

NOTES

Characterization of a Canine Homolog of Human Aichivirus[∇]

Amit Kapoor,^{1*} Peter Simmonds,² Edward J. Dubovi,³ Natasha Qaisar,¹ Jose A. Henriquez,¹
Jan Medina,¹ Shelly Shields,⁴ and W. Ian Lipkin¹

Center for Infection and Immunity, Columbia University, New York, New York 10032¹; Centre for Immunology, Infection, and Evolution, University of Edinburgh, Edinburgh EH9 3JT, United Kingdom²; College of Veterinary Medicine, Cornell University, Ithaca, New York 14853³; and Pfizer Veterinary Medicine Research and Development, New York, New York 10017⁴

Received 6 June 2011/Accepted 22 August 2011

Many of our fatal “civilization” infectious diseases have arisen from domesticated animals. Although picornaviruses infect most mammals, infection of a companion animal is not known. Here we describe the identification and genomic characterization of the first canine picornavirus. Canine kobuvirus (CKoV), identified in stool samples from dogs with diarrhea, has a genomic organization typical of a picornavirus and encodes a 2,469-amino-acid polyprotein flanked by 5′ and 3′ untranslated regions. Comparative phylogenetic analysis using various structural and nonstructural proteins of CKoV confirmed it as the animal virus homolog most closely related to human Aichivirus (AiV). Bayesian Markov chain Monte Carlo analysis suggests a mean recent divergence time of CKoV and AiV within the past 20 to 50 years, well after the domestication of canines. The discovery of CKoV provides new insights into the origin and evolution of AiV and the species specificity and pathogenesis of kobuviruses.

Picornaviruses are ubiquitous, infecting a wide range of vertebrate species, particularly mammals, and include a large number of clinically important human and animal pathogens. The family *Picornaviridae* is genetically highly diverse, currently comprising 12 classified genera and several more currently proposed genera, many of which include several species, subspecies, and serotypes/genotypes (<http://www.picornaviridae.com/>) (18). Aichivirus (AiV), discovered in stool samples from children with diarrhea in the Aichi Prefecture of Japan, is a human pathogen that causes acute gastroenteritis and is classified as a species of the genus *Kobuvirus* of the family *Picornaviridae* (31). Other recognized and recently identified kobuviruses (kobu-like viruses) include bovine kobuvirus (BKoV) (30) and porcine kobuvirus (PKoV) (23, 24). Although classified as separate genera, human klassevirus/salivirus (6, 9, 19) and avian turdivirus 1 (29) show greater sequence similarity to kobuviruses than to other picornavirus genera and group with the *Kobuvirus* genus on phylogenetic analysis.

AiV is globally distributed and has been identified at low incidence in sporadic gastroenteritis cases, suggesting fecal-oral transmission. However, serological studies indicate that up to 90% of the human population has been exposed to AiV by the age of 40 years (25). Information on the origin and ongoing circulation of AiV in humans and the existence of potential animal reservoirs for human infection remains elusive (18, 25, 30).

The majority of emerging human infectious viral diseases have originated through cross-species transmission (11, 21).

Nonetheless, successful host transfers of viruses appear rare because of biological and epidemiological barriers to transmission (22). Sustained contact between animal viruses and humans increases the likelihood of the stochastic emergence of a virus adapted to infect and replicate in humans (21). Canines and humans have cohabited for thousands of years; hence, we have begun to focus on characterizing unknown viruses at this human-animal interface (12). Although picornaviruses have been found in a wide range of mammalian species, none have previously been reported to infect domestic dogs (18). Here we identify the first picornavirus of canines and provide genetic evidence that it is the most closely related animal virus homolog of human AiV.

