding site of the 2009 H1N1pdm HA that bind to α 2,3-linked and α 2,6-linked sialic acid glycan receptor. a–c, Ribbon diagrams of the 2009 H1N1pdm HA receptor

binding pocket interacting with an α 2,6-linked sialic acid glycan in the pocket (a), an α 2,3 H1N1pdm HA with α 2,3-linked sialic acid glycan (b), an α 2,3 G222D revertant H1N1pdm HA with an α 2,6-linked sialic acid glycan (c).



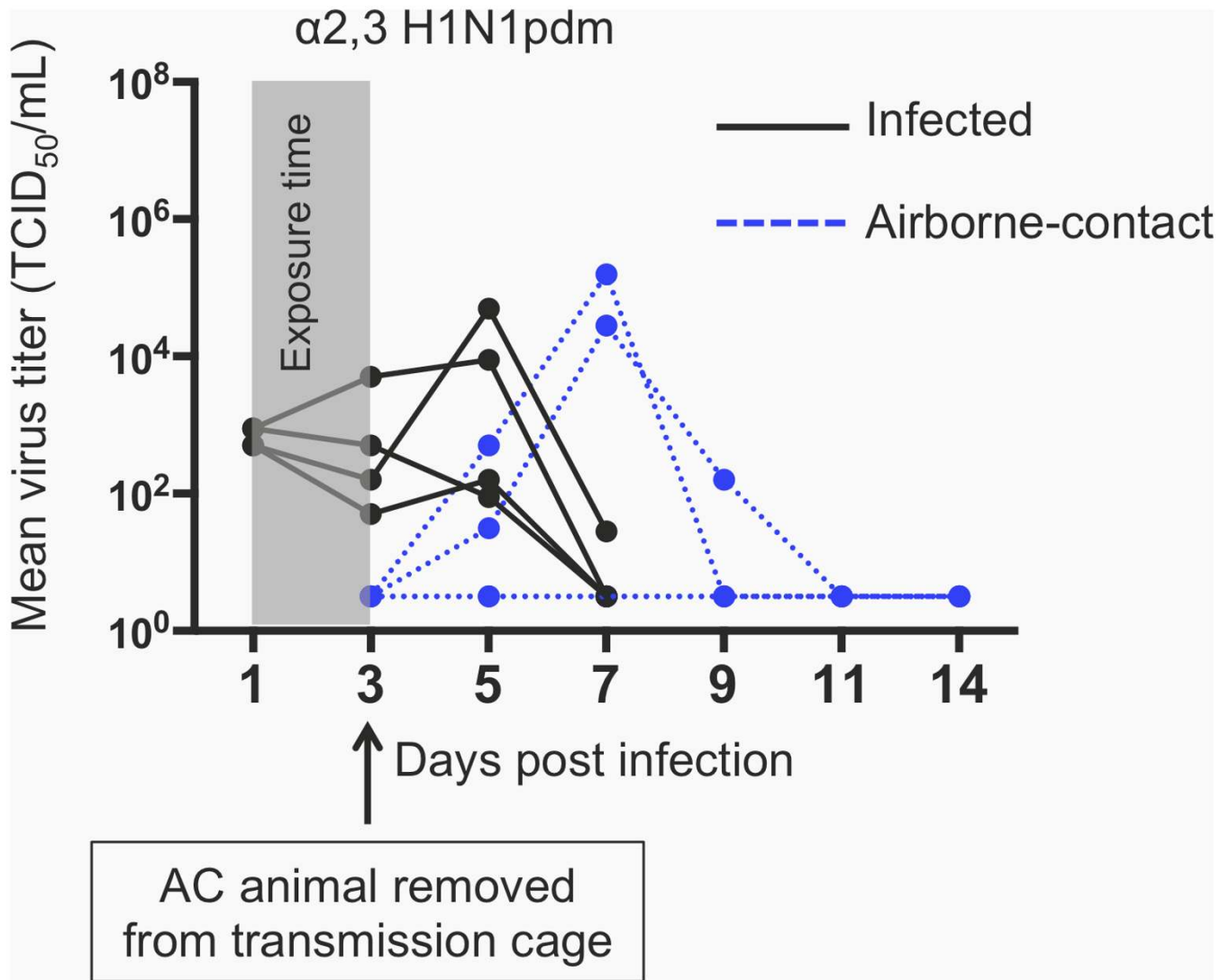
Extended Data Figure 2 | Replication of $\alpha 2,3$ H1N1pdm virus in ferret respiratory tract. a–f, We confirmed that the $\alpha 2,3$ H1N1pdm virus replicated to high titres on days 1, 3, 5 and 7 in different parts of the ferret respiratory tract.

