Journal of Virology

Intrahost Evolutionary Dynamics of Canine Influenza Virus in Naïve and Partially Immune Dogs

Karin Hoelzer, Pablo R. Murcia, Gregory J. Baillie, James L. N. Wood, Stephan M. Metzger, Nikolaus Osterrieder, Edward J. Dubovi, Edward C. Holmes and Colin R. Parrish

J. Virol. **2010, 84(10):5329. DOI: 10.1128/JVI.02469-09. Published Ahead of Print 10 March 2010.**

JOURNAL OF VIROLOGY, May 2010, p. 5329–5335 Vol. 84, No. 10
0022-538X/10/$12.00 doi:10.1128/JVI.02469-09

Intrahost Evolutionary Dynamics of Canine Influenza Virus in Naïve and Partially Immune Dogs▽

Karin Hoelzer,[1] Pablo R. Murcia,[2] Gregory J. Baillie,[2,3] James L. N. Wood,[2] Stephan M. Metzger,[4] Nikolaus Osterrieder,[4,5] Edward J. Dubovi,[6] Edward C. Holmes,[7,8] and Colin R. Parrish[1]*

Baker Institute for Animal Health, Department of Microbiology and Immunology, College of Veterinary Medicine, Cornell University, Ithaca, New York 14853[1]; Cambridge Infectious Diseases Consortium, Department of Veterinary Medicine, Cambridge CB3 0ES, United Kingdom[2]; Wellcome Trust Sanger Institute, Wellcome Trust Genome Campus, Hinxton, Cambridge CB10 1SA, United Kingdom[3]; Department of Microbiology and Immunology, College of Veterinary Medicine, Cornell University, Ithaca, New York 14853[4]; Institute für Virologie, Freie Universität Berlin, Philippstrasse 13, 10115 Berlin, Germany[5]; Department of Population Medicine and Diagnostic Sciences, College of Veterinary Medicine, Cornell University, Ithaca, New York 14853[6]; Center for Infectious Disease Dynamics, Department of Biology, Pennsylvania State University, State College, Pennsylvania 16802[7]; and Fogarty International Center, National Institutes of Health, Bethesda, Maryland 20892[8]

Received 23 November 2009/Accepted 4 March 2010

The patterns and dynamics of evolution in acutely infecting viruses within individual hosts are largely unknown. To this end, we investigated the intrahost variation of canine influenza virus (CIV) during the course of experimental infections in naïve and partially immune dogs and in naturally infected dogs. Tracing sequence diversity in the gene encoding domain 1 of the hemagglutinin (HA1) protein over the time course of infection provided information on the patterns and processes of intrahost viral evolution and revealed some of the effects of partial host immunity. Viral populations sampled on any given day were generally characterized by mean pairwise genetic diversities between 0.1 and 0.2% and by mutational spectra that changed considerably on different days. Some observed mutations may have affected antigenicity or host range, including reversions of CIV host-associated mutations. Patterns of sequence diversity differed between naïve and vaccinated dogs, with some presumably antigenic mutations transiently reaching high frequency in the latter. CIV populations are therefore characterized by the rapid generation and clearance of genetic diversity. Potentially advantageous mutations arise readily during the course of single infections and may give rise to antigenic escape or host range variants.

Cross-species transmissions of novel viral pathogens are commonly associated with the occurrence of major infectious disease epidemics in humans, yet the evolutionary mechanisms associated with cross-species transfers remain elusive. In most cases, multiple mutations contribute to the breaching of host range barriers, but the source of these mutations—whether they are preformed in the original host or arise in the new host—is generally unclear (3, 10, 11). Therefore, it is critical to determine the nature of evolutionary processes operating over the time scale of single infectious periods (1, 10).

In general, intrahost viral populations will reflect a balance between the generation of mutational diversity through error-prone replication and the loss of that diversity by genetic drift and positive or negative (purifying) selection. However, the strength and frequency of these processes, as well as their implications for viral emergence, are not well understood. Although there is a large body of work examining the intrahost sequence dynamics during prolonged or persistent viral infections (see, for example, references 4 and 16), far less is known about the evolutionary dynamics during acute infections by viruses such as influenza virus. An informative exception was a recent study on intrahost variation in three different avian influenza viruses which revealed levels of genetic variation of between 1×10^{-4} and 8×10^{-4} mutations per nucleotide (8). For influenza virus, it is also clear that antibody selection plays an important role in driving the evolution of hemagglutinin (HA) and to a lesser degree the neuraminidase (NA) sequence, particularly in populations where high levels of immunity occur and reinfection is frequent.

Influenza A viruses have their main reservoir in aquatic birds, with some strains infecting and spreading in mammals. Occasionally, influenza viruses transfer to new hosts, either by transfer from aquatic birds to terrestrial birds or mammals, or between mammalian hosts (reviewed, for example, in references 7 and 13). These host-transferred viruses are usually inefficiently transmitted and disappear after causing one infection or a small number of infections. In some cases, however, the novel virus gains the ability to spread efficiently and consequently can cause major epidemics or pandemics in new, immunologically naïve, populations. Canine influenza virus (CIV) is a new virus of dogs derived from A/H3N8 equine influenza virus (EIV). CIV was first recognized as a pathogen of dogs during 2004 when it caused disease in greyhounds in Florida, although serological data suggest an initial transfer to

* Corresponding author. Mailing address: Baker Institute for Animal Health, Department of Microbiology and Immunology, College of Veterinary Medicine, Cornell University, Hungerford Hill Road, Ithaca, NY 14853. Phone: (607) 256-5649. Fax: (607) 256-5608. E-mail: crp3@cornell.edu.
▽ Published ahead of print on 10 March 2010.

