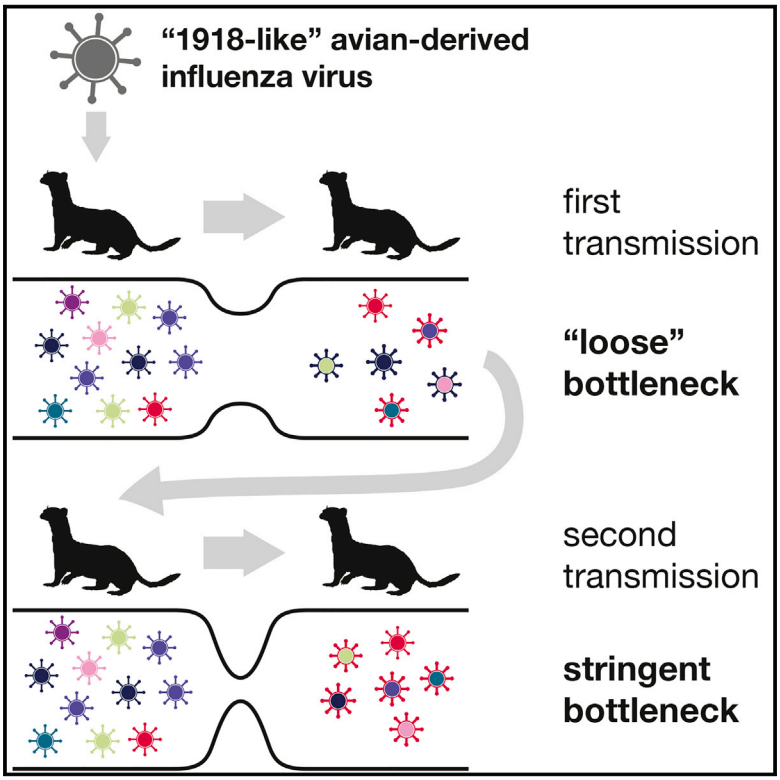
Since January 2020 Elsevier has created a COVID-19 resource centre with free information in English and Mandarin on the novel coronavirus COVID-19. The COVID-19 resource centre is hosted on Elsevier Connect, the company's public news and information website.

Article

Cell Host & Microbe

Selective Bottlenecks Shape Evolutionary Pathways Taken during Mammalian Adaptation of a 1918-like Avian Influenza Virus

Graphical Abstract



Highlights

- HA diversification arising during initial ferret adaption of avian flu virus is maintained
- Low-frequency transmissible polymerase variants arise subsequently
- Transmission bottleneck selects specific HA variants
- Mammalian transmission can evolve through multiple genetic pathways

Authors

Louise H. Moncla, Gongxun Zhong, Chase W. Nelson, ..., Tokiko Watanabe, Yoshihiro Kawaoka, Thomas C. Friedrich

Correspondence

kawaokay@vetmed.wisc.edu (Y.K.), thomasf@primate.wisc.edu (T.C.F.)

In Brief

Human pandemic influenza viruses can emerge unpredictably from avian reservoirs. Moncla et al. trace the evolutionary path of an avian-like H1N1 virus as it adapts to mammals. They find that initial adaption involves a loose bottleneck, which becomes selective as adaptation progresses, with mammalian transmission evolving via multiple genetic pathways.

Moncla et al., 2016, Cell Host & Microbe *19*, 169–180
February 10, 2016
http://dx.doi.org/10.1016/j.chom.2016.01.011

Cell Host & Microbe
Article

Selective Bottlenecks Shape Evolutionary Pathways Taken during Mammalian Adaptation of a 1918-like Avian Influenza Virus

Louise H. Moncla,[1,2] Gongxun Zhong,[1] Chase W. Nelson,[3] Jorge M. Dinis,[1,2] James Mutschler,[1,2] Austin L. Hughes,[3] Tokiko Watanabe,[1,4] Yoshihiro Kawaoka,[1,4,*] and Thomas C. Friedrich[1,2,*]
[1]Department of Pathobiological Sciences, University of Wisconsin School of Veterinary Medicine, Madison, WI 53706, USA
[2]Wisconsin National Primate Research Center, Madison, WI 53715, USA
[3]Department of Biological Sciences, University of South Carolina, Columbia, SC 29208, USA
[4]Division of Virology, Department of Microbiology and Immunology, Institute of Medical Science, University of Tokyo, 4-6-1, Shirokanedai, Minato-ku, Tokyo 108-8639, Japan
*Correspondence: kawaokay@vetmed.wisc.edu (Y.K.), thomasf@primate.wisc.edu (T.C.F.)
http://dx.doi.org/10.1016/j.chom.2016.01.011

SUMMARY

Avian influenza virus reassortants resembling the 1918 human pandemic virus can become transmissible among mammals by acquiring mutations in hemagglutinin (HA) and polymerase. Using the ferret model, we trace the evolutionary pathway by which an avian-like virus evolves the capacity for mammalian replication and airborne transmission. During initial infection, within-host HA diversity increased drastically. Then, airborne transmission fixed two polymerase mutations that do not confer a detectable replication advantage. In later transmissions, selection fixed advantageous HA1 variants. Transmission initially involved a "loose" bottleneck, which became strongly selective after additional HA mutations emerged. The stringency and evolutionary forces governing between-host bottlenecks may therefore change throughout host adaptation. Mutations occurred in multiple combinations in transmitted viruses, suggesting that mammalian transmissibility can evolve through multiple genetic pathways despite phenotypic constraints. Our data provide a glimpse into avian influenza virus adaptation in mammals, with broad implications for surveillance on potentially zoonotic viruses.

INTRODUCTION

In 1918, a novel H1N1 influenza virus, at least partially avian in origin, emerged and killed ~50 million people worldwide (Taubenberger et al., 2001, 2005; Vana and Westover, 2008; Worobey et al., 2014). Influenza viruses can quickly generate within-host genetic diversity due to their high mutation rates, rapid replication, and large population sizes, facilitating adaptation to new environments and host species (Baccam et al., 2006; Drake and Holland, 1999; Lauring and Andino, 2010; Nobusawa and Sato, 2006; Parvin et al., 1986; Sanjuán et al., 2010). Defining how natural selection and genetic bottlenecks shape avian influenza virus adaptation to mammals is critical for understanding how pandemics occur and identifying potentially zoonotic viruses in nature.

Transmission bottlenecks can define subsequent viral evolution. We showed that positive selection for specific hemagglutinin (HA) variants caused a strong bottleneck during evolution of ferret-transmissible reassortant H5N1 viruses (incorporating avian virus HA in a human H1N1 virus backbone) (Wilker et al., 2013). Another study showed that founder effects, random nonselective processes, drive bottlenecks during mammal-adapted seasonal influenza virus transmission in ferrets and guinea pigs (Varble et al., 2014). The evolutionary forces acting as fully avian-derived influenza viruses evolve toward mammalian transmission have not been assessed.

Recently, members of our team (Watanabe et al., 2014) examined the potential for a virus resembling the 1918 strain to re-emerge from wild bird reservoirs by generating an avian-derived reassortant virus with close sequence identity to the 1918 pandemic virus. In ferret passage, this "1918-like" virus acquired substitutions in HA and polymerase associated with altered tissue tropism, enhanced replication, and efficient transmissibility (Watanabe et al., 2014). Here, we use deep sequencing to evaluate the evolutionary processes by which these viruses adapted to replication and transmission in mammals. No new transmission experiments were conducted for the present work, and all ferret nasal swabs were collected prior to the gain-of-function research pause.

RESULTS

Transmissible HA Variants Arise in Ferrets

1918-like avian viral replication was limited in directly inoculated animals, and neither virus nor specific antibodies were detected in contact ferrets (Watanabe et al., 2014). Introducing the mammal-adapting substitution PB2 E627K enhanced replication (Figures S1A, S1B, and S1E), but not transmission. To assess the impact of known mammalian-adapting mutations in HA on transmission of this virus, HA E190D and G225D (H3, mature peptide numbering) were introduced into the 1918-like avian/PB2 E627K virus, generating what we will term the "avian-like" "HA190D225D" virus. Deep sequencing showed that both HA