Each tissue homogenate is highlighted with a dashed circle; the grey circles represent washes. Each point represents a single animal. The horizontal black line indicates the mean viral titre on a given day.



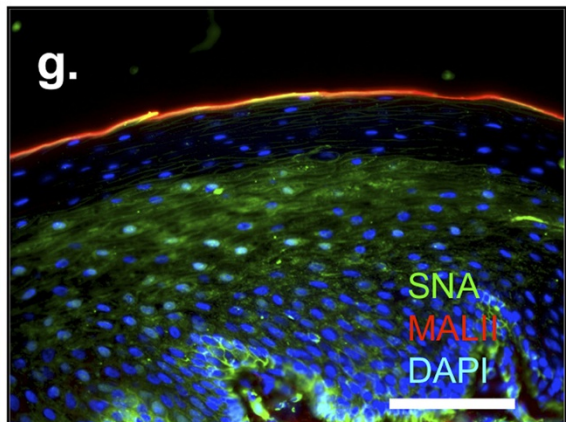
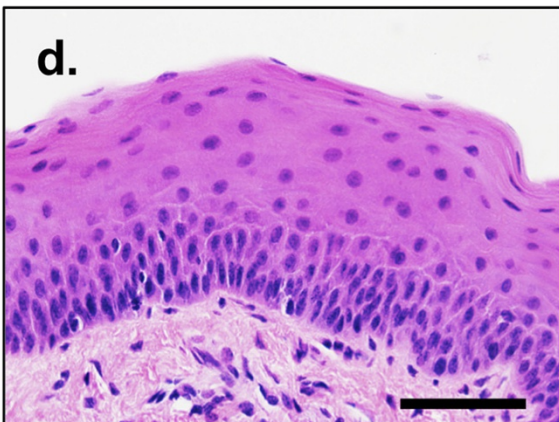
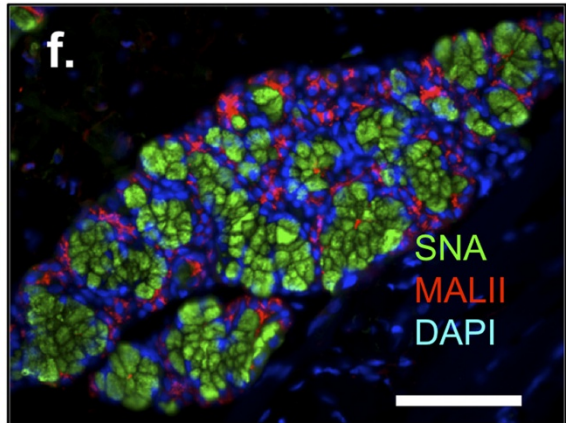
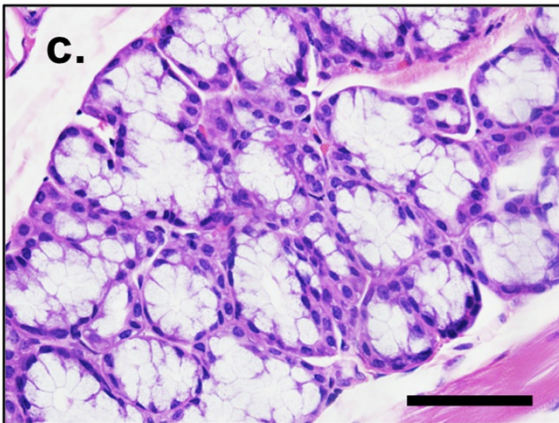
