Virology 477 (2015) 89–97

Contents lists available at ScienceDirect

Virology

journal homepage: www.elsevier.com/locate/yviro

Molecular characterization of canine kobuvirus in wild carnivores and the domestic dog in Africa

Ximena A Olarte-Castillo [a], Felix Heeger [a,b], Camila J Mazzoni [a,b], Alex D Greenwood [a], Robert Fyumagwa [c], Patricia D Moehlman [d], Heribert Hofer [a], Marion L East [a,*]

[a] Leibniz Institute for Zoo and Wildlife Research, Alfred-Kowalke-Strasse 17, D-10315 Berlin, Germany
[b] Berlin Center for Genomics in Biodiversity Research, Königin-Luise-Straße 6–8, 14195 Berlin, Germany
[c] Tanzania Wildlife Research Institute, P.O. Box 661, Arusha, Tanzania
[d] EcoHealth Alliance, 460 West 34th Street, New York, NY, USA

A R T I C L E I N F O

Article history:
Received 12 September 2014
Returned to author for revisions
23 December 2014
Accepted 9 January 2015
Available online 7 February 2015

Keywords:
Kobuvirus
Whole genome
ORF
IRES
Next-generation sequencing
Phylogeny
Domestic dog
Golden jackal
Side-striped jackal
Spotted hyena

A B S T R A C T

Knowledge of *Kobuvirus* (Family *Picornaviridae*) infection in carnivores is limited and has not been described in domestic or wild carnivores in Africa. To fill this gap in knowledge we used RT-PCR to screen fresh feces from several African carnivores. We detected kobuvirus RNA in samples from domestic dog, golden jackal, side-striped jackal and spotted hyena. Using next generation sequencing we obtained one complete *Kobuvirus* genome sequence from each of these species. Our phylogenetic analyses revealed canine kobuvirus (CaKV) infection in all four species and placed CaKVs from Africa together and separately from CaKVs from elsewhere. Wild carnivore strains were more closely related to each other than to those from domestic dogs. We found that the secondary structure model of the IRES was similar to the *Aichivirus*-like IRES subclass and was conserved among African strains. We describe the first CaKVs from Africa and extend the known host range of CaKV.

© 2015 Elsevier Inc. All rights reserved.

Introduction

Kobuvirus is a recently described and emerging genus in the family *Picornaviridae* (Knowles et al., 2012). Currently this genus is divided into three species, *Aichivirus A* (formerly *Aichi virus*), first described from gastroenteritis in humans in the Aichi Prefecture, Japan, in 1989 (Yamashita et al., 1991), *Aichivirus B* (formerly *Bovine kobuvirus*) and *Aichivirus C* (www.picornaviridae.com). Kobuviruses are small, non-enveloped viruses with a single stranded, polyadenylated, positive-sense RNA genome that encodes a single polyprotein (Reuter et al., 2011). The genome contains a long 5′ untranslated region (UTR) which contains the internal ribosome entry site (IRES), a single open reading frame (ORF) that encodes a polyprotein, and a 3′ UTR. Within the picornavirus family, there are four classes of IRES elements (Type I–IV) and a new IRES subclass was recently described in *Aichivirus* (Yu et al., 2011). The polyprotein is proteolytically cleaved into a leader protein (L), and the structural (P1) and non-structural (P2 and P3) regions (Fig. 1a, Reuter et al., 2011). The role of the 3′-UTR is currently unclear, although there is evidence of its involvement in replication (Brown et al., 2004). Furthermore, conserved barbell-like secondary structures in the 3′-UTR have been recently described in distantly related picornaviruses (Boros et al., 2012). Recombination has been widely reported in the *Picornaviridae* (Lukashev, 2010, Sweeney et al., 2012) and is considered to play a significant role in the diversification of this family (Oberste et al., 2004). Recombination is expected to occur in kobuviruses but assessment of its potential importance is hampered by the relatively small number of published complete *kobuvirus* genomes available (Lukashev, 2010).

Current knowledge of *Kobuvirus* infections in carnivores is extremely limited. *Kobuvirus* infection has been reported in three domesticated carnivores (the ferret, *Mustela putorius furo*, the domestic cat, *Felis catus* and the domestic dog, *Canis lupus familiaris*). The novel *Kobuvirus* described from ferrets in Europe belongs to the *Aichivirus B* species (Smits et al., 2013). Feline kobuvirus (FeKV) and canine kobuvirus (CaKV) belong to the *Aichivirus A* species. FeKV infection has only been reported in domestic cats in Asia (Chung et al., 2013). CaKV infection has been

* Corresponding author.
E-mail address: east@izw-berlin.de (M. East).

http://dx.doi.org/10.1016/j.virol.2015.01.010
0042-6822/© 2015 Elsevier Inc. All rights reserved.

