PLOS PATHOGENS

Characterization of changes in the hemagglutinin that accompanied the emergence of H3N2/1968 pandemic influenza viruses

Johanna West[1], Juliane Röder[1¤a], Tatyana Matrosovich[1], Jana Beicht[1¤b], Jan Baumann[1¤c], Nancy Mounogou Kouassi[1¤d], Jennifer Doedt[1¤e], Nicolai Bovin[2], Gianpiero Zamperin[3], Michele Gastaldelli[3], Annalisa Salviato[3], Francesco Bonfante[3], Sergei Kosakovsky Pond[4], Sander Herfst[5], Ron Fouchier[5], Jochen Wilhelm[6], Hans-Dieter Klenk[1†], Mikhail Matrosovich[1]*

1 Institute of Virology, Philipps University, Marburg, Germany, 2 Shemyakin-Ovchinnikov Institute of Bioorganic Chemistry of the Russian Academy of Sciences, Moscow, Russia, 3 Division of Comparative Biomedical Sciences, Istituto Zooprofilattico Sperimentale delle Venezie, Legnaro, Italy, 4 Institute for Genomics and Evolutionary Medicine, Temple University, Philadelphia, Pennsylvania, United States of America, 5 Department of Viroscience, Erasmus Medical Centre, Rotterdam, Netherlands, 6 Institute of Lung Health (ILH), Universities of Giessen and Marburg Lung Center (UGMLC), Giessen, Germany

† Deceased.
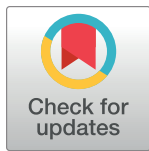¤a Current address: BioNTech SE, Mainz, Germany
¤b Current address: Research Center for Emerging Infections and Zoonoses, University of Veterinary Medicine Hannover, Hannover, Germany
¤c Current address: WHO Regional Office for Europe, Copenhagen Ø, Denmark
¤d Current address: Department for Viral Zoonoses-One Health, Heinrich Pette Institute, Leibniz Institute for Experimental Virology, Hamburg, Germany
¤e Current address: Myriad International GmbH, Köln, Germany
* m.matrosovich@gmail.com

Abstract

The hemagglutinin (HA) of A/H3N2 pandemic influenza viruses (IAVs) of 1968 differed from its inferred avian precursor by eight amino acid substitutions. To determine their phenotypic effects, we studied recombinant variants of A/Hong Kong/1/1968 virus containing either human-type or avian-type amino acids in the corresponding positions of HA. The precursor HA displayed receptor binding profile and high conformational stability typical for duck IAVs. Substitutions Q226L and G228S, in addition to their known effects on receptor specificity and replication, marginally decreased HA stability. Substitutions R62I, D63N, D81N and N193S reduced HA binding avidity. Substitutions R62I, D81N and A144G promoted viral replication in human airway epithelial cultures. Analysis of HA sequences revealed that substitutions D63N and D81N accompanied by the addition of N-glycans represent common markers of avian H3 HA adaptation to mammals. Our results advance understanding of genotypic and phenotypic changes in IAV HA required for avian-to-human adaptation and pandemic emergence.

Author summary

Pandemic influenza A viruses (IAVs) originate from animal IAVs and always contain the envelope protein hemagglutinin (HA) of the animal parent. Molecular changes essential for the animal-to-human adaptation of the HA are not well defined. To address this problem, we studied phenotypic effects of eight amino acid substitutions that accompanied evolution of the HA of the 1968 pandemic IAV from its inferred avian precursor. We also analysed patterns of evolution of these HA positions of IAVs in birds and mammals. Our results predict that the 1968 pandemic IAV acquired its HA from a wild duck virus. We showed that two substitutions were primarily responsible for increased HA binding to human-type receptors and, in addition, marginally decreased stability of the HA, whereas three other substitutions reduced binding avidity of the HA. We demonstrated that substitutions accompanied by the addition of N-glycan at either position 63 or 81 represent a previously unrecognized common marker of avian-to-mammalian adaptation of IAVs. Our findings highlight the importance of the studies on previous pandemic IAVs for the influenza risk assessment and pandemic preparedness.

Introduction

Wild aquatic birds represent the major natural reservoir of IAVs, which occasionally transmit, adapt and circulate for prolonged periods of time in domestic birds and mammals [1,2]. Because animal IAVs do not replicate efficiently in humans, zoonotic transmissions of IAVs are typically restricted to isolated cases of infection (for a recent review, see [3]). If a zoonotic IAV against which people have no protective immunity acquires the ability to transmit efficiently in humans, it may initiate an influenza pandemic. Genetic and virological data available for the four last pandemic IAVs (H1N1/1918, H2N2/1957, H3N2/1968, and H1N1/2009) indicate that they all contained antigenically novel hemagglutinin (HA) gene segments derived from animal IAVs; the other gene segments originated from either animal or contemporary human IAVs (for reviews, see [4,5]). Thus, it is particularly important to understand which adaptive changes in the HA were required for the emergence of previous pandemic viruses from their animal precursors.

The HA mediates attachment of IAVs to sialic acid-containing glycan receptors on cells. Tropism, replication efficiency and pathogenicity of IAVs in different host species strongly depend on the optimal interplay between viral receptor-binding properties and spectra of sialoglycans expressed in target tissues of these species (for reviews, see [6–8]). HAs of the previous pandemic IAVs differed from avian HAs by one or two amino acid substitutions in the conserved positions of the receptor-binding site (RBS). These substitutions were found to be essential for the switch of the HA receptor specificity from preferential binding to Neu5Acα2-3Gal-terminated glycans (avian-type receptors) to preferential binding to Neu5Acα2-6Gal-terminated glycans (human-type receptors). In the case of H2N2/1957 and H3N2/1968 IAVs, substitutions Q226L and G228S were responsible for this switch in receptor specificity. In the case of H1N1/1918 and H1N1/2009 IAVs, this role was played by substitutions E190D and G225D/E (for recent reviews, see [9,10]). It remains unexplored whether other substitutions in the HA of pandemic IAVs were required for adaptation to receptors in humans, for example, by adjusting HA interactions with sub-terminal oligosaccharide parts of the receptors and/or modulating binding avidity.

After endocytosis and acidification of endosomes, the HA of IAVs undergoes a low-pH-triggered conformational transition that mediates fusion between the viral and endosomal

membranes. The conformational stability of the HA determines both the pH range of viral-endosomal fusion and stability of the virus in the environment. There is growing evidence that the pH optimum of fusion and stability of the HA differ between IAVs from different host species and that these differences may affect viral host range, pathogenicity, airborne transmission and pandemic potential (for reviews, see [11,12]). Human IAVs typically have a lower fusion pH optimum (from 5.0 to 5.4) than swine IAVs and zoonotic poultry IAVs of the H5 and H7 subtypes (pH from 5.6 to 6.2). By contrast with IAVs in poultry, IAVs of wild aquatic birds with different HA subtypes including H3 HA often display high conformational stability [13–15]. The HAs of the H1N1/1918, H2N2/1957 and H3N2/1968 pandemic IAVs had a pH optimum of fusion typical for human viruses (5.1–5.4) [15,16]. The earliest isolates of the H1N1/2009 had a less stable HA (pH optimum of fusion 5.4–5.5), but more stable variants were selected during a few months of viral circulation in humans [17,18]. The fusion pH and stability of the immediate HA precursors of the pandemic viruses were not studied, and it remains obscure whether alterations of these properties played a role in pandemic emergence.
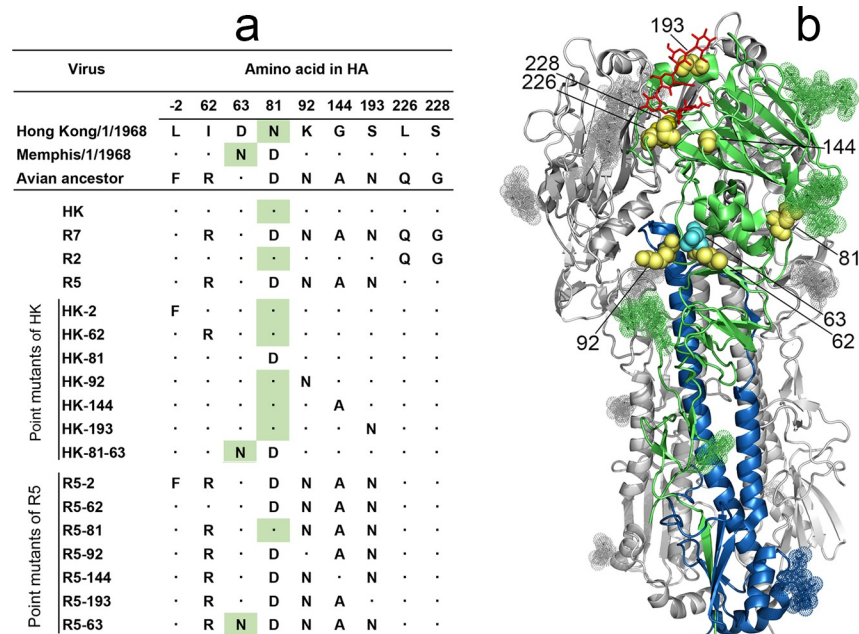
We previously studied adaptive changes in the HA of the pandemic IAV A/Hong Kong/1/1968 (H3N2), which differed from the inferred avian ancestor HA by eight amino acid substitutions. Introduction of avian-virus-like amino acids at positions 226 and 228 of the HA altered cell tropism, reduced replication efficiency in cultures of human airway epithelial cells and abolished transmission of the virus in experimentally infected pigs [19,20]. A combination of avian-type amino acid reversions at five other HA positions impeded replication in human airway cultures and markedly impaired transmissibility in pigs [21]. These results confirmed the critical role of substitutions Q226L and G228S in the avian-to-human transmission of the H3 HA and suggested that at least some of the other substitutions contributed to the emergence of the H3N2 pandemic virus.

In this study, we wished to further characterize changes in the HA that accompanied its avian-to-human adaptation during generation of the 1968 pandemic IAVs. We also wished to identify which substitutions, in addition to substitutions Q226L and G228S, played a role in the adaptation to humans. To address these questions, we prepared a panel of 18 recombinant variants of A/Hong Kong/1/1968 (H3N2) containing either human-type or avian-type amino acids at HA positions that separated the H3N2/1968 viruses from their inferred avian ancestor. We compared these IAVs for their membrane fusion activity and stability, receptor-binding properties, replication efficiency in MDCK cells and cultures of human airway epithelial cells. We also analyzed patterns of evolution of H3 HA codons in question in IAVs from different host species.

Results

Preparation of recombinant variants of A/Hong Kong/1/1968-PR8 (H3N2) with substitutions in the HA

The 1968 pandemic IAV HA differed from the avian precursor by 8 amino acid substitutions [21,22] (Figs 1 and S1, S5, and S6). Seven substitutions were shared by all viral strains isolated in the first year of the pandemic. One of these substitutions, F(-2)L, was located in the cleavable signal peptide and six substitutions were located in the HA1 subunit of the mature HA protein, with G228S and Q226L in the RBS, A144G and N193S at the rim of the RBS and R62I and N92K in the vestigial esterase subdomain. The eighth substitution from D to N occurred at either position 63 or position 81 in the vestigial esterase subdomain of the HA and generated a new glycosylation site, N_{63}-C_{64}-T_{65} or N_{81}-E_{82}-T_{83}, respectively. Either site was glycosylated with attached N-glycans detectable by X-ray analysis (see, for example, structures 4O58.pdb and 2YPG.pdb). Pandemic IAVs with these two HA variants differing solely by the location of

a

Virus	-2	62	63	81	92	144	193	226	228
Hong Kong/1/1968	L	I	D	N	K	G	S	L	S
Memphis/1/1968	·	·	N	D	·	·	·	·	·
Avian ancestor	F	R	·	D	N	A	N	Q	G
HK	·	·	·	·	·	·	·	·	·
R7	·	R	·	D	N	A	N	Q	G
R2	·	·	·	·	·	·	·	Q	G
R5	·	R	·	D	N	A	N	·	·
HK-2	F	·	·	·	·	·	·	·	·
HK-62	·	R	·	·	·	·	·	·	·
HK-81	·	·	·	D	·	·	·	·	·
HK-92	·	·	·	·	N	·	·	·	·
HK-144	·	·	·	·	·	A	·	·	·
HK-193	·	·	·	·	·	·	N	·	·
HK-81-63	·	·	N	D	·	·	·	·	·
R5-2	F	R	·	D	N	A	N	·	·
R5-62	·	·	·	D	N	A	N	·	·
R5-81	·	R	·	·	N	A	N	·	·
R5-92	·	R	·	D	·	A	N	·	·
R5-144	·	R	·	D	N	·	N	·	·
R5-193	·	R	·	D	N	A	·	·	·
R5-63	·	R	N	D	N	A	N	·	·

(Rows HK through HK-81-63 are grouped as "Point mutants of HK"; rows R5-2 through R5-63 are grouped as "Point mutants of R5".)

Fig 1. HA sequences and designations of 2:6 recombinant IAVs used in this study. (**a**) Amino acid differences between HAs of two 1968 pandemic virus lineages, their putative avian ancestor and 2:6 recombinant PR8-based viruses. Dots depict sequence identity with the HA of A/Hong Kong/1/1968. Numbering of amino acid positions starts from the N-terminus of the mature protein. Green background marks asparagine residues of glycosylation sites 63–65 and 81–83. (**b**) Location of amino acid substitutions shown as yellow space-filling models on the X-ray structure of the H3 HA complex with human receptor analogue LSTc (2YPG.pdb) [23]. Two HA monomers are colored gray, and the third monomer is colored green (HA1) and blue (HA2). LSTc is shown as red stick model, N-linked glycans are shown as dotted space-filling models. Cyan spheres show location of N_{63} present in the HA of A/Memphis/1/1968 lineage. The model was generated using PyMOL 2.0.6 (Schrödinger, LLC).

new N-linked glycan co-circulated during 1968 and a few years afterwards. The A/Memphis/1/1968-like IAVs containing an N-glycan at position 63 (NG_{63}) became extinct after 3 years of circulation; the A/Hong Kong/1/1968-like IAVs (NG_{81}) continued to cause seasonal influenza outbreaks until 1976 and were substituted by a drift lineage that lost NG_{81} and gained NG_{63} (for the evolution of HA glycosylation sites 63 and 81 in human H3N2 viruses, see S2 Fig).

To study phenotypic effects of the substitutions, we generated a panel of 2:6 recombinant IAVs that contained HA and NA of A/Hong Kong/1/1968 (H3N2) and the remaining 6 gene segments of the laboratory strain PR8. The panel included the virus with wild type HA (HK) and its HA variants with either human-type or avian-type amino acids at corresponding HA positions (Fig 1a). The R7 variant carried all seven avian-type substitutions in the mature HA, and thus mimicked the HA structure of the avian precursor of the 1968 pandemic IAVs. Two viruses were made to represent combined effects of substitutions at either positions 226 and 228 (variant R2) or at five other positions (variant R5). Single-point mutants of HK served to determine effects of reversions from human-type to avian-type amino acid at individual HA positions. The double mutant HK-81-63 represented the sequence of A/Memphis/1/1968 and was used to study the effect of the NG_{63}. The point mutants of R5 were made to study effects of individual reversions from avian-type to human-type amino acids in the context of the avian HA with human-type L_{226} and S_{228}. Finally, variants HK-2 and R5-2 were prepared to characterize the phenotype of the amino acid substitution in the signal peptide.

Effects of substitutions on HA conformational stability and membrane fusion activity

We first compared stability and fusion properties of HK, its avian precursor R7 and the intermediate variants R2 and R5 (Fig 2). As these characteristics critically depend on the low-pH-triggered conformational transition of the HA, we determined the pH at which the virus HA changed its conformation by studying pH-induced alteration of HA sensitivity to protease digestion (Fig 2a). R7 and R2 underwent conformational transition at a slightly lower pH than did HK and R5. In agreement with this finding, R7 and R2 initiated syncytia formation in MDCK cells at about 0.1 units of pH lower than did HK and R5 (Fig 2b). To corroborate observed differences in viral fusion pH, we compared inhibition of viral infection in MDCK cells by ammonium chloride which counteracts endosomal acidification (Fig 2c). R2 and R7 were more sensitive than HK and R5 to inhibition by NH_4Cl, confirming that R2 and R7 require a lower pH in endosomes for fusion and cell entry.

Acid stability of the HA correlates, at least partially, with HA resistance to heat and denaturing agents. To test effects of the latter two factors of environmental stability we studied
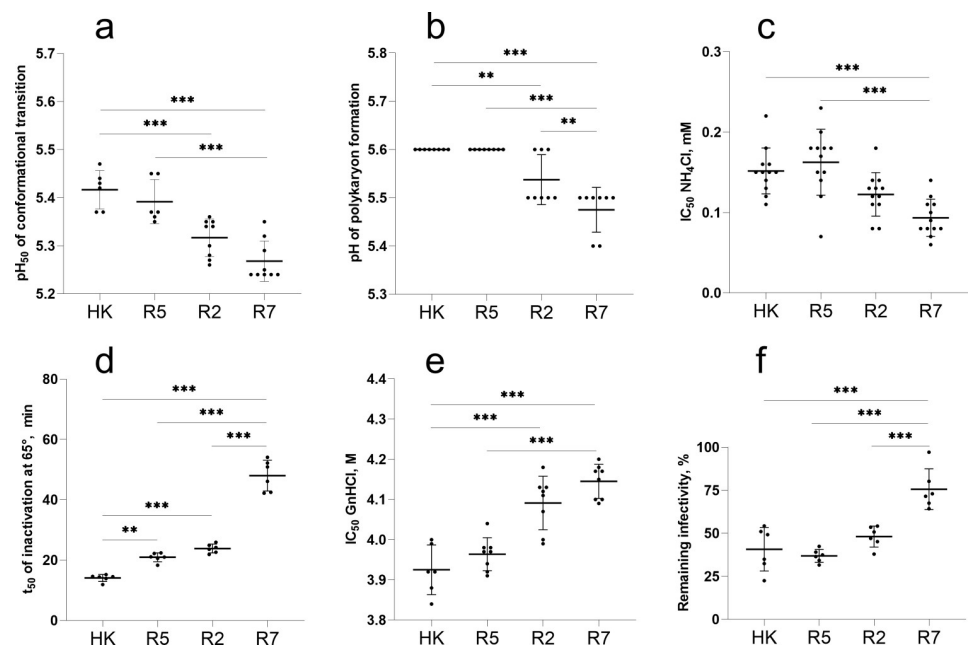
Fig 2. Conformational stability and membrane fusion properties of HK, R5, R2 and R7. (**a**) pH of acid-induced conformational transition of HA. Solid-phase adsorbed viruses were incubated in acidic buffers and treated with proteinase K. Viral binding of fet-HRP was assayed, and pH values that corresponded to 50% reduction of HA binding activity (pH_{50}) were determined from binding-versus-pH curves. (**b**) pH threshold of polykaryon formation. Inoculated MDCK cells were cultured for 16 h, treated with trypsin and exposed to different pH buffers. After returning to neutral medium and incubation for 3 h, the cells were fixed, stained and analysed under the microscope. The data show highest pH values at which polykaryon formation was detected. (**c**) Inhibition of viral infection by ammonium chloride. MDCK cells were inoculated in the presence of various concentrations of NH_4Cl, incubated overnight, fixed, and immunostained for NP. Concentrations of NH_4Cl that reduced numbers of infected cells by 50% (IC_{50}) were determined from dose-response curves. (**d**) HA stability at elevated temperature. Solid-phase adsorbed viruses were incubated in PBS at 65°C for different time periods and assayed for their binding to fet-HRP to determine incubation time required for 50% reduction of the binding activity (t_{50}). (**e**) HA stability in chaotropic buffer. Solid-phase adsorbed viruses were incubated in buffers containing GnHCl for 60 min at 4°C washed with PBS and assayed for binding to fet-HRP. Data show concentrations of GnHCl that reduced viral binding activity by 50%. (**f**) Reduction of infectivity after incubation of the viruses for 2 h at 45°C determined by focus assay in MDCK cells. All panels show data points, mean values and SDs from 1 to 4 independent experiments performed with 2 to 7 replicates. P values for the differences between the viruses were determined with Tukey's multiple comparison procedure.

inactivation of the HA receptor-binding activity by the chaotropic agent guanidinium chloride (GnHCl) (Fig 2e) and heat treatment (Fig 2d) as well as the effect of heat treatment on viral infectivity (Fig 2f). In general, the stability of the viruses in all three assays correlated with their acid stability. R7 was the most stable variant, HK was least stable, whereas R2 and R5 displayed intermediate stability.