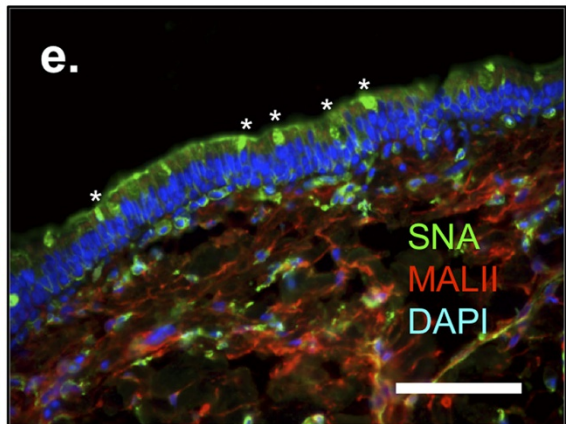
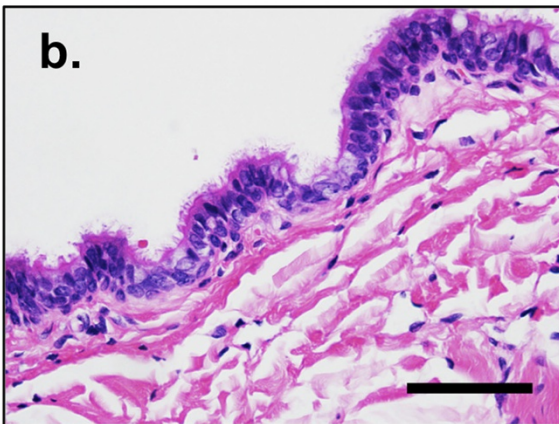
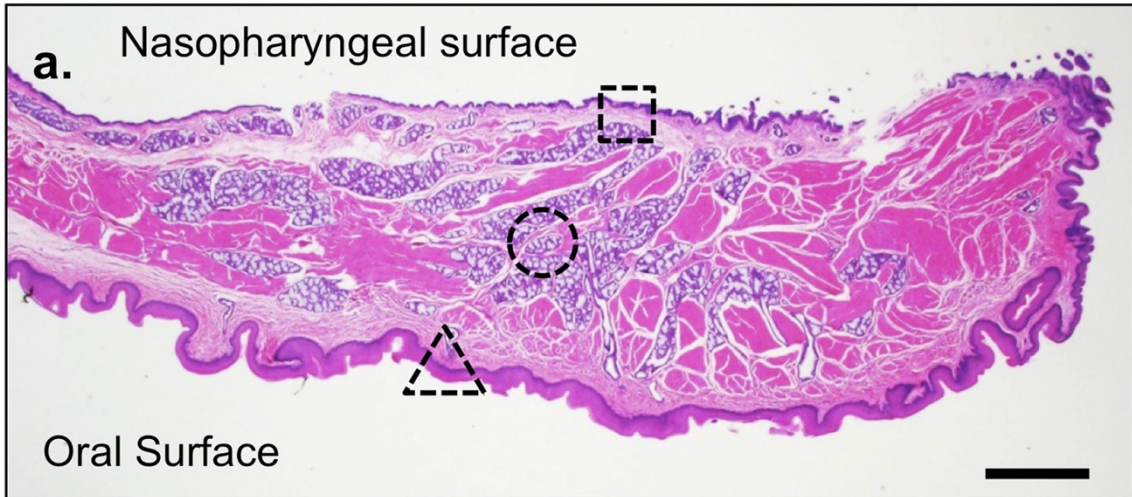
Extended Data Figure 3 | Stability of engineered mutations in viruses replicating in the soft palate. Deep sequencing of the HA gene segment from virus populations in the soft palate from 1, 3, 5 and 7 dpi reveals a rapid change