Stool samples from dogs with acute gastroenteritis were enriched for viral nucleic acids (15), randomly amplified, and subjected to unbiased high-throughput sequencing (5). Bioinformatic analysis of sequences at the predicted amino acid level revealed the presence of several sequences substantially similar to those of picornaviruses (15). Sequence fragments were mapped to prototypic kobuvirus genomes, and gaps in genomic sequences were filled by PCR using specific and degenerate primers. Preliminary phylogenetic analysis of approximately 2,800 nucleotides (nt) of continuous genomic sequence revealed the presence of a novel virus most closely related to human AiV and tentatively named canine kobuvirus (CKoV). Thereafter, specific primers targeting highly conserved RNA polymerase (RdRp) gene motifs in CKoV were used in PCRs to screen fecal samples or rectal swabs from 18 dogs representing three different outbreaks of acute gastroenteritis in canine shelters in the United States. CKoV variants were found in 0 of 7, 1 of 5, and 4 of 6 dogs. The variants from two different outbreaks showed only 5% nucleotide sequence differences in a 253-nt-long RdRp fragment and 4% nucleotide

* Corresponding author. Mailing address: Center for Infection and Immunity, Columbia University, 722 West 168th Street, New York, NY 10032. Phone: (212) 304-5690. Fax: (212) 342-9044. E-mail: ak3117@columbia.edu.

[∇] Published ahead of print on 31 August 2011.