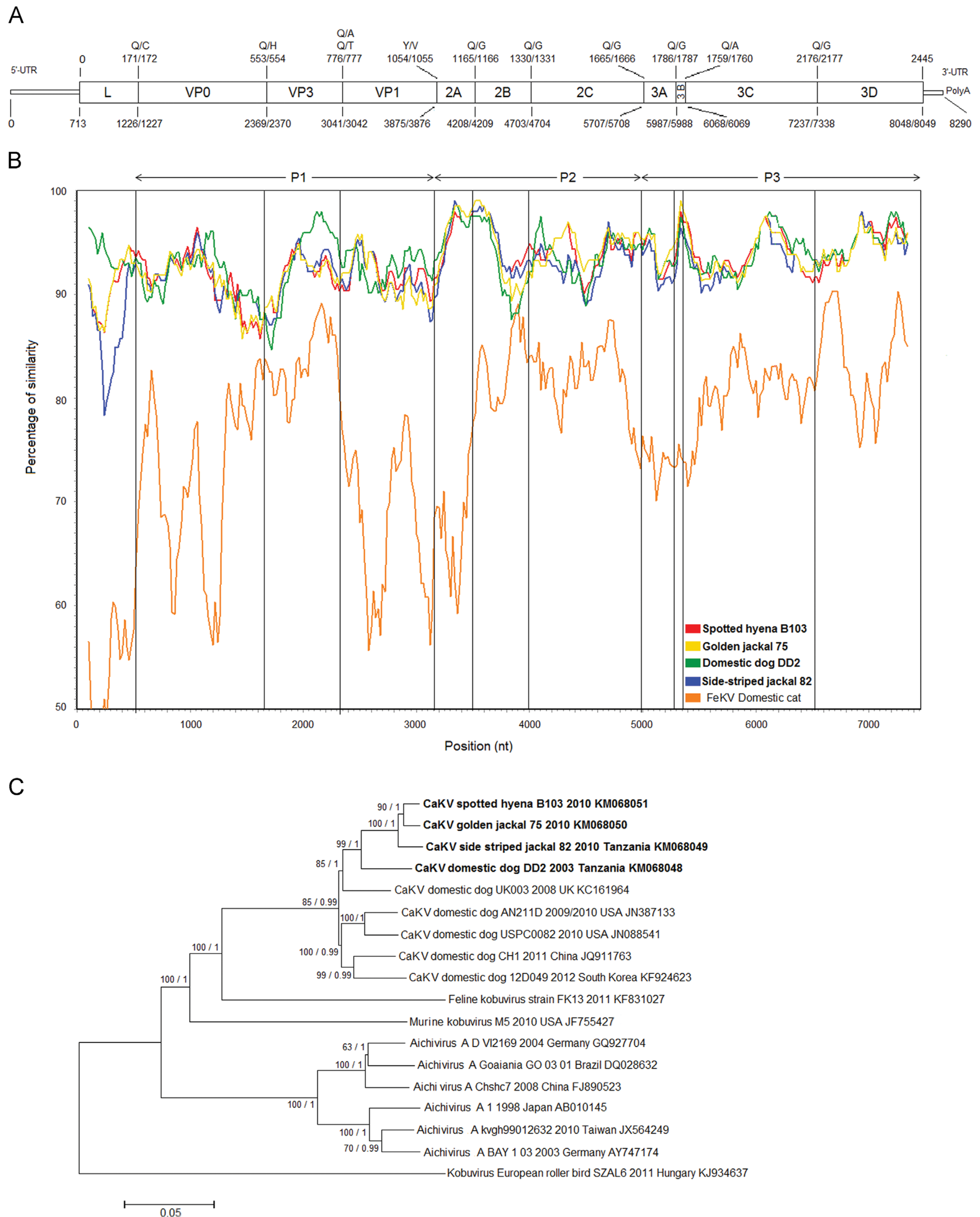

Fig. 1. Comparison of the complete single ORF of Tanzanian strains with known kobuvirus *Aichivirus A* strains. (a) Graphical depiction of the complete genome of canine kobuvirus (CaKV). The predicted positions of cleavage sites, amino-acid (above bar) and nucleotide (below bar) are shown. (b) Similarity plot analysis: The percentage similarity of complete ORF sequences from Tanzanian strains (plotted in different colors) compared against the most closely related strain UK003 (KC161964). CaKV ORF from Tanzanian host species: domestic dog (green); side-striped jackal (blue) golden jackal (yellow) spotted hyena (red). FeKV ORF, domestic cat from Asia (orange). (c) Phylogenetic relationships between kobuviruses from Tanzania, East Africa with strains within the *Aichivirus A* group based on complete ORF nucleotide sequences. Numbers at the nodes indicate bootstrap percentage values from 1000 replicates followed by Bayesian posterior probabilities. Strains obtained in this study are depicted in bold. GeneBank accession numbers for each complete kobuvirus ORF used in this analysis are quoted on their respective branches.

described in domestic dogs in a larger geographical area, including the USA (Kapoor et al., 2011; Li et al., 2011), Europe (Carmona-Vicente et al., 2013; Di Martino et al., 2013) and Asia (Oem et al., 2014). CaKV infection is often associated with diarrhea and other clinical signs (Carmona-Vicente et al., 2013; Li et al., 2011), even though asymptomatic infections also occur (Kapoor et al., 2011). To the best of our knowledge, currently the only described *Kobuvirus* infection in a wild carnivore is CaKV infection in the red fox (*Vulpes vulpes*) in Europe (Di Martino et al., 2014).

We are unaware of *Kobuvirus* infection reported in any wild carnivore or domestic dog in Africa. This study aimed to help fill this gap in knowledge by investigating *Kobuvirus* infection in five wild carnivore species and the domestic dog in East Africa. The wild carnivores included the bat-eared fox (*Otocyon megalotis*), golden jackal (*Canis aureus*), silver-backed jackal (*Canis mesomelas*) and side-striped jackal (*Canis adustus*) in the family Canidae, and the spotted hyena (*Crocuta crocuta*) in the family Hyaenidae. These wild carnivores inhabited protected areas within the Serengeti ecosystem whereas the domestic dogs lived in villages surrounding the ecosystem. Using next generation sequencing, genomic, and phylogenetic analyses we provide the first molecular characterization of the complete *Kobuvirus* genome from four African carnivore host species. We also compared partial sequences of the relatively conserved *Kobuvirus* 3D gene which encodes the RNA-dependent RNA polymerase, and the complete VP1 gene, which encodes the most variable *Kobuvirus* protein (Reuter et al., 2011).

Results

CaKV RNA was detected in feces from four of the six species screened, including the domestic dog, golden jackal, side-striped jackal and spotted hyena by RT-PCR. No CaKV RNA was detected in the small number of samples screened from bat-eared foxes and silver-backed jackals. Four of 15 (26.7%) samples from domestic dogs, one of eight (12.5%) samples from golden jackals and one of 14 (7.1%) samples from spotted hyena were positive for CaKV RNA. All domestic dog samples were from juvenile animals. Two positive samples from side-striped jackal were from juveniles belonging to the same litter.

The genetic analysis based on the four complete genome sequences obtained in this study revealed sequences of the same length for all strains: 8290 nucleotides (excluding the poly A tail), with a single ORF of 7335 nucleotides encoding a 2445 amino acid polyprotein. The predicted organization of genes in the genomes obtained by this study was identical to that previously described for CaKV (5′-UTR,VP0, VP3, VP1, 2A, 2B, 2C, 3A, 3B, 3C, 3D, 3′-UTR, polyA Fig. 1a). The length of the 5′- and 3′-UTR was 712 and 242 nucleotides, respectively. Between nucleotides 8118 and 8166 all African strains presented the conserved barbell-like structures (not shown) in the 3′-UTR. The pyrimidine-rich tract within this structure presented the short sequence UCUUU which started at nucleotide 8142 and was conserved in all strains except for the one domestic dog strain (which presented CUUUU). The cleavage site between VP3 and VP1 was not conserved among all strains. One strain from a spotted hyena and one strain from a side-striped jackal presented a Glutamine (Q)/Threonine (T) cleavage site (amino-acids 776/777) whereas one strain from a golden jackal and one strain from a domestic dog in Tanzania, presented a Q/Alanine (A) cleavage site (Fig. 1a). Our analysis to search for evidence of recombination events throughout the complete genome found none.

The similarity plot analysis (Fig. 1b) compared the four complete ORF nucleotide sequences from the Tanzanian strains (each plotted with a different color) and one FeKV ORF sequence (plotted in orange) with the complete ORF sequence from a domestic dog CaKV strain in the UK (UK003). This CaKV is currently the most closely related strain to the Tanzanian strains. In the L region and the VP3 gene, the domestic dog strain DD2 (plotted in green, Fig. 1b) was more similar to the reference sequence than the Tanzanian wild carnivore strains. In the rest of the genes all the strains presented a homogeneous range of similarity to the reference strain. In the L region, the strain from the side-striped jackal (plotted in blue, Fig. 1b) presented a considerably lower similarity than the other Tanzanian strains. The regions P1 and P2 presented a higher range of variability than the P3 region. The similarity of the Tanzanian strains to the reference CaKV was consistently greater than to the FeKV strain (plotted in orange, Fig. 1b). The average nucleotide/amino-acid similarity of the five available complete CaKV ORF sequences from outside Africa to each of the four complete ORF sequences obtained from African carnivores by this study was similar (93.3%/98.8% for the domestic dog strain; 92.6%/98.2% for the side-striped jackal strain; 92.9%/98.6% for the golden jackal strain; 93.0%/98.7% for the spotted hyena strain).