at position 222, but no change in the other engineered sites. The engineered sites are highlighted in blue, while the wild-type nucleotide is in orange. Each bar represents a single animal.



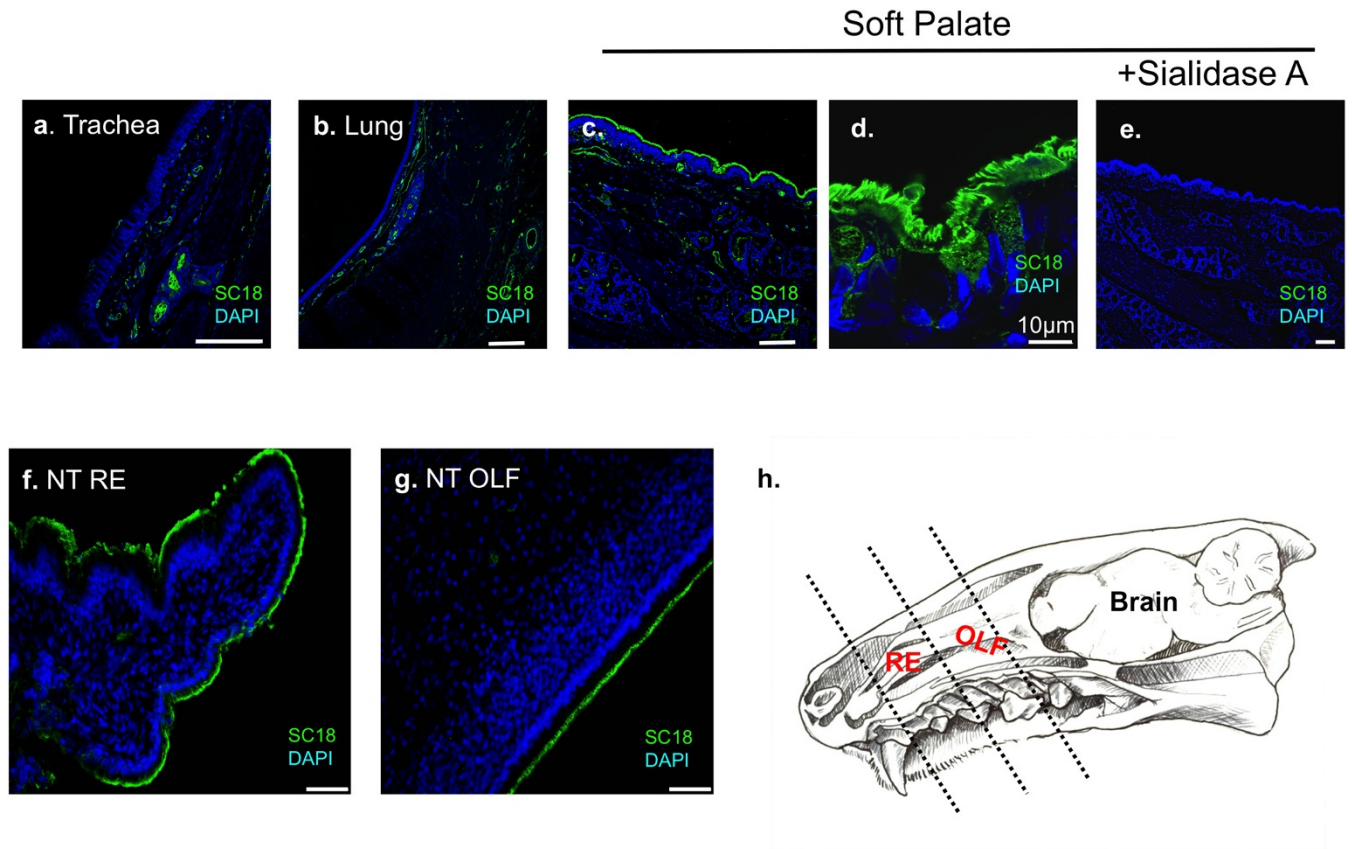
Extended Data Figure 4 | Airborne transmission of $\alpha 2,3$ H1N1pdm virus after 48 h exposure time. Transmission studies were performed with 4 pairs of animals (8 animals total) in double secure cages with perforated dividers. One ferret in each pair was infected with 10^6 TCID₅₀ of the indicated virus; a naive ferret (referred to as airborne-contact) was introduced into the adjacent compartment 24 h later. The airborne-contact animal was removed from the

transmission cage on day 3 post-infection as indicated by the black arrow. Nasal secretions were collected every other day for 14 days. Viral titres from the nasal secretions are graphed for each infected or airborne-contact animal. The grey shading indicates the exposure time between the infected and airborne-contact animals.