TABLE 1. Virus samples examined from each infected dog and a summary analysis of the results of sequencing the HA1 sequence each day after infection

Virus or dog	Day after infection	No. of sequences	Total no. of variant sites	Mean pairwise divergence	Mean d_N/d_S	No. of stop codons	No. of indels[a]	No. of sites with frequency of >1	No. of sites hit by >1 different mutation
Challenge virus		88	27	0.0012	0.674	1	0	7	3
Dogs									
Naïve									
8271	All	204	48	0.0010	0.902	1	0	8	2
	2	62	19	0.0013	0.793	0	0	3	1
	3	70	15	0.0006	1.664	1	0	1	0
	4	72	24	0.0010	0.853	0	0	3	0
7151	All	199	47	0.0014	0.758	0	0	11	1
	2	64	15	0.0012	1.137	0	0	3	0
	3	57	13	0.0009	0.285	0	0	3	0
	4	80	28	0.0018	1.03	0	0	8	0
Vaccinated									
5981	All	114	26	0.0011	0.310	1	0	2	1
	2	67	15	0.0011	0.38	1	0	2	1
	3	47	12	0.0010	0.285	0	0	1	0
6685	All	82	12	0.0009	0.853	1	0	3	0
	2	71	9	0.0007	1.137	0	0	3	0
	3	11	3		0.285	1	0	0	0
Naturally infected									
A1100		57	9	0.0004	0.214	1	1	1	0
38		46	6	0.0003	1.71	0	0	0	0
EIV sequences		158		0.0390	0.42	0	6		

[a] Indels, insertions or deletions.

dogs around 2000 (2, 20). CIV now circulates in several regions of the United States where the host population density is high and where susceptible dogs are replenished frequently (2, 20). In the HA gene (H3), up to five amino acid changes distinguish the currently sequenced CIV isolates from the EIV ancestors (12). CIV isolates are antigenically similar to contemporary EIV strains, but some antigenic differences are detected using panels of monoclonal antibodies (12). Herein we describe, for the first time, the genetic diversity and evolutionary dynamics of the HA1 gene segment (HA1 is the gene encoding domain 1 of hemagglutinin) of CIV within infected naïve, partially immune, and naturally infected dogs.

MATERIALS AND METHODS

Experimental dog infections. The experimental dog infections from which the virus samples were obtained have been reported previously (14). Seronegative beagles (Marshall Farms, North Rose, NY), approximately 8 weeks old, were placed in group housing prior to challenge inoculation. Some dogs were vaccinated subcutaneously (s.c.) with a commercial equine H3N8 subunit vaccine (FluVac Innovator 4; Fort Dodge), while others received only resuspension buffer (negative control). Revaccination with vaccine or buffer was performed 4 weeks later. Each dog was challenged 3 weeks after the time of the second vaccination with a nebulized sample that contained 1×10^6 PFU of an egg-grown influenza A virus A/canine/PA/10915-07 (GenBank accession number GQ280286). Two milliliters of virus-containing allantoic fluid was administered by aerosolization with flowthrough oxygen to each dog for ~10 min. Vaccinated animals developed antibody responses after vaccination, and both naïve and vaccinated dogs developed increased antibody titers after challenge (14). All dogs developed pyrexia (temperatures of >39.5°C) on day 2 postinoculation (p.i.), and except for dog 6685, all showed at least mild signs of respiratory disease on at least 1 day postinoculation. Two nasal swabs were taken from each dog on each day between days 0 and 10 p.i. and stored at −80°C until RNA isolation (14).

Naturally infected dog samples. Nasal swab elutes were collected from dogs naturally infected with CIV while present in animal shelters in the northeastern United States in 2008 and stored at −80°C.

RNA isolation, RT-PCR analysis, cloning, and sequencing. Viral RNA from swabs or challenge virus was isolated using the QIAamp viral RNA minikit (Qiagen, Valencia, CA), then reverse transcribed, and PCR amplified using a two-step reverse transcription-PCR (RT-PCR) protocol. Copy DNA (cDNA) was generated using Superscript III reverse transcriptase (Invitrogen, Carlsbad, CA) and primer Bm-HA1 (5′-TATTCGTCTCAGGGAGCAAAAGCAGGGG-3′) (5). RT was performed at 55°C for 90 min, followed by incubation at 70°C for 10 min. Platinum *Pfx* DNA polymerase (Invitrogen, Carlsbad, CA) was used to prepare the second strand using primers Bm-HA1 and EHA1007rw (5′-TTGG GGCATTTTCCATATGT-3′). This generated a 1,007-nucleotide (nt) fragment covering the first 956 nucleotides of the HA1 open reading frame (ORF) as well as the 43 nucleotides immediately upstream of the HA1 start codon. PCR was performed for 40 cycles (1 cycle consisted of 30 s at 94°C, 1 min at 55°C, and 1 min at 68°C), and then the reaction mixture was incubated at 68°C for 10 min. PCR products purified from agarose gels were cloned into pCR4Blunt-TOPO (Invitrogen) and transformed into *Escherichia coli* (One Shot TOP10 cells; Invitrogen). Clones were sequenced using M13 forward and M13 (19) reverse primers.

