Genomic characterization of TT viruses (TTVs) in pigs, cats and dogs and their relatedness with species-specific TTVs in primates and tupaias

Hiroaki Okamoto,1 Masaharu Takahashi,1 Tsutomu Nishizawa,1 Akio Tawara,2 Katsuhiko Fukai,3 Umetaro Muramatsu,3 Yoshihisa Naito4 and Akira Yoshikawa5

1 Immunology Division and Division of Molecular Virology, Jichi Medical School, Tochigi-Ken 329-0498, Japan
2 First Department of Internal Medicine, Yamanashi Medical University, Yamanashi-Ken 409-3898, Japan
3 Central Animal Hygiene Service Center of Tochigi Prefecture, Tochigi-Ken 321-0905, Japan
4 Department of Clinical Veterinary Medicine, Faculty of Agriculture, Iwate University, Iwate-Ken 020-8550, Japan
5 Japanese Red Cross Saitama Blood Center, Saitama-Ken 338-0001, Japan

Using PCR with primers derived from a non-coding region of the human TT virus (TTV) genome, the TTV sequence in serum samples obtained from pigs (Sus domesticus), dogs (Canis familiaris) and cats (Felis catus) was identified and the entire genomic sequence was determined for each representative isolate. Three TTV isolates (Sd-TTV31 from a pig, Cf-TTV10 from a dog and Fc-TTV4 from a cat) comprising 2878, 2797 and 2064 nucleotides, respectively, each had three open reading frames (ORFs) encoding 436–635 (ORF1), 73–105 (ORF2) and 224–243 (ORF3) aa but lacked ORF4, similar to tupaia TTV. ORF3 was presumed to arise from a splicing of TTV mRNA, similar to human prototype TTV. Although the nucleotide sequence of Sd-TTV31, Cf-TTV10 and Fc-TTV4 differed by more than 50% from each other and from previously reported TTVs of 3–4–3–9k b and TTV-like mini viruses (TLMVs) of 2–8–3–0 kb isolated from humans and non-human primates as well as tupaia TTVs of 2–2 kb, they resembled known TTVs and TLMVs with regard to genomic organization and presumed transcriptional profile rather than animal circoviruses of 1–7–2–3 kb. Phylogenetic analysis revealed that Sd-TTV31, Cf-TTV10 and Fc-TTV4 were closer to TTVs from lower-order primates and tupaias than to TTVs from higher-order primates and TLMVs. These results indicate that domestic pigs, cats and dogs are naturally infected with species-specific TTVs with small genomic size and suggest a wide distribution of TTVs with extremely divergent genomic sequence and length in animals.

Introduction

TT virus (TTV) is a recently discovered DNA virus that was originally isolated from a patient (initials, T. T.) with post-transfusion hepatitis of unknown aetiology (Nishizawa et al., 1997; Okamoto et al., 1998). TTV is an unenveloped, small, spherical particle with a diameter of 30–32 nm (Itoh et al., 2000) and has a circular, single-stranded DNA genome of negative polarity that is approximately 3–8 kb in length (Okamoto et al., 1999a; Miyata et al., 1999; Mushahwar et al., 1999). The most closely related known virus is chicken anaemia virus (CAV), a member of the family Circoviridae, genus Gyrovirus, which includes other animal circoviruses such as porcine circovirus (PCV) and psittacine beak and feather disease virus (BFDV) (genus Circovirus) (Todd et al., 2000). However, because of significant sequence divergence between TTV and animal circoviruses, it has been proposed that TTV belongs to a new virus family, provisionally designated Circinoviridae (Mushahwar et al., 1999). TTV has an extremely wide range of sequence divergence (Khudyakov et al., 2000; Mushahwar, 2001; Okamoto et al., 1999b; Tanaka et al., 2001). In addition, smaller members of this virus family, i.e. TTV-like mini viruses (TLMVs), whose genomic length is

Author for correspondence: Hiroaki Okamoto.
Fax +81 285 44 1557. e-mail hokamoto@jichi.ac.jp

The nucleotide sequence data reported herein have been assigned DDB/EMBL/GenBank accession nos AB076001–AB076003 for the entire sequences of the Sd-TTV31, Cf-TTV10 and Fc-TTV4 isolates and AB076004–AB076018 for the partial sequences of 15 isolates.
approximately 2.8–3.0 kb, have been identified recently (Takahashi et al., 2000a, b; Biagini et al., 2001).