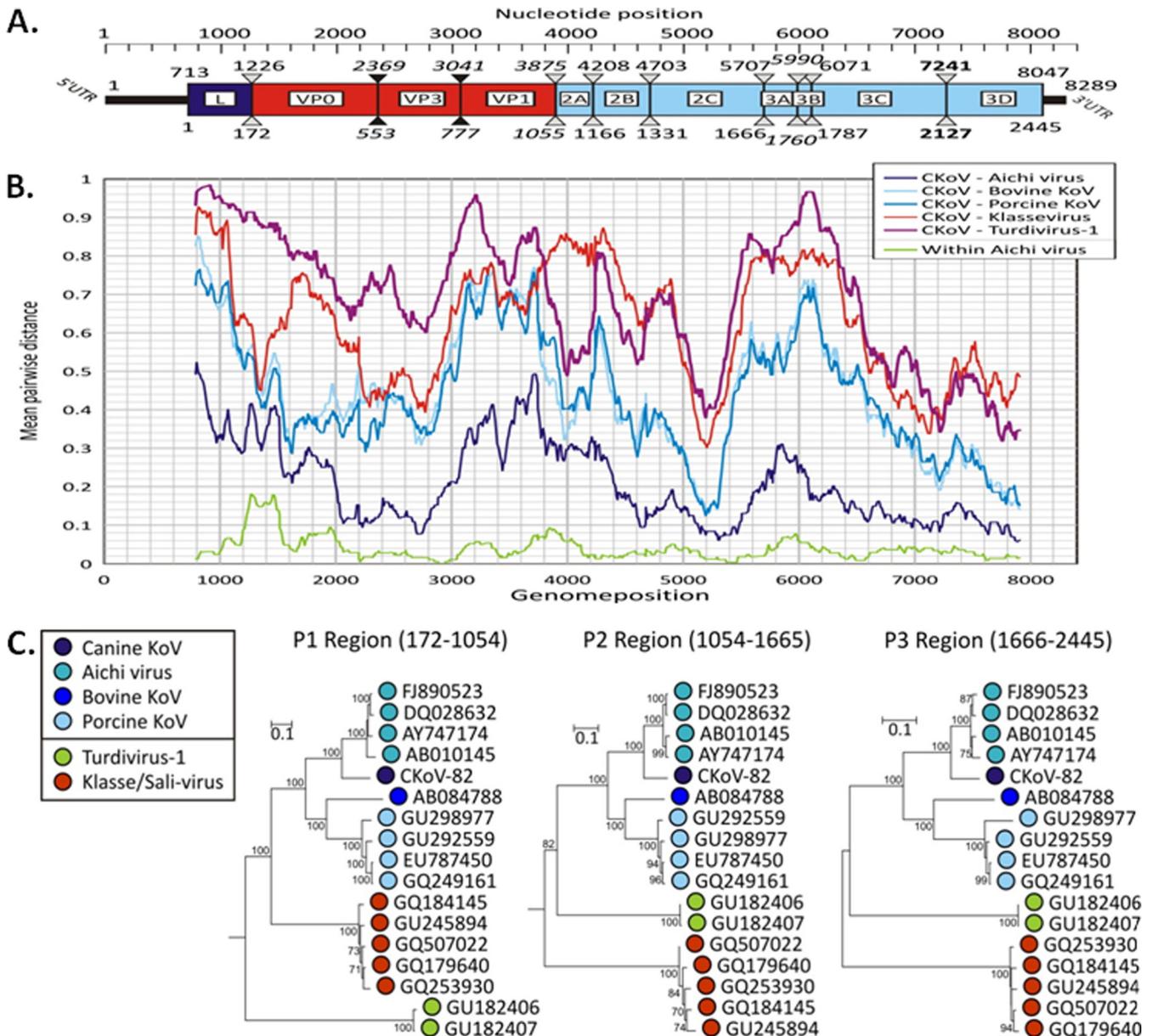


FIG. 1. Genome organization of CKoV and amino acid sequence divergence from other kobuviruses and kobu-like viruses. (A) Predicted genome organization of CKoV, showing amino acid and nucleotide positions of predicted cleavage sites in the polyprotein (numbering based on the CKoV genomic sequence). Sites were predicted by NetPicoRNA analysis (positions shown in normal type) and by alignment with known cleavage sites in AiV (italic type; VP0/VP3, VP3/VP1, and VP1/P2A junctions). Cleavage at the 3C/3D site predicted by NetPicoRNA analysis (bold type) differed from the homologous site in the sequence annotations of AiV sequences (accession no. FJ890523 and AB010145). (B) Mean divergence of kobuvirus and kobu-like translated amino acid sequences from CKoV scaled to the genome diagram. The divergence between AiV variants is shown for comparison. (C) Phylogenetic comparison of CKoV with other kobuviruses (PKoV and BKoV) and kobu-like viruses (klassevirus/salivirus and turdiviruses) in the P1, P2, and P3 gene regions. Translated amino acid sequences from the coding region of CKoV were aligned with other available complete genome sequences of kobuviruses and kobu-like viruses. Trees were constructed by the neighbor-joining method using Poisson-corrected p distances. Data were bootstrap resampled 100 times with values shown on branches. The trees were rooted with the turdivirus 2 sequence (accession no. GU182408) (not shown).

sequence differences in a 409-nt-long VP1 fragment; the majority of these differences occurred at synonymous sites.

The genomic sequence of CKoV was determined directly from a stool sample from one of the five dogs with gastroenteritis. The CKoV genome comprises at least 8,289 nt and encodes a 2,469-amino-acid polyprotein flanked on each side by untranslated regions (UTRs) (see Fig. 2A). The positions of

cleavage sites in the CKoV polyprotein were predicted by sequence alignment and identification of conserved sites homologous to those predicted or demonstrated for AiV. Sites for the 3C protease were independently predicted by NetPicoRNA analysis and produced largely concordant results (Fig. 1A). For the nonstructural gene region, NetPicoRNA predictions were consistent for the L/VP0, 2A/2B, 2B/2C, 2C/

TABLE 1. Pairwise divergence^a in the P1, P2, and P3 regions between CKoV and kobuviruses and kobu-like viruses^b

Region and virus	% Divergence from:					
	CKoV	AiV	BKoV	PKoV	Human klassevirus	Turdivirus 1
P1						
CKoV		27.7	54.2	51.6	61.9	79.8
AiV	28.3		54.8	51.9	62.1	80.3
BKoV	47.0	46.4		43.4	64.8	80.2
PKoV	46.1	45.7	40.2		64.8	80.8
Human klassevirus	50.4	50.4	52.9	52.8		78.5
Turdivirus 1	60.7	61.0	62.8	64.0	60.7	
P2						
CKoV		17.8	37.1	37.7	66.0	61.4
AiV	22.3		38.5	38.1	66.4	60.2
BKoV	37.0	38.3		27.4	63.1	60.9
PKoV	37.7	38.3	32.4		64.1	62.0
Human klassevirus	51.2	51.1	53.0	52.4		66.5
Turdivirus 1	50.6	49.5	50.5	52.6	54.4	
P3						
CKoV		14.7	38.3	38.6	57.5	59.6
AiV	19.4		37.9	38.2	57.6	58.9
BKoV	37.0	37.0		30.2	59.9	60.1
PKoV	39.1	37.6	32.4		60.5	59.5
Human klassevirus	47.0	47.5	49.9	50.1		63.4
Turdivirus 1	47.7	47.1	50.5	49.8	52.1	

^a Pairwise divergence of nucleotide (lower left quadrant) and amino acid (upper right quadrant) sequences.