Sequence analysis. Between 11 and 88 HA1 sequences from each dog on each day were aligned and then trimmed to cover amino acids 1 to 301 of the nascent HA1 (EMBL accession numbers FN423504 to FN423665), representing a total of 713,370 nucleotides and 790 ORFs. We also aligned the same HA1 region of 158 equine sequences and the 6 CIV sequences in the Influenza Virus Resource (http://www.ncbi.nlm.nih.gov/genomes/FLU/FLU.html).

The mean pairwise genetic diversity within each sample was calculated from the uncorrected pairwise distance matrix (*p*-distance) between taxa. Minimum spanning trees were estimated from the sequence data using Prim's algorithm in BioNumerics as described previously (15, 19). The single likelihood ancestor counting (SLAC) algorithm available in the Datamonkey web interface of the

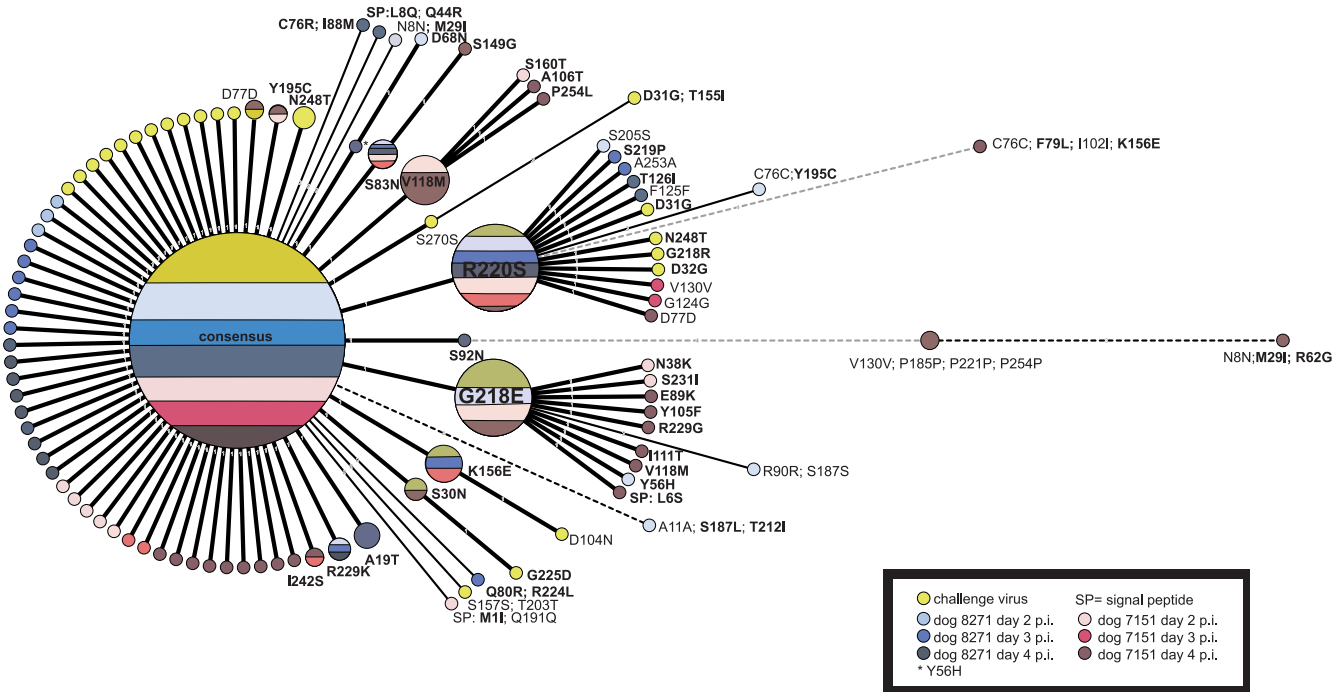

FIG. 1. Minimum spanning trees of the challenge virus (yellow) and virus sequences isolated from the two naïve dogs (dog 8271 [light, medium, and dark blue]) and dog 7151 [pink, red, and brown] on days 2 to 4 postinoculation. Mutations (except for singletons) are indicated on the respective branches, and nonsynonymous mutations are indicated in bold type. The size of each circle is proportional to the number of sequences representing the node. Hypothetical nodes are indicated by broken lines. SP, mutation located in signal peptide.

HyPhy software package (9) was used to estimate the mean numbers of synonymous (d_S) and nonsynonymous (d_N) substitutions per site (the ratio of synonymous substitutions to nonsynonymous substitutions per site is d_N/d_S).