We next studied the effects of non-226/228 single-point HA substitutions in HK and R5 using three assays (S3 Fig). None of the substitutions affected pH optimum of the HA conformational transition in the protease sensitivity assay. Four point mutants differed from the corresponding parental viruses by their sensitivity to ammonium chloride. Among them, mutants R5-63 and R5-81 containing N-linked glycan NG_{63} and NG_{81}, respectively, were less sensitive than R5, whereas the mutant HK-81 lacking NG_{81} was more sensitive than HK. We concluded that N-glycan-containing variants entered the cells from less acidic endosomal compartment. This finding agrees with the study in which the presence of N-glycans in the globular head of H1N1 IAVs reduced receptor-binding avidity and facilitated HA-mediated fusion [24]. In the polykaryon formation assay, the avian-type amino acid in the signal peptide of the variants HK-2 and R5-2 correlated with a minor (ΔpH, +0.1) but reproducible elevation of their pH threshold of fusion (S3b Fig). This effect, albeit small, was unexpected given that the signal peptide is not present in the mature HA and that HK-2 and R5-2 did not differ from their parents, HK and R5, in the conformational transition assay and in any other assays used (see below). Alteration of the signal peptide could potentially affect intracellular maturation, secretion and incorporation of the HA into viral particles, and recent bioinformatics analysis revealed that passaging of human IAVs in cell culture was occasionally accompanied by mutations in the signal peptide [25]. These notions prompt further work on potential effects of the HA signal peptide on replication and interspecies adaptation of IAVs.

Collectively, this part of the study revealed that a combination of substitutions Q226L and G228S increased the pH of the conformational transition of the pandemic virus HA by about 0.15 pH units and as a consequence marginally decreased its environmental stability. The effects of other avian-to-human point substitutions and of their combination on conformational stability and membrane fusion activity were either smaller or below the detection limit of the assays.

Receptor-binding profile of the avian precursor of the 1968 pandemic viruses and effect of amino acid substitutions on the HA preference for the type of Neu5Ac-Gal linkage

To characterize receptor-binding properties of the viruses we determined their binding to soluble synthetic sialoglycopolymers (SGPs) carrying multiple copies of sialyloligosaccharide moieties attached to a hydrophilic polymeric carrier. The structures of the sialyloligosaccharide parts and designations of SGPs are shown in the Materials and Methods section. The high molecular mass (1 MDa) SGPs contained about 50 times more copies of the sialoligand per macromolecule and bound to IAVs with much higher avidity than structurally identical 20-kDa SGPs. As a result, utilization of the 1-MDa SGPs was instrumental for comparison of IAVs with large differences in binding avidity, such as avian and human IAVs, whereas the 20-kDa SGPs were more useful than 1-MDa SGPs for characterization of IAVs with minor differences in the binding avidity [26,27].

Although most avian IAVs use Neu5Acα2-3Gal-terminated glycans as their cellular receptors, viruses adapted to species of the orders *Anseriformes*, *Charadriiformes* and *Galliformes* typically differ by their ability to recognize sub-terminal parts of the receptor glycans ([28] and references therein). We assumed that analysis of the fine receptor binding specificity of R7
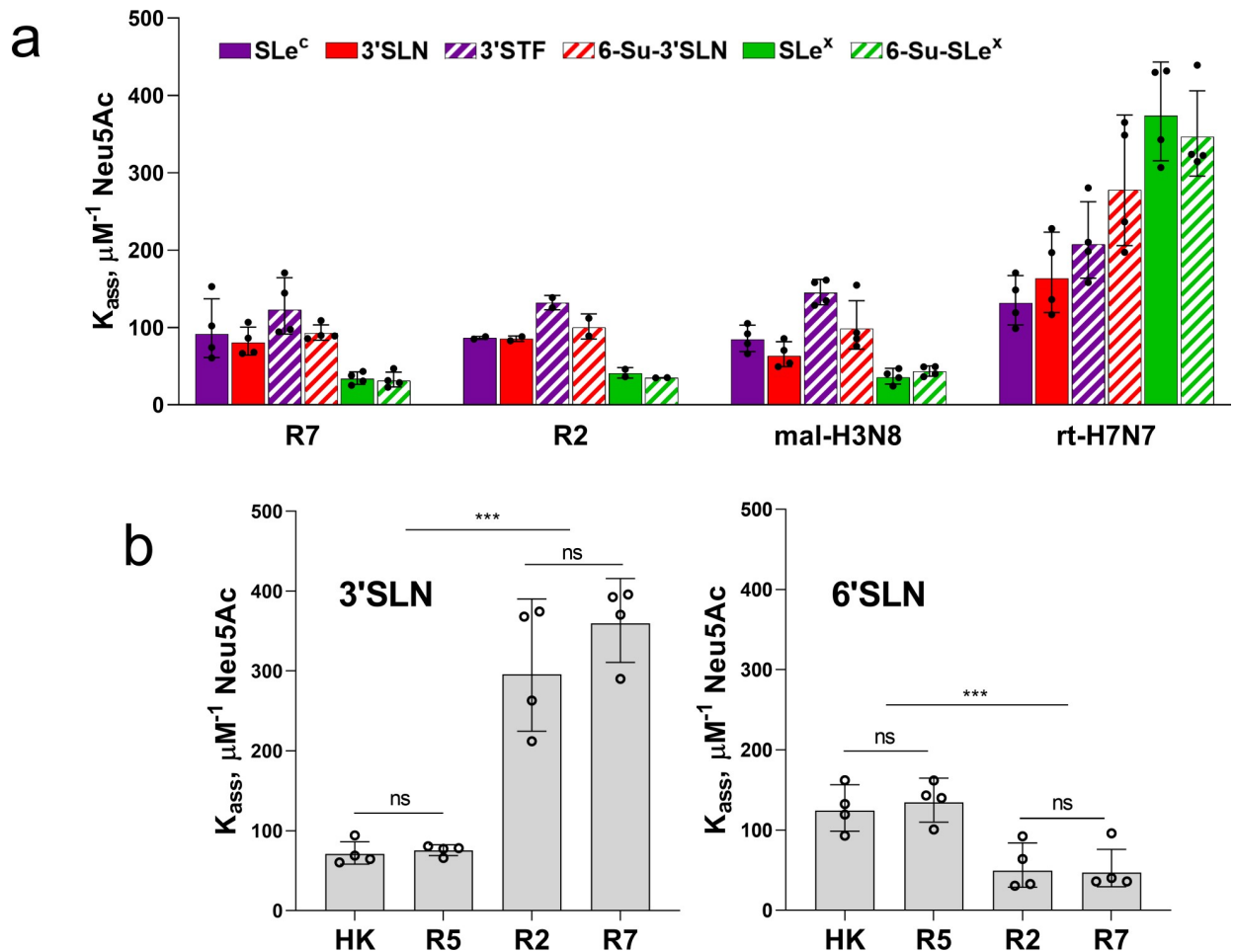
Fig 3. Binding of IAVs to biotinylated SGPs. The solid-phase adsorbed viruses were allowed to bind biotinylated SGPs from solution, and association constants of viral complexes with SGPs (K_{ass}) were determined as described in Materials and Methods. Higher values of K_{ass} reflect stronger binding. (**a**) Binding to six low molecular mass Neu5Acα2-3Gal-containing SGPs differing by structure of penultimate sugar moieties. Wild type IAVs A/mallard/Alberta/279/1998 (H3N8) and A/ruddy turnstone/ Delaware/2378/1988 (H7N7) were tested in parallel with R7 and R2. Two to 4 experiments were performed on different days with one replicate for each virus-SGP pair per experiment. (**b**) Binding to high molecular mass SGPs 3'SLN and 6'SLN. Two experiments were performed on different days with two replicates per each virus-SGP pair per experiment. All panels show the individual values adjusted for day as described in Materials and Methods with geometric mean (bars) and SDs. P values for the differences between Kass presented in the panel 3a are shown in S3 Table.

may predict which avian species perpetuated the precursor of the 1968 pandemic virus. To this end, we determined binding of R7 and R2 to a panel of Neu5Acα2-3Gal-containing 20-kDa SGPs. Two representative wild-type viruses, A/mallard/Alberta/279/1998 (H3N8) (mal-H3N8) and A/ruddy turnstone/ Delaware/2378/1988 (H7N7) (rt-H7N7), were used for a comparison (Fig 3a). R7 shared the binding profile with mal-H3N8 which represented typical receptor-binding properties of IAVs in ducks. Similar to other duck viruses [28], R7 and mal-H3N8 bound to fucosylated receptors SLe^x and 6-Su-SLe^x significantly less efficiently than to non-fucosylated receptors and bound to 3'SLN and its sulfated counterpart 6-Su-3'SLN with similar efficiency. A second control virus, rt-H7N7, displayed receptor-binding characteristics which are often shared by IAVs of *Charadriiformes and Galliformes* [28,29]. Namely, this virus displayed the highest binding avidity for fucosylated receptors SLe^x and 6-Su-SLe^x and bound to 6-Su-3'SLN significantly better than to 3'SLN and SLe^c. R7 did not display these features, thus showing no signs of adaptation of the R7 HA to gulls, shorebirds or gallinaceous poultry. The

binding profile of the variant R2 did not significantly differ from that of R7, suggesting that five human-type amino acid substitutions separating R2 from R7 had minor (if any) effect on HA binding to Neu5Acα2-3Gal-terminated receptors.

Substitutions L226Q and G228S in the HAs of pandemic H3N2/1968 viruses switched the viral recognition of the type of Neu5Ac-Gal linkage [8,10]. To determine whether and to what extent the other five substitutions in the mature HA contributed to this switch, we compared binding of HK, R5, R2 and R7 to 6'SLN and 3'SLN. High molecular mass SGPs were used to ensure measurable binding of each virus to both SGPs (Fig 3b). As expected, comparison of R7 with R5 and comparison of R2 with HK showed that a combination of substitutions Q226L and G228S strongly reduced HA binding to 3'SLN and increased HA binding to 6'SLN; the magnitude of the second effect was noticeably smaller. No significant differences in the viral binding profiles were observed in pairs R7/R2 and HK/R5. These results indicated that a combination of substitutions at positions 62, 81, 92, 193 and 144 of the pandemic virus HA had much lower (if any) effect than substitutions Q226L/G228S on viral recognition of the Neu5Ac-Gal linkage type.

Effects of non-226/228 substitutions on binding avidity of the HA

As limited experiments with 1-MDa SGPs (Fig 3b) did not reveal significant differences in their binding to HK and R5, we employed more sensitive receptor-binding assays to assess effects of amino acid substitutions separating these viruses. Fig 4a shows data on binding of HK, R5 and their point mutants to low molecular mass 6'SLN. HK bound to 6'SLN about 2-fold weaker than R5 indicating that a combination of 5 human-type amino acids decreased binding avidity. Three single-point HK mutants, HK-62, HK-81 and HK-193 displayed elevated binding avidity. Three other HK mutants, HK-92, HK-144 and HK-2, did not differ from the parental HK. Remarkably, the double mutant HK-81-63 bound to 6'SLN significantly weaker than HK-81 and showed binding avidity that was comparable to that of HK. Thus, whereas substitution N81D and loss of NG_{81} increased HA binding to 6'SLN, the substitution D63N and attachment of NG_{63} fully compensated for this effect. The effects of single human-type substitutions in R5 HA on viral binding to 6'SLN well correlated with the effects of corresponding avian-type substitutions in HK HA (Fig 4a). Namely, substitutions at positions 62, 81, 193 and 63 in R5 decreased binding avidity, whereas substitutions at positions -2, 92 and 144 had no significant effect. This correlation indicated that the effects of individual non-226/228 substitutions on HA binding to 6'SLN do not depend on the identity of the other 4 amino acids separating HK from R5. We therefore conclude that non-226/228 amino acids studied are not involved in substantial epistatic interactions.

Because of the relatively weak binding of HK, R5 and point mutants to Neu5Acα2-3Gal-containing receptors we could not reliably quantify binding of these IAVs to corresponding 20-kDa SGPs. Using more sensitive but less discriminative 1-MDa SGPs, we found that R5 bound stronger than HK to both 3'SLN and SLe^c and that most single-point mutants did not significantly differ in this respect from the parental IAVs (Fig 4b and 4c). The effects were only observed in the case of HK-193 (elevated binding to 3'SLN), R5-62 (reduced binding to SLe^c) and R5-81 (reduced binding to both SGPs); these effects directly correlated with the effects of corresponding substitutions on viral binding to 6'SLN. These results suggested that single-point substitutions had a weaker effect than their combination on HA binding to 3'SLN and SLe^c.

To further characterize receptor-binding properties of the HA mutants, we studied inhibition of viral single-cycle infection in MDCK cells in the presence of *Vibrio cholerae* sialidase which reduced levels of sialic acids on the cell surface (Fig 4d). A higher value of 50%
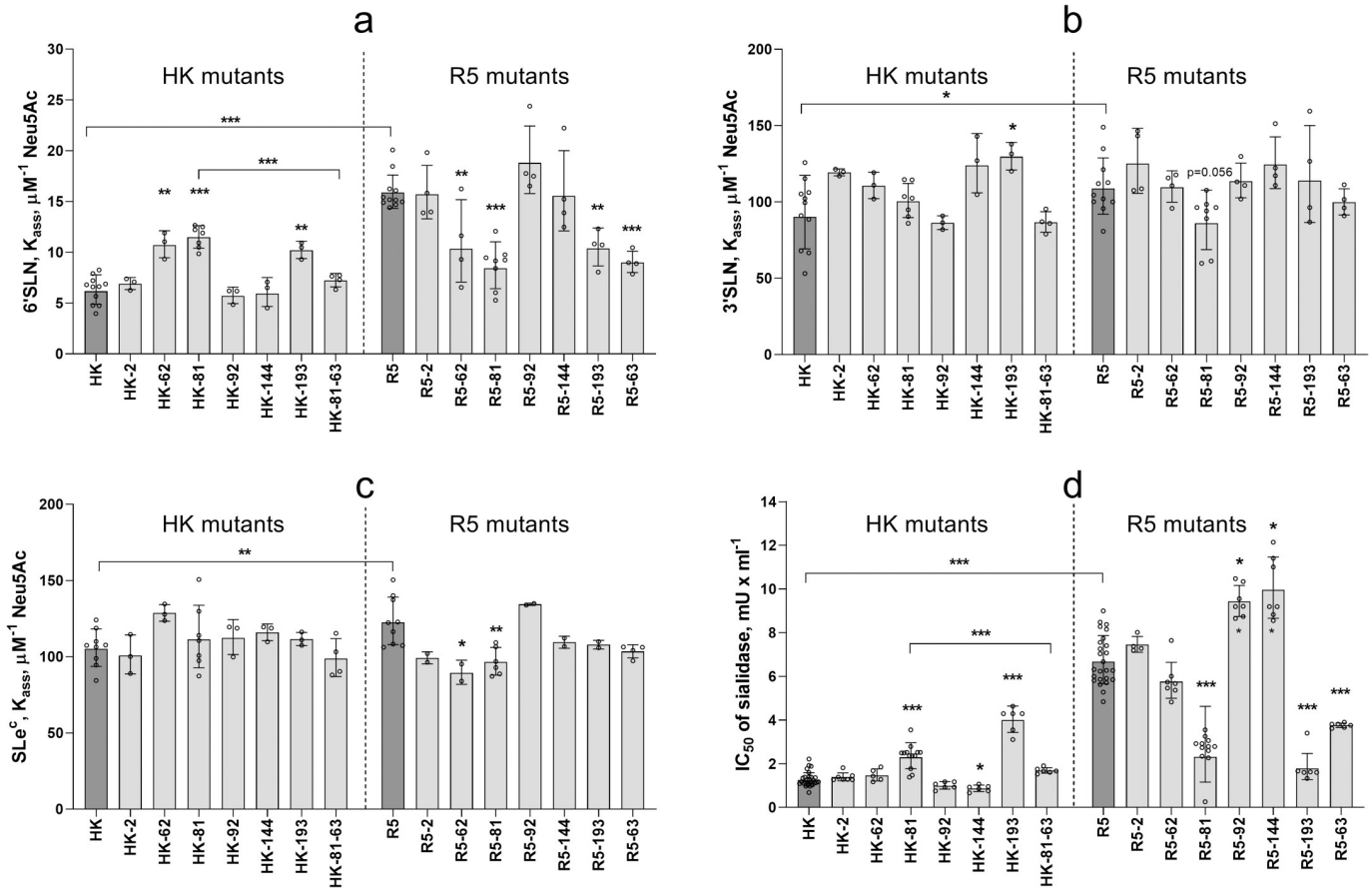
Fig 4. Receptor-binding properties of HA point mutants of HK and R5. (a-c) Association constants of viral complexes with biotinylated SGPs 6'SLN (20 kDa), 3'SLN and SLe^c (both 1 MDa) were determined as described in Materials and Methods. Data represent combined results from 4 to 11 experiments performed on different days with 1 replicate for each virus-SGP pair per experiment. (**d**) Inhibition of viral cell entry by *Vibrio cholerae* sialidase. MDCK cells were incubated with solutions of gradually diluted sialidase for 30 min, inoculated with 200 FFU of the viruses without removing sialidase, fixed after one cycle of replication and immunostained for viral NP. The figure shows concentrations of sialidase that reduced numbers of infected cells by 50% (IC$_{50}$). From 2 to 9 experiments were performed on different days using 3 to 4 replicates per virus. All panels show the individual values adjusted for day as described in Materials and Methods with geometric mean (bars) and SDs. Vertical dotted line separates point mutants of HK and point mutants of R5. Asterisks depict P values for the differences between single-point mutants and the corresponding parental virus, either HK or R5 (dark gray bars). Asterisks over horizontal lines depict differences between HK and R5 and between HK-81 and HK-81-63.

inhibitory concentration of sialidase (IC$_{50}$) suggested that the virus can infect cells expressing lower amounts of sialic acid; this effect was interpreted as an indication of a higher binding avidity. The avidity of the viruses for receptors on MDCK cells correlated to a large extent with viral binding to 6'SLN (compare Fig 4a and 4d). Thus, in both assays, i) R5, HK-81 and HK-193 bound to the cells stronger than HK, ii) HK-81-63 bound weaker than HK-81, iii) R5-81, R5-193 and R5-63 bound weaker than R5, iv) substitution at position -2 of the signal peptide affected binding of neither R5, nor HK. In contrast with a significant effect of substitutions at position 62 on binding to 6'SLN (Fig 4a), these substitutions showed no apparent effect on binding to MDCK cells. As another distinction from the 6'SLN binding data, R5-92 and R5-144 bound to MDCK cells somewhat stronger than R5 and HK-144 bound weaker than HK.