Increasing lines of evidence indicate that non-human primates and tupaias (tree shrews, Tupaiia belangeri chinensis) are infected with TTVs (Leary et al., 1999; Okamoto et al., 2000a; Verschoor et al., 1999). The entire nucleotide sequences of species-specific TTVs that infect non-human primates, such as the chimpanzee (Pan troglodytes), Japanese macaque (Macaca fuscata), cotton-top tamarin (Saguinus oedipus) and douroucouli (Aotes trivirgatus) as well as tupaias, have been determined (Okamoto et al., 2000b, 2001; Inami et al., 2000). Furthermore, TTV DNA has been detected in serum samples obtained from domesticated farm animals, such as chickens, pigs, cows and sheep (Leary et al., 1999). However, the TTVs in these farm animals have not been characterized fully as yet. Therefore, in the present study, we isolated TTVs from domestic pigs and the two most popular pets, cats and dogs, and determined their complete TTV DNA sequences to define their genomic characteristics and evolutionary relatedness by comparing them with species-specific TTVs and TLMVs from humans and non-human primates as well as tupaias TTV and animal circoviruses.

**Methods**

- **Animals.** Serum samples were obtained from 11 150-day-old, domesticated farm pigs (Sus domesticus), eight dogs (Canis familiaris) and seven cats (Felis catus) during their medical checkups. Samples were kept at −20 °C until testing. These cats and dogs had been kept for up to 10 years as pets in independent houses.

- **Extraction of nucleic acids and amplification by PCR.** Nucleic acids were extracted from 100 µl of serum using a High Pure Viral Nucleic Acid kit (Roche) and dissolved in 50 µl of nuclease-free distilled water. An amount equivalent to 20 µl of serum was subjected to the following PCR methods. In addition to the three PCR methods described previously with primers based on the untranslated region (UTR) (NG472 and NG352 for the first round and NG473 and NG351 for the second round) (Peng et al., 2002), N22 primers (NG059 and NG063 for the first round and NG061 and NG063 for the second round) (Okamoto et al., 1998) or set B primers (forward 1 and reverse 1 for the first round and forward 2 and reverse 2 for the second round) (Leary et al., 1999), conventional PCR was carried out using the primer pair NG343 and NG344, which had been performed for the detection of tupaias TTV DNA (Okamoto et al., 2001). The amplification product was electrophoresed on a 2.5% NuSieve 3:1 agarose gel (FMC BioProducts) to detect the band of the full-length TTV genome.

- **Determination and analysis of TTV sequences.** The products amplified by PCR were ligated into the pT7BlueT vector (Novagen) and both strands were sequenced with the BigDye Terminator Cycle Sequencing Ready Reaction kit on an ABI PRISM 3100 Genetic Analyser (Applied Biosystems). Sequence analysis was performed using GENETYX-MAC, version 10.1.6 (Software Development), and CLUSTAL W (Thompson et al., 1994). Phylogenetic trees were constructed by the neighbour-joining method (Saitou & Nei, 1987). The reliability of the phylogenetic results was assessed using 1000 bootstrap replicates (Felsenstein, 1985). The final tree was obtained using the TREEVIEW program, version 1.6.0 (Page, 1996).

**Results**

**Detection of TTV DNA in sera from pigs, dogs and cats**

When PCR was performed using the UTR primers, N22 or set B primers, no amplification signals were observed in the serum samples from any of the 26 animals tested. However, upon performing PCR with primers NG343 and NG344, which had been derived from the well-conserved area in the UTR of a human TTV genome (TA278), nine (82%) of 11 pigs, three (38%) of eight dogs and three (43%) of seven cats were found to be positive for TTV DNA.

**Full-length nucleotide sequences of the swine, canine and feline TTVs**

The amplified products by PCR with primers NG343 and NG344, which measured 69–99 bp (primer sequences at both ends excluded), were sequenced for all 15 animals testing positive and analysed phylogenetically (accession nos AB076004–AB076018). The tree revealed phylogenetic differences of TTV depending on the species. Two genetic groups of TTV with intergroup differences of 21–48, 23–25 and 56–57% were identified among all of the pigs, dogs and cats, respectively. Based on common 69–99 bp sequences, two pairs of inverse primers for amplification of TTV in each species (NG372–NG373 and NG384–NG385 for swine TTV; NG309–NG310 and NG560–NG561 for canine TTV; and NG493–NG494 and NG365–NG366 for feline TTV) were designed to amplify the entire TTV genomes.