Mutations detected within individual infected animals were further characterized as to their frequency and presence in other samples and were mapped to inferred amino acid sequences of the mature HA1. Mutations at a frequency of >1 were compared to the EIV and CIV sequences obtained at the epidemiological scale. Mutations that changed the coding sequence were mapped to antigenic sites A to E as reviewed in references 17 and 18.

Nucleotide sequence accession numbers. The egg-grown influenza A virus A/canine/PA/10915-07 sequence was deposited in GenBank and assigned accession number GQ280286, and the HA1 sequences from dogs were deposited in EMBL and assigned accession numbers FN423504 to FN423665.

RESULTS

Experimentally challenged animals were exposed to high titers of virus by aerosol, but we were unable to detect virus on day 1 after inoculation. We were able to monitor the viral populations over 3 days in the experimentally infected naïve dogs, sampling at least 57 individual viral sequences per time point. Infections were short lived, with high viral RNA levels being seen only on days 2 and 3 or 4 after inoculation (14). Sufficient RNA could be recovered for analysis from vaccinated dogs only on days 2 and 3 p.i., with lower numbers of sequences analyzed for day 3. The two viruses isolated in 2008 from naturally infected dogs in New York city, New York, were most closely related to each other and fell into a separate cluster from the virus used for experimental challenge, which had been collected in Pennsylvania in 2006 (results not shown).

Intrahost sequence variation and dynamics after experimental challenge. The mean estimates of pairwise genetic diversity ranged from 0.00060 to 0.00180 (mean, 0.00110) (Table 1) and were similar for samples from individual days in naïve and vaccinated dogs and in the challenge virus. Overall, we detected weaker purifying selection (d_N/d_S ratios close to the neutral expectation of 1.0) for both the challenge virus and for most samples from dogs than was seen in the epidemiological-scale EIV sequences collected from GenBank (Table 1). For samples from experimentally infected dogs, mean d_N/d_S ratios varied between 0.29 and 1.70 and averaged 0.83, while the mean d_N/d_S ratios for the challenge virus and the epidemiological-scale EIV data set were 0.69 and 0.41, respectively. No marked differences in d_N/d_S were detected between the naïve and vaccinated dogs. Insertions or deletions were not observed in the experimentally infected dog samples or the challenge virus, but some stop codons indicated defective viruses (Table 1).

After challenge of naïve dogs, a large proportion of sequences in all samples were identical to the consensus sequence, suggesting that the inoculation did not apply a strong bottleneck on the virus population (Fig. 1). Most mutation-bearing sequences had only a single mutation, while some mutations were sampled more than once (Fig. 1 and Tables 1 and 2). Of 16 nonsynonymous mutations sampled repeatedly in the naïve dogs, five mutations were also detected in the epidemiological-scale EIV sequences, while eight were located at sites that were variable among these sequences but involved different substitutions (Table 2). Four mutations changing HA residues 83, 118, 218, and 220 were present at a frequency of >1 in more than one sample (Fig. 1), and all were located in or near known antigenic sites. The mutations at residues 218 and 220 were detected on day 2 p.i. in all dogs examined and were

TABLE 2. Mutations detected at a frequency of >1 within at least one sample

Mutation (amino acid substitution)	No. of animals	Sample (dog or virus) with this mutation (day)[a]	Motif location[b]	Mutation seen in epidemiological-scale equine H3N8	Amino acid variable in equine H3N8?
A19T	1	8271 (day 4)***		No	No (A)
S30N	1	Challenge virus**, 7151 (day 4)**		No	Yes (S, T)
D31G	0	Challenge virus**		No	Yes (D, N)
Y56H	1	8271 (day 2)**, 8271 (day 4)*	Near Ag Ca	No	No (Y)
S83N	2	8271 (day 2)*, 8271 (day 3)*, 8271 (day 4)**, 7151 (day 2)*, 7151 (day 3)****	Ag Eb	Yes	Yes (N, S, K)
S92N	2	8271 (day 4)*, 7151 (day 4)***		Yes	Yes (S, N)
V118M	1	7151 (day 2)****, 7151 (day 4)**********	Near Ag A	No	No (L)
T126A	3	Challenge virus*, 8271 (day 4)*, 7151 (day 4)*, A1100**	Ag B	Yes	Yes (T, A, I)
V130V	1	7151 (day 3)*, 7151 (day 4)***	Near Ag A	No	
R135I	1	6685 (day 2)**********	Ag A	No	Yes (R, G, T, E, S)
K156E	2	Challenge virus**, 8271 (day 3)*, 7151 (day 3)****	Ag B	Yes	Yes (K, E, Q, N)
P185P	1	7151 (day 4)***	Near Ag B	No	No (P)
G218E	4	Challenge virus******, 8271 (day 2)*******, 7151 (day 2)******, 7151 (day 4)**********, 5981 (day 2)****, 6685 (day 2)**	Ag Db	No	No (G)
R220S	4	Challenge virus**********, 8271 (day 2)**********, 8271 (day 3)****, 8271 (day 4)*********, 7151 (day 2)**********, 7151 (day 3)********, 7151 (day 4)*, 5981 (day 2)**********, 5981 (day 3)**********, 6685 (day 2)**	Ag Db	No	No (R)
P221P	2	7151 (day 4)***, 6685 (day 2)*		No	No (P)
N248T	0	Challenge virus*****	Near Ag Dc	No	Yes (N, K)
P254P	2	7151 (day 4)***, 6685 (day 2)*		No	No (P)
S270S	0	Challenge virus**		No	No (S)