^b Reference 25.

3A, and 3B/3C sites, all containing Q/G residues on either side of the predicted site. A Q/S cleavage site between 3C and 3D of AiV (30, 31) was predicted at a position equivalent to residue 1966 of CKoV, but NetPicoRNA made a strong prediction of an alternative cleavage site at position 2177 (IDYM Q/GRPG); the latter is shown in Fig. 1A. Q/H and Q/A cleavage sites in the capsid protein were predicted by comparison with AiV sites, as was the more speculative VP1/2A site with the noncanonical sequence motif RPTY/VHWA at position 1055. The 3A/3B site was not predicted by netPicoRNA and was tentatively placed at position 1760 (SETQ/AAYS) based on alignment with the AiV sequence.

The genetic relationship of CKoV with other known kobuviruses was assessed by plotting amino acid sequence divergence over translated polyprotein sequences of other kobuviruses and kobu-like viruses. CKoV showed the greatest amino acid sequence similarity to AiV throughout the genome (dark blue line, Fig. 1B). Its divergence from AiV was, however, consistently greater than that observed within different isolates of AiV (mean values shown as a green line), confirming its identity as a separate type or species. The divergence scan revealed no evidence of genetic recombination among CKoV, AiV, other kobuviruses, or kobu-like viruses. Consistent with pairwise distances calculated for P1, P2, and P3 gene blocks (Table 1), CKoV clustered with AiV on phylogenetic analysis in all three genome regions (Fig. 1C) and showed consistently greater divergence throughout the PKoV and BKoV genomes and the genomes of the kobu-like viruses klassevirus and turdivirus 1. 2C, 3C, and 3D showed the greatest similarity across the group, with pronounced dips in variability corresponding to regions containing highly conserved helicase, protease, and polymerase motifs. The most divergent genes were for the leader protein, VP1, 3A, and 3B (Fig. 1B).

The 712-nt-long 5' UTR sequence of CKoV showed the greatest similarity to AiV in length, base composition (76% nucleotide identity with AiV), and predicted secondary structure (Fig. 2B). The 5' UTR sequence adjacent to the polyprotein sequence showed the presence of a putative internal ribosome entry site (IRES) similar in structure to that of AiV and with an arrangement of stem-loops, a GNRA motif, and a polypyrimidine tract most similar to those of the type II IRES found in other picornavirus genera (e.g., *Aphthovirus*, *Hepatovirus*, and *Cardiovirus*). Sequence similarity extended to the 5'-terminal region with predicted stem-loops comparable in size and position to those found in AiV. To date, most RNA structure predictions for AiV (and other kobuviruses) have concentrated on the three 5' stem-loops that have been shown to be essential for replication. These are conserved in CKoV (Fig. 2), with the terminal loops corresponding to stem-loops B and C of other kobuviruses (the CKoV sequence is incomplete at the 5' end and lacks the terminal 32 bases that would likely form the homolog of stem-loop A) (26, 30). The 3' UTR of CKoV was 242 nt long, showed >85% nucleotide sequence identity to the AiV 3' UTR, and contains a stable hairpin similar to that of AiV. In common with kobuviruses and kobu-like viruses, CKoV showed evidence of a large-scale structure in the coding region of the genome (Table 2). The folding energies of the native sequences were compared with those of the same sequence permuted in sequence order but retaining the same frequencies of dinucleotides as the original sequence (algorithm NDR) (1). A mean folding energy (MFE) difference (MFED) of over 20% for CKoV was the highest recorded for any positive-stranded mammalian RNA virus to date and predicted that a substantial proportion of the base pairing in the genome was sequence order dependent. Similarly high values were detected for other kobuviruses, each of these sub-