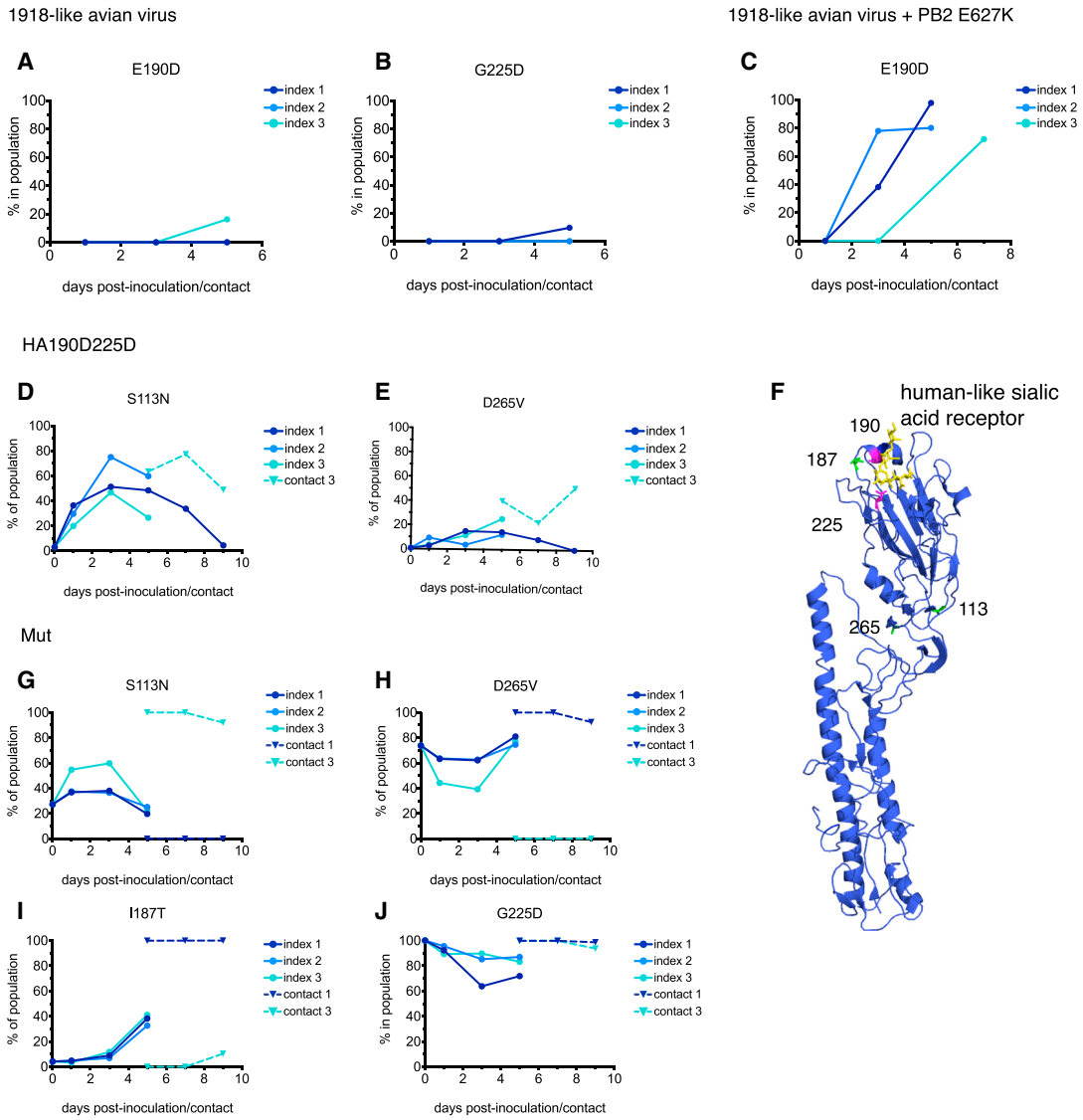


Figure 1. Enumerating HA Mutants in Inoculated and Transmitted Virus Populations

SNPs present in ≥1% of viruses shown for each index (solid traces) and contact animal (dashed traces).

(A and B) HA E190D (A) and G225D (B) became detectable in ferrets inoculated with the 1918-like avian virus.

(C) HA E190D arose independently in all three animals inoculated with 1918-like avian/PB2627K.

(D and E) HA mutations encoding S113N (D) and D265V (E) arose in the HA190D225D virus and were transmitted.

(F) HA amino acids discussed in this study highlighted on the crystal structure of the human 1918 virus A/Brevig Mission/1/1918 (PDB: 2wrg). Green, residues 190 and 225; magenta, residues 113, 187, and 265. A human-type receptor is shown in yellow.

(G–J) Frequencies of HA mutations in Mut virus encoding S113N (G), D265V (H), I187T (I), and G225D (J).

E190D and G225D arose naturally before being introduced: each was detectable in one animal 5 days after inoculation with the 1918-like avian virus (Figures 1A and 1B). HA E190D became dominant in all three animals inoculated with the 1918-like avian/PB2 E627K virus (Figure 1C). The HA190D225D virus was transmitted in 1 of 3 ferret pairs (Figures S1C and S1F). To characterize mutations occurring in HA190D225D, we analyzed RNA isolated from stock viruses and from nasal secretions from all three index animals and the infected contact. We included stock viruses in this analysis both to define "input" viral populations and because even viruses created by reverse genetics can accumulate mutations during passage in culture. We enumerated all SNPs occurring at ≥1% frequency in each virus sequenced; we previously established that this conservative cutoff ensures that only bona fide mutations are considered (Wilker et al., 2013).

Sequencing revealed that two additional HA mutations arose and were transmitted. HA G400A, encoding serine-to-asparagine at amino acid 113 (S113N), was present in 3% of viruses in the HA190D225D stock and rose to 46%–75% in all three index animals by 3 days post-inoculation, declining on day 5 (Figure 1D, solid traces). S113N was detected in 63% of viruses in

the contact soon after transmission (Figure 1D, broken trace). The second mutation, HA A868T, encoded aspartate-to-valine at position 265 (D265V). D265V was not detectable in the virus stock above 1% but increased in frequency in all three index animals and was present in 40% of viruses transmitted to the contact (Figure 1E). Both S113N and D265V are located near the base of the HA globular head (Figure 1F). Cloning and sequencing of full-length HA from day 5 showed that S113N and D265V were linked in a small fraction of viruses (2/20 [10%] clones in index and 1/32 [3.125%] in contact; Figures S2A and S2B).

Viruses were next isolated from the contact animal's nasal wash at day 9 post-contact and amplified on MDCK cells to produce the "Mut" stock virus, which was used to inoculate another three ferret pairs (Watanabe et al., 2014). The Mut virus exhibited similar replication kinetics and pathogenicity to human-derived 1918 viruses and was transmitted in 2/3 pairs (Figures S1D and S1F) (Watanabe et al., 2014). HA S113N and D265V were maintained at frequencies comparable to those in the inoculum (Figures 1G and 1H, solid traces). Another novel HA mutation, HA T634C, arose, encoding isoleucine-to-threonine at position 187 (I187T) near the receptor-binding pocket (Figure 1F). Present at 4% in the Mut stock virus, I187T rose to 32%–41% in all three index animals by day 5 post-inoculation and was fixed after transmission in contact animal 1 (Figure 1I). I187T was not transmitted to contact animal 3 but arose on day 9 post-contact. The rapid increase of I187T in all index animals and its independent re-generation in contact animal 3 strongly suggest that this mutation is advantageous. Indeed, all 17 publicly available human 1918 sequences have 187T, as do 310/338 full-length avian H1N1 HAs (Reid et al., 1999, 2003; Sheng et al., 2011; Taubenberger et al., 1997).

S113N, D265V, and I187T were not all transmitted together. S113N was fixed in contact animal 3 after transmission, while D265V and I187T were undetectable (Figures 1G–1I, light blue broken traces). By contrast, D265V and I187T were fixed after transmission in contact animal 1, and S113N was undetectable (Figures 1G–1I, dark blue broken traces). Cloning and Sanger sequencing showed that S113N and D265V were linked at $\leq$ 33% frequency in Mut index animals 1 and 3 (Figures S3C and S3D); we cannot determine from these data why S113N and D265V were not transmitted together. Our results suggest that in this genetic context, HA 190D and 225D may benefit transmission, but multiple other SNPs may also bestow increased transmission efficiency.

Selection Maintains HA Polymorphism

HA G225D is associated with a switch from α-2,3- to α-2,6-linked sialic acid receptor binding in H1N1 viruses (Tumpey et al., 2007; Lakdawala et al., 2015). Introducing HA190D and 225D into the 1918-like avian virus in the laboratory likely affected the evolutionary pathway these viruses followed, and we cannot know how evolution would have proceeded in their absence. Once introduced, HA 225D remained at or near fixation in all HA190D225D-infected animals (Figure S3A). Although G225D was present in 97% of the Mut stock virus, its frequency decreased to 72%–87% by day 5 post-inoculation in all three index animals (Figure 1J, solid traces). Sequencing at day 5 revealed glycine at position 225 (225G) in 9%–15% of viruses (Figures S3B–S3D). Also, an asparagine (225N) was now encoded by 8%–13% of viruses in index animals 1 and 3 (Figures S3B–S3D). HA 225D was fixed after transmission in both contacts, but its frequency declined to 93% in contact 3 by day 9 (Figure 1J). Indeed, by day 5 viruses encoded a mixture of 225D/225N in contact 1 and 225D/225N/225G in contact 3 (Figures S3B and S3C). Thus, G225D may be favored during transmission between hosts, but polymorphism at 225 seems to be tolerated within a host.

The upper respiratory tract contains a mixture of cell types with different sialic acid receptors (Chandrasekaran et al., 2008; Jayaraman et al., 2012; Matrosovich et al., 2004; Thompson et al., 2006). Various anatomical sites and cell types may exert differing degrees of selective pressure favoring α-2,3- or α-2,6-tropic viruses (Lakdawala et al., 2015). HA 225N was identified in severe human infections during the 2009 H1N1 pandemic (Kong et al., 2014; Mak et al., 2010; Matos-Patrón et al., 2015; Wu et al., 2013). It enables binding to both α-2,6 and α-2,3 receptors and increases binding to non-sialylated glycans (Matos-Patrón et al., 2015). Diversity at position 225 may therefore be maintained due to balancing selection, which preserves multiple beneficial alleles in a population. Alternatively, increasing diversity at site 225 could be due to relaxation of selection, but this hypothesis would not explain the significant increase in nonsynonymous nucleotide diversity we observed in this region, as described below.