Using the DNA extracted from the sera of 13 viraeic animals as template, the entire genomic sequence of TTV was amplified by PCR with the above-mentioned primers.
TT viruses in animals

The TTV genome was found to be approximately 2.9–2.9 kb in size in all eight samples from the two groups of pigs, 2.8 kb in two samples from dogs and only 2.1 kb in the three samples from the two groups of cats. Then, the PCR-amplified product from each of the swine, canine and feline serum samples with the strongest PCR signal was molecularly cloned and the TTV clones named Sd-TTV31, Cf-TTV10 and Fc-TTV4 were sequenced over the entire genome. All three isolates had a circular genomic structure. The Sd-TTV31 and Cf-TTV10 isolates had a similar genomic length of 2878 and 2797 nt, respectively, in contrast with the Fc-TTV4 isolate which had a genomic length of only 2064 nt. The three isolates were deduced to be single-stranded from the results of the PCR upon digestion with S1 nuclease or mung bean nuclease, similar to the genomic DNAs of known TTVs and TLMVs.

The Sd-TTV31, Cf-TTV10 and Fc-TTV4 isolates differed from each other by 54–56%. In addition, comparison of these three genomes against reported TTV and TLMV genomes from humans and non-human primates as well as the tupaia TTV genome, whose entire or partial nucleotide sequence is known (Biagini et al., 2000, 2001; Erker et al., 1999; Hallett et al., 2000; Hijikata et al., 1999; Inami et al., 2000; Muljono et al., 2001; Mushahwar et al., 1999; Okamoto et al., 1999a, 2000b, 2001; Takahashi et al., 2000a, b; Tanaka et al., 2001; Peng et al., 2002), revealed that the three genomes are less than 45% similar to the TTV and TLMV genomes reported previously.

Proposed genomic organization of the Sd-TTV31, Cf-TTV10 and Fc-TTV4 isolates

Three distinct species of TTV mRNAs (2.9–3.0, 1.2 and 1.0 kb) with three different splicings are observed in human TTVs of 3.8 kb (Kamahora et al., 2000; Okamoto et al., 2000c). However, the consensus motifs (Breathnach et al., 1978; Mount, 1982) for the short splicing that are shared by all three mRNAs of human TTVs and located at their 5’ termini were not recognized in the Sd-TTV31 sequence, although they were found in the Cf-TTV10 (nt 546–632) and Fc-TTV4 (nt 168–224) sequences. In addition, the consensus motifs for the third splicing, which are present in the shortest mRNA of human TTVs, were not identified in any of the Sd-TTV31, Cf-TTV10 and Fc-TTV4 sequences. Fig. 1 compares the proposed genomic organization of the Sd-TTV31, Cf-TTV10 and Fc-TTV4 isolates. They possessed in common three ORFs (ORF1–3) in exactly the same orientation but lacked ORF4.

UTR sequence of Sd-TTV31, Cf-TTV10 and Fc-TTV4 isolates

The UTR in Sd-TTV31 and Fc-TTV4 was defined as the sequence between the end of ORF3 and the beginning of ORF2 and spanned 823 and 492 nt, respectively, occupying 24–29% of the entire genome (Fig. 1). The UTR in Cf-TTV10 was defined as the sequence between the end of ORF1 and the beginning of ORF2 and spanned 881 nt occupying 31% of the entire genome.

The nucleotide sequences downstream of the TATA box that is located in the middle of the UTR of Sd-TTV31, Cf-TTV10 and Fc-TTV4 were compared with those in 13 TTVs and TLMVs of humans and non-human primates and a tupaia TTV as well as CAV (accession no. M55918) (Fig. 2). This particular region is the most conserved among the entire genomes of all TTV and TLMV isolates. Remarkably, all 17 TTV and TLMV isolates, including Sd-TTV31, Cf-TTV10 and Fc-TTV4, shared two highly conserved sequences of 15 nt each (CGAATGGCTGAGTTT and AGGGGCAATTCGGGC), which were located in the 3’-terminal parts of the NG343 (sense) and NG344 (antisense) primers used in the initial PCR amplification of Sd-TTV31, Cf-TTV10 and Fc-TTV4 in the present study. In addition to these two 15 nt sequences, a 12 nt sequence of GGGGGGTTGGCC, including the Sp1 site (underlined), was well preserved. Of note, the CAV sequence differed significantly from all TTV and TLMV isolates even in this most homologous region.