The average percentage similarity values of gene sequences across the entire ORF within Tanzanian strains and between these gene sequences and those in CaKV strains from elsewhere are tabulated in Table 1. The average percentage of nucleotide/amino acid similarity between the complete genome sequences from Tanzanian CaKV strains was higher than between Tanzanian strains and CaKV strains from outside Africa and between Tanzanian strains and FeKV (Table 1). Within Tanzanian strains the coding region of the single ORF (L, P1, P2 and P3) presented higher nucleotide similarity (98.3%–99.7%) than the 5′- and 3′-UTR (97.2%, 97.7% respectively). The average percentage similarity between the complete genome sequences of Tanzanian strains was mostly higher than between Tanzanian strains and complete genome sequences from strains outside Africa (Table 1). The average percentage nucleotide similarity of the complete genome of Tanzanian strains (Table 2) was lowest between the strains from the golden jackal and domestic dog (93.9%) and highest between the strains from a golden jackal and spotted hyena (98.3%). The average percentage amino acid amino-acid similarity of the polyprotein between Tanzanian strains ranged from 98.9% to 99.2% (Table 2).

Our phylogenetic inference based on four complete *Kobuvirus* ORF sequences was strongly supported (Fig. 1c). We also conducted analyses based on both a partial 3D gene fragment to permit comparison of our African strains with most CaKVs described from elsewhere (Fig. 2a) and on the complete VP1 gene

Table 1
Average percentage of amino-acid/nucleotide sequence similarity: within Tanzanian strains; between Tanzanian strains and CaKV from outside Africa; and between Tanzanian strains and one FeKV. Percentage similarities were not computed for amino acids in sequences outside coding regions (NA).

Average percentage of similarity (amino-acid/ nucleotide)	Within Tanzanian strains	Between Tanzanian strains and other CaKV	Between Tanzanian strains and FeKV
Complete single ORF	99.3/95.9	98.6/92.9	93.2/81.7
5′-UTR	97.2/NA	94.9[a]/NA	Missing data[b]
L	98.3/93.3	96.7/91.1	77.0/71.4
P1	98.9/95.0	97.5/90.8	92.0/80.3
P2	99.4/96.2	99.5/94.4	95.5/83.4
P3	99.7/97.3	99.4/94.5	96.3/84.1
3′-UTR	97.7/NA	95.1/NA	86.2/NA

[a] Comparison only with strains USPC0082 (JN088541) and 12D049 (KF924623) which have a complete 5′-UTR.
[b] Feline Kobuvirus FK13 (KF831027) has an incomplete 5′-UTR hence comparison not possible.

Table 2
The average percentage of amino acid (above the diagonal) and nucleotide (below the diagonal) similarity of the complete the 5′-UTR, 3′-UTR, L gene, and P1–P3 regions among Tanzanian strains.

	Domestic dog DD2	Side-striped jackal 82	Golden jackal 75	Spotted hyena B103
Complete genome/ polyprotein				
Domestic dog DD2	**100**	98.9	99.2	99.4
Side-striped jackal 82	94.8	**100**	99.4	99.2
Golden jackal 75	93.9	97.8	**100**	99.5
Spotted hyena B103	94.1	97.6	98.3	**100**
5′-UTR				
Domestic dog DD2	**100**	–	–	–
Side-striped jackal 82	95.6	**100**	–	–
Golden jackal 75	95.6	100	**100**	–
Spotted hyena B103	95.8	98.0	98	**100**
L				
Domestic dog DD2	**100**	97.1	99.4	99.4
Side-striped jackal 82	91	**100**	97.7	97.7
Golden jackal 75	91.2	94.5	**100**	98.8
Spotted hyena B103	91.4	94.7	97.1	**100**
P1				
Domestic dog DD2	**100**	98.3	98.5	99.1
Side-striped jackal 82	93.2	**100**	99.5	99.2
Golden jackal 75	91.9	97.3	**100**	99.2
Spotted hyena B103	92.6	97.5	97.6	**100**
P2				
Domestic dog DD2	**100**	99.5	99.3	99.3
Side-striped jackal 82	95.6	**100**	99.5	99.2
Golden jackal 75	94.5	97.4	**100**	99.7
Spotted hyena B103	93.9	97.4	98.6	**100**
P3				
Domestic dog DD2	**100**	99.6	99.9	99.9
Side-striped jackal 82	96.4	**100**	99.5	99.5
Golden jackal 75	95.5	98.5	**100**	99.7
Spotted hyena B103	95.8	98.2	99.1	**100**
3′-UTR				
Domestic dog DD2	**100**	–	–	–
Side-striped jackal 82	95.9	**100**	–	–
Golden jackal 75	97.5	98.3	**100**	–
Spotted hyena B103	96.7	98.3	99.2	**100**

(Fig. 2b) to compare strains from African wild carnivores with CaKV strains from red fox in Europe: the only wild carnivore host of *Kobuvirus* so far described. In all three phylogenies, the four sequences from Tanzanian carnivores (domestic dog, golden jackal, side-striped jackal and spotted hyena) were more closely related to CaKVs than to other kobuviruses, including FeKV. All sequences from Tanzania clustered together and separately from CaKVs from areas outside Africa (i.e., Europe, Asia and the USA). Furthermore, in all three phylogenies, the strains from wild carnivores inside the protected areas of the Serengeti ecosystem were more closely related to each other than to the strains from domestic dogs outside the protected area, which were placed on separate branches. In the analysis of the partial 3D gene (Fig. 2a), the placement of two domestic dog strains (DD2 and DD3) indicated that they were more closely related to the strains from wild carnivores than to two other domestic dog strains (DD20 and DD22) which were more distantly related to strains from wild carnivores in the Serengeti ecosystem. The analysis of the complete VP1 gene also revealed that strain DD2 was more closely related to strains from wild carnivores in the Serengeti ecosystem than to strain DD22. Interestingly, analysis of the VP1 gene (Fig. 2b) placed the cluster of sequences from Tanzania closer to gene sequences from red foxes and domestic dogs in Europe than to sequences from domestic dogs in Asia and the USA. Our analysis of the complete ORF reveals some evidence of geographical clustering, given that the two sequences from Asia (China and South Korea) cluster together, as do two sequences from USA and the one sequence from the UK is placed separately from those from Africa, Asia and the USA.

We investigated the predicted secondary structures of the core IRES of all our African strains. Using homologous sequences, the secondary structure model (Fig. 3) indicated that all strains presented a *Aichivirus*-like IRES with the presence of principal domains H–L. The initiation codon (AUG_{713}) was contained in the L domain and the functionally important GNRA tetraloop was in the apical subdomain of the J domain. Of the 541 nucleotides that form the core IRES, only 57 nucleotide variations occurred when considering sequences from both Tanzanian strains and published CaKV strains.