The inverse frequencies of S113N and D265V could similarly be explained by balancing selection. S113N was shown to compensate for a loss of HA heat stability due to substitutions E190D and G225D (Watanabe et al., 2014), a phenotype associated with transmissibility in other influenza viruses (de Vries et al., 2014; Imai et al., 2012; Linster et al., 2014). D265V is located proximal to S113N (Figure 1F) and may serve a similar function in modulating heat stability. We hypothesize that during the early stages of mammalian adaptation, selective pressures (e.g., diverse receptor profiles, temperatures, and host factors) may promote accumulation and maintenance of a large number of mutations in HA that later can be selected by the new host environment (Lauring et al., 2012; Parrish et al., 2008).

Low-Frequency Transmissible Polymerase Variants

Viral genome analysis revealed that two novel mutations arose in the polymerase complex. In one of the animals inoculated with HA190D225D, a G-to-A mutation occurred at PA nucleotide 781, encoding valine-to-methionine at amino acid 253 (V253M). PA V253M was present in only 9% of viruses on day 5 in the index animal but was detected at 73% in the contact shortly after transmission (Figure 2A).

A C-to-A mutation at PB2 nucleotide 2,078 encoded alanine-to-aspartate at amino acid 684 (PB2 A684D). PB2 A684D was never detectable in the transmitting index animal (Figure 2B) but was present in 75% of the contact's viruses shortly after transmission (Figure 2B). We cannot exclude the possibility that PB2 A684D was transmitted at very low frequency, but these data suggest that it likely arose de novo in the contact animal.

PA V253M and PB2 A684D were maintained at high frequencies in the Mut virus (Figures 2C and 2D). We also identified a mutation encoding threonine-to-isoleucine at position 232 of the nucleoprotein (NP) in contact animal 3 (Figure 2E, broken

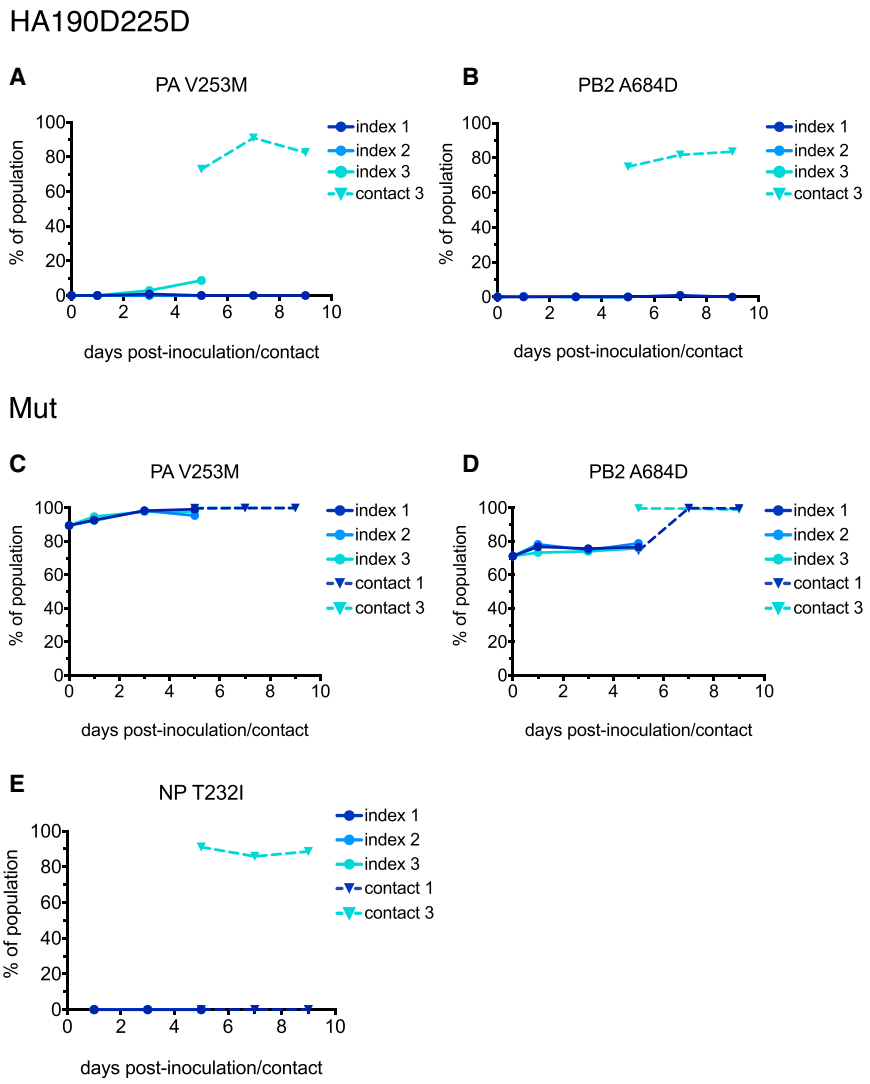


Figure 2. Enumerating Polymerase Variants in Inoculated and Transmitted Viruses

SNPs present in ≥1% of viruses shown for each index (solid traces) and contact animal (dashed traces).

(A and B) Frequencies of mutations in HA190D225D viruses encoding PA V253M (A) and PB2 A684D (B) are shown.

(C and D) Frequencies of these same mutations in Mut viruses are shown: PA V253M (C), PB2 A684D (D).

(E) A mutation encoding NP T252I became detectable at day 5 post-contact in contact animal 3.

trace). This mutation was not detectable in index animal 3 before transmission, indicating that it likely arose early after transmission. Interestingly, neither PA V253M nor PB2 A684D were found to enhance polymerase activity when introduced into the 1918-like avian virus backbone (with or without PB2 E627K) either separately or together (Watanabe et al., 2014). Thus, airborne transmission can greatly alter the viral population, allowing even minor variants to become dominant in subsequent generations despite lacking a measurable fitness benefit. Further, new variants can rapidly emerge de novo after transmission.

HA Nucleotide Diversity Increases during Replication

To more directly measure selection during mammalian adaptation, we calculated π, the nucleotide diversity of viral populations in all index animals (Nelson et al., 2015). π quantifies the average number of pairwise differences per nucleotide site among a set of sequences. Comparing the frequencies of synonymous (silent) and nonsynonymous (amino-acid-changing) polymorphisms, denoted πS and πN, respectively, allows us to evaluate the impact of natural selection. πS > πN indicates that, on average, purifying selection is acting to remove deleterious mutations, while πN > πS indicates that diversifying selection is favoring new mutations. Effectively equal levels of πN and πS suggest that allele frequencies are determined primarily by genetic drift.

The influenza genome in index animals infected with the HA190D225D and Mut viruses is generally under purifying selection: πS was significantly greater than πN for almost every viral gene, except for NS2/NEP in the Mut index animals ($p < 0.01$ or $p < 0.001$, Tables 1 and S2). Strikingly, the exception to this pattern was the 344-codon HA1 region of HA. πN/πS ratios for HA1 were >1 at every time point in each index animal (range = 1.128–3.077, Table S2), likely primarily due to high-frequency SNPs in codons 113, 187, 225, and 265, although there were also 6–7 nonsynonymous HA1 SNPs present between 1%–13% (Table S3). Average HA1 πN/πS > 1 in HA190D225D and Mut index animals when all time points are combined (mean πN/πS = 1.718 and 2.007, respectively; $p < 0.001$, Table 1). These data provide further indication that diversification in HA1 is strongly favored in index animals.

Sliding Window Analysis Reveals Selection Acting at Specific Sites in HA and Polymerase

To assess whether selection was promoting polymorphism across the entire HA gene, or acting at a few discrete sites, we re-calculated πN and πS using a sliding window of 150 nucleotides (50 codons) and a step size of 3 nucleotides (1 codon). In all index animals, HA1 had several prominent πN peaks (Figures 3A, 3B, and S4A–S4C) near S113N, I187T, G225D, and D265V (Figures 3C and 3D). Similar peaks were also apparent in several viruses at PA V253M and PB2 A684D (Figures 3A, 3B, and S5A–S5C). These data suggest that selection favors accumulation of mutations at a few specific sites during avian virus replication in mammals.

πN peaks also occurred in HA2 in HA190D225D index animals (Figures 3A and 3C). These peaks result from low-frequency nonsynonymous SNPs (between 1% and 10% in the population)

Table 1. Mean Synonymous and Nonsynonymous Nucleotide Diversity ± SE in Influenza Genes