The following conclusions could be made from the binding data. A combination of human-type substitutions separating HK from R5 reduced HA binding to both Neu5Acα2-6Gal- and Neu5Acα2-3Gal-containing receptors. Three of these substitutions, namely, R62I, N193S and either D81N or D63N, were primarily responsible for the reduction of binding avidity. The

other three substitutions either had a weak binding-enhancing effect (N92K and A144G) or no effect [F(-2)L].

The observed reduction of the avidity for human-type receptors during HA evolution from its avian precursor was unexpected. Since this result was obtained in experiments with MDCK-grown IAVs and with non-natural receptor analogues, we performed additional experiments using more natural experimental models. As shown previously, the relatively large N-glycans attached to the HA in MDCK cells could alter viral receptor-binding properties as compared to the same virus grown in another cell system [30,31]. To address this possibility, we re-grew a representative group of viruses in the human cell line Calu-3 and in differentiated cultures of primary human tracheal-bronchial epithelial cells (HTBE cultures). The Calu-3-grown HK, R5 and single-point mutants of HK displayed the same patterns of binding to 6'SLN and sialidase-treated cells (S4a and S4b Fig) as did their MDCK-grown counterparts (Fig 4a and 4d). The HTBE-grown HK, R5 and two glycosylation mutants R5-81 and R5-63 also showed the same relative binding avidity (S4c Fig) as did corresponding MDCK-grown variants (Fig 4d). These results indicated that studied MDCK-grown IAVs correctly represented relative receptor-binding avidity, which the viruses would have during their replication (and glycosylation) in human epithelial cells.

To test whether differences in binding avidity of R5 and HK for SGPs and MDCK cells correlate with viral binding to biologically relevant receptors in humans, we studied attachment of R5 and HK to the apical surface of HTBE cultures which closely mimic structure and functions of human target cells in vivo [32,33]. R5 attached to HTBE cells more efficiently than HK
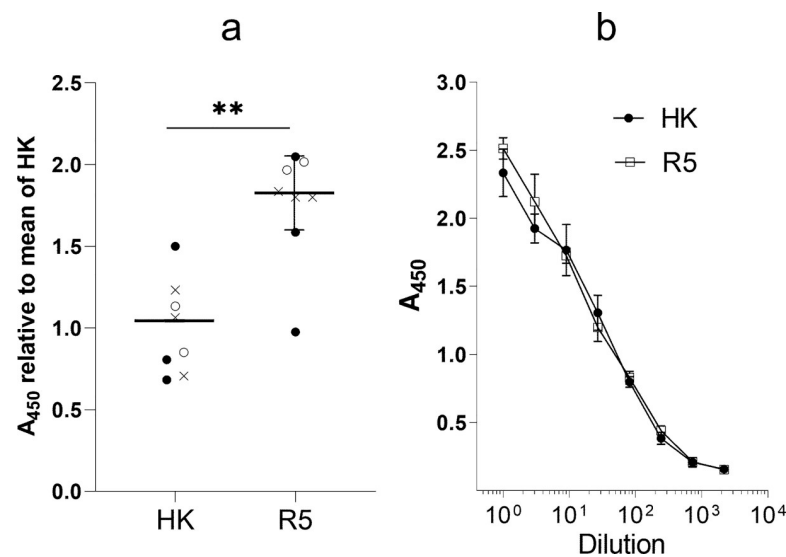
Fig 5. Attachment of HK and R5 to cells in HTBE cultures. (**a**) The apical sides of live HTBE cultures were washed with PBS+ to remove accumulated mucins and inoculated with 0.2 ml of DMEM-BSA containing 1.3×10^6 FFU of HK and R5. Control cultures were inoculated with DMEM-BSA. After 1-h incubation at 4°C the cultures were washed, fixed and immune-stained using anti-HK primary antibodies and HRP-labelled secondary antibodies. The mean absorbance in the control cultures was subtracted, and the results were expressed as the relative absorbance at 450 nm (A_{450}) in R5-treated and HK-treated cultures with respect to the mean absorbance in the latter. Open circles, closed circles and crosses depict individual data points from three experiments performed on different days. Mean, SD and P values were calculated using within-day averages. (**b**) Control of the concentrations of physical viral particles in suspensions of HK and R5 used for the HTBE attachment experiments. Suspensions were serially diluted in PBS and adsorbed in the wells of ELISA microplates. The wells were washed, fixed and immuno-stained as described above. Shown are the results of one experiment with 5 replicates per condition. The absorbance (A_{450}) reflects non-specific binding of HK and R5 to the plastic. Overlap of the A_{450} vs dilution curves indicate that suspensions used in the Fig 5a contained equal amounts of viral particles.
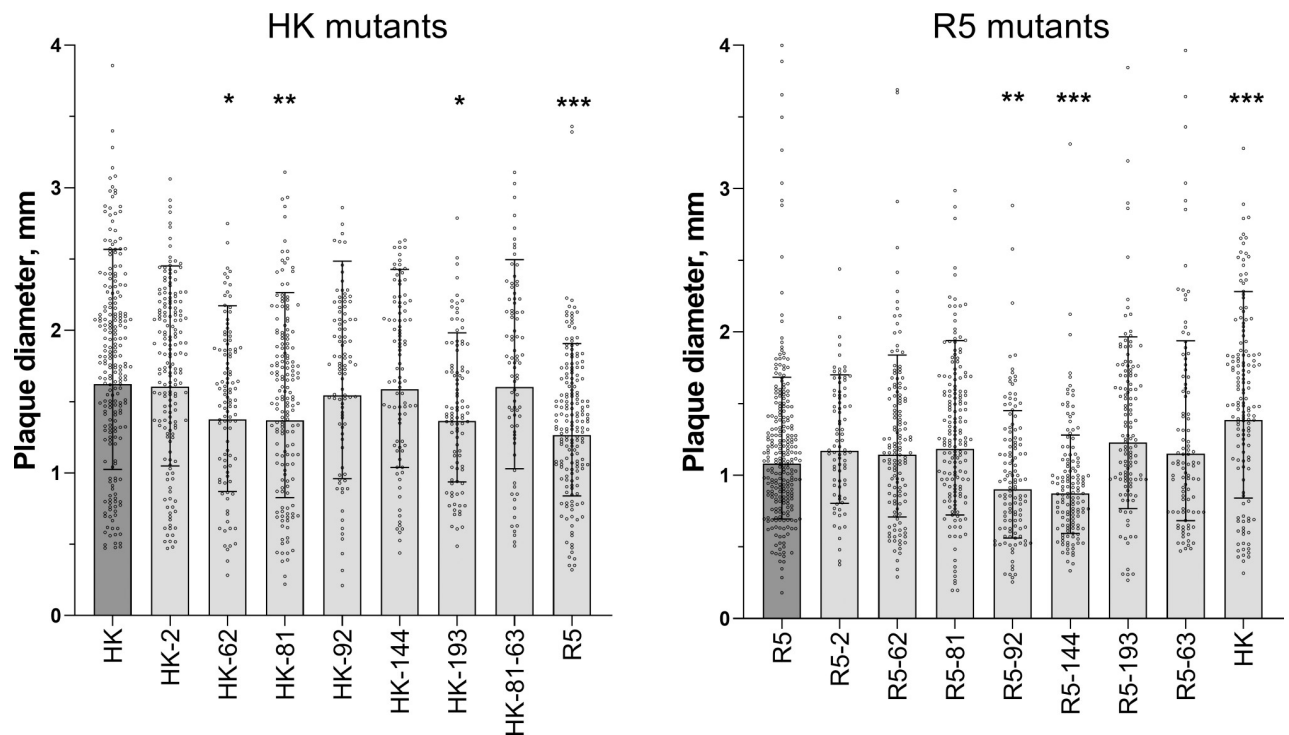
Fig 6. Diameter of plaques formed by viruses in MDCK cells. Cells in six-well plates were inoculated, incubated under semi-solid overlay medium for 48 h at 37°C, fixed and immunostained. Two panels represent two groups of viruses tested separately. Data were analysed after log-transformation using a general mixed model including a random intercept term accounting for day-to-day variation between experiments as described in Materials and Methods. Each panel shows diameters of individual plaques adjusted for day, geometric mean (bars) and geometric SDs from 1 to 4 experiments performed on different days. Asterisks depict P values for the differences between the mutants and the corresponding parental virus (HK in the left panel and R5 in the right panel).

(Fig 5) in agreement with relative binding efficiency of these viruses to soluble 6'SLN and to MDCK cells. These results confirmed that non-226/228 human-type amino acid substitutions in the precursor avian HA reduced efficiency of viral binding to receptors on airway epithelial cells in humans.

Effects of non-226/228 substitutions in the HA on viral infection in MDCK cells

In MDCK cells, R5 formed smaller plaques than did HK indicative of a less efficient multicycle replication of the former virus (Fig 6). The effects of point substitutions in the HA on plaque size was studied separately for HK mutants and R5 mutants (Fig 6). Although the resolving power of the assay was limited by substantial heterogeneity of the plaques formed by the same virus, we noticed reproducible effects of some of the point substitutions. Thus, HK-62, HK-81 and HK-193 formed smaller plaques in comparison with HK, whereas R5-92 and R5-144 formed smaller plaques in comparison with R5. For these five mutants, the reduced size of the plaques correlated with the elevated avidity of the virus for either 6'SLN (HK-62), MDCK cells (R5-92, R5-144), or both substrates (HK-81, HK-193) (compare Fig 6 and Fig 4). We previously studied dependence of replication efficiency of IAVs on their binding avidity and demonstrated that excessive avidity slowed down the release of viral progeny from infected cells and spread of the infection [34]. We assume that the same mechanism explains, at least in part, the smaller plaque size of the mutants of HK and R5 containing avidity-enhancing substitutions.

Effects of non-226/228 substitutions on viral infection in HTBE cultures

Paulson and colleagues postulated that changes in receptor-binding properties of avian IAVs during their adaptation to humans may serve to increase viral binding to and infection of human airway epithelial cells and to minimize its binding to and neutralization by respiratory mucus [35,36]. To determine whether non-226/228 substitutions in the HA contributed to these effects, we inoculated differentiated HTBE cultures with either R5 or HK and determined the numbers of infected cells 8 h post-infection. To focus on the role of viral interaction with receptors on cells, the cultures were extensively washed prior to infection to remove accumulated mucins. In parallel, replicate cultures were infected in the presence of the endogenous mucins. Fewer cells were infected with HK than with R5 in the mucus-deprived cultures (Fig 7a), this effect agreed with the higher binding avidity of R5 for receptor analogues and HTBE cells (Figs 4 and 5). As expected, the presence of HTBE mucins reduced infectivity of both viruses, with HK still infecting fewer cells than R5. The mean percentages of cells infected in the presence of mucins with respect to the infection without mucins were 15.1% for HK and 17.2% for R5, and the difference was not statistically significant. Thus, our results suggested that substitutions separating HK from R5 reduced efficiency of entry of HK into human airway epithelial cells and that these substitutions did not make HK less sensitive than R5 to neutralization by human airway mucins.

Reduced infectivity of HK compared to R5 in HTBE cultures was in apparent inconsistency with our previous observation of more efficient multicycle replication of HK in this cell system [21]. We therefore compared single- and multicycle replication of HK and R5 in the same experiment (Fig 7b). This experiment confirmed that R5 infected more cells in the first round of infection in HTBE cultures, whereas HK produced more viral progeny after multiple infection cycles. To infer which of the substitutions separating HK from R5 contributed to more efficient multicycle replication of HK in HTBE cultures, we compared replication of HK and its six single-point HA mutants (HK-2, HK-62, HK-81, HK-92, HK-144, and HK-193) under competitive conditions. Equivalent amounts of plaque-forming units of the viruses were mixed, and three different dilutions of this mixture were inoculated into the replicate HTBE cultures. The first group of cultures (group L) received 5 PFU of each of the seven viruses per culture, two other groups received 20 and 320 PFU per culture (groups M and H, respectively). The viral progeny was harvested 3 days post inoculation, and the proportions of each HA mutant in the inoculum and in the harvests were analysed by next generation sequencing (NGS) (Fig 8). HA segments of all 6 mutants were present in the harvests from all cultures in the group H. By contrast, the harvests in the group M and, especially, group L were highly heterogeneous, with proportions of the mutants varying between 0.00 and 1.00. No statistically significant changes with respect to the inoculum were observed in the case of HK-92 and HK-193. The proportions of HK-62, HK-81, and HK-144 were significantly reduced in the harvests in the group L as compared to their proportions in the inoculated mixture. The mutants HK-81 and HK-144 also displayed reduced frequencies in the group H. By contrast, the proportion of HK-2 was higher in the harvests than in the inoculum for all three infection doses used. These results indicated that HK-62, HK-81, and HK-144 replicated less efficiently than HK-2 and the other viruses in the mixture.

Whereas the point mutants of HK were readily quantified in the mixtures by NGS analysis of the single nucleotide polymorphisms in corresponding positions, the amount of the parent HK could not be directly quantified by sequencing. We therefore inferred proportions of HK in the mixtures by subtracting proportions of the mutants from the theoretical value of 1. Because of the high errors of such inference, the data could not be analyzed for statistical significance. One can see, however, that the proportion of HK in the harvests increased with
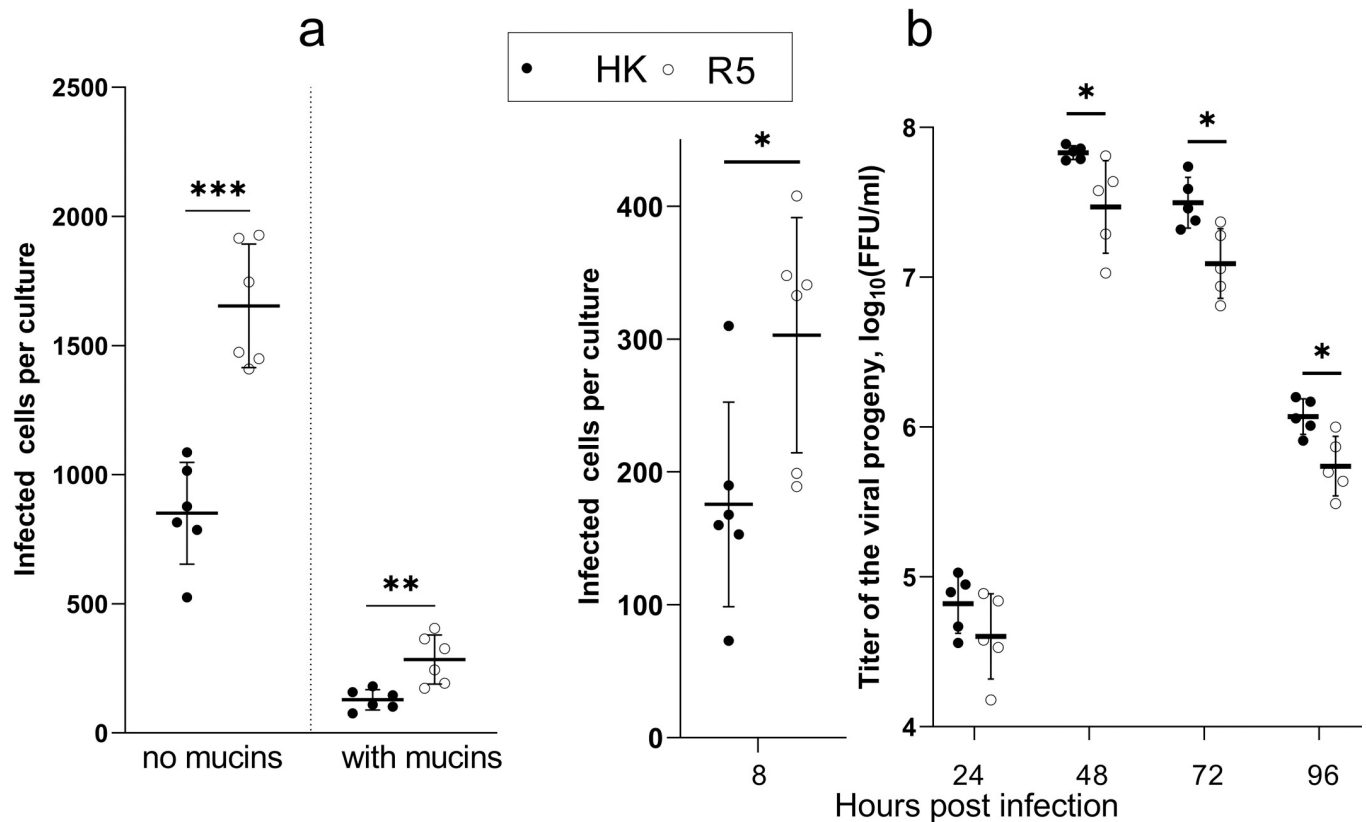
Fig 7. Infectivity and multicycle replication of HK and R5 in HTBE cultures. (**a**) The apical sides of HTBE cultures were washed with PBS+ to remove mucins and inoculated with 2×10^4 FFU of HK (closed circles) and R5 (open circles) with and without addition of mucins using 6 replicate cultures per condition. The inoculum was removed after 1 h. The cultures were incubated for 7 h under ALI conditions, fixed, immuno-stained for viral NP, and numbers of infected cells were counted. (**b**) The apical sides of washed HTBE cultures were inoculated with 7×10^4 FFU of the viruses without addition of mucins and processed as described above. Six cultures per virus were fixed 8 h post infection for immuno-staining and counting of infected cells. Viral progeny was periodically harvested by washing the apical sides of the remaining cultures, and the harvests were titrated simultaneously at the end of the experiment. P values were determined using Student's t-test.

respect to the inoculum (Fig 8). Remarkably, this pattern of HK resembled the pattern shown by HK-2 and differed from the patterns displayed by the other mutants, suggesting that HK and HK-2 have comparable high fitness and replicate more efficiently than the other 5 viruses studied. This conclusion is in keeping with the fact that HK-2 did not differ from HK in most other phenotypic assays. It is worth noting that the differences between viral proportions in the harvests vs inoculum increased with decreasing the multiplicity of infection, that is, from group H to group L. We assume that this pattern reflects a higher number of viral replication cycles and hence a more efficient selection of the best-fit viruses in group L compared to groups M and H.