Discussion

Our results provide the first genetic description of infection of domestic dogs in Africa with CaKV. Before our study, the known host range of CaKV included two species in the Canidae (domestic dog and red fox, Di Martino et al., 2014). Our results extend the known canid hosts to the golden jackal and side-striped jackal. We also identified the first non-canid CaKV host, the spotted hyena (family Hyaenidae, sub-order Feliformia). *Picornaviridae* has been described as a family of host-specific viruses, infecting single or closely-related hosts (Lewis- Rogers and Crandall, 2010). Our results provide evidence that kobuviruses may not be as host specific as previously thought. This suggestion is supported by *Aichivirus A* infection in humans and a growing number of mammal species (Phan et al., 2011) as well as a recently reported kobuvirus closely related to *Aichivirus A* in an avian species (Pankovics et al., in press): the European roller (*Coracias garrulus*). Also *Aichivirus B* infection has been reported in several members of the Bovidae, including cattle (*Bos Taurus*), sheep (*Ovis aries*) and black goats (*Capra aegagrus hircus*) (Lee et al., 2012; Reuter et al., 2010) and one domestic carnivore, the ferret (Smits et al., 2013).

The results of our phylogenetic analysis based on the relatively small number of available complete *Kobuvirus* single ORF sequences revealed that strains in Africa were distinct from CaKVs from geographical regions outside Africa and that strains from specific geographical regions (i.e., Africa, Asia and the USA) clustered together (Fig. 1c). Results of the analyses of both the partial 3D gene (Fig. 2a) and the complete VP1 gene (Fig. 2b) revealed that CaKV strains obtained from the three wild carnivores in the Serengeti ecosystem were placed separately from sequences obtained from domestic dogs, suggesting that strains circulating in wild carnivores in protected areas of the Serengeti ecosystem are distinct from those in domestic dogs in surrounding villages. Furthermore, results from the well-conserved 3D gene (Fig. 2a) revealed that two strains from Tanzanian domestic dogs (DD20 and DD22) were not placed together with the other two strains from Tanzanian domestic dogs (DD2 and DD3) or with the other four strains from the Tanzanian wild carnivores, suggesting that genetically distinct CaKV strains circulate in the local domestic dog population. This interpretation is supported by the phylogenetic results of the VP1 gene in which strain DD2 was placed more closely to Tanzanian strains from wild carnivores than DD22 (Fig. 2a).

The P2 and P3 regions in the picornavirus ORF are more highly conserved than the P1 region which contains genes encoding the capsid proteins (Brown et al., 2003). Comparison of the nucleotide sequences of the four strains obtained in this study and other CaKV strains revealed that the P1 region was more variable (Table 1, Fig. 1b) than either the P2 or P3 regions. Even so, this

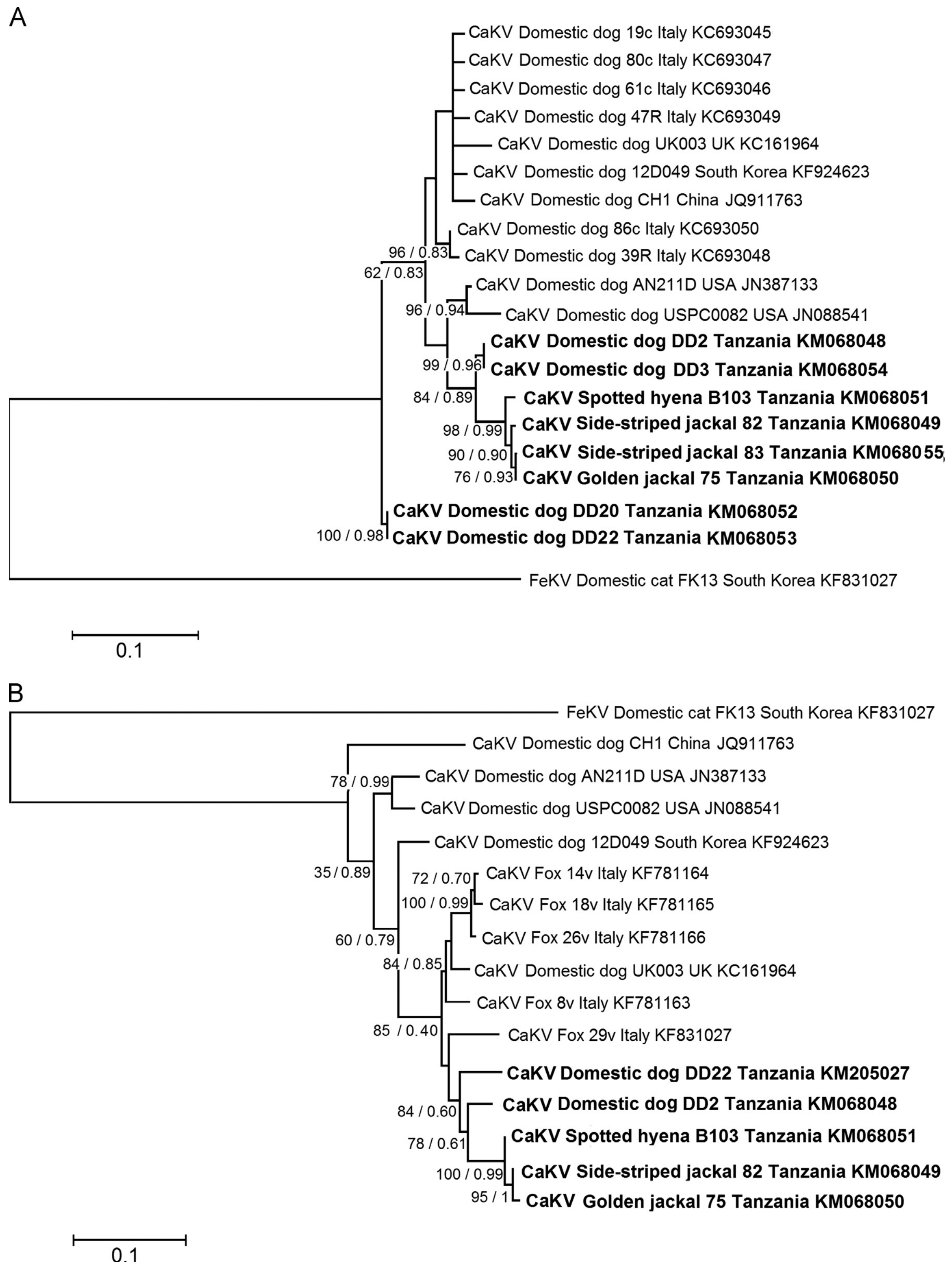

Fig. 2. Maximum likelihood (ML) phylogenies under the GTR+I+G model for (a) a fragment of 477 nucleotides of the 3D gene and (b) for the complete VP1 gene. Numbers at the nodes indicate bootstrap percentage values from 1000 replicates followed by Bayesian posterior probabilities. Only values greater than 60/0.6 are shown. Strains obtained in this study are depicted in bold. GeneBank accession numbers for each sequence used in this analysis are quoted on their respective branches.

variation appears to be moderate in comparison with the range of variability in the P1 region in the *Aichivirus C* group (Fan et al., 2013) and other enteroviruses (*Enterovirus B* species, genus *Enterovirus*, family *Picornaviridae*) (Oberste, 2008) than in CaKV. Taken together, this information indicates that CaKV nucleotide and amino-acid variability throughout the ORF is less than that in other picornaviruses. The high similarity throughout the ORF, and the relatively limited number of complete genome sequences currently available for CaKV, might prevent the detection of recombination events (Lukashev, 2010). More extensive sampling of carnivores, including felids, will help clarify the host range of CaKV and the occurrence of recombination in this group.