Extended Data Figure 5 | Influenza receptor distribution on ferret soft palate. Haematoxylin and eosin (H&E) staining of the soft palate from an uninfected ferret highlights the nasopharyngeal and oral surfaces. Scale bar, 1.25 mm. **a**, Areas highlighted in parts **b–g** are marked with dashed polygons: square, nasopharyngeal surface (**b** and **e**); circle, submucosal gland (**c** and

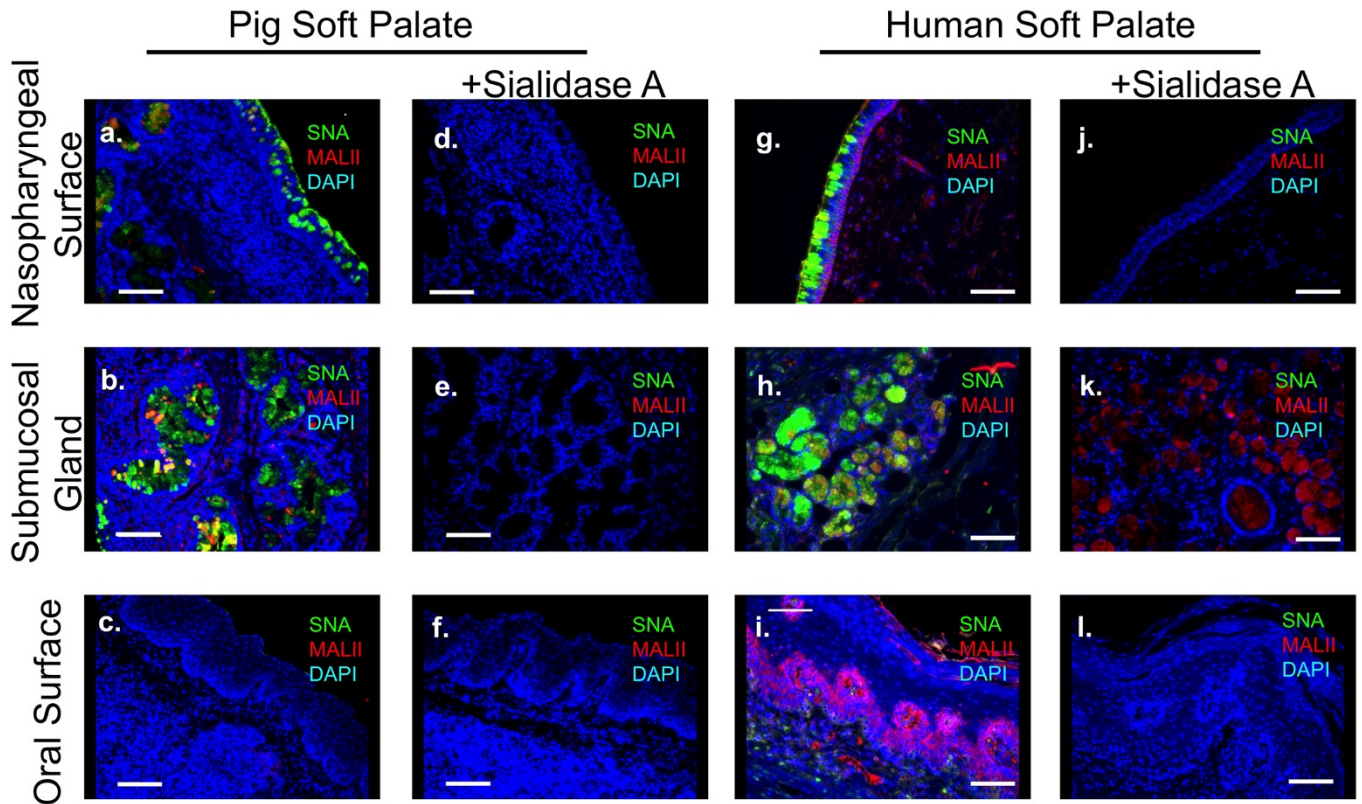
f); and triangle, oral surface (**d** and **g**). **b–d**, H&E staining of these regions, reproduced from Fig. 4a–c, are shown. **e–g**, Staining with plant lectins specific for α 2,6 sialic acid (SNA) and α 2,3 sialic acid (MAL II) are shown. Scale bars are 100 μ m in images **b–g**.



Extended Data Figure 6 | SC18 staining of ferret respiratory tissues.