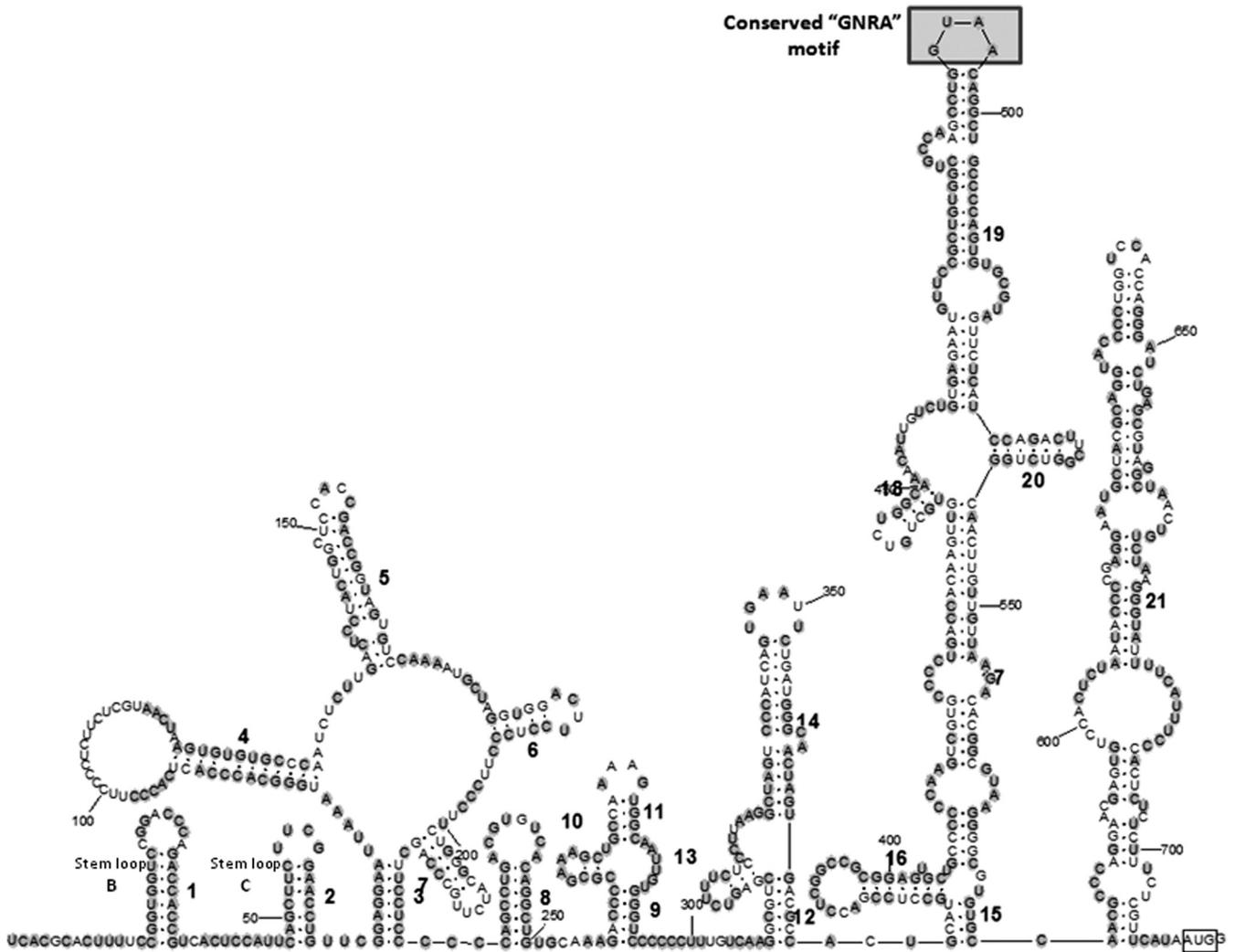


FIG. 2. Sequence and secondary structure of CKoV 5' UTRs. The 724-nt-long 5' UTR of CKoV folds similarly to the type II IRES elements described in the AiV 5' UTR. The nucleotides conserved between CKoV and AiV are shown using grey spheres.

stantially greater than for most other picornaviruses, flaviviruses, and caliciviruses (MFED range, 7% to 15%). Klassevirus/salivirus and turdivirus 1 showed similarly high MFED values, observations that contrast with the absence of RNA folding in the coding regions of the more divergent turdiviruses 2 and 3, which fall outside the kobuvirus/kobu-like virus phylogenetic grouping (Table 2).