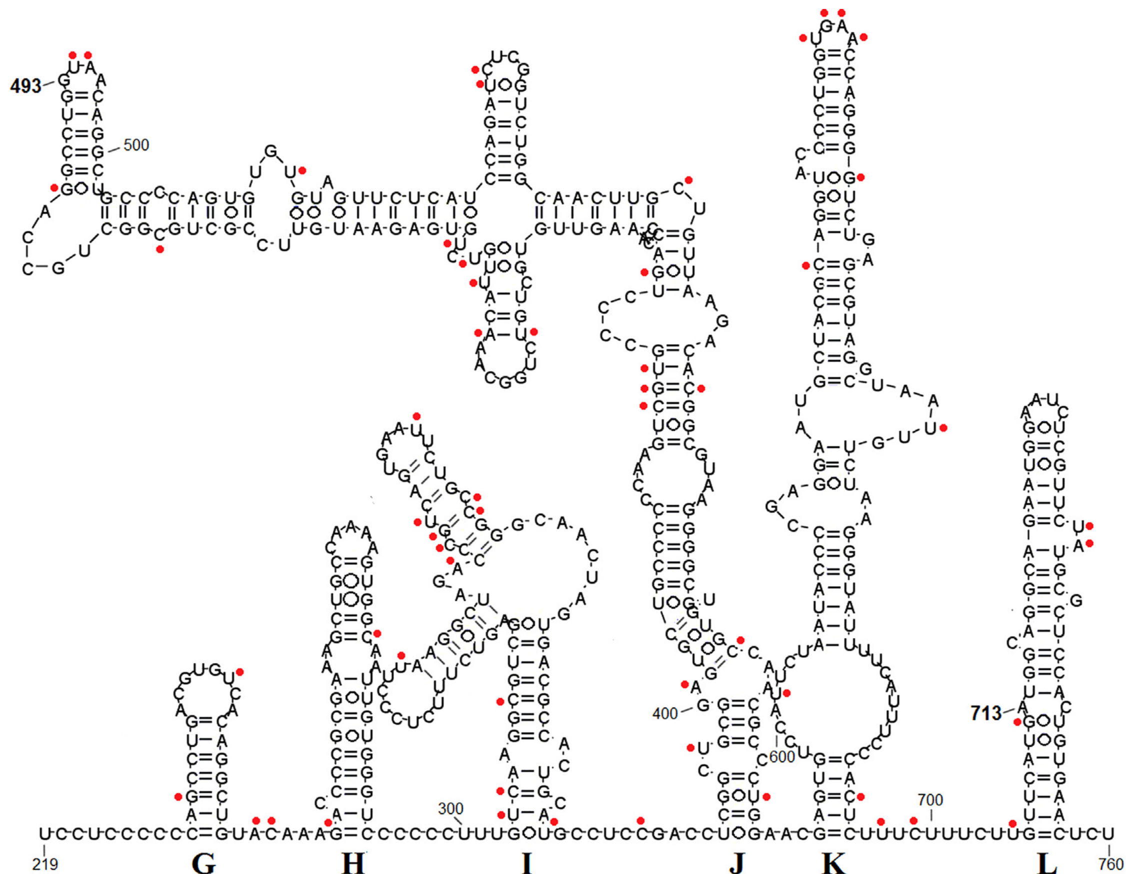

Fig. 3. Predicted RNA secondary structure of the designated *Aichivirus*-like IRES of a *Kobuvirus* strain in a spotted hyena B103 (KM068051). Each domain is labeled according to standard nomenclature (G-L). Red circles denote non-identical nucleotides in published CaKV. Non highlighted nucleotides are identical to those in published CaKV. The initiation codon starts at position 713. The GNRA motif starts at position 493.

In general, the size of the complete ORF as well as each gene of the four strains obtained in this study (7335 nucleotides/2445 amino-acids) was the same as that for other published CaKV (Kapoor et al., 2011; Li et al., 2011) and the polyprotein presented an identical organization to other kobuviruses (Fig. 1a). The predicted cleavage sites were conserved among published CaKV strains, including the proposed unusual cleavage site between VP1 and 2A (Kapoor et al., 2011). However, the predicted cleavage site between VP3 and VP1 in the strains obtained from the spotted hyena and the side-striped jackals was not typical compared to all other previously described CaKV from domestic dogs, as both presented the less frequent amino-acid combination Q/T. This atypical cleavage site was also found in human aichivirus via direct N-terminal amino-acid sequencing of purified virions (Yamashita et al., 1998) and a recently described *Kobuvirus* obtained from sewage in Japan presented the amino-acid combination Q/V at this cleavage site (Yamashita et al., 2014).

Overall, the Tanzanian strains and other CaKV strains presented a high nucleotide and amino-acid similarity. Interestingly, even though the 5′- and the 3′-UTR had a lower similarity than the coding region of the single ORF (Table 1), the secondary structure model of the *Aichi-virus*-like IRES (Fig. 3) in the 5′-UTR and the barbell-like structures in the 3′-UTR were conserved among all Tanzanian strains. Both these structures were also conserved in all published CaKoV. Moreover, the *Aichivirus*-like IRES has been described in *Murine Kobuvirus* and barbell-like structures in both *Murine Kobuvirus* and *Aichivirus* (Boros et al., 2012; Sweeney et al., 2012).

Conclusion

In this study we report for the first time the identification and genetic characterization of CaKV in four African carnivore species. We extend the known host range of CaKV to include two further canid species, the golden jackal and side-striped jackal, and one non-canid species, the spotted hyena. Our results considerably expand the number of complete CaKV genome sequences available and show that CaKV appears to be an unusually homogeneous picornavirus group.

Methods

Study site and sample collection

The study was conducted in the Tanzanian sector of the Serengeti ecosystem which covers approximately 25,000 km^2, and straddles the border between Tanzania and Kenya in East Africa. A large proportion of the ecosystem (14,763 km^2) is contained in the Serengeti National Park (Serengeti NP) in Tanzania, and most of the ecosystem is under some category of protection (East et al., 2012). From 2001 to 2012 fresh fecal samples were collected shortly after deposition by African carnivores, including four wild canid species (bat-eared foxes, $n=5$; golden jackals, $n=8$; silver-backed jackals, $n=2$; side-striped jackals, $n=2$) and one non-canid species (spotted hyena, $n=14$). Fresh feces were also collected from domestic dogs in 2003 ($n=15$) in villages surrounding the

Serengeti NP. Samples from bat-eared foxes, silver-backed jackals, side-striped jackals and spotted hyena were collected in the center of the Serengeti NP, where domestic dogs are prohibited (East et al., 2012). Samples from golden jackals were collected in the Serengeti NP, close to the south-western boundary with the Ngorongoro Conservation Area where domestic dogs owned by pastoralists occur. Samples were stored and transported frozen at -80 °C, or were preserved in RNAlater (Sigma-Aldrich Inc., St. Louis, MO, USA), initially stored at −10 °C, and finally stored at −80 °C until laboratory analyses were performed.