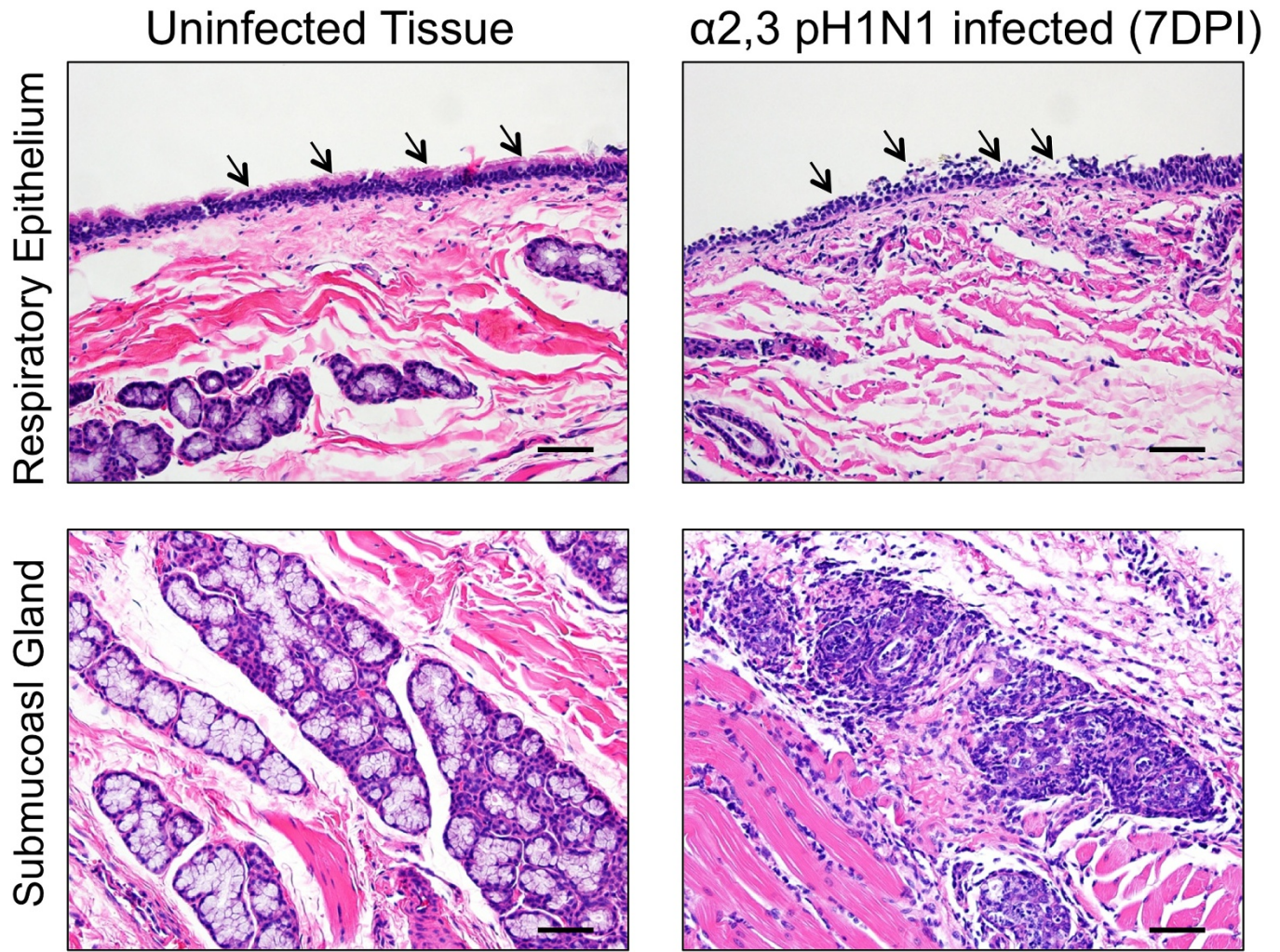
a–d, f, g. Sections of ferret trachea (**a**), lung (**b**), soft palate (**c, d**), biopsy of nasal turbinate (NT) tissue with respiratory epithelium (RE) (**f**) and olfactory epithelium (OLF) (**g**) were stained with purified SC18 HA protein to identify areas expressing long-chain $\alpha 2,6$ -linked sialic acids. **h.** Illustration of ferret head (sectioned along the midline) highlighting the anatomical locations of

respiratory epithelium and OLF tissues. Goblet cells on the respiratory epithelium of the soft palate (nasopharyngeal surface) also stained positive for SC18 (**d**). **e.** Absence of SC18 staining after sialidase A treatment indicates the high specificity of SC18 for the respiratory epithelium of the soft palate. All scale bars are 100 μm unless indicated.