Fig. 1. Comparison of the predicted genomic organization of the swine (a), canine (b) and feline (c) TTV isolates. The circumference of each circle represents the relative size of the genome. The closed arrows represent ORFs (ORF1–3). The open boxes located between an upstream closed box and downstream closed arrow in ORF3, which encodes a putative joint protein, represent an area corresponding to an intron in the shorter mRNA. The shaded box indicates the GC-rich stretch and the small closed circle represents the position of the TATA box.
Fig. 2. Alignment of the nucleotide sequence of TTVs and TLMVs obtained from humans, non-human primates and a tupaia as well as Sd-TTV31, Cf-TTV10 and Fc-TTV4 at a region that is the most conserved among TTVs and TLMVs. The 139–172 nt sequence downstream of the TATA box in the UTR is compared among representative human TTVs of five genetic groups (Peng et al., 2002) [TA278 of group 1 (accession no. AB017610), PMV of group 2 (AF261761), TUS01 of group 3 (AB017613), KC009 of group 4 (AB038621) and JT33F of group 5 (AB064606)], TLMVs [CBD231 (AB026930), TGP96 (AB041962) and Pt-TTV8-II (AB041963)] isolated from humans and a chimpanzee, TTVs [Pt-TTV6 (AB041957), Pt-TTV9 (AB041958), So-TTV2 (AB041959), So-TTV2 (AB041960), At-TTV3 (AB041961) and Tbc-TTV14 (AB057358)] isolated from non-human primates and a tupaia, and TTVs (Sd-TTV31, Cf-TTV10 and Fc-TTV4) isolated from a pig, a dog and a cat in the present study, as well as CAV (M55918). A dash indicates an identical nucleotide in comparison with the top sequence and a solidus indicates a deletion. A putative cap site with the sequence, GAG (Okamoto et al., 2000c), is dotted. The sequences corresponding to primers NG343 (sense) and NG344 (antisense) are boxed. The conserved 12 nt sequence, including the Sp1 site (GGGCGG), is overlined.

In the middle of the UTR of the Sd-TTV31 and Cf-TTV10 genomes, there is a GC-rich stretch of 159 and 162 nt, respectively (Fig. 1), which is longer than the GC-rich stretch (35–133 nt) in all known TTVs and TLMVs. The Sd-TTV31 genome contained five successive repeat sequences of 33–36 nt, forming GC-rich stem–loop structures starting at nt 2719, 2753, 2785, 2821 and 2856. Similarly, the Cf-TTV10 genome contained six GC-rich sequences of 29–30 nt, forming hairpins starting at nt 2634, 2665, 2694, 2750, 88 and 118, respectively. In contrast, the Fc-TTV4 genome had no GC-rich stretch in the UTR, although it had two stem–loop structures starting at nt 1929 and 2015, respectively.

Coding region sequence of Sd-TTV31, Cf-TTV10 and Fc-TTV4 isolates

The ORF1s of Sd-TTV31, Cf-TTV10 and Fc-TTV4 isolates encoded 436–635 aa and were rich in Arg at their N-termini. Only one of the four conserved motifs (motif 3) present in putative replication-associated proteins (Rep proteins), which are involved in rolling-circle replication (Mushahwar et al., 1999), was discernible in these three isolates: YPVR at aa 520–523 in Sd-TTV31, YLSK at aa 424–427 in Cf-TTV10 and YKLK at aa 309–312 in Fc-TTV4. Although significant sequence similarity was not observed in the amino acid sequence of ORF2 among Sd-TTV31, Cf-TTV10 and Fc-TTV4 isolates nor between these and TTVs or TLMVs, the conserved motif (WX(H)X3CX3CX3X(H) in the N terminus of the ORF2 protein of reported TTVs and TLMVs (Hijikata et al., 1999; Takahashi et al., 2000b) was also shared by Sd-TTV31, Cf-TTV10 and Fc-TTV4.

ORF3 of Sd-TTV31, Cf-TTV10 and Fc-TTV4 encoded a putative joint protein of 224, 243 and 231 aa, respectively, which included the same amino acid sequence of 72, 104 and 104 aa, respectively, encoded by ORF2. In Sd-TTV31, Cf-TTV10 and Fc-TTV4, the C-terminal portion of ORF3 was rich in Ser (15, 12 and 18 residues, respectively).

Phylogenetic analysis of TTVs and TLMVs

Due to difficulties in constructing a phylogenetic tree using the full-genome sequences of TTVs and TLMVs, which
markedly differ in sequence and size, trees were constructed using the entire amino acid sequences of ORF1 (436–770 residues) and ORF2 (59–131 residues) in the TTV and TLMV isolates (Fig. 3). In both trees, Sd-TTV31, Cf-TTV10 and Fc-TTV4 were closest to each other but genetic distances between them were much greater than those among human TTVs of five genetic groups. When compared with known TTVs and TLMVs, these three animal TTVs were closer to the tupaia TTV of Tbc-TTV14 and the TTVs from primates of lower order (tamarin and douroucouli) than to the TTVs from humans and higher-order non-human primates (Japanese macaques and a chimpanzee) and the TLMVs from humans and a chimpanzee.