Group	Index/Contact		Gene										
			PB2	PB1	PA	HA1	HA2	NP	NA	M1	M2	NS1	NS2/NEP
HA190D225D	Index	π_N	0.00100 ± 0.00008	0.00099 ± 0.00010	0.00092 ± 0.00006	0.00196 ± 0.00011	0.00171 ± 0.00010	0.00093 ± 0.00004	0.00109 ± 0.00005	0.00085 ± 0.00005	0.00059 ± 0.00005	0.00094 ± 0.00008	0.00055 ± 0.00005
		π_S	0.00143 ± 0.00013	0.00142 ± 0.00013	0.00147 ± 0.00012	0.00127 ± 0.00009	0.00139 ± 0.00011	0.00154 ± 0.00010	0.00156 ± 0.00012	0.00198 ± 0.00013	0.00351 ± 0.00034	0.00140 ± 0.00019	0.00093 ± 0.00012
		π_N/π_S	0.699***	0.697***	0.626***	1.543***[a]	1.230**[a]	0.604***	0.699***	0.429***	0.168***	0.671*	0.591***
	Contact	π_N	0.00118 ± 0.00029	0.00111 ± 0.00003	0.00116 ± 0.00020	0.00213 ± 0.00027	0.00093 ± 0.00019	0.00096 ± 0.00020	0.00108 ± 0.00026	0.00087 ± 0.00019	0.00063 ± 0.00017	0.00081 ± 0.00023	0.00053 ± 0.00014
		π_S	0.00141 ± 0.00044	0.00194 ± 0.00009	0.00146 ± 0.00036	0.00124 ± 0.00032	0.00163 ± 0.00040	0.00168 ± 0.00025	0.00137 ± 0.00041	0.00112 ± 0.00040	0.00116 ± 0.00024	0.00123 ± 0.00036	0.00103 ± 0.00039
		π_N/π_S	0.837	0.572**	0.794	1.718[a]	0.571	0.571**	0.788	0.777	0.543*	0.659	0.515
Mut	Index	π_N	0.00098 ± 0.00004	0.00078 ± 0.00008	0.00081 ± 0.00007	0.00277 ± 0.00020	0.00075 ± 0.00008	0.00075 ± 0.00005	0.00088 ± 0.00008	0.00072 ± 0.00005	0.00063 ± 0.00006	0.00079 ± 0.00005	0.00065 ± 0.00012
		π_S	0.00132 ± 0.00008	0.00139 ± 0.00010	0.00139 ± 0.00013	0.00138 ± 0.00011	0.00170 ± 0.00013	0.00148 ± 0.00010	0.00161 ± 0.00017	0.00139 ± 0.00009	0.00092 ± 0.00010	0.00156 ± 0.00021	0.00085 ± 0.00013
		π_N/π_S	0.742***	0.561***	0.583***	2.007***[a]	0.441***	0.507***	0.547**	0.518***	0.685**	0.506**	0.765
	Contact	π_N	0.00079 ± 0.00015	0.00076 ± 0.00010	0.00081 ± 0.00015	0.00093 ± 0.00026	0.00076 ± 0.00012	0.00080 ± 0.00008	0.00065 ± 0.00012	0.00063 ± 0.00007	0.00036 ± 0.00005	0.00072 ± 0.00009	0.00042 ± 0.00005
		π_S	0.00103 ± 0.00019	0.00133 ± 0.00016	0.00131 ± 0.00024	0.00105 ± 0.00016	0.00132 ± 0.00010	0.00126 ± 0.00016	0.00113 ± 0.00020	0.00101 ± 0.00012	0.00073 ± 0.00005	0.00108 ± 0.00013	0.00071 ± 0.00005
		π_N/π_S	0.767	0.571**	0.618**	0.886	0.576**	0.635**	0.575**	0.624**	0.493***	0.667**	0.592**

Mean πN (synonymous) and πS (nonsynonymous) ± SE are shown for each viral gene in each group. All time points for each group were combined. Results of paired t tests of the null hypothesis that $\pi_S = \pi_N$: *p < 0.05, **p < 0.01, ***p < 0.001.

[a]Genes for which π_N is significantly greater than π_S.

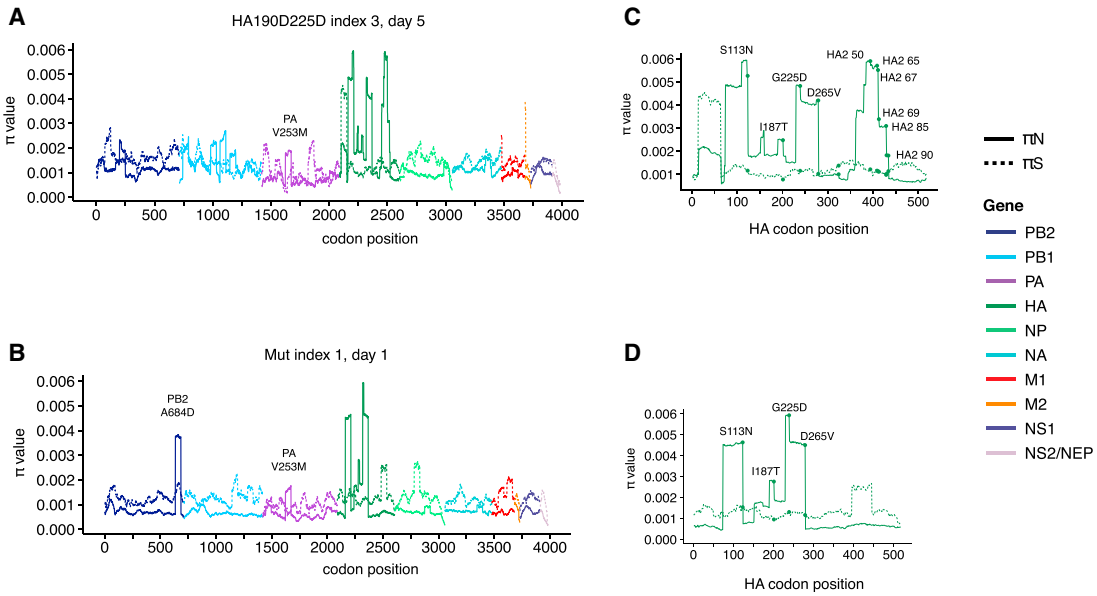


Figure 3. Sliding Window Analysis to Detect Sites Likely under Positive Selection

πN (solid lines) and πS (dashed lines) were calculated for the genome using a 50-codon window and 1-codon step. Gene segments are shown as a contiguous "coding region" and denoted in different colors. Two example time points shown; see Figure S3 for more information.

(A) On day 5 in HA190D225D index animal 3, the largest πN peaks are in HA.

(B) On day 1 in Mut index animal 1, distinct πN peaks appear in PB2 (corresponding to PB2 A684D), PA (corresponding to PA V263M), and several sites in HA.

(C and D) Close views of HA in HA190D225D index animal 3, day 5 (C) and index animal 1, day 1 (D) show πN peaks near indicated sites in HA1 and/or HA2.

at HA2 amino acid positions 50, 65, 67, 69, 85, and 90, corresponding to elevated πN/πS (range = 1.237–2.400) at several time points (Table S2). This signal disappears after the first transmission event. We speculate that diversification may be favored in the HA190D225D HA stalk domain as a means of modulating stability. Once additional HA mutations accumulate, as in the Mut virus, diversification may no longer be favored. The functional impact of these HA2 mutations is not clear, but these data further support the hypothesis that the early stages of mammalian adaptation are marked by HA diversification.

Haplotype Diversity Increases during Viral Replication

Our analyses thus far have considered individual mutations in isolation, but complex traits like receptor binding are likely to be conferred by combinations of mutations. We use the term "haplotype" to denote a set of mutations that co-occur on the same DNA molecule. If a given set of mutations is beneficial, we would expect the frequency of haplotypes containing those variants to increase. To test whether selection favored specific HA haplotypes, we used LinkGE, a computational tool we developed and validated previously (Wilker et al., 2013), which enumerates combinations of nucleotides at specified sites in deep sequencing data. We assessed all polymorphic sites in HA1 detected more than once, resulting in a final query of 12 sites (see Supplemental Experimental Procedures and Tables S4 and S5).

Haplotype diversity expanded dramatically during viral replication. Two haplotypes were present during initial HA190D225D infection in the index animal, but five were present by day 5 (Figure 4A). Similarly, 3–5 HA haplotypes were present in the Mut virus on day 1, rising to 9–10 on day 3, and 12–14 on day 5 (Figures 4B and 4C). By day 5, no HA haplotype was present above 30% in index animals. In agreement, Shannon entropy in HA dramatically increased over time in index animals (Figures 4E and 4F).

HA190D225D Virus Transmission Involves a Loose Bottleneck

The number and diversity of viruses founding infection provide the starting point for further within-host evolution. Transmission of only a few virions would result in a founder effect, with reduced diversity across the entire viral genome; this effect does not necessarily involve natural selection. Selection favoring transmission of specific variants might be observed as a reduction in diversity localized to the selected region or gene segment. Fixation of advantageous alleles by selection and the resulting reduction in linked nucleotide diversity is called a selective sweep (Nielsen, 2005). Strong selection during transmission could favor the emergence of mammalian-transmissible viruses, while strong founder effects might impose a barrier to host switching. To differentiate among these processes, we calculated πN and πS across the viral genome at time points surrounding transmission. As we cannot know the exact day on which contact animals became infected, we evaluated viruses in the contact animal at the first time point at which virus was detectable and the corresponding time point in the index animal. To ensure that this choice did not significantly impact our results, we also performed these analyses using the index animal day 3 viruses.