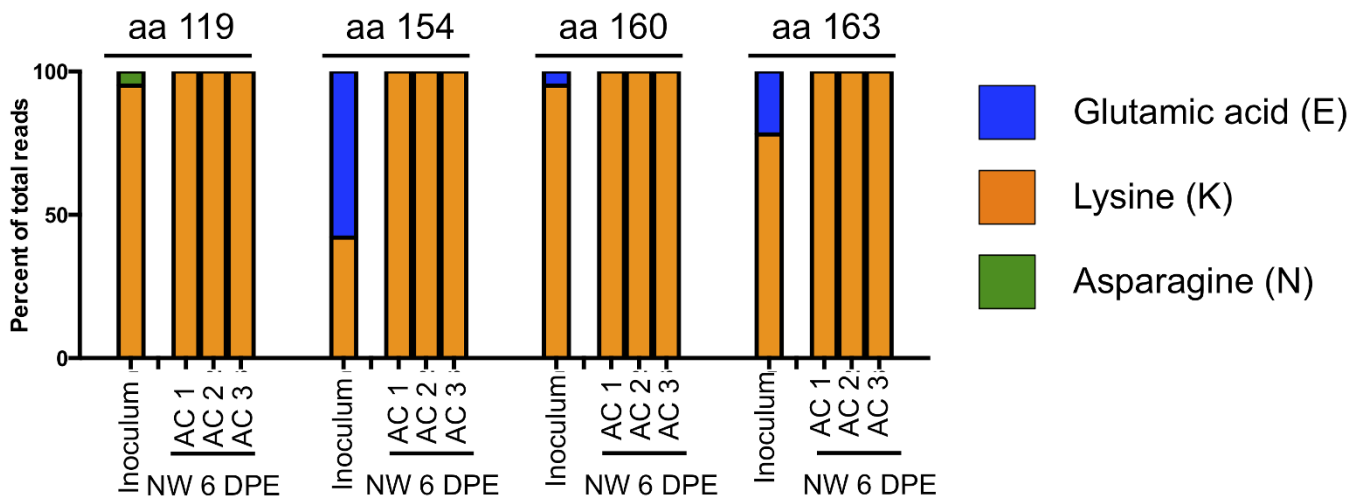
Extended Data Figure 7 | Influenza receptor distribution on pig and human soft palate. a–c, g–i, Pig (a–c) and human (g–i) soft palate tissues were stained with plant lectins SNA and MALII which are commonly used as markers for α 2,6 and α 2,3 sialic acids, respectively. d–f, j–l, Sialidase-A-treated control was run for each sample to ensure specificity of plant lectins and are displayed. Expression of α 2,6 sialic acids (SNA staining) is found on the ciliated respiratory epithelium and goblet cells of the nasopharyngeal surface and in the

submucosal glands of both the pig and human soft palate. Expression of α 2,3-linked sialic acids is low in the pig soft palate and found primarily in goblet cells and submucosal glands. In the human soft palate, MALII (α 2,3-linked sialic acids) staining sensitive to sialidase A treatment is found in the goblet cells and respiratory epithelium of the nasopharyngeal surface and in the basal cells of the oral surface. MALII staining in the submucosal glands was not sensitive to sialidase A treatment. Scale bars, 100 μ m.