Detection of Kobuvirus RNA

Using the metagenomics protocol described by Li et al. (2011), we detected sequences which mapped in the 3D gene region of the CaKV genome in one fecal sample from a domestic dog. To design primers to screen carnivore fecal samples for the presence of CaKV RNA we used these initial sequence data to target a partial fragment of the conserved 3D gene (KobVF 5′-CTCRCGYCGATCYCTCTTTG-3′, KobVR 5′-GTCRCCRTACGCGAGGATCT-3′). For the positive samples obtained, we also amplified the highly variable VP1 gene using previously described primers (Di Martino et al., 2014). Total RNA was extracted from 200 µl of 10% (w/v) fecal suspension in phosphate buffered saline solution using the QIAamp Viral RNA Mini Kit (Qiagen GmbH, Hilden, Germany), according to manufacturer's instructions. RT-PCR was performed using SuperScript™ III One-Step RT-PCR System (Life Technologies GmbH, Darmstadt, Germany) following the manufacturer's instructions, in a total reaction volume of 25 µl. Amplicons were purified using MSB® Spin PCRapace (Stratec Molecular GmbH, Berlin, Germany). In order to avoid RNases, all surfaces were cleansed with RNase away (Thermo Fisher Scientific GmbH, Ulm, Germany). The purified products were cloned using the QIAGEN PCR Cloning*plus* Kit (Qiagen GmbH, Hilden, Germany) according to manufacturer's instructions. Clones were confirmed by PCR with M-13 universal primers and subsequent sequencing using the 3130 Genetic Analyzer (Applied Biosystems, Darmstadt, Germany). Both strands of a total of five clones per virus strain per gene were sequenced. Vector sequences were trimmed using VecScreen (http://www.ncbi.nlm.nih.gov/VecScreen/VecScreen.html). Subsequently, sequences were assembled in the CLC Main Workbench 7.0 (CLC Bio, Aarhus, Denmark, http://www.clcbio.com/).

Genome sequencing

From the samples that tested positive when screened, one sample per species ($n=4$) was selected for sequencing of the complete genome. For this, cDNA was synthesized from total RNA using the Expand Reverse Transcriptase (Roche Diagnostics Deutschland GmbH, Mannheim, Germany) and random primers (Life Technologies GmbH, Darmstadt, Germany). Long-range PCRs were performed using Expand Long Range dNTPack (Roche Diagnostics Deutschland GmbH, Mannheim, Germany) in a final volume of 50 µl with the primers KoV61F (5′-RGACCACCGTYACTCCATTC-3′)/ KoV4750R (5′-GGGTYRCCYGTTCACAYTCT-3′) and KoV4584F (5′-HACYCCCATRTCYATCTCYG-3′) / KoV8232R (5′-ATGGCTTAGGGRCTCACC-3′). Amplicons were purified using QIAquick PCR Purification Kit (Qiagen, Hilden, Germany). Using hybrid capture (Maricic et al., 2010; Tsangaras et al., 2014) and 3′-RACE the 5′-UTR and the 3′UTR were obtained, respectively. The 3′-RACE (Life Technologies GmbH, Darmstadt, Germany) was performed according to the manufacturer´s instructions using the primer KEndF (5′-GTCATTTGGCCTTCCACAGY-3′) for the first PCR and the primer KEnd2F (5′-CTCACCTGTGGTGCAATACC-3′) for the nested PCR. Purified products were cloned using the QIAGEN PCR Cloning*plus* Kit and sequenced using a 3130 Genetic Analyzer (Applied Biosystems, Darmstadt, Germany). For the hybrid capture cDNA was synthesized from total RNA using the SuperScript® reverse transcriptase (Life Technologies GmbH, Darmstadt, Germany) according to the manufacturer´s protocol, followed by Klenow Fragment (5U). The long-range PCR products were used as baits. Purified double-stranded cDNA and long-range PCR products were subsequently sheared in a M220 ultrasonicator™ (Life Technologies GmbH, Darmstadt, Germany) using the standard program to generate fragments of 200 bp and 300 bp respectively. Libraries were indexed by sample and were built as described previously (Meyer and Kircher, 2010) and DNA quantity and quality was monitored using the Agilent 2200 TapeStation (Agilent Technologies, Böblingen, Germany). Pooled libraries of the long-range PCR products as well as enriched libraries from the hybrid capture were deep-sequenced on the Illumina MiSeq™ platform (Illumina Inc., San Diego, USA). Quality assessment of the sequences was assessed with FastQC 0.10.1 (freely available at http://www.bioinformatics.babraham.ac.uk/projects/fastqc). The adapters and any following sequence were removed with Cutadapt 1.3 (Martin, 2011) and read pairs where one read was shorter than 60 bp after this were also removed. Read pairs overlapping for at least 10 bases were combined into one sequence with PANDAseq 2.7 (Masella et al., 2012). Reads that could not be overlapped were trimmed with Trimmomatic 0.32 (Bolger et al., 2014) using a sliding window approach (window size 8, minimum average quality 20) and leading and trailing bases with a quality below 20 were removed. Trimmed reads from each sample were assembled *de novo* using the medium sensitivity method in Geneious 7.1.2 (Biomatters Ltd., Auckland, New Zealand).

Genetic analyses

We obtained four complete genomes from CaKVs, including three sequences from wild carnivores (golden jackal, side-striped jackal and spotted hyena) and one from a domestic dog. For each of the four genomes obtained in this study, we predicted picornaviral polyprotein cleavage sites within the ORF using NetPicoRNA 1.0 (Blom et al., 1996). Similarities along the genome were visualized with SimPlot 5.1 (Lole et al., 1999) using a Kimura-2-parameter distance model with a sliding window of 200 nucleotides and a step size of 20 bases. Average nucleotide and amino-acid similarities were calculated within the four strains described by this study, and between these Tanzanian CaKVs and the five complete CaKV genome sequences (accession numbers USA JN088541, JN387133; UK KC161964; China JQ911763; South Korea KF924623) and the one available nearly complete FeKV genome sequence (accession number KF831027), using Discovery Studio Visualizer 4.0 (Accelrys Software Inc, San Diego, USA). We searched for evidence of recombination including all of the 10 sequences mentioned above plus the 11 available complete genome sequences from *Aichi virus* from humans (accession numbers JX564249, GQ927712, GQ927711, GQ927706, GQ927705, GQ927704, FJ890523, AB010145, DQ028632, AB040749, AY747174). These sequences were aligned using the MUSCLE algorithm (Edgar, 2004) in Geneious 7.1.2 (Biomatters Ltd, Auckland, New Zealand) and possible recombination break points were then determined using the Recombination Detection Program 3.44 (Martin et al., 2010) using the RDP method (Martin and Rybicki, 2000).

The secondary structure of the 5′- and 3′-UTRof the four strains obtained in this study were reconstructed by aligning their sequences with six homologous sequences previously published for *Kobuvirus* (*Aichivirus* AB040749, *Murine Kobuvirus* JF755427, CaKV JQ911763, JN088541, JN387133, KC161964). A manual alignment was performed including restrictions and constrains from published secondary structures (Sweeney et al., 2012). RNA was folded in MFOLD (Zuker 2003) at a fixed temperature of 37 °C.