Collectively, we conclude from the replication experiment in HTBE cultures that the avian-type substitutions at positions 62, 81 and 144 decrease in-vitro fitness of the pandemic virus, substitution at position -2 has no effect on fitness, and that effects of substitutions 92 and 193 (if any) were not statistically significant under assay conditions.

Airborne transmission of HK and R5 in ferrets

Transmission through the air is essential for the pandemic spread of IAVs in humans. To evaluate effects of non-226/228 substitutions in the HA of the 1968 pandemic viruses on
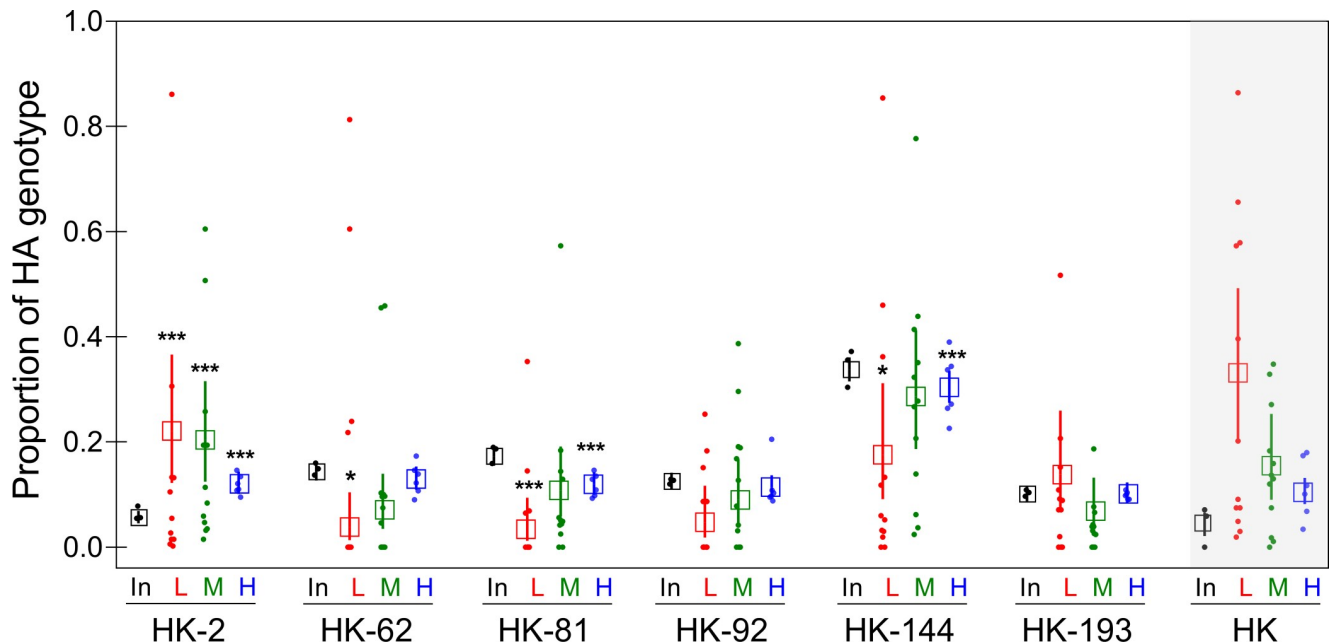
Fig 8. Competitive replication of HK and its 6 single-point HA mutants in HTBE cultures. HTBE cultures were inoculated with the mixtures of HK and its 6 HA mutants containing 5 PFU (L, 12 replicate cultures), 20 PFU (M, 12 replicates) and 320 PFU (H, 6 replicates) of each virus. After 1-h incubation, the inoculum was removed, the cultures were incubated under ALI conditions, and the apical material was harvested at 72 h post-inoculation. Small filled circles show proportions of each HA genotype determined by next generation sequencing in the inoculated mixture (In) and in each replicate harvest in the L, M and H groups. Empty squares and error bars represent the mean values and confidence intervals of proportions. Asterisks depict statistical significance of the differences between proportion of the corresponding genotype in the harvest and in the inoculum. The proportions of the parent HK virus (shown on gray background) were inferred by subtracting proportions of the six HA variants from a theoretical value of 1. Because of the high intrinsic errors of this approach the data for HK were not analysed for statistical significance.

https://doi.org/10.1371/journal.ppat.1009566.g008

transmissibility, we employed the ferret airborne transmission model. To avoid potential undesirable effects of the PR8-derived gene segments on viral fitness in ferrets, recombinant HK and R5 containing all 8 gene segments of A/Hong Kong/1/1968 were used in the transmission experiments. All directly inoculated ferrets shed the viruses from the nose and throat starting from day 1 after infection, the duration of shedding and peak titers did not significantly differ between HK and R5 (Fig 9). Two of the airborne contact ferrets in the HK group became infected and shed the virus in high titers starting from the day 3 after the contact, the third contact ferret produced one virus-positive swab on the day 3. All four contact animals in the HK group seroconverted. In the R5 group, only one of four contact animals shed the virus and seroconverted. We sequenced the HA of the transmitted HK and R5 viruses present in the throat swabs of all positive indirect contact animals at 3 days post exposure and found no substitutions. Although only small numbers of animals were used in the transmission studies, a careful interpretation of these results indicated that R5 transmitted less efficiently than HK, although not significant, suggesting that substitutions other than those at the RBS were required for efficient transmission of HK virus. Although it was tempting to study effects of individual substitutions separating HK and R5 on viral transmissibility, these studies were not pursued for ethical reasons given the small differences in transmissibility of parental viruses and, hence, necessity to use large groups of animals for statistically significant detection of even smaller effects.
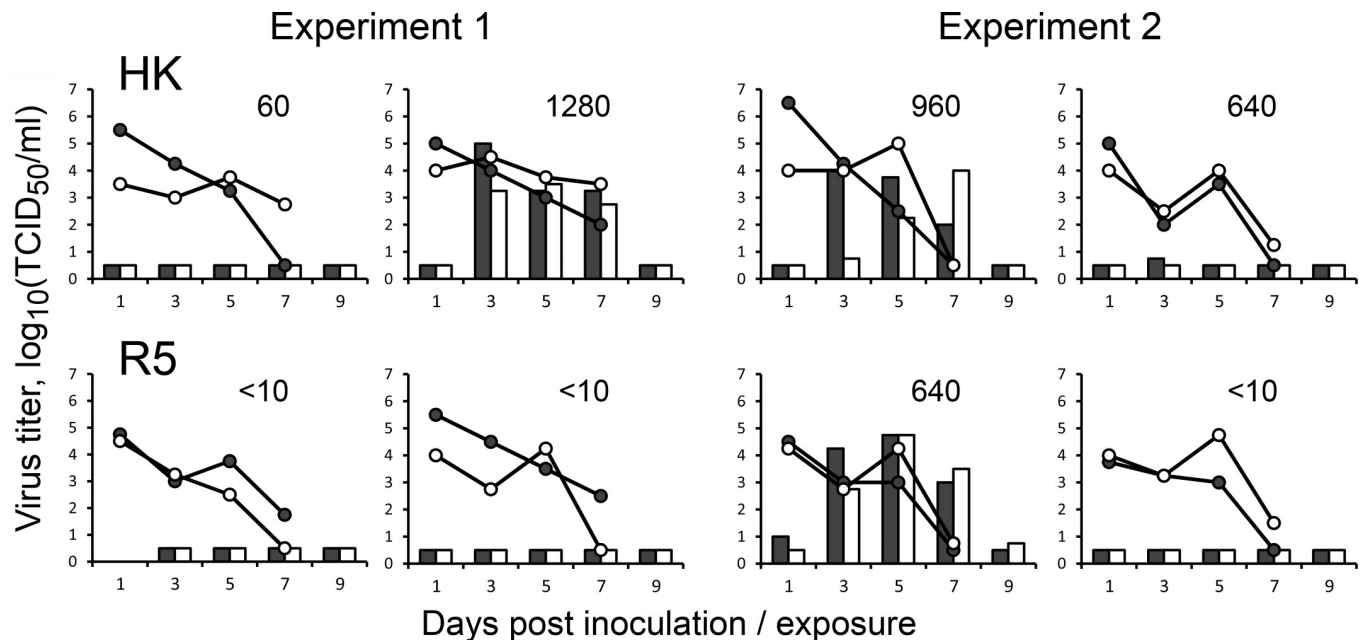
Fig 9. Comparison of airborne transmission of HK and R5 in ferrets. Groups of two ferrets were inoculated intranasally with 10^6 $TCID_{50}$ of recombinant viruses HK (top panels) and R5 (bottom panels) containing all eight gene segments of A/Hong Kong/1/1968. One naïve ferret was co-housed with each inoculated ferret in a separate transmission cage starting from one day after inoculation. Data show results of two replicate experiments performed on different days. Lines depict viral titers in nasal swabs (empty circles) and throat swabs (closed circles) collected from inoculated ferrets. White and black bars depict viral titers in nasal and throat swabs, respectively, of the indirect contact ferrets. Numbers show titers of hemagglutination inhibiting antibodies in the blood collected from the indirect contact animals 2 weeks post exposure.

Analysis of H3 HA sequences in different viral host species

The H3 HA sequences available from the GISAID EpiFlu database include sequences of avian IAVs (predominantly from wild aquatic birds) and of several stable mammalian-adapted virus lineages that emerged from the avian reservoir via interspecies transmission (Figs 10a and S5). Among them, the H3N8 equine IAVs were first recognized in 1963 [37]; they continuously circulate and evolve in horses until now. The IAVs of an independent H3N8 equine lineage caused epizootic in China in 1989, circulated in horses for a few years and became extinct [38]. Only one virus of this lineage, A/equine/Jilin/1/1989, was sequenced. The first of two canine H3 lineages originated from a contemporary equine H3N8 virus in Florida around 2003; the second canine lineage (H3N2) emerged from an avian precursor in Asia and was first recognized in 2006 (for reviews, see [2,39]). In addition to stable mammalian-adapted lineages, the database contains a small number of IAVs isolated from mammals which cluster with avian viruses and do not form persistent mammalian lineages (S5 Fig). The human H3N2 lineage includes H3N2/1968 pandemic viruses and their permanently evolving descendants that cause seasonal influenza epidemics. Finally, multiple lineages of H3N2 swine IAVs all originated from independent human-to-swine transmissions of seasonal IAVs followed by evolution in pigs [40].

We analysed prevalence of specific amino acids at nine HA positions in question in avian IAVs and IAVs of different mammalian lineages (Figs 10b, S5 and S6 and S1 Table). We also estimated selective pressures on these positions in different hosts using codon-based likelihood methods FEL, Contrast-FEL and MEME, which compare rates of synonymous and non-synonymous substitutions [41–43], and DEPS/FADE to identify directional evolution along human internal branches [44] (Table 1). Finally, we compared HA sequences of the earliest
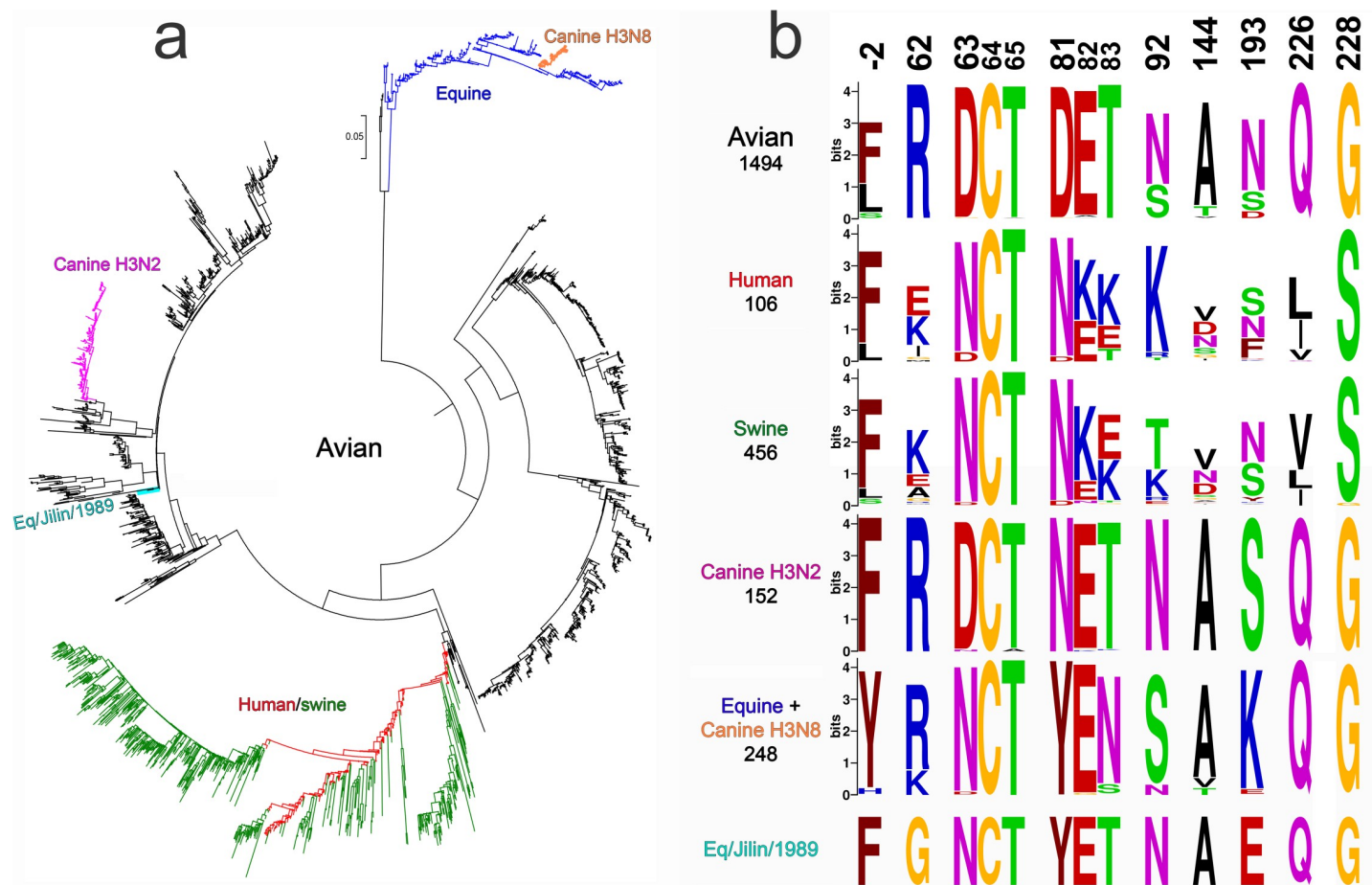
Fig 10. Host-specific lineages of IAVs with H3 HA and variation of amino acids at selected HA positions. (**a**) Phylogenetic tree for the H3 HA nucleotide sequences of representative sequences of human and swine viruses and all unique sequences of other mammalian and avian viruses available from GISAID EpiFlu database. The numbers of analysed sequences are shown in panel 10b below the lineage name. S5 Fig shows the same tree with strain names, accession numbers and amino acids at 9 HA positions under study. (**b**) Protein logos for indicated HA positions of the viral lineages shown in panel 10a. The overall height of each stack of letters depicts sequence conservation measured in bits. The height of each letter is proportional to the frequency of the corresponding amino acid in the alignment, the letters are ordered from most to least frequent. Only one sequence was available for the Eq/Jilin/1989 lineage.

equine and canine isolates with sequences of their closest avian counterparts to identify potential amino acid substitutions at 9 HA positions and their correlation with substitutions in the H3N2/1968 pandemic viruses (Table 2 and S5 Fig). Human-origin swine viruses were not included in the latter two analyses as we were interested in the avian-to-mammalian shifts. The results of these studies are summarized below.

Position -2. FEL predicts that this position is under pervasive negative selection (dN/dS = 0.49) in avian IAVs; the site exhibits a large number of substitutions (both synonymous and non-synonymous) on avian branches (see S6 Fig). No significant selection effects were detected by other likelihood methods (Table 1). Amino acids F and L are typically present at this position in the HAs of avian, human and swine IAVs (Figs 10b, S5 and S6). The earliest isolates of the A/equine/Miami/1963-like (H3N8) IAVs had H in position -2, whereas closest avian IAVs carry F, L and Y (Table 2 and S5 and S6 Figs). However, given a significant divergence of this equine lineage from other IAVs, the identities of amino acids at HA position -2 of an avian precursor and first equine-adapted variants remain unclear. No changes at this position occurred during emergence of other mammalian lineages. Collectively, these analyses

Table 1. Selection pressure analysis of 9 sites of the HA of avian, mammalian and human IAVs [a].

Site	Substitutions (Syn: Non-Syn) [b]			Pervasive selection dN/dS, FEL significance for dN≠dS [c]			Episodic positive selection number of branches, MEME significance [d]			Directional selection (target in H, empirical Bayes factor) [e]	Comparative selection [f]
	A	M	H	A	M	H	A	M	H		
-2	13:14	0:1	2:5	0.491 *	0.354	1.89	0	0	0		
62	18:1	1:2	2:5	0.0237 ***	0.548	1.57	0	0	0		H>A ***, overall ***
63	4:5	1:2	1:2	0.145 ***	0.767	0.782	0	0	0	N (37)	
81	3:7	0:0	0:1	0.378 *	0	0.675	5	0	0		overall *
92	10:16	1:2	0:2	0.504 *	0.616	0.734	0	0	0	T (41)	
144	13:24	0:3	1:8	0.807	1.05	3.04 *	0	0	0 *		H>A *
193	16:42	0:3	0:7	0.679	0.464	1.01	0	3	0		
226	14:0	0:0	3:15	0 ***	0 *	6.98 ***	0	0	7 ***		H>A ***, H>M ***, overall ***
228	17:0	1:1	2:0	0 ***	0.23	0 **	0	0	0		

[a] Groups A, M and H contained, respectively, 1492 sequences of avian IAVs, 406 sequences of equine, canine, feline and seal IAVs and 803 sequences of human IAVs isolated in the years from 1968 to the end of 1999 (set H). We restricted analyses only to internal branches (see Materials and Methods). Asterisks depict significance levels for the tests as follows: *, P< 0.05; **, P < 0.01; ***, P < 0.001.

[b] The number of synonymous and non-synonymous substitutions inferred to have occurred on internal branches in the corresponding group set with the SLAC method.