CKoV shares a relatively high G+C content (59%) with kobuviruses (52 to 59%), and all have an unusual overrepresentation of cytosine residues (38% in the case of CKoV), reflecting at least in part the previously reported preference for third-codon position C residues rather than the U seen in turdivirus 1 (29). Kobu-like viruses have a modest reduction in the CpG ratio (compared to expected values) and a much greater proportional reduction in UpA. Each of these compositional features contrasts markedly with those of turdiviruses 2 and 3 and more distantly related picornaviruses. The compositional bias additionally leads to effective codon numbers (14) in all of the kobu-like virus genomes (38% [CKoV] to 47%) that are substantially lower than those of the composi-

tionally distinct turdivirus 2 and 3 genomes (51 and 54%) and other picornavirus genomes (all greater than 50%) (27). Nucleotide composition analysis (14) was performed to predict the host origin; first- and second-component scores of -4.369 and 0.467 placed CKoV within the vertebrate cluster, consistent with a canine origin (data not shown).

To determine the evolutionary relationship between CKoV and other kobuviruses, a Bayesian Markov chain Monte Carlo estimation of the time to the most recent common ancestor (TMRCA) (3, 4) of AiV, BKoV, PKoV, and CKoV was performed by using the RdRp region sequence and an external rate calibration based on the evolutionary rates estimated for PKoV (20) and nonenteric picornaviruses (7). The TMRCA was estimated by using a relaxed molecular clock with an uncorrelated log-normal distribution. The Tamura-Nei (TN93) nucleotide substitution method was used, with rate heterogeneity among sites modeled by a discrete gamma distribution with four rate categories, as determined by a best-fit substitution model implemented in MEGA5.05 (28). All analyses were preceded by several Yule speciations (3, 4, 17). The mean

TABLE 2. Mean folding energy differences and base compositions of CKoV and other kobuviruses and kobu-like viruses^d

Virus	RNA folding		Base composition (%)					% Dinucleotides ^c	
	MFE ^a	% MFED ^b	fA	fC	fG	fU	fG+C	rCG	rUA
Kobuviruses									
CKoV	-65.19	20.53	19.9	37.9	20.8	21.3	58.8	0.7288	0.4258
AiV	-62.82	16.74	19.8	38.7	20.8	20.8	59.4	0.8002	0.4471
BKoV	-63.86	13.42	20.3	33.1	21.8	24.8	54.9	0.7156	0.5433
PKoV	-62.69	17.14	20.5	31.4	20.6	27.4	52.1	0.6287	0.5653
Kobu-like viruses									
Human klassevirus	-60.27	15.08	17.9	36.1	20.3	25.6	56.5	0.7224	0.5027
Turdivirus 1	-62.65	15.83	19.7	37.3	20.8	22.2	58.1	0.7216	0.5509
Non-kobu-like viruses									
Turdivirus 2	-56.18	1.15	23.4	23.6	23.8	29.2	47.4	0.3758	0.6736
Turdivirus 3	-55.06	-1.34	25.3	23.3	23.6	27.7	46.9	0.3424	0.6775

^a In kcal/mol.^b MFED of the native sequence from a control with a scrambled sequence order.^c Ratio of the observed dinucleotide frequency to the predicted frequency based on mononucleotide composition.^d Reference 25.

estimated TMRCA for the AiV group and CKoV is 13 years before the present (ybp) (95% highest posterior density, 6 to 20 ybp) based on the PKoV calibration and 87 ybp (range, 42 to 138 ybp) based on the nonenteric picornavirus calibration. Thus, we estimate that the common ancestor of the AiV genotypes and CKoV probably existed between 25 and 50 ybp. However, observation of time dependency in substitution rates in host genes (8) and recent evidence from endogenous viral elements for substantially lower long-term substitution rates in a variety of animal viruses argue against simple extrapolation of substitution rates measured over short observation periods (2, 10, 13, 16).