Transmission of HA190D225D resulted in a slight reduction in both πN and πS across most of the genome (Figures 5A, S5A, and S5D) that was significant overall ($F_{10,\ 22}$ = 2.50, p = 0.035 for πN; $F_{10,\ 22}$ = 3.32, p = 0.009 for πS). The mean difference in πN between index and contact animals was significantly greater

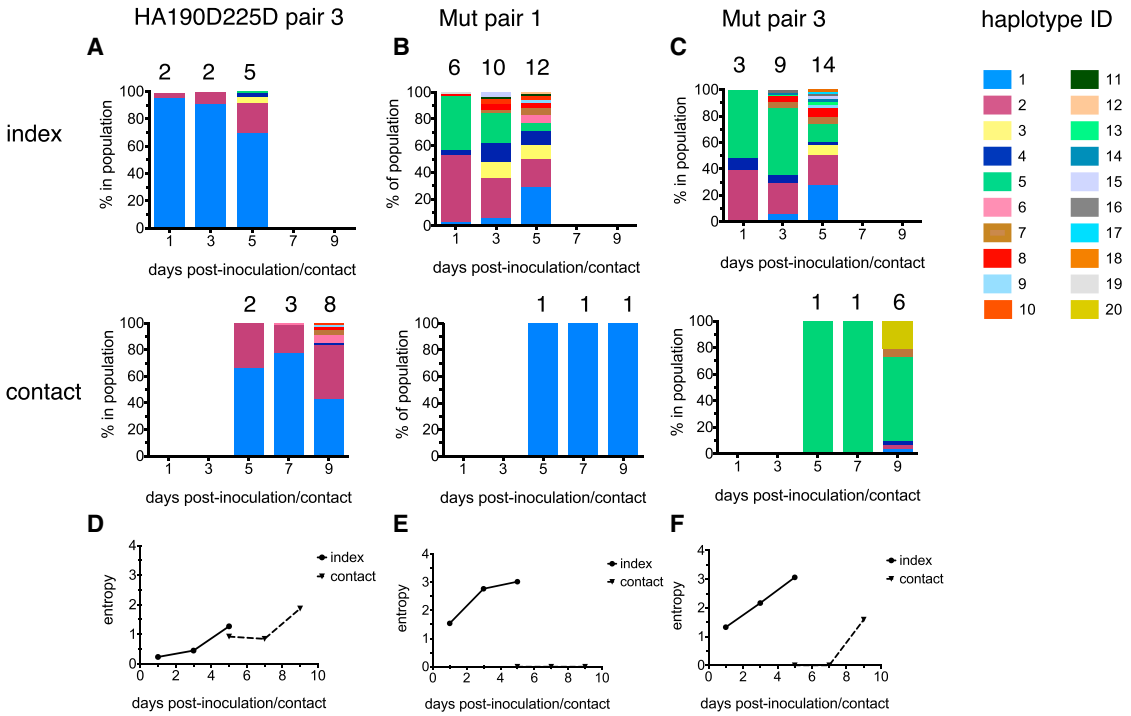


Figure 4. HA Genetic Diversity Assessed by Haplotype Analysis

Haplotypes were defined using LinkGE to query 12 nucleotide positions corresponding to amino acids 113, 145, 159, 187, 189, 190, 192, 225, and 265. Each color represents a unique, arbitrarily numbered haplotype. Numbers above each bar indicate number of haplotypes detected in that time point. Due to limits on read length, sites were queried with two overlapping windows. Window 1 data are shown; see Tables S4 and S5 for more information.

(A) Haplotypes detected in HA190D225D animals.

(B and C) Haplotypes detected in Mut pairs 1 (B) and 3 (C).

(D–F) Haplotype diversity was quantified using Shannon entropy as described in the Supplemental Experimental Procedures for HA190D225D pair 3 (D), Mut pair 1 (E), and Mut pair 3 (F). Solid lines, index animals; dashed lines, contact animals.

for HA2 than for all other genes (Dunnett's test, 5% family error rate; Figure 5B), most likely due to the loss of a large number of low-frequency SNPs in HA2 present in the index animal before transmission. Selection may favor diversification in HA2 during viral replication, but the loss of low-frequency SNPs suggests that they are not favored for transmission.

The two most prevalent HA haplotypes in the HA190D225D index animal were present after transmission in the contact (Figure 4A), and there was only a minimal reduction in haplotype diversity (entropy = 1.27 versus 0.922, Figure 4D). These data together suggest that HA190D225D transmission involved a loose genetic bottleneck. This could indicate that infection in the contact was founded by a large viral inoculum. The relatively low transmission efficiency of HA190D225D may also suggest that no one variant was most fit for mammalian transmission and that a larger, heterogeneous population of "less fit" viruses may have been required to found airborne infection.

Patterns of selection in the HA190D225D contact animal were similar to those observed in the index. $\pi N/\pi S$ ratios in the contact remained <1 in almost all genes except HA1, which retained its signature of diversifying selection throughout infection (Table S2). However, differences in mean πN and πS values did not attain statistical significance except in PB1, NP, and M2, where $\pi S > \pi N$, perhaps due to low sample numbers (Table 1). Consistent with these findings, HA entropy increased over time in the contact, surpassing haplotype diversity levels present in the index before transmission (Figure 4D). Together, these data show that much of the genetic diversity in the HA190D225D virus was maintained through transmission and that selection continued to promote HA1 diversification afterward.

Mut Virus Transmission Involves Selective Sweeps on HA1

During the next set of experiments, transmission frequency increased to 2/3 ferret pairs. Although these sample sizes are low, this suggests that the Mut virus may have evolved increased fitness for mammalian infection. Patterns of viral sequence diversity were strikingly different during Mut transmission. In Mut pair 1, πN and πS were higher or only slightly reduced across the majority of the genome after transmission (Figures 5C, S5B, and S5E). In contrast, HA1 experienced a reduction in πN that was significantly larger than the reduction in the rest of the genome; mean πS changes were not significantly different between index and contact animals ($F_{10, 22} = 0.27$; $p = 0.982$; Figure 5D). Accordingly, we detected only one HA haplotype shortly after transmission (Figure 4B), and entropy was drastically reduced (Figure 4E).

In Mut pair 3, unlike pair 1, transmission reduced diversity in multiple gene segments. However, similar to Mut pair 1, the most striking reduction occurred in HA1 (Figures 5E, S5C, and S5F), in which πN was 86% lower after transmission. Mean πN

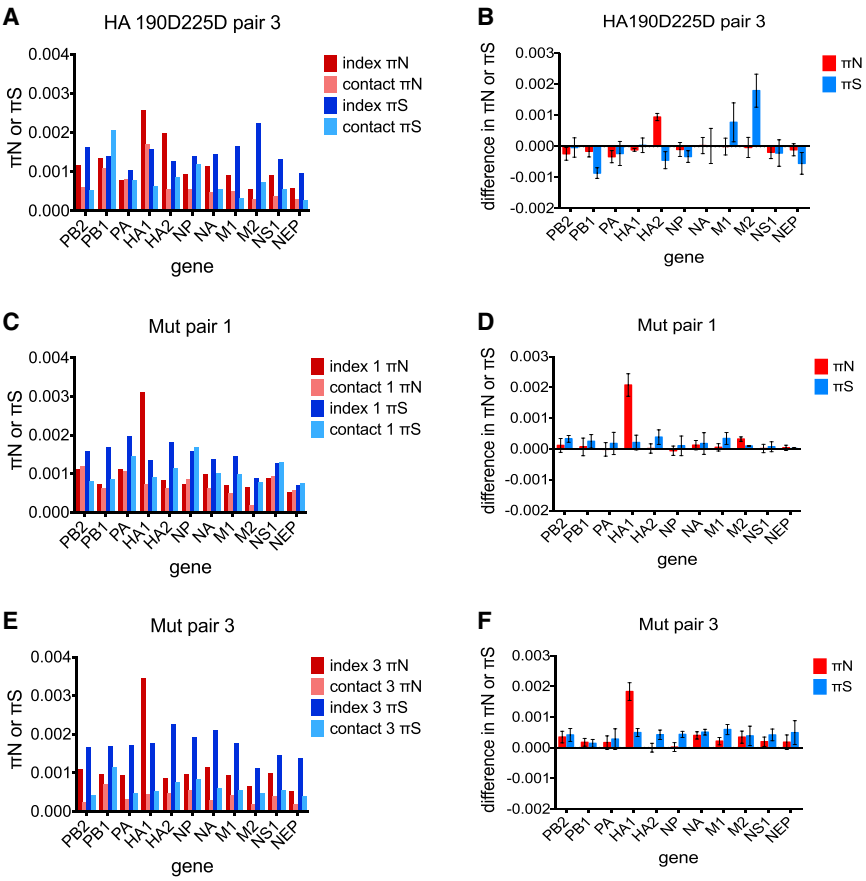


Figure 5. Characterizing Transmission Bottlenecks

(A–F) πN and πS were calculated for each viral gene in each animal in the transmitting pairs; values at day 5 after infection/contact are shown for HA190D225D pair 3 (A). Differences in πN and πS between index and contact animals were compared using pooled time points; mean differences ± SEM are shown. See the Supplemental Experimental Procedures for more information. Differences for HA190D225D pair 3 are shown in (B). πN and πS values at day 5 for Mut pair 1 (C) and differences in πN and πS between index and contact animals for Mut pair 1 (D). πN and πS values at day 5 for Mut pair 1 (E) and differences in πN and πS between index and contact animals for Mut pair 1 (F).

difference was significantly greater in HA1 than in other genes ($F_{10, 22} = 7.56$; $p < 0.001$; Dunnett's test, 0.1% level; Figure 5F). Again, there were no significant differences in the change in πS across genes ($F_{10, 22} = 0.29$; $p = 0.976$; Figure 5F). As in pair 1, only a single HA1 haplotype was detectable immediately after transmission, and entropy dropped from 3.06 to 0 (Figure 4F). Notably, different haplotypes were transmitted in pairs 1 and 3: HA D265V and I187T were fixed in pair 1 while HA S113N became fixed in pair 3 (Figures 4B and 4C). The transmitted haplotype in pair 1 was present in only 5% of the viral population in the index animal on day 3 (Figure 4B), underscoring our previous finding (Wilker et al., 2013) that variants present at very low frequencies in a source host can be transmitted by respiratory droplets.