[c] Site-level dN/dS estimate along internal branches in the corresponding group inferred using the FEL method. Asterisks show significance levels for the test that dN≠dS. When significant result is obtained, negative selection is inferred if dN/dS < 1, otherwise diversifying positive selection is inferred.

[d] Site-level characterization of episodic positive selection in the corresponding group inferred using the MEME method. Asterisks show significance levels for the test that dN>dS for some fraction of branches. The number of individual branches where episodic selection may have acted is indicated as well (empirical Bayes factor ≥ 100)

[e] Site-level characterization of directional selection using the DEPS/FADE method. The residue that is the putative target of directional selection is indicated, with empirical Bayes factor supporting directional selection shown.

[f] Site-level characterization of differences in selective pressures between branch groups using the Contrast-FEL method. Asterisks show significance levels for the test that dN/dS differ between pairs of branch groups, or among all branch groups (overall). Notation like H>A means that dN/dS for "human" branches is greater than dN/dS for "avian" branches. Only significant results are listed.

provided no indications of the association of the substitution F(-2)L with viral host range and interspecies transmission.

Position 62. The codon is negatively selected in birds overall (FEL) with 99.3% of analysed avian HAs containing R_{62} (S1 Table). Contrast-FEL detects a higher dN/dS in human IAVs compared to avian IAVs; the point estimate of dN/dS in humans is 1.57, but this is not significantly different from 1. This finding is consistent with the location of the amino acid in the antibody-binding site E of the H3 HA [45] and its evolution under immune selection pressure in humans. In fact, all positions analysed here, with the exclusion of position -2, are also located in the antibody-binding sites A (residue 144), B (193), D (226, 228) and E (62,63,81,92). Avian-type R_{62} is conserved among canine and most equine IAVs, with conservative substitution R to K in equine viruses isolated after 2008 (Figs 10b, S5 and S6). The H3N2/1968 pandemic viruses acquired non-conservative substitution R62I. Another independent host switch event, transmission of an avian IAV to horses in Asia in 1989, was also accompanied by non-conservative substitution R62G (Table 2). Two independent non-conservative substitutions of conserved avian-type residue R_{62} suggest potential adaptive role of these substitutions during avian-to-mammalian transmission.

Table 2. Alteration of amino acids at 9 positions of the H3 HA during host shifts [a].

Host-specific virus lineage and the earliest isolate	Position								
	-2	62	63	81	92	144	193	226	228
Human/swine Hong Kong/1/1968 Memphis/1/1968	F→L	R→I	D→N*	D→N*	N→K	A→G	N→S	Q→L	G→S
Equine Equine/Miami/1/1963 (H3N8)	F/L/Y→H		D→N*	D→Y			N→K		
Canine H3N2 Canine/Guangdong/1/2006				D→N*					
Eq/Jilin/1989 Equine/Jilin/1/1989 (H3N8)		R→G	D→N*	D→Y			N/D→E		
Sporadic isolates of avian-like IAVs from mammals[b]					N→S		N→S		

[a] Table shows substitutions at the indicated positions that separate the earliest isolates of stable mammalian lineages from the phylogenetically closest avian viruses. Empty cells indicate a lack of changes. Asterisk next to N indicates that substitution generates glycosylation site. The analysis was performed using 2489 HA sequences depicted in the Figs 10, S5 and S6.

[b] The HAs of 10 swine, canine and seal IAVs which clustered with HAs of avian viruses but did not form stable lineages. Substitutions were observed in two viruses, A/harbour seal/New Hampshire/179629/2011 (H3N8) (N92S) and A/seal/Massachusetts/3911/1992 (H3N3) (N193S).

Positions 63 and 81. Avian HAs contained D_{63} and D_{81} in 98.3% and 98.6% of analysed sequences, respectively (Fig 10b and S1 Table). Both codons are negatively selected in birds overall (FEL). The substitution from D to N at either position 63 or position 81 of two co-circulating pandemic virus lineages generated glycosylation sites. Remarkably, the substitution D63N and acquisition of glycosylation site accompanied emergence of both equine IAV lineages, whereas substitution D81N with new glycosylation site occurred during transmission of an avian H3N2 virus to dogs (Table 1). The glycosylation sites became fixed in H3N8 equine and H3N2 canine lineages, moreover, the equine-origin H3N8 canine and human-origin H3N2 swine IAVs inherited and preserved the glycosylation sites of their mammalian precursors. As a result, all known mammalian IAVs with H3 HA differ from H3 avian IAVs by the presence of N-glycan at either position 63 or position 81 (Figs 10b, S2, S5 and S6). Observed parallel evolution of amino acids at positions 63 and 81 during avian-to-mammalian adaptation represents a strong indication of their adaptive role in H3N2/1968 pandemic IAVs. There is evidence of directional evolution towards N on the human branches at site 63. Furthermore, substitution D81Y occurred in both H3N8 equine lineages. Thus, alteration of the properties of amino acid in position 81 may play an adaptive role in interspecies transmission irrespectively from and/or in addition to the effect of the substitution on HA glycosylation.

Position 92. Avian IAVs contain either N_{92} or S_{92}. The codon is negatively selected in birds overall (FEL, dN/dS = 0.504) (Table 1). Whereas none of the avian HA sequences contained K_{92}, non-conservative substitution N92K altering change of the amino acid side chain accompanied emergence of the pandemic IAVs. Of note, this substitution is located in the close proximity of another charged human-type substitution R62I (see Fig 1b) and could compensate for the effect of the latter substitution on the net surface charge of the protein in this area. No substitutions in position 92 occurred during emergence of other stable mammalian lineages, however, conservative substitution N92S was present in H3N8 IAVs that caused epizootic with cases of fatal pneumonia in New England harbour seals in 2011 [46] (Table 2 and S5 Fig). Along the human branches, there is evidence of directional selection towards T, with two clades showing N→T substitutions which were then maintained.

Position 144. In accord with location of the amino acid in the antigenic site, codon 144 is under positive selection in humans (FEL dN/dS = 3.04 and MEME). dN/dS on human

branches is significantly higher than on the avian branches (Contrast-FEL). Avian IAVs typically contain A, I, or V but never G. By contrast, the H3N2/1968 pandemic IAVs carried a non-conservative substitution A144G, which could affect the structure of polypeptide loop 140–145 located in the vicinity of the RBS. These notions suggest potential functional significance of this substitution for the avian-to-human adaptation. There is no evidence of parallel evolution at this site in other instances of avian-to-mammalian transmissions (Table 2).

Position 193. This amino acid is located at the upper rim of the RBS. The codon is predicted to be under negative selective pressure in horses and dogs, but not in other species (Table 1). Avian HAs contain N and, less frequently, S or D. Unique substitutions to K_{193} and E_{193} occurred during independent transmissions of avian precursors to horses (A/equine/Miami/1963-like lineage and A/equine/Jilin/1/1989-like lineage); conservative substitution N-to-S was found in H3N3 avian-like IAVs isolated from seals in 1992 (Table 2 and S5 Fig). These findings suggest a potential functional role of the substitution in position 193 during interspecies transmission.

Amino acids Q_{226} and G_{228} are critical for the avian HA binding to avian-type receptor motif Neu5Acα2-3Gal [8,9,47]. Both codons are under purifying selection in birds, horses and dogs, which share binding preference for Neu5Acα2-3Gal-terminated receptors. By contrast, the codon 226 is under pervasive and episodic positive selection in humans (FEL dN/dS = 6.98, MEME), and significantly higher dN/dS in humans compared to both avian and mammalian lineages.

Discussion

Unavailability of immediate animal precursors of pandemic IAVs hampers understanding of genetic and phenotypic changes in the HA that were essential for the viral animal-to-human adaptation and pandemic spread. To mitigate this problem, we generated and characterized the recombinant IAV R7 containing the HA of the hypothetical most recent common ancestor of H3 avian and H3N2/1968 pandemic IAVs. R7 displayed receptor-binding profile typical for duck viruses and differed in this respect from IAVs perpetuated by gulls, shorebirds and gallinaceous land-based poultry ([28,29] and references therein). The HA of R7 showed high conformational stability and low pH optimum of fusion compatible with the aquatic bird origin of the H3N2/1968 HA [13–15]. These properties of R7 agreed with the hypothesis that the H3N2/1968 pandemic IAVs originated from a duck virus [22,48]. IAVs with the HA sequences highly similar to R7 were isolated from both wild migratory ducks captured on a Pacific flyway in Japan and domestic ducks in Southern China [48,49] (S1b Fig). It seems likely that the precursor wild duck virus was transmitted to humans via domestic ducks either with or without additional intermediate host species.

The HAs of all four characterized pandemic IAVs (H1N1/1918, H2N2/1957, H3N2/1968 and H1N1/2009) were relatively stable, whereas swine IAVs, highly pathogenic H5 and H7 IAVs from gallinaceous poultry and some IAVs of aquatic birds display low conformational stability [11,15,16]. These observations suggested that adaptation of animal IAVs to humans may require stabilizing substitutions in the HA [11,18], however, it remained unclear whether this mechanism contributed to the emergence and initial pandemic spread of any known pandemic virus. We found that the HA of HK was in fact slightly less stable than the precursor HA of R7 with a pH_{50} of conformational transition of 5.4 and 5.25, respectively (Fig 2). Reduced HA stability of HK was primarily associated with substitutions at positions 226 and 228; substitutions at other HA positions had lower if any effects. Thus, the duck precursor of the H3N2/1968 IAVs had a sufficiently stable HA and was able to adapt to humans without elevation of its conformational stability.

Analysis of the receptor-binding specificity of HK and its HA variants (Fig 3b) confirmed the concept that preferential binding of the H3N2/1968 viruses to Neu5Acα2-6Gal-terminated receptors is primarily determined by substitutions Q226L and G228S [50,51]. The combination of other human-type substitutions in the HA decreased binding of HK to both Neu5Acα2-6Gal- and Neu5Acα2-3Gal-terminated receptor analogues, reduced its attachment to apical surfaces of HTBE cultures and lowered infectivity for HTBE cells without affecting efficiency of viral neutralization by human airway mucus (Figs 4, 5 and 7). These results indicated that non-226/228 substitutions lowered the avidity of HA binding to receptors on human target cells. Although HK infected fewer cells than R5 during initial inoculation into the HTBE cultures, HK produced more infectious viral particles after multicycle replication (Fig 7), suggesting that it outperforms R5 during post-entry replication stage(s). The reduced binding avidity can increase fitness of HK by facilitating release of viral progeny from cells and preventing its receptor-mediated self-aggregation [34,52]. In addition, some of the non-226/228 substitutions could, in principle, promote replication of HK relative to R5 by avidity-independent mechanisms, such as facilitation of synthesis and intracellular processing of the HA protein or assembly of viral particles. Further studies are needed to clarify potential roles of these mechanisms.

Our observation of reduced HA avidity of H3N2/1968 IAVs is in line with the observed lower avidity of the HA of swine-origin H1N1/2009 pandemic IAV as compared to its closest available swine counterparts [53,54]. The NA catalytic activity of H1N1/2009 was also lower than that of swine IAV NAs [54], in agreement with the concept that a functional balance between HA and NA is essential for efficient replication and transmission of IAVs [55,56]. The H3N2/1968 pandemic virus was a reassortant containing the H3 HA of an avian parent and the N2 NA of a human parent. This NA differed from typical avian N2 NAs by substrate specificity [57,58] and by substitutions in the second sialic acid binding site that reduced catalytic activity [59,60]. We speculate that the reduction of binding avidity of the avian-origin HA of the H3N2/1968 IAVs could have simplified its functional match with the human-origin NA.

Reduced avidity of the HK HA was primarily associated with the avian-to-human substitutions R62I, N193S and either D81N or D63N (Fig 4). The substitution R62I is located relatively far from the receptor binding site and decreases the local and the net positive charges of the HA. The negative effect of this substitution on binding avidity could be partially associated with the reduction of electrostatic attraction of IAV particles to negatively charged soluble sialoglycans and cell membranes [34,61]. Amino acid 193 is located at the upper rim of the receptor-binding site. Charged substitutions at this position, such as N/S ➔ K/D, were shown to affect receptor-binding properties of avian and equine viruses with different HA subtypes [28,29,51,62,63]. In the crystal structures of the avian H5 HA and canine H3 HA complexed with avian-type receptor glycans 6-Su-3'SLN and 6-Su-SLe^x, the side chain of K_{193} interacts with the sulfogroup attached to GlcNAc-3 [64,65]. In the H3N2/1968 HA complexes with human-type receptor analogues LSTc and 6SLN-LN, the side chain of S_{193} contacts the Gal-4 residue of the glycan [66,67]. These observations suggest that substitution N193S affects binding avidity of HK by altering HA interactions with sub-terminal saccharide residues of both avian-type and human-type receptor glycans. The N-glycans on the HA globular head typically decrease binding avidity with the effect being dependent on glycan structure and location with respect to the receptor-binding site (for review, see [8]). Substitutions D81N and D63N are located in the same area of the HA and result in addition of structurally similar complex type N-linked glycans containing up to 4 antennae [68]. We assume that bulky NG_{63} and NG_{81} either reach the lower rim of the RBS and directly interfere with HA-receptor interactions or have some yet undefined allosteric negative effect on binding.

The essential role of HA substitutions Q226L and G228S in the emergence of H2N2/1957 and H3N2/1968 pandemic IAVs is well established. To test whether other substitutions separating the HA of HK from its avian precursor were at all required for the avian-to-human adaptation, we compared transmission of HK and R5 via airborne droplets in ferrets, the currently preferred animal model for prediction of IAV replication and transmissibility in humans [69]. R5 transmitted less efficiently than HK (Fig 9) supporting the concept that some of the non-226/228 substitutions in the HA contributed to the human adaptation and pandemic spread of H3N2/1968 IAVs [21]. Unfortunately, the low statistical power of the current ferret transmission model with small group sizes did not allow us to study effects of individual substitutions on viral fitness and transmissibility. Additional experiments are needed to address this question, for example, analyses of viral replication and transmission in ferrets assisted by deep mutational scanning of the positions of interest [70]. Alternatively, improved methods to assess viral transmissibility need to be developed.

To rank the substitutions in the order of their potential importance for the avian-to-human adaptation of the H3N2/1968 HA we took into account various adaptation-related characteristics, such as significant effect of the substitution on the HA phenotype, its location in the functional region of the HA, and dissimilar patterns of evolution of corresponding positions in birds and mammals (Table 3).

The substitutions Q226L/G228S displayed the maximal total score in such combined analysis, supporting the validity of this approach. Among the non-226/228 substitutions, D81N and D63N showed the highest score. Moreover, these HA positions were characterized by a remarkable parallel evolution during interspecies transmission events (Table 2 and Figs 10, S5 and S6). In the context of HK HA, either substitution reduced the binding avidity and slightly elevated the pH optimum of viral fusion within endosomes (Figs 4 and S3). However, it seems unlikely that these effects alone could explain addition of a novel N-glycan in all independent cases of avian H3 HA adaptation to such distinctive hosts as humans, dogs and horses. It is also unlikely that NG_{63}/NG_{81} served to mask HA antigenic epitopes, given a lack of herd immunity in mammals during emergence and initial epidemic spread of a novel IAV. The HA of the HK-like strains A/X31 and A/Aichi/2/1968 containing NG_{81} was often used as a model in the general research on the role of N-glycans in protein folding, quality control and intracellular transport [72–74]. These studies showed that folding, formation of disulfide bonds and quality control of the nascent HA chain in the ER is largely regulated by concerted interactions of N-glycans attached in critical HA regions with lectin chaperones calnexin and calreticulin. NG_{81} was found to engage calreticulin, and point mutants lacking NG_{81} displayed delay in folding due to a less efficient formation of the critical intrachain disulfide bond C64-C76 located in vicinity of this glycan [73,74]. We therefore hypothesize that the addition of either NG_{81} or the structurally equivalent NG_{63} increases fitness of avian-origin HA in mammals by ensuring its interactions with mammalian chaperones. Of note, two other pandemic viruses, H1N1 from 1918 and 2009, contained NG_{94} in the same area of the HA. Although fitness-enhancing mechanisms of substitutions D63N/D81N remain to be fully characterized, we conclude that these substitutions represent a previously unrecognized important marker of avian-to-mammalian adaptation and pandemic potential of IAVs.

The relative importance of the other substitutions for the adaptation is less clear. R62I shows a high score (Table 3) and represents particular interest because it reduces HA avidity, increases viral replicative fitness in MDCK cells and HTBE cultures and because codon 62 displays distinctive evolution in avian and mammalian IAVs. Substitutions N193S and A144G are located in the functionally important region on the opposite rims of the RBS, show phenotypes in receptor-binding assays and could serve to fine-tune HA interactions with receptors in humans. The substitution F(-2)L in the signal peptide was neutral in most phenotypic and

Table 3. Characteristics of avian-to-human amino acid substitutions in pandemic H3N2/1968 viruses.

	F(-2)L	R62I	D63N	D81N	N92K	A144G	N193S	226+228
Phenotypical effects[a]								
Preference for Neu5Ac2-6Gal-terminated receptors								+
Binding avidity		+	+	+	+	+	+	+
pH of conformational transition								+
Fusion activity (polykarion/NH4Cl)	+		+	+				+
Replication in MDCK		+		+	+	+	+	+[b]
Replication in HTBE		+	NT	+		+		+[b]
Structural features								
Location in the functional region of the HA	+					+	+	+
Alteration of charge/non-conservative substitution		+	+	+	+	+		+
Addition of N-glycan			+	+				
Variation of the codon in H3 HAs								
Purifying selection in avian IAVs	+	+	+	+	+			+
Conserved amino acid in avian IAVs		+	+	+				+
Human-type amino acid is not found in avian IAVs		+	+	+	+	+		+
Parallel evolution in mammalian IAVs		+	+	+			+	+[b]
Total[c]	3	8	8	10	5	6	4	12

[a] Plus indicates that point substitution in corresponding position affects properties of either HK, R5 or both viruses in one or more phenotypical assays used. Empty cells depict a lack of statistically significant effect. NT, not tested.

[b] These characteristics were described in the literature [19,50,71]

[c] Total number of plusses in the column. This number reflects probability of the adaptive role of the substitution during the avian-to-human transmission of the HA.

https://doi.org/10.1371/journal.ppat.1009566.t003

genotypic analyses performed, however, this substitution showed weak effect on viral membrane fusion activity, thus deserving attention in the future studies.

The research on mammalian adaptation of avian IAVs was strongly stimulated and advanced by two independent reports on HA substitutions that allowed airborne transmission of avian H5N1 IAVs in ferrets [75,76]. In each study, two substitutions in the RBS changed HA binding preference from Neu5Acα2-3Gal motif to Neu5Acα2-6Gal motif, one substitution removed N-glycan from the tip of the HA thus increasing binding avidity and one substitution increased HA conformational stability. Remarkably, apart from the alteration of the Neu5Ac-Gal linkage specificity, other ferret-adaptation changes in the H5N1 HA (alterations of HA avidity, stability and N-glycosylation) are discordant with the changes that accompanied emergence of the pandemic H3N2/1968. This discrepancy can be explained, at least in part, by the differences between properties of poultry-adapted H5N1 IAVs and duck-origin precursor of H3N2/1968 and between factors required for airborne transmission in ferrets and pandemic spread in humans. In any case, our results highlight the importance of the studies on previous pandemic IAVs for the influenza risk assessment and preparedness.