[a] The number of sequences detected in each sample is indicated by the number of asterisks. The sampling day postinoculation is shown in parentheses.
[b] Ag, antigenic site.

also present in the challenge virus and reached high frequencies in several samples. Residues 218 and 220 are invariant in the epidemiological-scale EIV sequences, suggesting that changes at these sites reduce viral fitness, at least in horses (Table 2). Mutations changing residues 83 and 118 were not detected in the challenge virus but were repeatedly observed in samples from the infected dogs on different days (Fig. 1 and Table 3), indicating that these mutations remained in the viral populations. Notably, residue 83 reverts to the codon observed in EIV, and while L118V distinguishes older from more-recent CIV isolates, the V118M mutation represents an additional mutation at the same codon.

After challenge of vaccinated dogs, several mutations were detected in the viral populations that were also present in the naïve dogs (Tables 1 and 2). In addition, a high proportion of viral sequences in both vaccinated dogs harbored mutations on day 2 postinoculation (Fig. 2 and Table 3). For example, 30 of 67 sequences isolated on day 2 p.i. from dog 5981 possessed the R220S mutation, and 30 of 47 sequences from dog 5981 had that mutation on day 3 p.i. This mutation was also present in the challenge virus and the naïve dogs but at a much lower relative frequency. Strikingly, 75% of sequences isolated from vaccinated dog 6685 on day 2 contained an R135I mutation (Fig. 2 and Table 3), but none of the 11 sequences from that same animal on day 3 p.i. contained the mutation. This mutation was not detected in the other animals or in the challenge virus. Interestingly, residue 135 is located in antigenic site A and is highly variable in EIV isolates, encoding five different residues, and is also subject to putative positive selection (d_N/d_S of >1) in the epidemiological-scale EIV sequences (results not shown).

Genetic diversity in naturally infected dogs. HA1 sequences from naturally infected dogs showed mean estimates of pairwise genetic diversity similar to those in the experimentally infected animals (Table 1) and with similar population structures, in which most sequences were identical to the consensus sequence, and most variants contained only single mutations (results not shown). Seven variable sites were detected in the sample from dog 38, while 8 variable sites were detected in the sample from dog A1100, two of which (residues 111 and 126) harbored mutations also detected in samples from experimental infections. In dog A1100, one mutation changing residue 126 was present at a frequency of >1, and single sequences contained a deletion of residue 47 or a stop codon, and a change at codon 135 (K135N) which affected a residue mutated at a high frequency in dog 6685 on day 2 p.i.

Three residues in HA1 (N54K, N83S, and W222L) distinguish all CIVs from related EIV sequences (12), and reversions of the N83S mutation were detected in several samples from the CIV-infected dogs (Table 2). However, the N83S mutation never reached high frequency in any sample and was generally not sampled later in the infection, compatible with either a lower fitness or random loss through genetic drift.

DISCUSSION

The emergence of CIV through transfer of EIV from horses to dogs represents an opportunity to study intrahost evolution