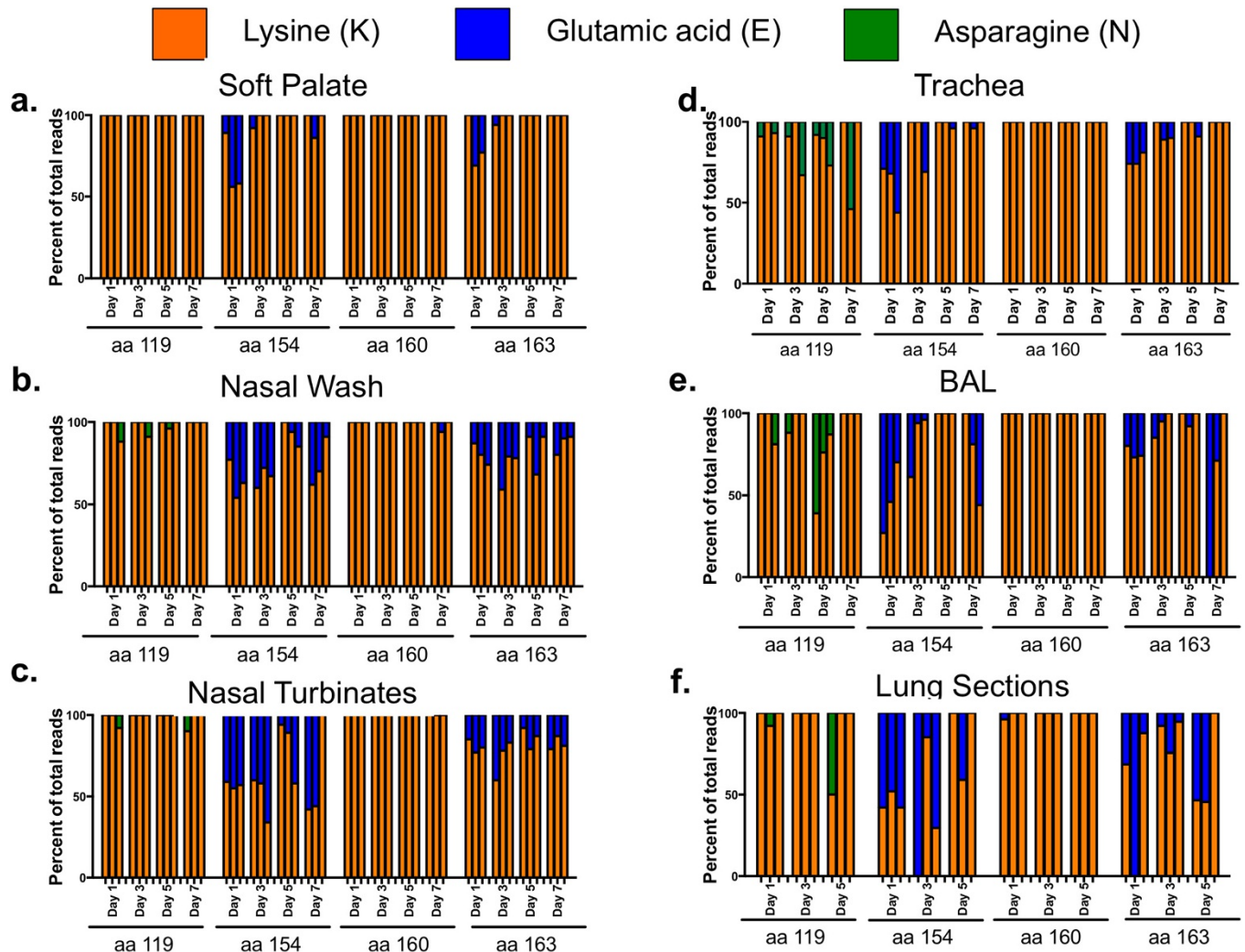


Extended Data Figure 8 | Pathology of the soft palate during infection with α 2,3 H1N1 pdm virus. The soft palate was removed from three ferrets infected with α 2,3 H1N1 pdm virus on 7 dpi. The tissue sections were stained with

haematoxylin and eosin. Black arrows indicate the ciliated respiratory epithelium of the soft palate tissue (nasopharyngeal surface). Scale bars, 100 μ m in all images.



Extended Data Figure 9 | Quasi-species in putative lysine fence. Deep sequencing analysis of the α 2,3 H1N1pdm inoculum revealed a mixed population at four lysine residues surrounding the receptor binding site of the HA protein. The lysine fence was restored in viruses from the nasal wash of airborne-contact animals from 6 days post-exposure (DPE). Each bar represents a single animal, and each amino acid (aa) that contained a quasi-species is indicated.



Extended Data Figure 10 | Quasi-species of lysine fence in various ferret respiratory tissue sections. Deep sequencing of viruses from respiratory tissues of ferrets infected with α 2,3 H1N1pdm virus. **a–f**, Virus populations from the soft palate (**a**), nasal wash (**b**), nasal turbinates (**c**), trachea

(**d**), bronchoalveolar lavage (BAL) (**e**), or lung sections (**f**) were analysed and the proportion of lysine, glutamic acid, or asparagine are presented. Each bar represents a single animal. The lung section is an average of the right middle lung lobe and a portion of the left caudal lung tissue.