Materials and methods

Ethics statement

Animal experiments were conducted at Erasmus MC, in strict compliance with European guidelines (EU directive on animal testing 2010/63/EU) and Dutch legislation (Experiments on Animals Act, 1997). The study protocol was approved by Stichting Dier Experimenten Commissie Consult (DEC Consult, permit number EMC3381), a Dutch independent animal experimentation ethics review board.

Cells and wild type IAVs

Cultivation of all non-infected and virus-infected cell cultures was performed at 37˚C in 5% CO_2. MDCK cells, human embryonic kidney 293T cells, and human bronchial adenocarcinoma Calu-3 cells were propagated using Dulbecco's modified Eagle medium (DMEM; Gibco) supplemented with 10% fetal calf serum (FCS; Gibco), 100 IU/ml penicillin and 100 μg/ml streptomycin (pen-strep), and 2 mM glutamine. DMEM containing pen-strep, 2 mM glutamine and 0.1% bovine serum albumin (PAA Laboratories) (DMEM-BSA) was used for the viral infections.

Differentiated cultures of primary human tracheobronchial cells (HTBE) were prepared as described previously [32]. In brief, primary HTBE cells (Lonza) were expanded on plastic in BEGM growth medium (Lonza) and stored in aliquots in liquid nitrogen. Thawed passage-1 cells were grown on membrane supports (12-mm Transwell-Clear; pore size, 0.4 μm; Corning) in a 1:1 mixture of BEGM with DMEM. After 1 week, the medium was removed from the upper compartment and cells were maintained in BEGM/DMEM mixture under air-liquid interface (ALI) conditions. Fully differentiated 5- to 8-week-old cultures were used for the experiments.

A/Hong Kong/1/1968 (H3N2) was provided by Earl Brown, University of Ottawa, Ottawa, Ontario, Canada and grown in MDCK cells. A/Mallard/Alberta/279/1998 (H3N8) and A/Ruddy turnstone/Delaware/2378/1988 (H7N7) were provided by Robert Webster, St. Jude Children's Research Hospital, Memphis, TN, USA. The avian viruses were grown in 11-days-old embryonated hen's eggs.

Plasmids and recombinant IAVs

Reverse genetics plasmid pHW2000 and pHW2000 plasmids containing gene segments of A/Puerto Rico/8/1934 (H1N1) (PR8) were provided by Richard Webby and Robert Webster, St. Jude Children's Research Hospital, Memphis, TN, USA. The eight pHW2000 plasmids containing gene segments of HK/68, modified HA plasmids R2 and R5 and corresponding recombinant viruses were prepared previously [19,20].

Mutations were introduced into the HA plasmid of A/Hong Kong/1/1968 using a site-directed mutagenesis kit (QuikChange; Stratagene). 2:6 recombinant IAVs containing wt and modified HA of A/Hong Kong/1/1968, NA of A/Hong Kong/1/1968 and the remaining six gene segments of PR8 were generated by reverse genetics [77] as described before [78]. These viruses and their designations are listed in the Fig 1a. They were amplified in MDCK cells using DMEM-BSA medium containing 1 μg/ml of TPCK-treated trypsin (Sigma), clarified by low-speed centrifugation, and stored in aliquots at -80˚C. The identities of the HA- and NA-encoding genes of all viruses were confirmed by sequencing.

Virus titration and plaque size

Viruses were titrated in MDCK cells using single-cycle focus formation assay in 96-well plates [19] and plaque formation assay under Avicel RC/CL overlay medium in 6-well plates [79]. Infected cells were detected by immunostaining for viral nucleoprotein (NP). The viral concentrations were expressed in focus forming units (FFU) and plaque forming units (PFU) per ml, respectively. To determine the size of the plaques, plate wells containing from 5 to 50 plaques were scanned with a flat-bed scanner. The plaque diameters were measured with the Ruler Tool of Adobe Photoshop CS3 software version 10.0.1.

Low-pH-induced conformational transition of HA

Alteration of the HA sensitivity to protease digestion that accompany acid-induced conformational transition was determined using a solid-phase receptor binding assay as described

previously [21,26]. In brief, viruses were adsorbed in the wells of fetuin-coated microtiter plates and incubated with buffers containing 25 mM 2-(N-morpholino)ethanesulfonic acid (MES), 150 mM NaCl, 0.9 mM $CaCl_2$ and 0.5 mM $MgCl_2$ (MES-NaCl). The pH of the buffers varied from 4.8 to 6.0 in 0.1 steps. After incubation with MES-NaCl buffers for 15 min at 37˚C, the plates were washed with 25 mM phosphate buffered saline pH 7.2 (PBS) and incubated with 0.1 mg/ml of proteinase K in PBS for 1 h at 37˚C. After washing with PBS containing 0.01% tween 80, binding of peroxidase-labelled fetuin (fet-HRP) was determined and expressed in percentages of binding to low-pH-exposed virus with respect to that of the virus exposed to pH 7. Binding-versus-pH curves were plotted, and pH values that corresponded to HA inactivation by 50% (pH_{50}) were determined by linear interpolation.

Inactivation of HA by chaotropic agent and heat treatment

The effects of guanidinium hydrochloride (GnHCl) and elevated temperature on HA inactivation were quantified using the solid phase receptor binding assay described above. Viruses absorbed in the wells of fetuin-coated microtiter plates were either incubated for 1 h at 4˚C with PBS containing variable concentrations of GnHCl (6.0, 4.6, 3.6, 2.7, 2.1, 1.6 and 0 M) or incubated with PBS for different time periods at 65˚C. After washing, the binding of fet-HRP to GnHCl-treated and heat-treated viruses were determined and expressed in percentages with respect to the binding to control viruses incubated at 4˚C with PBS. Concentration of GnHCl and incubation time at 65˚C that reduced fet-HRP binding by 50% (IC_{50} and t_{50}, respectively) were determined by linear interpolation.

Reduction of viral infectivity after 2-h incubation at 45 ˚C

Viral stocks were diluted in DMEM-BSA to a concentration of 4000 FFU per ml. Replicate 0.7-ml aliquots were incubated in closed Eppendorf tubes for 2 h either in a water bath at 45˚C or on ice (control). All samples were next titrated using single-cycle focus formation assay in MDCK cells. Five technical replicates were used for the titration of each sample, and the results were averaged. The titers of heat-treated viruses were expressed as percentages of the corresponding control titers.

Low pH-induced polykaryon formation in virus-infected cells

pH-dependence of virus-induced cell-cell fusion was assayed as described [80] using MDCK cells instead of VERO cells. In brief, MDCK cultures in 96-well plates were inoculated with 1 FFU of the virus per cell in DMEM-BSA and incubated overnight. The medium was discarded, and the cultures were incubated for 15 min at 37˚C with the DMEM-BSA containing 1 μg/ml of TPCK trypsin. The trypsin-containing medium was substituted by pre-warmed MES-NaCl pH-buffers (pH range from 5.3 to 7.0 in 0.1 steps), incubated for 10 min at 37˚C and washed once with PBS containing 0.9 mM $CaCl_2$ and 0.5 mM $MgCl_2$ (PBS+). After 3-h incubation with DMEM-BSA at 37˚C, the cells were fixed with 70% ethanol, stained with Giemsa stain (Sigma) and analysed under the microscope. The highest pH at which more than 10 syncytia with more than 5 nuclei/syncytium were observed was taken as the pH threshold of polykaryon formation.

Infection inhibition by ammonium chloride

The assay determined viral dependence on endosomal acidification during infection as described previously [15]. In brief, MDCK cells in 96-well plates were inoculated with 200 FFU of the virus in either 0.1 ml DMEM-BSA containing various concentrations of NH_4Cl

(from 2 to 0.063 mM in 2-fold steps) or DMEM-BSA alone. The cells were incubated overnight, fixed and immuno-stained for viral NP. Concentrations of NH_4Cl that reduced numbers of infected cells by 50% (IC_{50}) were determined from dose-response curves by linear interpolation. Absolute values of IC_{50} could vary depending on the lot of cultures, lot and pH of culture medium, and other minor variations in the assay conditions, however, the relative sensitivities to NH_4Cl of the viruses tested in the same experiment were reproducible.

Infection inhibition by *Vibrio cholerae* sialidase

Binding avidity of the viruses for receptors on MDCK cells was compared using gradual desialylation of receptors with bacterial sialidase as described previously [21]. In brief, MDCK cells in 96-well plates were incubated with 0.05 ml per well of either serial two-fold dilutions of sialidase in DMEM-BSA (from 11.4 to 0.36 mU/ml) or DMEM-BSA alone for 30 min at 37˚C. Two hundred FFU of the viruses in 0.05 ml of DMEM-BSA were added per well without removing sialidase. No trypsin was added to the medium to avoid multicycle replication. The cultures were incubated overnight, fixed and immuno-stained for viral NP. Concentrations of sialidase that reduced numbers of infected cells by 50% (IC_{50}) were determined from dose-response curves by linear interpolation.

Viral binding to sialoglycopolymers

Receptor-binding specificity of the viruses was characterized using soluble synthetic sialoglycopolymers (SGPs) (GlycoNZ, Auckland, New Zealand) [27]. The SGPs contained 20 mol% of sialyloligosaccharide moieties and 5 mol% of biotin attached to either the low-molecular-mass (20-kDa) or high-molecular-mass (1000-kDA) poly-N-(2-hydroxyethyl)acrylamide backbone.

The structures of the sialyloligosaccharide moieties and designations of SGPs are shown below.

Neu5Acα2-3Galβ1-4GlcNAcβ	3'SLN
Neu5Acα2-3Galβ1-4(6-Su)GlcNAcβ	6-Su-3'SLN
Neu5Acα2-3Galβ1-4(Fucα1–3)GlcNAcβ	SLe^x
Neu5Acα2-3Galβ1-4(Fucα1–3)(6-Su)GlcNAcβ	6-Su-SLe^x
Neu5Acα2-3Galβ1-3GlcNAcβ	SLe^c
Neu5Acα2-3Galβ1-3GalNAcα	3'STF
Neu5Acα2-6Galβ1-4GlcNAcβ	6'SLN

The binding of the viruses to SGPs was determined in a direct solid-phase binding assay as described previously [26]. In brief, viruses adsorbed in the wells of fetuin-coated 96-well plates were allowed to interact with serially diluted SGPs followed by incubation with peroxidase–labelled streptavidin and tetramethylbenzidine (TMB) substrate solution. The association constants of viral complexes with SGPs (K_{ass}) were determined from the slopes of A_{450}/C versus A_{450} plots, where C is the concentration of the sialic acid in solution and A_{450} is the absorbance in the corresponding well.

Viral attachment and single-cycle infection in differentiated HTBE cultures

One day before the experiments, apical sides of the cultures were incubated with 0.15 ml of DMEM for 1 h at 37˚C to collect secreted mucus. The mucus suspension was clarified by

centrifugation at 6000x*g* for 5 min and stored at 4˚C. Immediately before the experiments, the cultures were washed 10 times with PBS+.

To study viral attachment, the cultures were incubated with viral suspensions in DMEM-BSA ($1.3x10^6$ FFU per culture) for 1 h at 4˚C. Control cultures were incubated with DMEM-BSA. The cultures were washed with PBS+ and fixed with 4% paraformaldehyde for 30 min at 4˚C. Attached viruses were quantified by immunostaining of the apical sides of the cultures directly on Transwell-Clear supports. The cultures were blocked with 5% normal donkey serum (NDS, Dianova) at 4˚C overnight, followed by sequential 1-h incubation at room temperature with in-house made rabbit polyclonal antibodies against HK and peroxidase-labelled donkey anti-rabbit antibodies (Dianova). Both antibodies were diluted in PBS buffer containing 10% normal horse serum (Dianova), 1% BSA, 1% NDS, 2% of the HTBE mucus suspension and 0.05% tween 80. After washing with 0.05% tween 80 in PBS, peroxidase activity was determined using TMB substrate. The mean substrate absorbency at 450 nm in the control cultures was subtracted from the absorbencies in virus-treated cultures.

To confirm that suspensions used in HTBE attachment experiments contained equal amounts of physical viral particles, non-specific viral binding to plastic was measured in control experiments [30]. Viral stocks were serially diluted in PBS and incubated in the wells of the immunoassay 96-well microplate (Greiner) overnight at 4˚C (0.05 ml/well). The wells were washed, fixed and immuno-stained with anti-HK antibodies and TMB substrate as described above for the HTBE experiments.

To study ability of the virus to enter into cell and initiate the first round of infection, replicate HTBE cultures were inoculated with $2x10^4$ FFU of the viruses in 0.2 ml of either DMEM-BSA or DMEM-BSA mixture with the mucus suspension collected the day before (3:1, vol/vol). The inoculum was removed 1 h post inoculation. The cultures were incubated for an additional 7 h at 37˚C under ALI conditions, fixed, immuno-stained for viral NP, and infected cells were counted under an inverted microscope as described elsewhere [78].

Competitive replication in HTBE cultures

Simultaneous competition between HK and 6 single-point HA mutants (HK-2, HK-62, HK-81, HK-92, HK-144, HK-193) was studied by inoculating HTBE cultures with a mixture containing equivalent amounts of these seven viruses based on infectious titers. Three different dilutions of this mixture were inoculated into HTBE cultures using 5 PFU of each virus per culture (low dose, L, 12 replicate cultures), 20 PFU per culture (medium dose, M, 12 cultures), and 320 PFU per culture (high dose, H, 6 cultures). After 1-h incubation at 37˚C, the inocula were removed, and the cultures were incubated under ALI conditions. At 72 h post-inoculation, 0.3 ml DMEM was added to the apical sides of the cultures for 30 min. The apical medium was collected, stored at −80˚C, and analysed together with the stored aliquot of the original inoculated viral mixture. The RNA was extracted from the virus-containing samples, processed and analysed by next generation sequencing as described previously [81]. Proportions of each mutant HA were calculated as the ratio of the number of reads coding for a specific mutation and the total number of reads covering the corresponding nucleotide position. To cope with the overdispersion of the values of proportions, they were assumed to follow a beta binomial distribution and analysed using a generalized linear model for location (mean), scale (variance) and shape (skeweness of distribution). For details of model construction see S1 Method. P values reflecting differences between the harvest and the inoculum were adjusted according to Dunnett's method [82]. Whereas each of the 6 HA mutants were unambiguously identified and quantified in the mixture via corresponding unique single nucleotide polymorphisms, the parent HK lacked a mutation that could serve as unique identifier for this virus.

We therefore inferred proportions of HK in the mixtures by subtracting experimentally determined proportions of all mutants from a theoretical value of 1 (for details see S1 Method).

Airborne transmission between ferrets

Respiratory droplet transmission experiments were performed as described previously [83]. In brief, groups of two seronegative female adult ferrets were inoculated intranasally with 10^6 $TCID_{50}$ of each virus by applying 0.25 ml of viral suspension to each nostril. One day after inoculation, one naive ferret was placed opposite to each inoculated ferret in a transmission cage that prevented direct contact but allowed airflow from the inoculated to the naïve ferret. The distance between the stainless steel grids of the ferret cages was 10 cm, and a directional airflow was present from the donor to the recipient cage of approximately 100 l/min. Nose and throat swabs were collected from inoculated and contact ferrets on days 1, 3, 5, and 7 post-inoculation and days 1, 3, 5, 7 and 9 post-exposure, respectively. Viral titers in swabs were determined by end-point titration in MDCK cells. Blood was collected from all ferrets on day 14 post exposure, and the presence of antibodies against the tested viruses was analysed by hemagglutination inhibition assay using standard procedures [84]. All animals were humanely killed at the end of the in-vivo phase of the study.

HA sequences and phylogenetic analyses

Full-length nucleotide sequences of the H3 HAs were downloaded from the GISAID EpiFlu database [85] accessed on March 11, 2020. Sequences were aligned using the MAFFT multiple alignment program implemented in the Unipro UGENE package [86], version 35. Sequences containing gaps and ambiguities, non-unique sequences and sequences of laboratory-derived IAVs were removed manually using Bio-Edit version 7.1.11 [87]. Jalview version 2.11 [88] was used to select representative sequences of swine IAVs by discarding all but one sequences with identity >99% (redundancy threshold 99%). The evolutionary history was inferred using IQ-TREE 2 with ModelFinder [89,90], the tree was plotted using MEGA7 [91]. Protein logos were generated using web-based application WebLogo (Crooks et al., 2004).

Selection pressure analyses

Three groups of host-specific HA sequences were analyzed, which included 1492 sequences of avian IAVs, 406 sequences of equine, canine, feline and seal IAVs and 803 sequences of human IAVs isolated in the years from 1968 to 1999. We partitioned the maximum likelihood tree into three groups of *internal* branches: human (801 branches), avian (1496 branches) and mammalian (394 branches). Our analyses used dN/dS techniques (for a review, see [92]). Because internal branches encompass at least one transmission event, we can assume that changes occurring along these branches have been "seen" by selection. Not including changes occurring along terminal branches we reduce the biasing effect of intra-host variation, which may be maladaptive on the population level [92], and tends to inflate dN/dS estimates [93]. For each site in the HA alignment, we addressed the following four questions based on tests available in the HyPhy 2.5 package [94].

1. What is the mean dN/dS at a site along the branches of interest? Does the site evolve subject to pervasive negative (dN/dS < 1) or positive diversifying (dN/dS > 1) selection? This test uses the Fixed Effects Likelihood (FEL) method [41], and significance was established using a likelihood ratio test (LRT), at p ≤ 0.05. In addition, we inferred the number of synonymous and non-synonymous changes and the most likely character at each internal node of the tree using the SLAC method.

2. Does the site evolve subject episodic positive diversifying selection (dN/dS > 1 along some fraction of the tree)? This test used the Mixed Effects Model of Evolution (MEME) method [42], and significance was established using LRT, at $p \leq 0.05$.

3. Does the site evolve under different selective pressures between groups of branches (dN/dS differ between some or all of the four sets of branches)? This test used the Contrast Fixed Effects Likelihood (Contrast-FEL) method [43], and significance was established using a collection of seven LRT (one for each pair of branch sets, and an omnibus test), with corrected $p \leq 0.01$.

4. Is there evidence of directional evolution on "human" branches, where specific amino-acids are being selected for? This is based on an improved version of the Directional Evolution of Protein Sequences (DEPS) test [44], and uses empirical Bayes Factors ≥ 100 to identify, which, if any residues at a given site are being selected for / against. This test is not based on dN/dS and is more suited to detect "sweeping" changes which involve only a few substitutions.