When more than one structure was possible, the structure with the highest negative free energy was selected. Secondary structures were viewed in VARNA (Darty et al., 2009).

Phylogenetic analyses

To investigate the phylogenetic relationship between CaKV strains from Tanzania with strains worldwide, we constructed a phylogeny based on the nucleotide sequences of the partial 3D gene (477 nucleotides), the complete VP1 gene (834 nucleotides) and the single ORF (i.e., the complete coding part of the genome excluding the 5′ and 3′ UTRs). The single ORF was used instead of the complete genome, because it allowed for comparison with more sequences. *Kobuvirus* sequences obtained in this study, together with those retrieved from Genbank were aligned using the MUSCLE algorithm (Edgar, 2004) in Geneious 7.1.2 (Biomatters Ltd., Auckland, New Zealand). For the analysis of the partial 3D gene we included eight CaKV sequences from carnivore hosts in Tanzania (domestic dogs $n=4$, golden jackal $n=1$, side-striped jackal $n=2$, spotted hyena $n=1$) produced by this study, and publically available CaKV sequences ($n=11$) from domestic dogs in Europe, Asia and the USA as well as one available FeKV sequence from Asia. Our analysis of the complete VP1 gene included one sequence from each of the three wild carnivore hosts in Tanzania (side-striped jackal, golden jackal, spotted hyena), two sequences from domestic dogs in Tanzania, plus published CaKV sequences from red foxes in Europe ($n=5$) and from domestic dogs in Europe, Asia and the USA ($n=5$) as well as one FeKV from a domestic cat in Asia. For the partial 3D gene, the complete VP1 gene and the complete single ORF sequences we reconstructed the phylogenetic relationships using ML and Bayesian MCMC phylogenetic inferences. The ML analysis was performed in PAUP*4.0b10 (Swofford, 2002) using 1000 bootstrap replicates to estimate the statistical support of the nodes. The Bayesian analysis was carried out using MrBayes 3.1 (Huelsenbeck and Ronquist, 2001; Ronquist, Huelsenbeck, 2003). The MCMC search was set to 10,000,000 iterations, with trees sampled every 1000th iteration. The nucleotide substitution model used in the ML analysis was obtained using ModelTest 3.7 (Posada and Crandall, 1998) and for the Bayesian analysis using MrModeltest 2.3 (Nylander, 2004). For both cases the Akaike information criterion was used to select the best-fitting model.

All the sequences obtained in this study were deposited in GenBank with the following accession numbers: four complete genome sequences numbered from KM068048 to KM068051, four partial 3D sequences numbered from KM068052 to KM068055 and one VP1 sequence as KM205027.

Acknowledgments

We are grateful to the Tanzania Commission of Science and Technology, the Tanzania Wildlife Research Institute, the Tanzania National Parks and the Ngorongoro Conservation Area Authority for permission to conduct this study. We thank the editor and two anonymous referees for their helpful comments and Malvina Andris, Nelly Boyer, Bjørn Figenschou, Annie Francis, Janine Helms, Thomas Shabani and Dagmar Thierer for their assistance. The study was conducted within the DFG-GRAKO 1121 “Genetic and Immunologic Determinants of Pathogen–Host-Interactions” at the Center for Infection Biology and Immunity (ZIBI) of the Humboldt University Berlin and at the Berlin Center for Genomics in Biodiversity Research (BeGenDiv) and was financed by the DFG, the Leibniz Institute for Zoo & Wildlife Research, Berlin, Germany and the EcoHealth Alliance. This is publication number 16 of the BeGenDiv.

References

Blom, N., Hansen, J., Blaas, D., Brunak, S., 1996. Cleavage site analysis in picornaviral polyproteins: discovering cellular targets by neural networks. Protein Sci. 5, 2203–2216.

Bolger, A.M., Lohse, M., Usadel, B., 2014. Trimmomatic: a flexible trimmer for Illumina sequence data. Bioinformatics 1, 2114–2120.

Boros, A., Nemes, C., Pankovics, P., Kapusinszky, B., Delwart, E., Reuter, G., 2012. Identification and complete genome characterization of a novel picornavirus in turkey (*Meleagris gallopavo*). J. Gen. Virol. 93, 2171–2182.

Brown, B.A., Oberste, M.S., Maher, K., Pallansch, M.A., 2003. Complete genomic sequencing shows that polioviruses and members of human enterovirus species C are closely related in the non-capsid coding region. J. Virol. 77, 8973–8984.

Brown, D.M., Kauder, S.E., Cornell, C.T., Jang, G.M., Racaniello, V.R., Semler, B.L., 2004. Cell-dependent role for the poliovirus 3′ noncoding region in positive-strand RNA synthesis. J. Virol. 78, 1344–1351.

Carmona-Vicente, N., Buesa, J., Brown, P.A., Merga, J.Y., Darby, A.C., Stavisky, J., Sadler, L., Gaskell, R.M., Dawson, S., Radford, A.D., 2013. Phylogeny and prevalence of kobuviruses in dogs and cats in the UK. Vet. Microbiol. 164, 246–252.

Chung, J.Y., Kim, S.H., Kim, Y.H., Lee, M.H., Lee, K.K., Oem, J.K., 2013. Detection and genetic characterization of feline kobuviruses. Virus Genes 47, 559–562.

Darty, K., Denise, A., Ponty, Y., 2009. VARNA: interactive drawing and editing of the RNA secondary structure. Bioinformatics 25, 1974–1975.

Di Martino, B., Di Profio, F., Melegari, I., Robetto, S., Di Felicem, E., Orusa, R., Marsilio, F., 2014. Molecular evidence of Kobuvirus in free-ranging red foxes (*Vulpes vulpes*). Arch. Virol. 159, 1803–1806.

Di Martino, B., Di Felice, E., Ceci, C., Di Profio, F., Marsilio, F., 2013. Canine kobuviruses in diarrhoeic dogs in Italy. Vet. Microbiol. 166, 246–249.

East, M.L., Nyahongo, J.W., Goller, K.V., Hofer, H., 2012. Does the vastness of the Serengeti limit human–wildlife conflicts? In: Somers, M.J., Hayward, M.W. (Eds.), Fencing for Conservation: Restriction of Evolutionary Potential or a Riposte to Threatening Processes. Springer, Heidelberg, New York, pp. 125–151.

Edgar, R.C., 2004. MUSCLE: a multiple sequence alignment method with reduced time and space complexity. BMC Bioinform. 5, 113.

Fan, S., Sun, H., Ying, Y., Gao, X., Wang, Z., Yu, Y., Li, Y., Wang, T., Yu, Z., Yang, S., 2013. Identification and characterization of porcine kobuvirus variants isolated from suckling piglet in Gansu Province, China. Viruses 5, 2548–2560.

Huelsenbeck, J.P., Ronquist, F., 2001. MRBAYES: Bayesian inference of phylogeny. Bioinformatics 17, 754–755.

Kapoor, A., Simmonds, P., Dubovi, E.J., Qaisar, N., Henriquez, J.A., Medina, J., Shields, S., Lipkin, W.I., 2011. Characterization of a canine homolog of human Aichi virus. J. Virol. 85, 11520–11525.