TABLE 3. Mutations present in viral sequences on multiple days

Mutation		Mutation seen in epidemiological-scale equine H3N8	Amino acid position variable in equine H3N8?	Motif location[a]	Sample (dog or virus) with this mutation (day)[b]
Nucleotide substitution	Amino acid substitution				
Synonymous mutations					
T24C	N8N	Yes	Yes (N, K)		8271 (day 2)*, 7151 (day 4)*
T228C	C76C	Yes	Yes (C, Y)		Challenge virus*, 827 (day 2)*, 7151 (day 3)*
C231T	D77D	Yes	Yes (D, N)		Challenge virus*, 7151 (day 4)*
T333C	I111I	No	Yes (I,L)		7151 (day 2)*, A1100*
A384T	T128T	No	No (T)		7151 (day 2)*, 5981 (day 3)*
A390G	V130V	No	No (V)	Near Ag A	7151 (day 3)*, 7151 (day 4)***
T435C	D145D	Yes	Yes (D, S)	Ag A	7151 (day 4)*, 6685 (day 3)*
T663G	P221P	No	No (P)		7151 (day 4)***, 6685 (day 2)*
A681G	S227S	No	Yes (S, P)		Challenge virus*, 8271 (day 2)*, 8271 (day 4)*
G762T	P254P	No	No (P)		7151 (day 4)***, 6685 (day 2)*
Nonsynonymous mutations					
G87A	M29I	No	Yes (I, L)		8271 (day 2)*, 7151 (day 4)*
G89A	S30N	No	Yes (S, T)		Challenge virus**, 7151 (day 4)**
A95G	D32G	Yes	Yes (D, G)		Challenge virus*, 7151 (day 2)*, 5981 (day 2)*
A131G	Q44R	No	No (Q)		8271 (day 4)*, 5981 (day 2)*
T166C	Y56H	No	No (Y)	Near Ag Ca	8271 (day 2)**, 8271 (day 4)*
G248A	S83N	Yes	Yes (S, N, K)	Ag Eb	8271 (day 2)*, 8271 (day 3)*, 8271 (day 4)**, 7151 (day 2)*, 7151 (day 3)****
G275A	S92N	Yes	Yes (S, N)		8271 (day 4)*, 7151 (day 4)***
G316A	A106T	No	No (T)	Near Ag Eb	7151 (day 4)*, 5981 (day 3)*
T332C	I111T	No	Yes (I,L)		7151 (day 4)*, 6685 (day 3)*
G352A	V118 M	No	No (L)	Near Ag A	7151 (day 2)****, 7151 (day 4)**********
A376G	T126A	Yes	Yes (T, A, I)	Ag B	Challenge virus*, 7151 (day 4)*, A1100**
A466G	K156E	Yes	Yes (K, E, Q, N)	Ag B	Challenge virus**, 8271 (day 3)*, 7151 (day 3)****
A584G	Y195C	No	No (Y)	Stabilizes receptor binding pocket?	8271 (day 2)*, 8271 (day 4)*, 5981 (day 3)*
G653A	G218E	No	No (G)	Ag Db	Challenge virus******, 8271 (day 2)*******, 7151 (day 2)******, 7151 (day 4)**********, 5981 (day 2)****, 6685 (day 2)**
A660C	R220S	No	No (R)	Ag Db	Challenge virus**********, 8271 (day 2)**********, 8271 (day 3)****, 8271 (day 4)*********, 7151 (day 2)**********, 7151 (day 3)*********, 7151 (day 4)*, 5981 (day 2)**********, 5981 (day 3)**********, 6685 (day 2)**
G686A	R229K	No	No (R)	Stabilizes receptor binding pocket?	8271 (day 2)*, 8271 (day 3)*, 8271 (day 4)*
T725G	I242S	No	Yes (I, V, T)	Ag Dc	8271 (day 3)*, 8271 (day 4)*
C761T	P254L	No	No (P)		8271 (day 4)*, 7151 (day 4)*

[a] Ag, antigenic site.
[b] The frequency with which the mutation is present in virus from animals or in the challenge virus is indicated by the number of asterisks. The sampling day postinoculation is shown in parentheses.

in an influenza virus that has recently acquired a new host range and in the face of low levels of vaccine-induced immunity. Mutations arose readily in the infected animals and reached high frequencies in some vaccinated dogs, but they were mostly transient and often were not detected on subsequent days. Hence, CIV populations are highly dynamic and characterized by a rapid turnover of likely deleterious mutations. In contrast to the long-term evolutionary dynamics observed during persistent RNA virus or retrovirus infections, such as HIV, the consensus sequences in naïve animals remained unaltered over the course of infection. Such viral dynamics would be expected, given the short duration of influenza infection and the small number of replication cycles. However, the finding that predicted antigenic mutations came to transiently dominate the viral populations in the vaccinated animals shows that advantageous mutations can be rapidly selected.