Analysis of πN and πS in viruses infecting Mut contact animals revealed that all the gene segments, including HA1, are under purifying selection after transmission. One exception was in pair 1, where πN/πS was 1.471 in PB2 on day 5, after which the entire genome remained under purifying selection for the duration of infection. In pair 3, a signal of diversifying selection is only detectable on day 5 in HA1 (πN/πS = 1.520, Table S2), the effects of which are apparent in an expansion of HA haplotypes and increase in entropy on day 9 (Figures 4C–4F).

DISCUSSION

Elucidating the evolutionary processes that allow certain viruses to cross species barriers is critical for evaluating the pandemic potential of avian influenza viruses. It is becoming increasingly clear that transmission bottlenecks can have a strong impact on influenza virus evolution. One recent study showed that bottleneck stringency was primarily determined by transmission route and that droplet transmission was characterized by non-selective founder effects (Varble et al., 2014). That study used a human-adapted virus containing synonymous barcodes, minimizing the potential role of selection. In contrast, H5N1 virus transmission between ferrets involved a selective bottleneck favoring mutations in HA1 (Wilker et al., 2013). Transmission of H7N9 viruses involved tight, attenuating bottlenecks, although the contributions of founder effects versus selection were unknown (Zaraket et al., 2015).

Replication and transmission of the HA190D225D and Mut viruses in this study can be conceptualized as modeling the progressive adaptation of an avian-like H1N1 virus in mammals. Here we observed two different types of bottlenecks. Much of the genetic variation in HA190D225D was transmitted, and selection continued to favor HA diversification in the contact animal. In contrast, Mut virus transmission involved reductions in genetic diversity that were sharpest in HA1. Patterns of diversity in both Mut pairs suggest that selection for HA1 variants occurred during transmission; the genome-wide reduction in diversity in Mut pair 3 may be a consequence of the combined action of selection on HA1 and a smaller infecting dose than in pair 1. Importantly, the data we observe in the Mut pairs cannot be explained by a founder effect. The increase or maintenance of diversity across the rest of the genome strongly suggests that multiple virions were transmitted, each of which harbored a single HA haplotype packaged with different variants of the other 7 gene segments. Our data therefore indicate that Mut virus transmission involved a selective sweep acting on HA1. We postulate that founder effects could have occurred in this study but that selection for HA1 variants had a greater impact on genetic diversity.

Successful replication and transmission in a novel host is likely accompanied by selection on processes including receptor binding, membrane fusion, interaction with host proteins, escape from restriction factors, and modulation of immune responses. We have shown that selection reduces HA nucleotide diversity during H5N1 reassortant virus transmission in ferrets (Wilker et al., 2013), similar to what we observe during transmission of the Mut virus. Although the H5N1 viruses we studied previously had an avian HA, the rest of their genome was derived from a 2009 H1N1 human pandemic isolate, limiting our ability to assess selection acting outside of HA. The present work afforded the opportunity to characterize selection across the entire avian-like influenza genome. An important caveat is that three amino acid substitutions were artificially introduced into the transmissible HA190225D virus. Although both HA E190D and G225D arose naturally before their introduction, PB2 E627K did not, and it is unknown whether these mutations would have become fixed had evolution proceeded organically. The timing of mutation emergence, the identity of adaptive substitutions, and the number of rounds of replication required may have been different if the evolutionary pathway of this particular virus had been completely unaltered. Despite these caveats, we find that selection is primarily confined to HA1, where diversification is strongest at a few sites encoding amino acids in the globular head (113 and 265) and near the sialic acid binding pocket (187, 190, and 225).

The lack of positive selection in the rest of the genome suggests that once the virus is able to replicate efficiently, e.g., once PB2 E627K is introduced, receptor binding and host cell entry may be the most substantial barriers to efficient mammalian transmission. Mammalian adaptation is likely a multi-factorial process and may proceed along different pathways in different genetic backgrounds, but the early, independent generation of HA E190D and G225D and the strong diversifying selection on HA1 observed here are consistent with this hypothesis. Selective sweeps on attachment proteins have been observed in transmission of other viruses, including hepatitis C virus (Brown et al., 2012) and HIV (Edo-Matas et al., 2012); positive selection on the SARS coronavirus surface glycoprotein may facilitate transmission from palm civets to humans (Chinese SARS Molecular Epidemiology Consortium, 2004; Song et al., 2005; Zhang et al., 2006). Diversification of receptor binding proteins followed by fixation of advantageous variants may be a common pathway by which viruses evolve transmissibility in new hosts.

Two polymerase mutations arose and became fixed during our study. PB2 A684D was first detectable immediately after transmission (Figure 2B), suggesting that it arose de novo in the contact animal. PA V253M, present in only 9% of viruses in the index animal on day 5, was at 73% in the contact immediately after transmission (Figure 2A). We cannot conclusively determine whether PA V253M was transmitted as a minor variant or arose de novo afterward. What is clear is that these two polymerase mutations drastically changed in frequency during the first transmission event and became fixed for the rest of the study, even though previous functional analyses suggested that these substitutions reduce the processivity of the 1918-like avian polymerase (without PB2 E627K) in vitro (Watanabe et al., 2014). In the presence of PB2 E627K, these mutations did not alter polymerase activity, suggesting that PB2 E627K was the primary determinant of improved viral replication. Viruses reached similar titer in vivo whether PB2 A684D and PA V253M were at low frequency (HA190D225D index animals) or nearly fixed (Mut index animals; Figure S1E). These mutations could confer an advantage that we did not directly test (e.g., RNA stability, RNA folding, codon bias). Given our phenotypic data and the known potential for bottlenecks to alter diversity in populations without selection (Hamilton et al., 2012; Varble et al., 2014), perhaps these mutations hitchhiked to high frequency during transmission. These data underscore that fixation of nonsynonymous variants after transmission is not in itself evidence that they are advantageous and agree with a recent study showing that tight bottlenecks during H7N9 virus transmission allowed deleterious variants to remain at high frequency (Zaraket et al., 2015). Together, our data suggest that for an avian virus to cross species barriers, overcoming the stochastic fluctuation of mutations during transmission bottlenecks may be just as essential as generating a constellation of mammalian-adapting mutations. Fixation or maintenance of non- or maladaptive mutations during airborne transmission may be a previously unappreciated barrier to influenza cross-species transmission.

Although HA S113N, I187T, and D265V are all transmitted, there is no single viral genotype that is shared among all three transmitting pairs. Therefore, in the presence of HA 190D and 225D, there are multiple available evolutionary pathways to mammalian transmission. For any virus lineage, evolving human transmissibility may not require convergence on one specific set of amino acid substitutions. We propose that, because within-host selection appears to maintain polymorphism at a few potentially adaptive sites in HA, the viral population is able to simultaneously explore multiple evolutionary pathways toward a transmissible phenotype. Transmission then selects for beneficial variants, allowing different variants to fix in different hosts.

Our study is limited to a small number of transmission events with one avian-like H1N1 virus lineage evolving in ferrets. Synthesizing our data and previous studies, we offer the following conceptual framework to explain our observations, acknowledging that expanded studies are necessary to confirm its broader application. We propose that the stringency of transmission bottlenecks and the main evolutionary forces driving them may change throughout the host adaptation process. We speculate that at earlier stages of host adaptation, there may not be one "most fit" virus that is able to outcompete the rest of the population after transmission to localized regions of the respiratory tract (Karlsson et al., 2015; Lakdawala et al., 2015; Matrosovich et al., 2004). A non-mutually exclusive hypothesis is that a larger inoculum may be required to found infection in the absence of a single highly fit variant. These hypotheses could explain the surprising finding that the first transmission bottleneck was loose, even though relatively few "avian-like" viruses would be expected to be fit for transmission in mammals. Indeed, we speculate that viruses that are well adapted for transmission in a particular host species would again exhibit loose transmission bottlenecks, as most variants would be of similar transmission fitness; the recent observation of loose genetic bottlenecks in transmission of seasonal influenza viruses among human household contacts supports this hypothesis (Poon et al., 2016).

Acquisition of advantageous alleles during the course of initial infection may generate a highly fit variant upon which selection can act during transmission. A tight bottleneck could result from numerous mechanisms, including a small founding population or the ability of a particularly fit variant to outcompete other transmitted virions very early after infection. Initial localized infection by one or a few variants in the contact animal (Karlsson et al., 2015) could also allow for the early dominance of a particular HA variant that then spreads through the upper respiratory tract. Airborne transmission involves numerous steps that impact the makeup of the viral population in the contact: to found infection, virions expelled by the source host must remain viable in a droplet, be inhaled in the contact, and then successfully replicate. Selection could act on any or all of these stages to influence the bottleneck we observe. Should these hypotheses be supported by future investigations, then defining within-host selection and transmission bottlenecks occurring in a spillover infection could be a way to assess the pandemic potential of emerging viruses: tight bottlenecks that include mammal-adapting variants would be expected to characterize viruses undergoing a transition to efficient transmissibility. It is also plausible that the size and stringency of influenza transmission bottlenecks vary widely due to chance alone, rather than progressive host adaptation. Multiple productive exposures, different routes of transmission, and/or differing infectious doses could have this effect without involving selection. We find this scenario less likely for our data given the signatures of selection in both contact animals in the second transmission event, but it too remains a possibility that merits further exploration.

EXPERIMENTAL PROCEDURES

Transmission Experiments

No new transmission experiments were performed for this study. All samples from ferrets infected with 1918-like avian influenza viruses were generated prior to the gain-of-function research pause. This study used viral RNA isolated from nasal wash samples collected from index ferrets on days 1, 3, and 5 post-inoculation and from contact ferrets on days 5, 7, and 9 post-contact as part of a previously described study (Watanabe et al., 2014). See Supplemental Experimental Procedures for more details.