Knowles, N.J., Hovi, T., Hyypiä, T., King, A.M.Q., Lindberg, A.M., Pallansch, M.A., Palmenberg, A.C., Simmonds, P., Skern, T., Stanway, G., Yamashita, T., Zell, R., 2012. Picornaviridae. In: King, A.M.Q., Adams, M.J., Carstens, E.B., Lefkowitz, E.J. (Eds.), Virus Taxonomy: Classification and Nomenclature of Viruses: Ninth Report of the International Committee on Taxonomy of Viruses. Elsevier Academic Press, San Diego, pp. 855–880.

Lee, M.H., Jeoung, H.Y., Lim, J.A., Song, J.Y., Song, D.S., An, D.J., 2012. Kobuvirus in South Korean black goats. Virus Genes 45, 186–189.

Lewis-Rogers, N., Crandall, K.A., 2010. Evolution of *Picornaviridae*: an examination of phylogenetic relationships and cophyhlogeny. Mol. Phylogenet. Evol. 54, 995–1005.

Li, L., Pesavento, P.A., Shan, T., Leutenegger, C.M., Wang, C., Delwart, E., 2011. Viruses in diarrhoeic dogs include novel kobuviruses and sapoviruses. J. Gen. Virol. 92, 2534–2541.

Lole, K.S., Bollinger, R.C., Paranjape, R.S., Gadkari, D., Kulkarni, S.S., Novak, N.G., Ingersoll, R., Sheppard, H.W., Ray, S.C., 1999. Full-length human immunodeficiency virus type 1 genomes from subtype C-infected seroconverters in India, with evidence of intersubtype recombination. J. Virol. 73, 152–160.

Lukashev, A.N., 2010. Recombination among picornaviruses. Rev. Med. Virol. 20, 327–337.

Maricic, T., Whitten, M., Pääbo, S., 2010. Multiplexed DNA sequence capture of mitochondrial genomes using PCR products. PLoS One 5, e14004.

Martin, D.P., Lemey, P., Lott, M., Moulton, V., Posada, D., Lefeuvre, P., 2010. RDP3: a flexible and fast computer program for analyzing recombination. Bioinformatics 26, 2462–2463.

Martin, D.P., Rybicki, E., 2000. RDP: detection of recombination amongst aligned sequences. Bioinformatics 16, 562–563.

Martin, M., 2011. Cutadapt removers adapter sequences from high-throughput sequencing reads. EMBnet 17, 10–12.

Masella, A.P., Bartram, A.K., Truszkowski, J.M., Brown, D.G., Neufeld, J.D., 2012. PANDAseq: paired-end assembler for Illumina sequences. BMC Bioinform. 13, 31.

Meyer, M., Kircher, M., 2010. Illumina sequencing library preparation for highly multiplexed target capture and sequencing. Cold Spring Harb. Protoc. , http://dx.doi.org/10.1101/pdb.prot5448.

Nylander, J.A.A., 2004. MrModeltest v2. Program distributed by the author. Evolutionary Biology Centre, Uppsala University.

Oberste, M.S., 2008. Comparative genomics of the coxsackie B viruses and related enteroviruses. Curr. Top. Microbiol. Immunol. 323, 33–47.

Oberste, M.S., Maher, K., Pallansch, M.A., 2004. Evidence for frequent recombination within species human enterovirus B based on complete genomic sequences of all thirty-seven serotypes. J. Virol. 78, 855–867.

Oem, J.K., Choi, J.W., Lee, M.H., Lee, K.K., Choi, K.S., 2014. Canine kobuvirus infections in Korean dogs. Arch. Virol. 159, 2751–2755.

Pankovics, P., Boros, A., Kiss, T., Reuter, G., 2015. Identification and complete genome analysis of kobuvirus in faecal samples of European roller (*Coracias garrulus*): for the first time in bird. Arch. Virol. , http://dx.doi.org/10.1007/s00705-014-2228-7 (in press).

Phan, T.G., Kapusinszky, B., Wang, C., Rose, R.K., Lipton, H.L., et al., 2011. The fecal viral flora of wild rodents. PLoS Pathog. 7 (9), e1002218.

Posada, D., Crandall, K.A., 1998. MODELTEST: testing the model of DNA substitution. Bioinformatics 14, 817–818.

Reuter, G., Boros, A., Pankovics, P., 2011. Kobuviruses – a comprehensive review. Rev. Med. Virol. 21, 32–41.

Reuter, G., Boros, A., Pankovics, P., Egyed, L., 2010. Kobuvirus in domestic sheep, Hungary. Emerg. Infect. Dis. 16, 869–870.

Ronquist, F., Huelsenbeck, J.P., 2003. MRBAYES 3: Bayesian phylogenetic inference under mixed models. Bioinformatics 19, 1572–1574.

Smits, S.L., Raj, V.S., Oduber, M.D., Schapendonk, C.M., Bodewes, R., Provacia, L., Stittelaar, K.J., Osterhaus, A.D., Haagmans, B.L., 2013. Metagenomic analysis of the ferret fecal viral flora. PLoS One 8, e71595.

Sweeney, T.R., Dhote, V., Yu, Y., Hellen, C.U., 2012. A distinct class of internal ribosomal entry site in members of the *Kobuvirus* and proposed *Salivirus* and *Paraturdivirus* genera of the *Picornaviridae*. J. Virol. 86, 1468–1486.

Swofford, D.L., 2002. PAUP: Phylogenetic Analysis using Parsimony (and other Methods), Version 4.0 beta. Sinauer Associates Incorporated, Sunderland, Massachusetts.

Tsangaras, K., Siracusa, M.C., Nikolaidis, N., Ishida, Y., Cui, P., Vielgrader, H., Helgen, K.M., Roca, A.L., Greenwood, A.D., 2014. Hybridization capture reveals evolution and conservation across the entire Koala retrovirus genome. PLoS One 21, e95633.

Yamashita, T., Sakae, K., Tsuzuki, H., Suzuki, Y., Ishikawa, N., Takeda, N., Miyamura, T., Yamazaki, S., 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.

Yamashita, T., Adachi, H., Hirose, E., Nakamura, N., Ito, M., Yaui, Y., Kobayashi, S., Minagawa, H., 2014. Molecular detection and nucleotide sequence analysis of a new Aichi virus closely related to canine kobuvirus in sewage samples. J. Med. Micriobiol. 63, 715–720.

Yamashita, T., Kobayashi, S., Sakae, K., Nakata, S., Chiba, S., Ishihara, Y., Isomura, S., 1991. Isolation of cytopathic small round viruses with BS-C-1 cells from patients with gastroenteritis. J. Infect. Dis. 164, 954–957.

Yu, Y., Sweeney, T.R., Kafasla, P., Jackson, R.J., Pestova, T.V., Hellen, C.U., 2011. The mechanism of translation initiation on Aichi virus RNA mediated by a novel type of picornavirus IRES. EMBO J. 30, 4423–4436.

Zuker, M., 2003. Mfold web server for nucleic acid folding and hybridization prediction. Nucl. Acids Res. 31, 3406–3415.