The viral population structures detected in the naïve animals

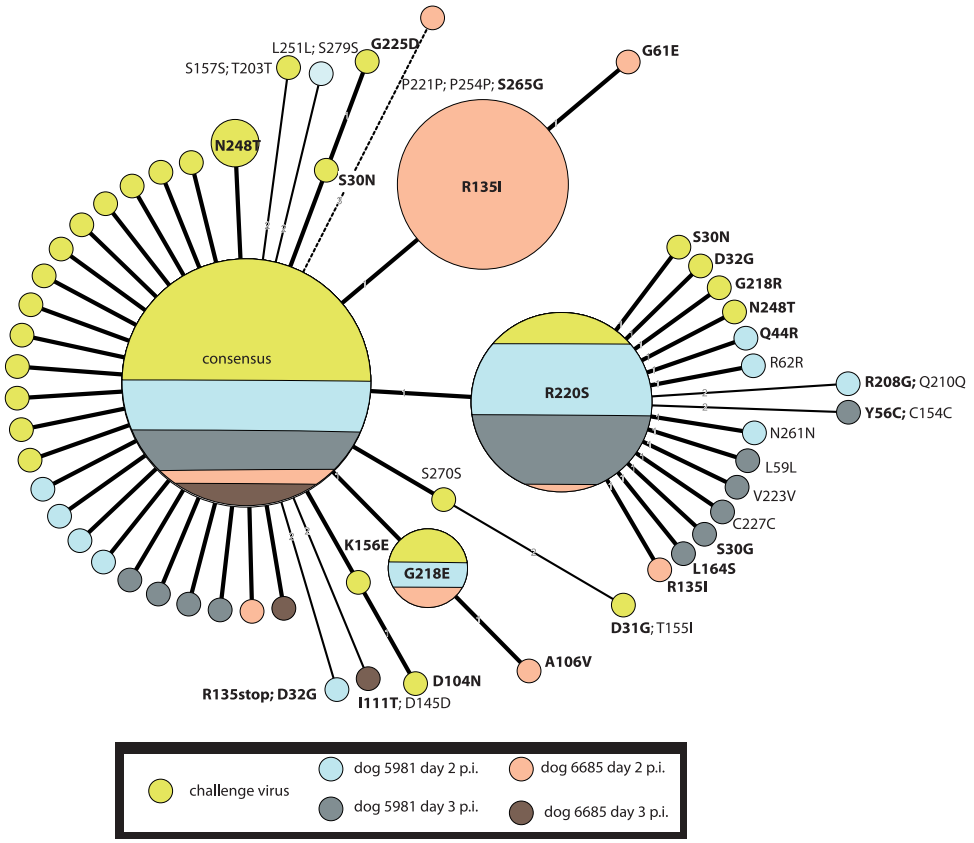

FIG. 2. Minimum spanning trees of the challenge virus (yellow) and virus sequences isolated from the two vaccinated dogs (dog 5981 [light and dark blue] and dog 6685 [red and brown]) on days 2 and 3 postinoculation. Mutations (except for singletons) are indicated on the respective branches, and nonsynonymous mutations are indicated in bold type. The size of each circle is proportional to the number of sequences representing the node. Hypothetical nodes are indicated by broken lines. SP, mutation located in signal peptide.

were largely comparable to those of the challenge virus. However, the lack of detectable virus on day 1 postinoculation indicates that virus recovered at later times had replicated in the dogs and did not represent persisting inoculum, and it also appears that no significant egg adaptation had occurred. In addition, sequences from naturally infected dogs were comparable to those observed in our naïve, experimentally challenged dogs, indicating that the experimental inoculation did not strongly distort the viral population structure.

A notable observation from this analysis was the generally weaker purifying selection detected among viruses sampled from a single animal than among viruses taken from different animals, which manifested as reduced d_N/d_S ratios in the latter. A similar observation has been made for dengue virus (6) and suggests that intrahost viral populations contain transient deleterious mutations, although the rapid turnover of variants detected here suggests that such selection often occurs over the short time periods between different sampling days.

Mutations in some sites appear selectively advantageous, if only transiently. For example, R135I reached high frequency on day 2 p.i. in vaccinated dog 6685, yet that mutation was not detected among the 11 sequences sampled on day 3 p.i. In addition, many mutations that were apparently generated *de novo* in the dogs map to known antigenic sites, and some affect residues that are under positive selection at the epidemiological scale in EIV. The highest variant frequencies were detected in the vaccinated dogs with prior immunity, while in the naïve animals, most viral sequences remained identical to the challenge virus consensus sequence. While a number of mutations in antigenic sites might be expected to occur by chance (17, 18), the location of high-frequency mutations at these sites in vaccinated dogs suggests a role of antigenic escape. As escape from host immunity and antigenic drift are a key feature of influenza A virus evolution, at least in humans, this is an area that needs to be explored further.

Finally, an undetermined proportion of the mutations observed here are clearly artifactual and invariably introduced during the experimental procedure used, particularly at the reverse transcription step. However, the fact that we consistently see mutations at frequencies of >1 within each sample and repeatedly observe mutations at the same sites (Table 1), even though such mutations have a very low probability of occurring independently at random (P. Murcia et al., submitted for publication), suggests that the major features of the mutational spectra we describe are likely to be correct.

The origin of host-range mutations is a key question in understanding the evolutionary dynamics of viral host switching. Our study suggests that such mutations can readily be produced during infections of individuals and maintained in the virus populations for short time periods. Indeed, one site in

HA1 that distinguishes CIV from its ancestor EIV (at residues 83) reverted in some viral sequences, and viral sequences with this mutation were detected repeatedly in the naïve dogs. This residue is variable in EIV sequences, but it has not been seen to revert in other CIV sequences, suggesting that it is associated with a viral fitness cost in dogs. This observation suggests that mutations that facilitate adaptation to a new host species might occur transiently in the donor host despite any associated fitness costs and provide a transient reservoir of preadapted mutations.

Overall, this study highlights the dynamic nature of influenza virus evolution, revealing how readily influenza viruses generate mutations of adaptive significance. Monitoring the spectrum of influenza virus intrahost genetic diversity may therefore provide important information for predicting both virus emergence and antigenic escape.

ACKNOWLEDGMENTS

We thank Virginia Scarpino, Wendy S. Weichert, and Shelagh M. Johnston for excellent technical support.

This study was funded in part by National Institutes of Health grant R01 GM080533 to E.C.H and C.R.P. K.H. was supported by a graduate research assistantship from the College of Veterinary Medicine at Cornell University.