**Discussion**

In the present study, a new TTV species with the smallest known genome of 2·1 kb was isolated from the sera of cats and two new TTV species of 2·8–2·9 kb were isolated from the sera of domestic pigs and houndogs. The genomic DNA of swine (Sd-TTV31), canine (Cf-TTV10) and feline (Fc-TTV4) TTV was presumed to be circular and single-stranded, similar to that of the human prototype TTV (Miyata et al., 1999; Mushahwar et al., 1999; Okamoto et al., 1999a, 2000d). The genomic length of the Fc-TTV4 isolate was 2064 nt and is comparable with those of animal circoviruses (PCV, BFDV and CAV), which have a circular, single-stranded DNA of 1·8–2·3 kb (Bassami et al., 1998; Niagro et al., 1998; Noteborn et al., 1999; Todd et al., 2000) and that of tupaia TTV of 2·2 kb (Okamoto et al., 2001). On the other hand, the genomic length of Sd-TTV31 and Cf-TTV10 was 2878 and 2797 nt, respectively, comparable with that of the TLMVs from humans and chimpanzees (Takahashi et al., 2000b; Okamoto et al., 2000b). The genomic organization of Sd-TTV31, Cf-TTV10 and Fc-TTV4 resembled those of TTVs and TLMVs that were isolated from humans and non-human primates as well as those of tupaia TTV (Okamoto et al., 2001), although the sequence similarity of the Sd-TTV31, Cf-TTV10 and Fc-TTV4 isolates against known TTV and TLMV isolates was less than 50%.

TTVs and TLMVs that infect humans and non-human primates are presumed to have a unique transcriptional profile (Okamoto et al., 2001) that is not observed among known members of the Circoviridae family (Niagro et al., 1998; Noteborn & Koch, 1995; Todd et al., 2000). However, the Sd-TTV31, Cf-TTV10 and Fc-TTV4 genomes obtained in the present study lacked either one or two of the three splicings (Kamahora et al., 2000; Okamoto et al., 2000c): they possessed in common only one splicing, which is involved in the creation of ORF3. The complete preservation of coding capacity for ORF1 and ORF2 proteins as well as one joint protein (ORF3), as illustrated in Fig. 1, indicates that the proposed fundamental genomic organization and transcriptional profile are characteristic to all TTVs and TLMVs isolated from humans and non-human primates as well as TTVs from lower-order mammals, non-human primates.
such as tupaia, swine, canines and felines. Strict conservation of the putative ORF3 in swine, canine and feline TTVs suggest that the ORF3 protein is indispensable for all TTVs and TLMVs. In fact, ORF3 has a Ser-rich tract in its C terminus and it has been demonstrated in vitro that ORF3 generates two forms of proteins with a different phosphorylation state, suggesting that ORF3 protein has function(s) similar to phosphorylated viral proteins, such as the hepatitis C virus NS5A protein (Asabe et al., 2001). The function and virological significance of ORF4, which has been found only in TTVs and TLMVs isolated from humans and non-human primates, remain unknown.

The Sd-TTV31, Cf-TTV10 and Fc-TTV4 isolates were clearly separate from CAV, which is most closely related to TTVs among known animal viruses. Phylogenetic analysis revealed that the Sd-TTV31, Cf-TTV10 and Fc-TTV4 isolates were closer to the TTVs from tupaia and primates of lower order (tamarin and douroucouli) than to known TTVs and TLMVs from humans and non-human primates of higher order, suggesting that the phylogenetic relatedness of TTV may reflect the evolutionary relationship of the infected hosts. The genomic length of TTV tends to be smaller the lower the order of the infected animal. However, of interest, the only exception is that humans and chimpanzees are infected with TTVs that were unique in that they were infected with TTVs that were lacking ORF4 as well as the GC-rich stretch in the UTR. In

In addition, the primers (NG343 and NG344) used for the initial detection of swine, canine and feline TTVs in the present study would be instrumental in extended research on TTVs in unexamined animal species in an attempt to further define their virological characteristics and evolutionary relationships.

This work was supported in part by grants from the Ministry of Education, Culture, Sports, Science and Technology of Japan and the Ministry of Health,Labour and Welfare of Japan. The authors are grateful to M. Mayumi for advice and encouragement during this study.

References


prevalence in humans, non-human primates and farm animals. *Journal of General Virology* 80, 2115–2120.


Received 4 January 2002; Accepted 6 February 2002