To conclude, we identified and genetically characterized a novel CKoV which is also the first picornavirus that infects domestic dogs. CKoV was found in stool samples of dogs with gastroenteritis; however, with the limited sampling available in the present study, we were unable to confirm or refute its pathogenic role. According to the eighth report of the International Committee on Taxonomy of Viruses, members of a species of the genus *Kobuvirus* are similar in genomic organization, with greater than 70% amino acid identity in P1 and greater than 70% amino acid identity in 2C and 3CD combined. Although CKoV has a genomic organization similar to that of AiV and shows greater than 70% amino acid identity with the P1, P2, and P3 regions of AiV, CKoV has consistently higher sequence divergence from any reported AiV sequence over the entire polyprotein sequence than individual AiVs have from one another (Fig. 1B). Moreover, the natural host of CKoV is distinct; therefore, we propose that CKoV be provisionally considered the prototype of a new species of the genus *Kobuvirus*.

Observations reported here are consistent with three scenarios: one where AiV and CKoV evolved from a common ancestor and two others where AiV originated from CKoV or vice versa. The presence of highly divergent AiV or CKoV would provide support for the latter two scenarios. Only minimal sequence diversity has been described in AiV; however, it is too early to assess the level of genetic diversity in CKoVs. Our findings suggest that picornavirus infections are wide-

spread in canines. We anticipate that subsequent studies will uncover kobuviruses infecting a range of other mammalian species and, conversely, other picornaviruses infecting dogs, including canine homologues of highly prevalent human enteroviruses and rhinoviruses. The recent identification of canine hepacivirus (12) and CKoV, canine viruses closely related to human viruses, emphasizes the need to focus virus discovery efforts on this human-animal interface.

Nucleotide sequence accession numbers. The nucleotide sequences determined in this study have been submitted to the GenBank database and assigned accession no. JN394542 to JN394545, JN394546 to JN394549, and JN088541.

This work is supported by awards from the National Institutes of Health (AI090196, AI081132, AI079231, AI57158 [Northeast Biodefense Center-Lipkin], AI070411, AI090055, AI072613, and EY017404) and Department of Defense and USDA 58-1275-7-370.

REFERENCES

- Albariño, C. G., et al. 2010. High diversity and ancient common ancestry of lymphocytic choriomeningitis virus. *Emerg. Infect. Dis.* **16**:1093–1100.
- Belyi, V. A., A. J. Levine, and A. M. Skalka. 2010. Unexpected inheritance: multiple integrations of ancient bornavirus and ebolavirus/marburgvirus sequences in vertebrate genomes. *PLoS Pathog.* **6**:e1001030.
- Drummond, A. J., and A. Rambaut. 2007. BEAST: Bayesian evolutionary analysis by sampling trees. *BMC Evol. Biol.* **7**:214.
- Drummond, A. J., A. Rambaut, B. Shapiro, and O. G. Pybus. 2005. Bayesian coalescent inference of past population dynamics from molecular sequences. *Mol. Biol. Evol.* **22**:1185–1192.
- Epstein, J. H., et al. 2010. Identification of GBV-D, a novel GB-like flavivirus from Old World frugivorous bats (*Pteropus giganteus*) in Bangladesh. *PLoS Pathog.* **6**:e1000972.
- Greninger, A. L., et al. 2009. The complete genome of klassevirus—a novel picornavirus in pediatric stool. *Virol. J.* **6**:82.
- Hicks, A. L., and S. Duffy. 2011. Genus-specific substitution rate variability among picornaviruses. *J. Virol.* **85**:7942–7947.
- Ho, S. Y., M. J. Phillips, A. Cooper, and A. J. Drummond. 2005. Time dependency of molecular rate estimates and systematic overestimation of recent divergence times. *Mol. Biol. Evol.* **22**:1561–1568.
- Holtz, L. R., et al. 2009. Klassevirus 1, a previously undescribed member of the family Picornaviridae, is globally widespread. *Virol. J.* **6**:86.
- Horie, M., et al. 2010. Endogenous non-retroviral RNA virus elements in mammalian genomes. *Nature* **463**:84–87.
- Jones, K. E., et al. 2008. Global trends in emerging infectious diseases. *Nature* **451**:990–993.
- Kapoor, A., et al. 2011. Characterization of a canine homolog of hepatitis C virus. *Proc. Natl. Acad. Sci. U. S. A.* **108**:11608–11613.