Sequence Data Generation

Viral RNA was extracted from nasal wash samples using the RNeasy mini kit and RNase-free DNase set (QIAGEN). RNA was reverse transcribed using the one-step SuperScript III one-step RT-PCR kit (Invitrogen) and amplified by segment-specific PCR (primers listed in Table S1) using Phusion high-fidelity DNA polymerase (New England Biolabs). PCR products were purified from a 1% agarose gel using the QIAquick Gel Extraction Kit (QIAGEN) and eluted in water. Cleaned PCR product was quantified with the Qubit dsDNA HS Assay Kit (Invitrogen) and diluted in DEPC-treated water to a concentration of 0.2 ng/μl, reverse transcribed, amplified, and prepared for sequencing using the Nextera XT Sample Preparation kit. Samples were processed on the Illumina MiSeq with a 500-cycle kit as an 8 pM library with 1% PhiX control.

Sequence Data Analysis

Fastq files were imported into CLC Genomics Workbench Version 6 (CLC Bio) for analysis. Reads were trimmed using a quality score threshold of Q30 and a minimum length of 100 bases. All reads were mapped to a reference sequence comprised of the cDNA plasmid sequences from the original 1918-like avian stock virus. Variant calling was performed using a frequency threshold of 1% and required a minimum coverage of 100 reads and a central base quality score of Q30 or higher. The average sequence depth acquired for each gene in this dataset was 7,680 ± 6,074.

Diversity Calculation

The π statistic for measuring nucleotide diversity was calculated using two different methods. The first, as displayed in the manuscript (Figures 3 and 5 and Tables 1 and S2), calculated the synonymous (πS) and nonsynonymous (πN) nucleotide diversity using SNPGenie (Nelson and Hughes, 2015), which adapts Nei and Gojobori's (1986) method of estimating synonymous and nonsynonymous substitutions for next-generation sequencing data (Bailey et al., 2014; Nei and Gojobori, 1986; Wilker et al., 2013). πN and πS were also independently estimated using PoPoolation version 1.2.2 for comparison (Kofler et al., 2011), which is shown in Table S2 and Figures S5A–S5C. Statistical analyses were performed in Minitab or Prism 6.0b for Mac (GraphPad, www.graphpad.com).

Haplotype Diversity and Shannon Entropy

Haplotypes were enumerated using the LinkGE software as previously described (Wilker et al., 2013). We queried a total of 12 nucleotide positions (400, 507, 550, 634, 640, 644, 648, 653, 747, 748, 749, and 868), which correspond to 9 amino acid positions (113, 145, 159, 187, 189, 190, 192, 225, and 265). We used our haplotype data to calculate a measure of Shannon entropy as previously described (Rogers et al., 2015). Entropy, H, for a given time, t, and a given segment, s, was calculated as:

$$H(t,s) = -\sum_{i=1}^{n} f_i(t,s)\log_2[f_i(t,s)],$$

where n is the total number of haplotypes for time t for segment s. f_i is the frequency for haplotype i at time t. This analysis was performed only for HA in the transmitting pairs.

ACCESSION NUMBERS

The accession number for the sequences reported in this paper is Sequence Read Archive (SRA): SRP068956.

SUPPLEMENTAL INFORMATION

Supplemental Information includes Supplemental Experimental Procedures, five figures, and five tables and can be found with this article online at http://dx.doi.org/10.1016/j.chom.2016.01.011.

AUTHOR CONTRIBUTIONS

L.H.M. performed research, analyzed and interpreted results, and wrote the manuscript. G.Z., T.W., and Y.K. performed previous ferret studies and provided RNA. G.Z. and J.M. performed research. G.Z., C.W.N., and A.L.H. analyzed data and interpreted results. J.M.D. interpreted results. Y.K. directed the study and interpreted results. T.C.F. directed the study, analyzed and interpreted results, and wrote the manuscript.

ACKNOWLEDGMENTS

We dedicate this work to the memory of Dr. Austin L. Hughes, our colleague, collaborator, and mentor, who passed away suddenly as this paper was in revision. Dr. Hughes was an outstanding scientist and generous collaborator. He made critical contributions to this work and will be deeply missed.

This work was supported by NIH grants AI084787 to T.C.F., AI077376 to A.L.H., and a supplement to NIH grant OD011106 to T.C.F. L.H.M. was supported in part by T32 GM07215. J.M.D. was supported by NSF Graduate Research Fellowship DGE-0718123, a Diversity Supplement to NIH R01 AI069274 to Y.K., and by a Ford Foundation Dissertator Fellowship. C.W.N. received support from a Presidential Fellowship and a Dept. of Biological Sciences Kathryn Hinnant-Johnson, M.D. Memorial Fellowship from the University of South Carolina and from NSF Graduate Research Fellowship DGE-0929297. Y.K. gratefully acknowledges support from a grant-in-aid for Specially Promoted Research and a contract research fund for the Program for Funding Research Centers for Emerging and Reemerging Infectious Diseases from the Ministries of Education, Culture, Sports, Science, and Technology; from grants-in-aid of Health, Labor, and Welfare of Japan; and from NIAID

grants. Samples from ferrets infected with 1918-like avian influenza viruses were generated prior to the gain-of-function research pause with support from NIH grant AI080598 (Molecular Mechanisms for the High Pathogenicity of 1918 Influenza Virus) to Y.K. We thank Bret Payseur for intellectual guidance.

Received: July 6, 2015
Revised: October 26, 2015
Accepted: January 25, 2016
Published: February 10, 2016

REFERENCES

Baccam, P., Beauchemin, C., Macken, C.A., Hayden, F.G., and Perelson, A.S. (2006). Kinetics of influenza A virus infection in humans. J. Virol. *80*, 7590–7599.

Bailey, A.L., Lauck, M., Weiler, A., Sibley, S.D., Dinis, J.M., Bergman, Z., Nelson, C.W., Correll, M., Gleicher, M., Hyeroba, D., et al. (2014). High genetic diversity and adaptive potential of two simian hemorrhagic fever viruses in a wild primate population. PLoS ONE *9*, e90714.

Brown, R.J., Hudson, N., Wilson, G., Rehman, S.U., Jabbari, S., Hu, K., Tarr, A.W., Borrow, P., Joyce, M., Lewis, J., et al. (2012). Hepatitis C virus envelope glycoprotein fitness defines virus population composition following transmission to a new host. J. Virol. *86*, 11956–11966.

Chandrasekaran, A., Srinivasan, A., Raman, R., Viswanathan, K., Raguram, S., Tumpey, T.M., Sasisekharan, V., and Sasisekharan, R. (2008). Glycan topology determines human adaptation of avian H5N1 virus hemagglutinin. Nat. Biotechnol. *26*, 107–113.

Chinese SARS Molecular Epidemiology Consortium (2004). Molecular evolution of the SARS coronavirus during the course of the SARS epidemic in China. Science *303*, 1666–1669.

de Vries, R.P., Zhu, X., McBride, R., Rigter, A., Hanson, A., Zhong, G., Hatta, M., Xu, R., Yu, W., Kawaoka, Y., et al. (2014). Hemagglutinin receptor specificity and structural analyses of respiratory droplet-transmissible H5N1 viruses. J. Virol. *88*, 768–773.

Drake, J.W., and Holland, J.J. (1999). Mutation rates among RNA viruses. Proc. Natl. Acad. Sci. USA *96*, 13910–13913.

Edo-Matas, D., Rachinger, A., Setiawan, L.C., Boeser-Nunnink, B.D., van 't Wout, A.B., Lemey, P., and Schuitemaker, H. (2012). The evolution of human immunodeficiency virus type-1 (HIV-1) envelope molecular properties and coreceptor use at all stages of infection in an HIV-1 donor-recipient pair. Virology *422*, 70–80.

Hamilton, M.B., Matrosovich, T.Y., Gray, T., Roberts, N.A., and Klenk, H.D. (2012). Population Genetics (Oxford, UK: Wiley-Blackwell).

Imai, M., Watanabe, T., Hatta, M., Das, S.C., Ozawa, M., Shinya, K., Zhong, G., Hanson, A., Katsura, H., Watanabe, S., et al. (2012). Experimental adaptation of an influenza H5 HA confers respiratory droplet transmission to a reassortant H5 HA/H1N1 virus in ferrets. Nature *486*, 420–428.

Jayaraman, A., Chandrasekaran, A., Viswanathan, K., Raman, R., Fox, J.G., and Sasisekharan, R. (2012). Decoding the distribution of glycan receptors for human-adapted influenza A viruses in ferret respiratory tract. PLoS ONE *7*, e27517.

Karlsson, E.A., Meliopoulos, V.A., Savage, C., Livingston, B., Mehle, A., and Schultz-Cherry, S. (2015). Visualizing real-time influenza virus infection, transmission and protection in ferrets. Nat. Commun. *6*, 6378.

Kofler, R., Orozco-terWengel, P., De Maio, N., Pandey, R.V., Nolte, V., Futschik, A., Kosiol, C., and Schlötterer, C. (2011). PoPoolation: a toolbox for population genetic analysis of next generation sequencing data from pooled individuals. PLoS ONE *6*, e15925.

Kong, W., Liu, L., Wang, Y., Gao, H., Wei, K., Sun, H., Sun, Y., Liu, J., Ma, G., and Pu, J. (2014). Hemagglutinin mutation D222N of the 2009 pandemic H1N1 influenza virus alters receptor specificity without affecting virulence in mice. Virus Res. *189*, 79–86.