REFERENCES

1. **Antia, R., R. R. Regoes, J. C. Koella, and C. T. Bergstrom.** 2003. The role of evolution in the emergence of infectious diseases. Nature **426:**658–661.
2. **Crawford, P. C., E. J. Dubovi, W. L. Castleman, I. Stephenson, E. P. Gibbs, L. Chen, C. Smith, R. C. Hill, P. Ferro, J. Pompey, R. A. Bright, M. J. Medina, C. M. Johnson, C. W. Olsen, N. J. Cox, A. I. Klimov, J. M. Katz, and R. O. Donis.** 2005. Transmission of equine influenza virus to dogs. Science **310:**482–485.
3. **Domingo, E., and J. J. Holland.** 1997. RNA virus mutations and fitness for survival. Annu. Rev. Microbiol. **51:**151–178.
4. **Farci, P., A. Shimoda, A. Coiana, G. Diaz, G. Peddis, J. C. Melpolder, A. Strazzera, D. Y. Chien, S. J. Munoz, A. Balestrieri, R. H. Purcell, and H. J. Alter.** 2000. The outcome of acute hepatitis C predicted by the evolution of the viral quasispecies. Science **288:**339–344.
5. **Hoffmann, E., J. Stech, Y. Guan, R. G. Webster, and D. R. Perez.** 2001. Universal primer set for the full-length amplification of all influenza A viruses. Arch. Virol. **146:**2275–2289.
6. **Holmes, E. C.** 2003. Patterns of intra- and interhost nonsynonymous variation reveal strong purifying selection in dengue virus. J. Virol. **77:**11296–11298.
7. **Horimoto, T., and Y. Kawaoka.** 2005. Influenza: lessons from past pandemics, warnings from current incidents. Nat. Rev. Microbiol. **3:**591–600.
8. **Iqbal, M., H. Xiao, G. Baillie, A. Warry, S. C. Essen, B. Londt, S. M. Brookes, I. H. Brown, and J. W. McCauley.** 2009. Within-host variation of avian influenza viruses. Philos. Trans. R. Soc. Lond. B Biol. Sci. **364:**2739–2747.
9. **Kosakovsky Pond, S. L., A. F. Poon, A. J. Leigh Brown, and S. D. Frost.** 2008. A maximum likelihood method for detecting directional evolution in protein sequences and its application to influenza A virus. Mol. Biol. Evol. **25:**1809–1824.
10. **Kuiken, T., E. C. Holmes, J. McCauley, G. F. Rimmelzwaan, C. S. Williams, and B. T. Grenfell.** 2006. Host species barriers to influenza virus infections. Science **312:**394–397.
11. **Parrish, C. R., E. C. Holmes, D. M. Morens, E. C. Park, D. S. Burke, C. H. Calisher, C. A. Laughlin, L. J. Saif, and P. Daszak.** 2008. Cross-species virus transmission and the emergence of new epidemic diseases. Microbiol. Mol. Biol. Rev. **72:**457–470.
12. **Payungporn, S., P. C. Crawford, T. S. Kouo, L. M. Chen, J. Pompey, W. L. Castleman, E. J. Dubovi, J. M. Katz, and R. O. Donis.** 2008. Influenza A virus (H3N8) in dogs with respiratory disease, Florida. Emerg. Infect. Dis. **14:**902–908.
13. **Reid, A. H., and J. K. Taubenberger.** 2003. The origin of the 1918 pandemic influenza virus: a continuing enigma. J. Gen. Virol. **84:**2285–2292.
14. **Rosas, C., G. R. Van de Walle, S. M. Metzger, K. Hoelzer, E. J. Dubovi, S. G. Kim, C. R. Parrish, and N. Osterrieder.** 2008. Evaluation of a vectored equine herpesvirus type 1 (EHV-1) vaccine expressing H3 haemagglutinin in the protection of dogs against canine influenza. Vaccine **26:**2335–2343.
15. **Roumagnac, P., F. X. Weill, C. Dolecek, S. Baker, S. Brisse, N. T. Chinh, T. A. Le, C. J. Acosta, J. Farrar, G. Dougan, and M. Achtman.** 2006. Evolutionary history of Salmonella typhi. Science **314:**1301–1304.
16. **Shankarappa, R., J. B. Margolick, S. J. Gange, A. G. Rodrigo, D. Upchurch, H. Farzadegan, P. Gupta, C. R. Rinaldo, G. H. Learn, X. He, X. L. Huang, and J. I. Mullins.** 1999. Consistent viral evolutionary changes associated with the progression of human immunodeficiency virus type 1 infection. J. Virol. **73:**10489–10502.
17. **Skehel, J. J., and D. C. Wiley.** 2000. Receptor binding and membrane fusion in virus entry: the influenza hemagglutinin. Annu. Rev. Biochem. **69:**531–569.
18. **Wiley, D. C., and J. J. Skehel.** 1987. The structure and function of the hemagglutinin membrane glycoprotein of influenza virus. Annu. Rev. Biochem. **56:**365–394.
19. **Wirth, T., D. Falush, R. Lan, F. Colles, P. Mensa, L. H. Wieler, H. Karch, P. R. Reeves, M. C. Maiden, H. Ochman, and M. Achtman.** 2006. Sex and virulence in Escherichia coli: an evolutionary perspective. Mol. Microbiol. **60:**1136–1151.
20. **Yoon, K. J., V. L. Cooper, K. J. Schwartz, K. M. Harmon, W. I. Kim, B. H. Janke, J. Strohbehn, D. Butts, and J. Troutman.** 2005. Influenza virus infection in racing greyhounds. Emerg. Infect. Dis. **11:**1974–1976.