13. **Kapoor, A., P. Simmonds, and W. I. Lipkin.** 2010. Discovery and characterization of mammalian endogenous parvoviruses. *J. Virol.* **84**:12628–12635.
14. **Kapoor, A., P. Simmonds, W. I. Lipkin, S. Zaidi, and E. Delwart.** 2010. Use of nucleotide composition analysis to infer hosts for three novel picorna-like viruses. *J. Virol.* **84**:10322–10328.
15. **Kapoor, A., et al.** 2008. A highly prevalent and genetically diversified Picornaviridae genus in South Asian children. *Proc. Natl. Acad. Sci. U. S. A.* **105**:20482–20487.
16. **Katzourakis, A., and R. J. Gifford.** 2010. Endogenous viral elements in animal genomes. *PLoS Genet.* **6**:e1001191.
17. **Lemey, P., A. Rambaut, A. J. Drummond, and M. A. Suchard.** 2009. Bayesian phylogeography finds its roots. *PLoS Comput. Biol.* **5**:e1000520.
18. **Lewis-Rogers, N., and K. A. Crandall.** 2010. Evolution of Picornaviridae: an examination of phylogenetic relationships and cophylogeny. *Mol. Phylogenet. Evol.* **54**:995–1005.
19. **Li, L., et al.** 2009. A novel picornavirus associated with gastroenteritis. *J. Virol.* **83**:12002–12006.
20. **Park, S. J., H. K. Kim, D. S. Song, H. J. Moon, and B. K. Park.** 2011. Molecular detection and genetic characterization of kobuviruses in fecal samples collected from diarrheic cattle in Korea. *Infect. Genet. Evol.* **11**: 1178–1182.
21. **Parrish, C. R., et al.** 2008. Cross-species virus transmission and the emergence of new epidemic diseases. *Microbiol. Mol. Biol. Rev.* **72**:457–470.
22. **Parrish, C. R., and Y. Kawaoka.** 2005. The origins of new pandemic viruses: the acquisition of new host ranges by canine parvovirus and influenza A viruses. *Annu. Rev. Microbiol.* **59**:553–586.
23. **Reuter, G., A. Boldizsar, I. Kiss, and P. Pankovics.** 2008. Candidate new species of Kobuvirus in porcine hosts. *Emerg. Infect. Dis.* **14**:1968–1970.
24. **Reuter, G., A. Boldizsar, and P. Pankovics.** 2009. Complete nucleotide and amino acid sequences and genetic organization of porcine kobuvirus, a member of a new species in the genus Kobuvirus, family Picornaviridae. *Arch. Virol.* **154**:101–108.
25. **Reuter, G., A. Boros, and P. Pankovics.** 2011. Kobuviruses—a comprehensive review. *Rev. Med. Virol.* **21**:32–41.
26. **Sasaki, J., et al.** 2001. Construction of an infectious cDNA clone of Aichi virus (a new member of the family *Picornaviridae*) and mutational analysis of a stem-loop structure at the 5' end of the genome. *J. Virol.* **75**:8021–8030.
27. **Simmonds, P.** 2006. Recombination and selection in the evolution of picornaviruses and other mammalian positive-stranded RNA viruses. *J. Virol.* **80**:11124–11140.
28. **Tamura, K., et al.** 4 May 2011, posting date. MEGA5: molecular evolutionary genetics analysis using maximum likelihood, evolutionary distance, and maximum parsimony methods. *Mol. Biol. Evol.* [Epub ahead of print.] doi: 10.1093/molbev/msr121.
29. **Woo, P. C., et al.** 2010. Comparative analysis of six genome sequences of three novel picornaviruses, turdiviruses 1, 2 and 3, in dead wild birds, and proposal of two novel genera, Orthoturdivirus and Paraturdivirus, in the family Picornaviridae. *J. Gen. Virol.* **91**:2433–2448.
30. **Yamashita, T., et al.** 2003. Isolation and characterization of a new species of kobuvirus associated with cattle. *J. Gen. Virol.* **84**:3069–3077.
31. **Yamashita, T., et al.** 1998. Complete nucleotide sequence and genetic organization of Aichi virus, a distinct member of the Picornaviridae associated with acute gastroenteritis in humans. *J. Virol.* **72**:8408–8412.