Multiple sequence alignments, trees, and machine readable analysis results files are available at https://github.com/spond/IAV-H3N2-analysis.

Statistics

Statistical tests were performed using Graphpad Prism 8.4 and R 3.6.0 (www.r-project.org). Unless stated otherwise, figures show data from individual biological replicates. The bars or horizontal lines indicate the group means, the length of the error bars is one standard deviation. The details are explained in the table footnotes and figure legends. Student's t test was used to compare two groups. Dunnett's or Tukey's multiple comparison tests were performed to compare more than two groups. More sophisticated statistical models were fitted by generalized linear models in R. If not stated otherwise, strictly positive variables were log-transformed before analysis, and percentage data were analyzed using quasi-binomial models with logit link. If the data consisted of experiments made on different days, day was included as a random intercept. Multiple tests of coefficients or contrasts within these models were done using simultaneous tests for general linear hypotheses, and P values were adjusted using the single step method [95] (comparable to Dunnett's and Tukey's procedure for multiple-to-one and all-pairwise comparisons). If mixed models were used to account for day-to-day variations between experiments, figures show data adjusted for day (that is, the variance attributed to day-to-day variation is removed from the data to better show how the values depend on the fixed factor). Details are explained in the methods sections of the respective experiments. Observed statistical significance is indicated in the figures as follows: *, P <0,05; **, P <0,01; ***, P <0,001.

Supporting information

S1 Method. Construction of a GAMLSS model to analyse the frequency of observation of the HA mutants in the competitive replication assay.
(DOCX)

S1 Data. Excel spreadsheet containing, in separate sheets, the underlying numerical data for Fig panels 2a, 2b, 2c, 2d, 2e, 2f, 3a, 3b, 3e, 4a, 4b, 4c, 4d, 5a, 5b, 6-HK, 6-R5, 8, 9, S3a, S3b, S3c, S4a, S4b, S4c.
(XLSX)

S1 Table. Prevalence of amino acids at indicated positions of the H3 HA of avian IAVs.
(DOCX)

S2 Table. Originating and submitting laboratories of the sequences from GISAID's EpiFlu Database on which this research is based.
(XLSX)

S3 Table. P values for the differences between Kass of viral complexes with SGPs presented in the Fig 3A.
(PDF)

S1 Fig. Inference of amino acid substitutions separating HAs of H3N2/1968 pandemic IAVs from their avian ancestor. (**a**) Phylogenetic relationships between H3N2/1968 and avian IAVs. All full-length HA sequences of avian H3 IAVs and of pandemic IAVs isolated in 1968–1969 were downloaded from the GISAID EpiFlu database and processed as described in Materials and Methods. The final dataset contained sequences of 1494 avian and 36 human IAV HAs. The evolutionary history was inferred using MEGA7 with the maximum likelihood method based on the Kimura 2-parameter model. The tree is drawn to scale, with branch lengths measured in the number of nucleotide substitutions per site. Colours depict avian IAVs from North America (*blue*), Eurasia and Oceania (*black*), and H3N2/1968 pandemic viruses (*red*). (**b**) Detailed view of the branch, which includes pandemic and the closest avian IAVs marked by the dashed box in the panel S1a. Subtypes of avian IAVs and accession numbers of all sequences are shown next to the strain names. HA amino acid sequence of the common avian-human ancestor (node A) was inferred using the Ancestors program of MEGA7. The ancestral sequence together with the sequences of representative avian and human IAVs depicted by green diamonds are shown in the panel S1c. (**c**) HA amino acid sequences of H3N2 pandemic IAVs viruses isolated in 1968, the closest avian IAVs and the inferred common avian-human ancestor. Numbering starts from the N-terminus of mature HA protein, the signal peptide is numbered from -15 to 0. Dots depict identity with the sequence of the avian ancestor. Glycosylation sites at HA positions 63–65 and 81–83 are highlighted by *yellow*. Arrow depicts virus strain A/Hong Kong/1/1968 (H3N2) used to make recombinant viruses in this study. The figure was generated using Bio-Edit.
(PDF)

S2 Fig. Evolution of glycosylation sites in HA positions 63 and 81 of human H3N2 IAVs. (**a**) Phylogenetic relationships between the HA of IAVs isolated from 1968 to 1995 were inferred using MEGA7 with the maximum likelihood method. GISAID EpiFlu accession numbers and amino acids in positions 63–65 and 81–83 are shown next to the strain names. *Coloured circles* depict presence of glycosylation sites 63–65 (*blue*) and 81–83 (*red*). One virus strain (A/Port Chalmers/1/1973) contained glycosylation sites at both positions (*magenta*). The branch containing sequences of the viruses isolated after 1975 is collapsed for clarity (*blue triangle*). Prototype strains used to define two glycosylation lineages of pandemic viruses are highlighted by *yellow*. (**b**) Protein logos show frequencies of amino acids at positions 63–65 and 81–83 of the HA. Years of virus isolation and numbers of analysed sequences (in parentheses) are shown on the left. The figure was generated using Phylo-mLogo software (Shih ACC, Lee DT, Peng CL, Wu YW. Phylo-mLogo: an interactive and hierarchical multiple-logo visualization tool for alignment of many sequences. BMC bioinformatics. 2007; 8: 63).
(PDF)

S3 Fig. Conformational stability and membrane-fusion properties of the point HA mutants of HK and R5. pH of acid-induced conformational transition (**a**), pH threshold of

polykarion formation (**b**), and inhibition of viral infection by ammonium chloride (**c**) were assayed as described in Materials and Methods. Panels show data points, geometric mean and SDs from 2 to 8 experiments performed on different days with 2 to 6 replicates each. Data in panels S3a and S3c were analysed using general linear mixed models in R 3.6.0. Concentration was log-transformed before analysis. Day was included as a random intercept. Figures show data adjusted for day. In panel S3b, due to the low resolution of the assay (with only 3 distinct pH values), pH was analysed as an ordinal variable using an ordered logistic regression model. Multiple tests of contrasts within these models were done using simultaneous tests for general linear hypotheses, and P values were adjusted using the single step method. In all panels, vertical dotted line separates point mutants of HK and point mutants of R5. Red asterisks depict point mutants that were significantly different from their parental viruses, either HK or R5. No significant difference was observed between HK and R5 in these assays.
(PDF)

S4 Fig. Receptor-binding properties of HK, R5 and their mutants grown in Calu-3 cells and HTBE cultures. (**a**) Association constants of viral complexes with 6'SLN (20 kDa). (**b,c**) Virus avidity for receptors on MDCK cells expressed as concentrations of the *Vibrio cholerae* sialidase that reduced numbers of infected cells by 50% (IC_{50}). The higher IC_{50}, the higher binding avidity. All panels show replicates, mean values (bars) and SDs. Asterisks in panels S4a and S4b show P values for the differences with HK (dark grey bars) determined by one-way Anova with Dunnett's multiple comparisons test. Asterisks in panel S4c depict differences between individual viruses determined by one-way Anova with Tukey's test.
(PDF)

S5 Fig. Variation of amino acids in selected positions of H3 HA during evolution in different host species (scalable vector graphics). The figure shows maximum likelihood tree for the nucleotide sequences of H3 HA inferred using IQ-TREE 2 and plotted using Mega 7. *Taxon labels* (for example, A/equine/Miami/1/1963-A/H3N8-129722-H-R-NCT-YEN-S-A-K-QSG) include virus name, subtype, GISAID EpiFlu accession number and amino acids in HA positions -2, 62, 63–65, 81–83, 92, 144, 193, 226–228. Colours depict the following stable host-specific lineages. *Black*, avian; *red*, human; *green*, swine; *blue*, equine; *cyan*, equine/Jilin/1/1989; *orange*, canine H3N8; *magenta*, canine H3N2. The earliest virus isolates from mammalian lineages are highlighted with *yellow*. *Green dots* depict 10 sporadic avian-like mammalian isolates. The tree includes the following full-length non-redundant sequences available from GISAID EpiFlu database: 106 representative sequences of human IAVs isolated from 1968 to 2020 (2 sequences per year); 456 representative sequences of swine IAVs selected from the total 2663 sequences using JalView; all 402 sequences of other mammalian IAVs (mainly equine and canine); all 1494 sequences of avian IAVs; 31 sequences of avian, swine and human IAVs sporadically isolated from a heterologous host species.
(PDF)

S6 Fig. Amino acid composition at the nine sites of H3 HA (scalable vector graphics). The tree is based on HA sequences used for selection pressure analyses and includes 1492 sequences of avian IAVs, 406 sequences of equine, canine, feline and seal IAVs and 803 sequences of human IAVs isolated in the years from 1968 to 1999. Unobserved ancestral codons were inferred using the SLAC method. Host-specific clades are depicted in the tree display for the site -2.
(PDF)

Acknowledgments

Author Contributions

Conceptualization: Hans-Dieter Klenk, Mikhail Matrosovich.

Formal analysis: Gianpiero Zamperin, Michele Gastaldelli, Francesco Bonfante, Jochen Wilhelm, Mikhail Matrosovich.

Funding acquisition: Johanna West, Sander Herfst, Ron Fouchier, Jochen Wilhelm, Mikhail Matrosovich.

Investigation: Johanna West, Juliane Röder, Tatyana Matrosovich, Jana Beicht, Jan Baumann, Nancy Mounogou Kouassi, Jennifer Doedt, Annalisa Salviato, Sergei Kosakovsky Pond, Sander Herfst.

Methodology: Gianpiero Zamperin, Michele Gastaldelli, Francesco Bonfante, Sergei Kosakovsky Pond, Sander Herfst, Ron Fouchier, Jochen Wilhelm, Mikhail Matrosovich.

Resources: Nicolai Bovin.

Supervision: Mikhail Matrosovich.

Visualization: Johanna West, Sergei Kosakovsky Pond, Sander Herfst, Jochen Wilhelm, Mikhail Matrosovich.

Writing – original draft: Johanna West, Mikhail Matrosovich.

Writing – review & editing: Johanna West, Nicolai Bovin, Francesco Bonfante, Sergei Kosakovsky Pond, Sander Herfst, Ron Fouchier, Jochen Wilhelm, Hans-Dieter Klenk, Mikhail Matrosovich.

References

1. Olsen B, Munster VJ, Wallensten A, Waldenström J, Osterhaus AD, Fouchier RA. Global patterns of influenza a virus in wild birds. Science. 2006; 312:384–388. https://doi.org/10.1126/science.1122438 PMID: 16627734

2. Yoon SW, Webby RJ, Webster RG. Evolution and ecology of influenza A viruses. Curr Top Microbiol Iimmunol. 2014; 385:359–375. https://doi.org/10.1007/82_2014_396 PMID: 24990620

3. Wang D, Zhu W, Yang L, Shu Y. The epidemiology, virology, and pathogenicity of human infections with avian influenza viruses. Cold Spring Harb Perspect Med. 2020:a038620. https://doi.org/10.1101/cshperspect.a038620 PMID: 31964651

4. Taubenberger JK, Kash JC. Influenza virus evolution, host adaptation, and pandemic formation. Cell Host Microbe. 2010; 7:440–451. https://doi.org/10.1016/j.chom.2010.05.009 PMID: 20542248

5. Guan Y, Vijaykrishna D, Bahl J, Zhu H, Wang J, Smith GJ. The emergence of pandemic influenza viruses. Protein Cell. 2010; 1:9–13. https://doi.org/10.1007/s13238-010-0008-z PMID: 21203993

6. de Graaf M, Fouchier RA. Role of receptor binding specificity in influenza A virus transmission and pathogenesis. EMBO J. 2014; 33:823–841. https://doi.org/10.1002/embj.201387442 PMID: 24668228

7. Byrd-Leotis L, Cummings RD, Steinhauer DA. The interplay between the host receptor and influenza virus hemagglutinin and neuraminidase. Int J Mol Sci. 2017; 18:1541. https://doi.org/10.3390/ijms18071541 PMID: 28714909

8.   Matrosovich MN, Klenk H-D, Kawaoka Y. Receptor specificity, host range and pathogenicity of influenza viruses. In: Kawaoka Y, editor. Influenza Virology: Current Topics. Wymondham, England: Caister Academic Press; 2006. p. 95–137.

9.   Gamblin SJ, Vachieri SG, Xiong X, Zhang J, Martin SR, Skehel JJ. Hemagglutinin structure and activities. Cold Spring Harb Perspect Med. 2020:a038638. https://doi.org/10.1101/cshperspect.a038638 PMID: 32513673

10. Thompson AJ, Paulson JC. Adaptation of influenza viruses to human airway receptors. J Biol Chem. 2020:100017. https://doi.org/10.1074/jbc.REV120.013309 PMID: 33144323

11. Russell CJ, Hu M, Okda FA. Influenza hemagglutinin protein stability, activation, and pandemic risk. Trends Microbiol. 2018; 26:841–853. https://doi.org/10.1016/j.tim.2018.03.005 PMID: 29681430

12. Russell CJ. Hemagglutinin stability and its impact on influenza A virus infectivity, pathogenicity, and transmissibility in avians, mice, swine, seals, ferrets, and humans. Viruses. 2021; 13:746. https://doi.org/10.3390/v13050746 PMID: 33923198

13. Scholtissek C. Stability of infectious influenza A viruses to treatment at low pH and heating. Arch Virol. 1985; 85:1–11. https://doi.org/10.1007/BF01317001 PMID: 4015405

14. Daidoji T, Watanabe Y, Ibrahim MS, Yasugi M, Maruyama H, Masuda T, et al. Avian influenza virus infection of immortalized human respiratory epithelial cells depends upon a delicate balance between hemagglutinin acid stability and endosomal pH. J Biol Chem. 2015; 290:10627–10642. https://doi.org/10.1074/jbc.M114.611327 PMID: 25673693

15. Baumann J, Kouassi NM, Foni E, Klenk HD, Matrosovich M. H1N1 swine influenza viruses differ from avian precursors by a higher pH optimum of membrane fusion. J Virol. 2016; 90:1569–1577. https://doi.org/10.1128/JVI.02332-15 PMID: 26608319

16. Galloway SE, Reed ML, Russell CJ, Steinhauer DA. Influenza HA subtypes demonstrate divergent phenotypes for cleavage activation and pH of fusion: implications for host range and adaptation. PLoS Pathog. 2013; 9:e1003151. https://doi.org/10.1371/journal.ppat.1003151 PMID: 23459660

17. Cotter CR, Jin H, Chen ZY. A single amino acid in the stalk region of the H1N1pdm influenza virus HA protein affects viral fusion, stability and infectivity. PLoS Pathog. 2014; 10:e1003831. https://doi.org/10.1371/journal.ppat.1003831 PMID: 24391498

18. Russier M, Yang G, Rehg JE, Wong SS, Mostafa HH, Fabrizio TP, et al. Molecular requirements for a pandemic influenza virus: An acid-stable hemagglutinin protein. Proc Natl Acad Sci USA. 2016; 113:1636–1641. https://doi.org/10.1073/pnas.1524384113 PMID: 26811446

19. Matrosovich M, Matrosovich T, Uhlendorff J, Garten W, Klenk HD. Avian-virus-like receptor specificity of the hemagglutinin impedes influenza virus replication in cultures of human airway epithelium. Virology. 2007; 361:384–390. https://doi.org/10.1016/j.virol.2006.11.030 PMID: 17207830

20. Van Poucke S, Uhlendorff J, Wang Z, Billiau V, Nicholls J, Matrosovich M, et al. Effect of receptor specificity of A/Hong Kong/1/68 (H3N2) influenza virus variants on replication and transmission in pigs. Influenza Other Respir Viruses. 2013; 7:151–159. https://doi.org/10.1111/j.1750-2659.2012.00376.x PMID: 22564359

21. Van Poucke S, Doedt J, Baumann J, Qiu Y, Matrosovich T, Klenk HD, et al. Role of substitutions in the hemagglutinin in the emergence of the 1968 pandemic influenza virus. J Virol. 2015; 89:12211–12216. https://doi.org/10.1128/JVI.01292-15 PMID: 26378170

22. Bean WJ, Schell M, Katz J, Kawaoka Y, Naeve C, Gorman O, et al. Evolution of the H3 influenza virus hemagglutinin from human and nonhuman hosts. J Virol. 1992; 66:1129–1138. https://doi.org/10.1128/JVI.66.2.1129-1138.1992 PMID: 1731092

23. Lin YP, Xiong X, Wharton SA, Martin SR, Coombs PJ, Vachieri SG, et al. Evolution of the receptor binding properties of the influenza A(H3N2) hemagglutinin. Proc Natl Acad Sci USA. 2012; 109:21474–21479. https://doi.org/10.1073/pnas.1218841110 PMID: 23236176

24. Ohuchi M, Ohuchi R, Sakai T, Matsumoto A. Tight binding of influenza virus hemagglutinin to its receptor interferes with fusion pore dilation. J Virol. 2002; 76:12405–12413. https://doi.org/10.1128/jvi.76.24.12405-12413.2002 PMID: 12438566

25. Lee RTC, Chang HH, Russell CA, Lipsitch M, Maurer-Stroh S. Influenza A hemagglutinin passage bias sites and host specificity mutations. Cells. 2019; 8:958. https://doi.org/10.3390/cells8090958 PMID: 31443542

26. Matrosovich MN, Gambaryan AS. Solid-phase assays of receptor-binding specificity. Methods Mol Biol. 2012; 865:71–94. https://doi.org/10.1007/978-1-61779-621-0_5 PMID: 22528154

27. Tuzikov A, Chinarev A, Shilova N, Gordeeva E, Galanina O, Ovchinnikova T, et al. 40 years of glycopolyacrylamide in glycobiology. Glycoconj J. 2021; 38:89–100. https://doi.org/10.1007/s10719-020-09965-5 PMID: 33443721

28. Gambaryan AS, Matrosovich TY, Boravleva EY, Lomakina NF, Yamnikova SS, Tuzikov AB, et al. Receptor-binding properties of influenza viruses isolated from gulls. Virology. 2018; 522:37–45. https://doi.org/10.1016/j.virol.2018.07.004 PMID: 30014856

29. Gambaryan AS, Matrosovich TY, Philipp J, Munster VJ, Fouchier RA, Cattoli G, et al. Receptor-binding profiles of H7 subtype influenza viruses in different host species. J Virol. 2012; 86:4370–4379. https://doi.org/10.1128/JVI.06959-11 PMID: 22345462

30. Gambaryan AS, Marinina VP, Tuzikov AB, Bovin NV, Rudneva IA, Sinitsyn BV, et al. Effects of host-dependent glycosylation of hemagglutinin on receptor-binding properties on H1N1 human influenza A virus grown in MDCK cells and in embryonated eggs. Virology. 1998; 247:170–177. https://doi.org/10.1006/viro.1998.9224 PMID: 9705910

31. Inkster MD, Hinshaw VS, Schulze IT. The hemagglutinins of duck and human H1 influenza viruses differ in sequence conservation and in glycosylation. J Virol. 1993; 67:7436–7443. https://doi.org/10.1128/JVI.67.12.7436-7443.1993 PMID: 8230464

32. Matrosovich MN, Matrosovich TY, Gray T, Roberts NA, Klenk HD. Human and avian influenza viruses target different cell types in cultures of human airway epithelium. Proc Natl Acad Sci U S A. 2004; 101:4620–4624. https://doi.org/10.1073/pnas.0308001101 PMID: 15070767

33. Davis AS, Chertow DS, Moyer JE, Suzich J, Sandouk A, Dorward DW, et al. Validation of normal human bronchial epithelial cells as a model for influenza a infections in human distal trachea. J Histochem Cytochem. 2015; 63:312–328. https://doi.org/10.1369/0022155415570968 PMID: 25604814

34. Gambaryan AS, Matrosovich MN, Bender CA, Kilbourne ED. Differences in the biological phenotype of low-yielding (L) and high-yielding (H) variants of swine influenza virus A/NJ/11/76 are associated with their different receptor-binding activity. Virology. 1998; 247:223–231. https://doi.org/10.1006/viro.1998.9274 PMID: 9705915

35. Couceiro JN, Paulson JC, Baum LG. Influenza virus strains selectively recognize sialyloligosaccharides on human respiratory epithelium; the role of the host cell in selection of hemagglutinin receptor specificity. Virus Research. 1993; 29:155–165. https://doi.org/10.1016/0168-1702(93)90056-s PMID: 8212857

36. Baum LG, Paulson JC. Sialyloligosaccharides of the respiratory epithelium in the selection of human influenza virus receptor specificity. Acta Histochem Suppl. 1990; 40:35–38. PMID: 2091044.