Lakdawala, S.S., Jayaraman, A., Halpin, R.A., Lamirande, E.W., Shih, A.R., Stockwell, T.B., Lin, X., Simenauer, A., Hanson, C.T., Vogel, L., et al. (2015). The soft palate is an important site of adaptation for transmissible influenza viruses. Nature *526*, 122–125.

Lauring, A.S., and Andino, R. (2010). Quasispecies theory and the behavior of RNA viruses. PLoS Pathog. *6*, e1001005.

Lauring, A.S., Acevedo, A., Cooper, S.B., and Andino, R. (2012). Codon usage determines the mutational robustness, evolutionary capacity, and virulence of an RNA virus. Cell Host Microbe *12*, 623–632.

Linster, M., van Boheemen, S., de Graaf, M., Schrauwen, E.J., Lexmond, P., Mänz, B., Bestebroer, T.M., Baumann, J., van Riel, D., Rimmelzwaan, G.F., et al. (2014). Identification, characterization, and natural selection of mutations driving airborne transmission of A/H5N1 virus. Cell *157*, 329–339.

Mak, G.C., Au, K.W., Tai, L.S., Chuang, K.C., Cheng, K.C., Shiu, T.C., and Lim, W. (2010). Association of D222G substitution in haemagglutinin of 2009 pandemic influenza A (H1N1) with severe disease. Euro Surveill. *15*, 19534, author reply pii/19535.

Matos-Patrón, A., Byrd-Leotis, L., Steinhauer, D.A., Barclay, W.S., and Ayora-Talavera, G. (2015). Amino acid substitution D222N from fatal influenza infection affects receptor-binding properties of the influenza A(H1N1)pdm09 virus. Virology *484*, 15–21.

Matrosovich, M.N., Matrosovich, T.Y., Gray, T., Roberts, N.A., and Klenk, H.D. (2004). Human and avian influenza viruses target different cell types in cultures of human airway epithelium. Proc. Natl. Acad. Sci. USA *101*, 4620–4624.

Nei, M., and Gojobori, T. (1986). Simple methods for estimating the numbers of synonymous and nonsynonymous nucleotide substitutions. Mol. Biol. Evol. *3*, 418–426.

Nelson, C.W., and Hughes, A.L. (2015). Within-host nucleotide diversity of virus populations: insights from next-generation sequencing. Infect. Genet. Evol. *30*, 1–7.

Nelson, C.W., Moncla, L.H., and Hughes, A.L. (2015). SNPGenie: estimating evolutionary parameters to detect natural selection using pooled next-generation sequencing data. Bioinformatics *31*, 3709–3711.

Nielsen, R. (2005). Molecular signatures of natural selection. Annu. Rev. Genet. *39*, 197–218.

Nobusawa, E., and Sato, K. (2006). Comparison of the mutation rates of human influenza A and B viruses. J. Virol. *80*, 3675–3678.

Parrish, C.R., Holmes, E.C., Morens, D.M., Park, E.C., Burke, D.S., Calisher, C.H., Laughlin, C.A., Saif, L.J., and Daszak, P. (2008). Cross-species virus transmission and the emergence of new epidemic diseases. Microbiol. Mol. Biol. Rev. *72*, 457–470.

Parvin, J.D., Moscona, A., Pan, W.T., Leider, J.M., and Palese, P. (1986). Measurement of the mutation rates of animal viruses: influenza A virus and poliovirus type 1. J. Virol. *59*, 377–383.

Poon, L.L., Song, T., Rosenfeld, R., Lin, X., Rogers, M.B., Zhou, B., Sebra, R., Halpin, R.A., Guan, Y., Twaddle, A., et al. (2016). Quantifying influenza virus diversity and transmission in humans. Nat. Genet. Published online January 4, 2016. http://dx.doi.org/10.1038/ng.3479.

Reid, A.H., Fanning, T.G., Hultin, J.V., and Taubenberger, J.K. (1999). Origin and evolution of the 1918 "Spanish" influenza virus hemagglutinin gene. Proc. Natl. Acad. Sci. USA *96*, 1651–1656.

Reid, A.H., Janczewski, T.A., Lourens, R.M., Elliot, A.J., Daniels, R.S., Berry, C.L., Oxford, J.S., and Taubenberger, J.K. (2003). 1918 influenza pandemic caused by highly conserved viruses with two receptor-binding variants. Emerg. Infect. Dis. *9*, 1249–1253.

Rogers, M.B., Song, T., Sebra, R., Greenbaum, B.D., Hamelin, M.E., Fitch, A., Twaddle, A., Cui, L., Holmes, E.C., Boivin, G., and Ghedin, E. (2015). Intrahost dynamics of antiviral resistance in influenza A virus reflect complex patterns of segment linkage, reassortment, and natural selection. MBio *6*, 6.

Sanjuán, R., Nebot, M.R., Chirico, N., Mansky, L.M., and Belshaw, R. (2010). Viral mutation rates. J. Virol. *84*, 9733–9748.

Sheng, Z.M., Chertow, D.S., Ambroggio, X., McCall, S., Przygodzki, R.M., Cunningham, R.E., Maximova, O.A., Kash, J.C., Morens, D.M., and Taubenberger, J.K. (2011). Autopsy series of 68 cases dying before and during the 1918 influenza pandemic peak. Proc. Natl. Acad. Sci. USA *108*, 16416–16421.

Song, H.D., Tu, C.C., Zhang, G.W., Wang, S.Y., Zheng, K., Lei, L.C., Chen, Q.X., Gao, Y.W., Zhou, H.Q., Xiang, H., et al. (2005). Cross-host evolution of severe acute respiratory syndrome coronavirus in palm civet and human. Proc. Natl. Acad. Sci. USA *102*, 2430–2435.

Taubenberger, J.K., Reid, A.H., Krafft, A.E., Bijwaard, K.E., and Fanning, T.G. (1997). Initial genetic characterization of the 1918 "Spanish" influenza virus. Science *275*, 1793–1796.

Taubenberger, J.K., Reid, A.H., Janczewski, T.A., and Fanning, T.G. (2001). Integrating historical, clinical and molecular genetic data in order to explain the origin and virulence of the 1918 Spanish influenza virus. Philos. Trans. R. Soc. Lond. B Biol. Sci. *356*, 1829–1839.

Taubenberger, J.K., Reid, A.H., Lourens, R.M., Wang, R., Jin, G., and Fanning, T.G. (2005). Characterization of the 1918 influenza virus polymerase genes. Nature *437*, 889–893.

Thompson, C.I., Barclay, W.S., Zambon, M.C., and Pickles, R.J. (2006). Infection of human airway epithelium by human and avian strains of influenza a virus. J. Virol. *80*, 8060–8068.

Tumpey, T.M., Maines, T.R., Van Hoeven, N., Glaser, L., Solórzano, A., Pappas, C., Cox, N.J., Swayne, D.E., Palese, P., Katz, J.M., and García-Sastre, A. (2007). A two-amino acid change in the hemagglutinin of the 1918 influenza virus abolishes transmission. Science *315*, 655–659.

Vana, G., and Westover, K.M. (2008). Origin of the 1918 Spanish influenza virus: a comparative genomic analysis. Mol. Phylogenet. Evol. *47*, 1100–1110.

Varble, A., Albrecht, R.A., Backes, S., Crumiller, M., Bouvier, N.M., Sachs, D., García-Sastre, A., and tenOever, B.R. (2014). Influenza A virus transmission bottlenecks are defined by infection route and recipient host. Cell Host Microbe *16*, 691–700.

Watanabe, T., Zhong, G., Russell, C.A., Nakajima, N., Hatta, M., Hanson, A., McBride, R., Burke, D.F., Takahashi, K., Fukuyama, S., et al. (2014). Circulating avian influenza viruses closely related to the 1918 virus have pandemic potential. Cell Host Microbe *15*, 692–705.

Wilker, P.R., Dinis, J.M., Starrett, G., Imai, M., Hatta, M., Nelson, C.W., O'Connor, D.H., Hughes, A.L., Neumann, G., Kawaoka, Y., and Friedrich, T.C. (2013). Selection on haemagglutinin imposes a bottleneck during mammalian transmission of reassortant H5N1 influenza viruses. Nat. Commun. *4*, 2636.

Worobey, M., Han, G.Z., and Rambaut, A. (2014). Genesis and pathogenesis of the 1918 pandemic H1N1 influenza A virus. Proc. Natl. Acad. Sci. USA *111*, 8107–8112.

Wu, C., Cheng, X., Wang, X., Lv, X., Yang, F., Liu, T., Fang, S., Zhang, R., and Jinquan, C. (2013). Clinical and molecular characteristics of the 2009 pandemic influenza H1N1 infection with severe or fatal disease from 2009 to 2011 in Shenzhen, China. J. Med. Virol. *85*, 405–412.

Zaraket, H., Baranovich, T., Kaplan, B.S., Carter, R., Song, M.S., Paulson, J.C., Rehg, J.E., Bahl, J., Crumpton, J.C., Seiler, J., et al. (2015). Mammalian adaptation of influenza A(H7N9) virus is limited by a narrow genetic bottleneck. Nat. Commun. *6*, 6553.

Zhang, C.Y., Wei, J.F., and He, S.H. (2006). Adaptive evolution of the spike gene of SARS coronavirus: changes in positively selected sites in different epidemic groups. BMC Microbiol. *6*, 88.