37. Webster RG, Bean WJ, Gorman OT, Chambers TM, Kawaoka Y. Evolution and ecology of influenza A viruses. Microbiol Rev. 1992; 56:152–179. https://doi.org/10.1128/mr.56.1.152-179.1992 PMID: 1579108

38. Guo Y, Wang M, Zheng GS, Li WK, Kawaoka Y, Webster RG. Seroepidemiological and molecular evidence for the presence of two H3N8 equine influenza viruses in China in 1993–94. J Gen Virol. 1995; 76:2009–2014. https://doi.org/10.1099/0022-1317-76-8-2009 PMID: 7636481

39. Parrish CR, Murcia PR, Holmes EC. Influenza virus reservoirs and intermediate hosts: dogs, horses, and new possibilities for influenza virus exposure of humans. J Virol. 2015; 89:2990–2994. https://doi.org/10.1128/JVI.03146-14 PMID: 25540375

40. Anderson TK, Chang J, Arendsee ZW, Venkatesh D, Souza CK, Kimble JB, et al. Swine influenza A viruses and the tangled relationship with humans. Cold Spring Harb Perspect Med. 2020:a038737. https://doi.org/10.1101/cshperspect.a038737 PMID: 31988203

41. Kosakovsky Pond SL, Frost SDW. Not so different after all: A comparison of methods for detecting amino acid sites under selection. Mol Biol Evol. 2005; 22:1208–1222. https://doi.org/10.1093/molbev/msi105 PMID: 15703242

42. Murrell B, Wertheim JO, Moola S, Weighill T, Scheffler K, Kosakovsky Pond SL. Detecting individual sites subject to episodic diversifying selection. PLoS Genet. 2012; 8:e1002764. https://doi.org/10.1371/journal.pgen.1002764 PMID: 22807683

43. Kosakovsky Pond SL, Wisotsky SR, Escalante A, Magalis BR, Weaver S. Contrast-FEL—a test for differences in selective pressures at individual sites among clades and sets of branches. Mol Biol Evol. 2021; 38:1184–1198. https://doi.org/10.1093/molbev/msaa263 PMID: 33064823

44. Kosakovsky Pond SL, Poon AFY, Leigh Brown AJ, Frost SDW. A maximum likelihood method for detecting directional evolution in protein sequences and its application to influenza A virus. Mol Biol Evol. 2008; 25:1809–1824. https://doi.org/10.1093/molbev/msn123 PMID: 18511426

45. Wiley DC, Skehel JJ. The structure and function of the hemagglutinin membrane glycoprotein of influenza virus. Annu Rev Biochem. 1987; 56:365–394. https://doi.org/10.1146/annurev.bi.56.070187.002053 PMID: 3304138

46. Anthony SJ, St Leger JA, Pugliares K, Ip HS, Chan JM, Carpenter ZW, et al. Emergence of fatal avian influenza in New England harbor seals. mBio. 2012; 3:e00166–12. https://doi.org/10.1128/mBio.00166-12 PMID: 22851656

47. Shi Y, Wu Y, Zhang W, Qi J, Gao GF. Enabling the 'host jump': structural determinants of receptor-binding specificity in influenza A viruses. Nat Rev Microbiol. 2014; 12:822–831. https://doi.org/10.1038/nrmicro3362 PMID: 25383601

48. Kida H, Kawaoka Y, Naeve CW, Webster RG. Antigenic and genetic conservation of H3 influenza virus in wild ducks. Virology. 1987; 159:109–119. https://doi.org/10.1016/0042-6822(87)90353-9 PMID: 2440178

49. Yasuda J, Shortridge KF, Shimizu Y, Kida H. Molecular evidence for a role of domestic ducks in the introduction of avian H3 influenza viruses to pigs in southern China, where the A/Hong Kong/68 (H3N2) strain emerged. J Gen Virol. 1991; 72:2007–2010. https://doi.org/10.1099/0022-1317-72-8-2007 PMID: 1875195

50. Connor RJ, Kawaoka Y, Webster RG, Paulson JC. Receptor specificity in human, avian, and equine H2 and H3 influenza virus isolates. Virology. 1994; 205:17–23. https://doi.org/10.1006/viro.1994.1615 PMID: 7975212

51. Matrosovich M, Tuzikov A, Bovin N, Gambaryan A, Klimov A, Castrucci MR, et al. Early alterations of the receptor-binding properties of H1, H2, and H3 avian influenza virus hemagglutinins after their introduction into mammals. J Virol. 2000; 74:8502–8512. https://doi.org/10.1128/jvi.74.18.8502-8512.2000 PMID: 10954551

52. Kaverin NV, Matrosovich MN, Gambaryan AS, Rudneva IA, Shilov AA, Varich NL, et al. Intergenic HA-NA interactions in influenza A virus: postreassortment substitutions of charged amino acid in the hemagglutinin of different subtypes. Virus Research. 2000; 66:123–129. https://doi.org/10.1016/s0168-1702(99)00131-8 PMID: 10725545

53. de Vries RP, de VE, Moore KS, Rigter A, Rottier PJ, de Haan CA. Only two residues are responsible for the dramatic difference in receptor binding between swine and new pandemic H1 hemagglutinin. J Biol Chem. 2011; 286:5868–5875. https://doi.org/10.1074/jbc.M110.193557 PMID: 21173148

54. Xu R, Zhu X, McBride R, Nycholat CM, Yu W, Paulson JC, et al. Functional balance of the hemagglutinin and neuraminidase activities accompanies the emergence of the 2009 H1N1 influenza pandemic. J Virol.2012; 86:9221–9232. https://doi.org/10.1128/JVI.00697-12 PMID: 22718832

55. Wagner R, Matrosovich M, Klenk HD. Functional balance between haemagglutinin and neuraminidase in influenza virus infections. Rev Med Virol. 2002; 12:159–166. https://doi.org/10.1002/rmv.352 PMID: 11987141

56. de Vries E, Du W, Guo H, de Haan CAM. Influenza A virus hemagglutinin-neuraminidase-receptor balance: Preserving virus motility. Trends Microbiol. 2020; 28:57–67. https://doi.org/10.1016/j.tim.2019.08.010 PMID: 31629602

57. Baum LG, Paulson JC. The N2 neuraminidase of human influenza virus has acquired a substrate specificity complementary to the hemagglutinin receptor specificity. Virology. 1991; 180:10–15. https://doi.org/10.1016/0042-6822(91)90003-t PMID: 1984642

58. Kobasa D, Kodihalli S, Luo M, Castrucci MR, Donatelli I, Suzuki Y, et al. Amino acid residues contributing to the substrate specificity of the influenza A virus neuraminidase. J Virol. 1999; 73:6743–6751. https://doi.org/10.1128/JVI.73.8.6743-6751.1999 PMID: 10400772

59. Uhlendorff J, Matrosovich T, Klenk HD, Matrosovich M. Functional significance of the hemadsorption activity of influenza virus neuraminidase and its alteration in pandemic viruses. Arch Virol. 2009; 154:945–957. https://doi.org/10.1007/s00705-009-0393-x PMID: 19458903

60. Du W, Guo H, Nijman VS, Doedt J, van der Vries E, van der Lee J, et al. The 2nd sialic acid-binding site of influenza A virus neuraminidase is an important determinant of the hemagglutinin-neuraminidase-receptor balance. Plos Pathog. 2019; 15:e1007860. https://doi.org/10.1371/journal.ppat.1007860 PMID: 31181126

61. Hensley SE, Das SR, Bailey AL, Schmidt LM, Hickman HD, Jayaraman A, et al. Hemagglutinin receptor binding avidity drives influenza A virus antigenic drift. Science. 2009; 326:734–736. https://doi.org/10.1126/science.1178258 PMID: 19900932

62. Medeiros R, Naffakh N, Manuguerra JC, van der Werf S. Binding of the hemagglutinin from human or equine influenza H3 viruses to the receptor is altered by substitutions at residue 193. Arch Virol. 2004; 149:1663–1671. https://doi.org/10.1007/s00705-003-0287-2 PMID: 15290389

63. Peng W, Bouwman KM, McBride R, Grant OC, Woods RJ, Verheije MH, et al. Enhanced human-type receptor binding by ferret-transmissible H5N1 with a K193T mutation. J Virol. 2018; 92:e02016–17. https://doi.org/10.1128/JVI.02016-17 PMID: 29491160

64. Collins PJ, Vachieri SG, Haire LF, Ogrodowicz RW, Martin SR, Walker PA, et al. Recent evolution of equine influenza and the origin of canine influenza. Proc Natl Acad Sci USA. 2014; 111:11175–11180. https://doi.org/10.1073/pnas.1406606111 PMID: 25024224

65. Xiong X, Tuzikov A, Coombs PJ, Martin SR, Walker PA, Gamblin SJ, et al. Recognition of sulphated and fucosylated receptor sialosides by A/Vietnam/1194/2004 (H5N1) influenza virus. Virus research. 2013; 178:12–14. https://doi.org/10.1016/j.virusres.2013.08.007 PMID: 24036174

66. Wu NC, Wilson IA. Structural biology of influenza hemagglutinin: An amaranthine adventure. Viruses. 2020; 12:1053. https://doi.org/10.3390/v12091053 PMID: 32971825

67. Eisen MB, Sabesan S, Skehel JJ, Wiley DC. Binding of the influenza A virus to cell-surface receptors: structures of five hemagglutinin–sialyloligosaccharide complexes determined by X-ray crystallography. Virology. 1997; 232:19–31. https://doi.org/10.1006/viro.1997.8526 PMID: 9185585

68. An Y, McCullers JA, Alymova I, Parsons LM, Cipollo JF. Glycosylation analysis of engineered H3N2 influenza a virus hemagglutinins with sequentially added historically relevant glycosylation sites. J Proteome Res. 2015; 14:3957–3969. https://doi.org/10.1021/acs.jproteome.5b00416 PMID: 26202417

69. Belser JA, Barclay W, Barr I, Fouchier RAM, Matsuyama R, Nishiura H, et al. Ferrets as models for influenza virus transmission studies and pandemic risk assessments. Emerg Infect Dis. 2018; 24:965–971. https://doi.org/10.3201/eid2406.172114 PMID: 29774862

70. Soh YS, Moncla LH, Eguia R, Bedford T, Bloom JD. Comprehensive mapping of adaptation of the avian influenza polymerase protein PB2 to humans. eLife. 2019; 8:e45079. https://doi.org/10.7554/eLife.45079 PMID: 31038123

71. Bateman AC, Busch MG, Karasin AI, Bovin N, Olsen CW. Amino acid 226 in the hemagglutinin of H4N6 influenza virus determines binding affinity for alpha2,6-linked sialic acid and infectivity levels in primary swine and human respiratory epithelial cells. J Virol. 2008; 82:8204–8209. https://doi.org/10.1128/JVI.00718-08 PMID: 18550676

72. Gallagher PJ, Henneberry JM, Sambrook JF, Gething MJ. Glycosylation requirements for intracellular transport and function of the hemagglutinin of influenza virus. J Virol. 1992; 66:7136–7145. https://doi.org/10.1128/JVI.66.12.7136-7145.1992 PMID: 1331514

73. Hebert DN, Zhang JX, Chen W, Foellmer B, Helenius A. The number and location of glycans on influenza hemagglutinin determine folding and association with calnexin and calreticulin. J Cell Biol. 1997; 139:613–623. https://doi.org/10.1083/jcb.139.3.613 PMID: 9348279

74. Daniels R, Kurowski B, Johnson AE, Hebert DN. N-linked glycans direct the cotranslational folding pathway of influenza hemagglutinin. Mol Cell. 2003; 11:79–90. https://doi.org/10.1016/s1097-2765(02)00821-3 PMID: 12535523

75. Imai M, Watanabe T, Hatta M, Das SC, Ozawa M, Shinya K, et al. Experimental adaptation of an influenza H5 HA confers respiratory droplet transmission to a reassortant H5 HA/H1N1 virus in ferrets. Nature. 2012; 486:420–428. https://doi.org/10.1038/nature10831 PMID: 22722205

76. Herfst S, Schrauwen EJ, Linster M, Chutinimitkul S, de Wit E, Munster VJ, et al. Airborne transmission of influenza A/H5N1 virus between ferrets. Science. 2012; 336:1534–1541. https://doi.org/10.1126/science.1213362 PMID: 22723413

77. Hoffmann E, Neumann G, Kawaoka Y, Hobom G, Webster RG. A DNA transfection system for generation of influenza A virus from eight plasmids. Proc Natl Acad Sci USA. 2000; 97:6108–6113. https://doi.org/10.1073/pnas.100133697 PMID: 10801978

78. Gerlach T, Hensen L, Matrosovich T, Bergmann J, Winkler M, Peteranderl C, et al. pH optimum of hemagglutinin-mediated membrane fusion determines sensitivity of influenza A viruses to the interferon-induced antiviral state and IFITMs. J Virol. 2017; 91:e00246–17. https://doi.org/10.1128/JVI.00246-17 PMID: 28356532

79. Matrosovich M, Matrosovich T, Garten W, Klenk HD. New low-viscosity overlay medium for viral plaque assays. Virol J. 2006; 3:63. https://doi.org/10.1186/1743-422X-3-63 PMID: 16945126

80. Reed ML, Yen HL, DuBois RM, Bridges OA, Salomon R, Webster RG, et al. Amino acid residues in the fusion peptide pocket regulate the pH of activation of the H5N1 influenza virus hemagglutinin protein. J Virol. 2009; 83:3568–3580. https://doi.org/10.1128/JVI.02238-08 PMID: 19193808

81. Wade A, Jumbo SD, Zecchin B, Fusaro A, Taiga T, Bianco A, et al. Highly pathogenic avian influenza A (H5N8) virus, Cameroon, 2017. Emerg Infect Dis. 2018; 24:1367–1370. https://doi.org/10.3201/eid2407.172120 PMID: 29912710

82. Lenth RV, Buerkner P, Herve M, Love J, Riebl H, Singmann H. Emmeans:Estimated marginal means, aka least-squares means. R package version 1.3. 2. Retrieved from https://cran.r-project.org/package=emmeans. 2019.

83. Munster VJ, de Wit E, van den Brand JM, Herfst S, Schrauwen EJ, Bestebroer TM, et al. Pathogenesis and transmission of swine-origin 2009 A(H1N1) influenza virus in ferrets. Science. 2009; 325:481–483. https://doi.org/10.1126/science.1177127 PMID: 19574348

84. WHO manual on animal influenza diagnosis and surveillance No. WHO/CDS/CSR/NCS/2002.5. Department of Communicable Diseases Surveillance and Control, WHO, Geneva, 2002.

85. Shu Y, McCauley J. GISAID: Global initiative on sharing all influenza data—from vision to reality. Euro Surveill. 2017; 22:30494. https://doi.org/10.2807/1560-7917.ES.2017.22.13.30494 PMID: 28382917

86. Okonechnikov K, Golosova O, Fursov M, team U. Unipro UGENE: a unified bioinformatics toolkit. Bioinformatics. 2012; 28:1166–1167. https://doi.org/10.1093/bioinformatics/bts091 PMID: 22368248

87. Hall T. BioEdit: An important software for molecular biology. GERF Bulletin of Biosciences. 2011; 2:60–61.

88. Waterhouse AM, Procter JB, Martin DMA, Clamp M, Barton GJ. Jalview Version 2—a multiple sequence alignment editor and analysis workbench. Bioinformatics. 2009; 25:1189–1191. https://doi.org/10.1093/bioinformatics/btp033 PMID: 19151095

89. Kalyaanamoorthy S, Minh BQ, Wong TKF, von Haeseler A, Jermiin LS. ModelFinder: fast model selection for accurate phylogenetic estimates. Nat Methods. 2017; 14:587–589. https://doi.org/10.1038/nmeth.4285 PMID: 28481363

90. Minh BQ, Schmidt HA, Chernomor O, Schrempf D, Woodhams MD, von Haeseler A, et al. IQ-TREE 2: New models and efficient methods for phylogenetic inference in the genomic era. Mol Biol Evol. 2020; 37:1530–1534. https://doi.org/10.1093/molbev/msaa015 PMID: 32011700

91. Kumar S, Stecher G, Tamura K. MEGA7: Molecular evolutionary genetics analysis version 7.0 for bigger datasets. Mol Biol Evol. 2016; 33:1870–1874. https://doi.org/10.1093/molbev/msw054 PMID: 27004904

92. Pond SL, Frost SD, Grossman Z, Gravenor MB, Richman DD, Brown AJ. Adaptation to different human populations by HIV-1 revealed by codon-based analyses. PLoS Comput Biol. 2006; 2:e62. https://doi.org/10.1371/journal.pcbi.0020062 PMID: 16789820

93. Kryazhimskiy S, Plotkin JB. The population genetics of dN/dS. PLoS Genet. 2008; 4:e1000304. https://doi.org/10.1371/journal.pgen.1000304 PMID: 19081788

94. Kosakovsky Pond SL, Poon AFY, Velazquez R, Weaver S, Hepler NL, Murrell B, et al. HyPhy 2.5-A customizable platform for evolutionary hypothesis testing using phylogenies. Mol Biol Evol. 2020; 37:295–299. https://doi.org/10.1093/molbev/msz197 PMID: 31504749

95. Hothorn T, Bretz F, Westfall P. Simultaneous inference in general parametric models. Biom J. 2008; 50:346–363. https://doi.org/10.1002/bimj.200810425 PMID: 18481363