Journal of General Virology (2000), **81**, 2273–2279. Printed in Great Britain

Characterization of a highly divergent TT virus genome

R. L. Hallett,[1] J. P. Clewley,[1] F. Bobet,[1] P. J. McKiernan[2] and C. G. Teo[1]

[1] Hepatitis and Retrovirus Laboratory, Central Public Health Laboratory, 61 Colindale Avenue, London NW9 5HT, UK
[2] The Liver Unit, Birmingham Children's Hospital, Whittal Street, Birmingham B4 6NH, UK

A novel TT virus (TTV)-like DNA sequence was detected in the serum of a patient (PM) with acute non-A–E hepatitis. The full-length genome sequence, referred to here as PM virus (PMV), was obtained and its relationship to other full or near full-length TTV sequences examined. Although it shares a common genomic arrangement and short conserved regions, the majority of the genome is extremely divergent, displaying an average genetic distance of 0·60 from all other TTV sequences. By comparing PMV with TTV genomes representing the most divergent types so far described, six major groups can be distinguished. The level of genetic diversity seen between these genomes is higher than would be expected within a single virus species. Indeed, PMV could be considered the prototype of an independent taxonomic group within the *Circoviridae* family. A genoprevalence study of sera from blood donors and patients with acute hepatitis suggests that PMV is rare.

Introduction

Efforts continue to isolate agents responsible for viral hepatitis that cannot be attributed to hepatitis viruses A to E. As part of these investigations, DNA from a novel virus was identified in the serum of a patient with post-transfusion hepatitis of unknown aetiology (Nishizawa *et al.*, 1997). The virus was named TT virus, and has subsequently been characterized as an unenveloped single-stranded DNA virus with a caesium chloride density of 1·31–1·34 g/cm^3 (Mushahwar *et al.*, 1999). Several complete TTV DNA sequences have now been published, revealing a circular genome which ranges from 3808 to 3852 nucleotides in length (Erker *et al.*, 1999; Okamoto *et al.*, 1999*b*; Hijikata *et al.*, 1999). Nuclease protection assays suggest that the genome is negative-stranded (Mushahwar *et al.*, 1999).

TTV was initially reported to share some physical characteristics with the *Parvoviridae*, although no nucleotide similarity was demonstrated. A resemblance between TTV and chicken anaemia virus (CAV), a circovirus, has been described by Takahashi *et al.* (1998*a*) on the basis of arginine-rich regions in the open reading frames (ORFs) from both viruses, and on similarities in their genomic arrangements (Miyata *et al.*, 1999). However, no significant sequence similarity to any other member of the *Circoviridae* has been shown, and this has led Mushahwar *et al.* (1999) to propose that TTV be placed in a new virus family called *Circinoviridae*. Recently, three novel human DNA virus sequences (CBD203, CBD231 and CBD279) have been described that appear to be intermediately related to TTV and CAV (Takahashi *et al.*, 2000). They have been named TTV-like mini virus (TLMV). It is not clear from these reports whether TTV, CAV and TLMV belong to one family of viruses.

Considerable genetic variability of TTV has been demonstrated, with a 222 bp fragment of the genome from the longest ORF (ORF-1) showing as much as 65% divergence between sequences. Several groups have assigned up to 16 genotypes based on the sequence variability in short TTV PCR products (Tanaka *et al.*, 1998; Okamoto *et al.*, 1999*b*). In contrast, phylogenetic analysis of full-length or near full-length TTV sequences reveals division into three main types, represented by the prototype TA278 in group 1, US35 in group 2 and JA10 in group 3 (Erker *et al.*, 1999). Most of the sequence variability occurs in the coding regions. In addition to these three TTV groups there are two other highly divergent complete sequences, TUS01 and SANBAN (Okamoto *et al.*, 1999*b*; Hijikata *et al.*, 1999).

In this study we have sequenced the full genome of another TTV-like virus, named PMV, with a view to characterizing its genetic organization and to ascertaining its taxonomic status. We have also conducted a survey of its prevalence.

Author for correspondence: Rachel Hallett.
Fax +44 20 8200 1569. e-mail rachel.hallett@virgin.net

The GenBank accession number of the sequence reported in this paper is AF261761.

0001-7145 © 2000 SGM

Downloaded from www.sgmjournals.org by
IP: 93.91.26.97
On: Tue, 28 Jul 2015 19:38:15

Methods

■ **Specimens.** Serum samples referred to our laboratory from patients with acute non-A–E hepatitis and UK blood donors were examined for the presence of TTV DNA by PCR. In some cases of acute non-A–E hepatitis, archived liver samples were also investigated.

■ **DNA extraction.** DNA was extracted from serum (100 µl volumes) by a guanidinium thiocyanate–silica method (Boom *et al.*, 1990). DNA was also extracted from paraffin-embedded liver biopsy samples by sequential treatments with Tween 20, proteinase K, Chelex-100 and chloroform, as described by Coombs *et al.* (1999).

■ **PCR.** A semi-nested PCR was used to amplify part of ORF-1 as previously described (Okamoto *et al.*, 1998). A PCR using primers located within the conserved non-coding region (set B; Leary *et al.*, 1999) was also used. All positive PCR products were gel-purified and sequenced as described below.

■ **PCR product cloning and sequencing.** PCR products were recovered from 2% agarose gels after staining with ethidium bromide and visualization under UV, and were purified with an Igenie DNA Extraction Kit (Helena BioSciences). They were sequenced in both directions using an ABI Prism DNA Sequencing Kit and an ABI 373 automated sequencer (PE Applied Biosystems). ORF-1 PCR products were cloned using an Invitrogen TOPO TA Cloning Kit and colony PCRs performed using the TTV-specific inner primers NG061 and NG063 (Okamoto *et al.*, 1998). Ten colony PCR products from each sample were sequenced.

■ **PMV genome sequencing.** PCR products covering the whole genome were generated using an Expand High Fidelity PCR System (Roche Diagnostics) according to the manufacturer's protocol for amplification of products up to 3 kb, and with a Clontech Advantage-GC 2 PCR kit. All PCRs were carried out in a 50 µl volume. First or single round PCRs contained 10 µl of extracted DNA as template; second round PCRs used 2 µl of first round product. The PCR primers used to generate the entire PMV genome sequence were: for fragment A, T801 and NG063 (Takahashi *et al.*, 1998*b*; Okamoto *et al.*, 1998); for fragment B, NG059, NG061 and BR1 (Okamoto *et al.*, 1998; Leary *et al.*, 1999); for fragment C, first round sense INV1 (5′ CCTTACAGACACCCCTT-ACTACCCT), first round antisense INV2 (5′ CAGTGGCACTTTCC-TTTTCTTTC), and second round sense INV3 (5′ ACTAAGCACT-CCGAGCGAAGC) with second round antisense INV4 (5′ ATAACC-CTAAGAGCCTTGCCCATAG).

The positions of the three PCR products are indicated in Fig. 1. All three were cloned using an Invitrogen TOPO TA Cloning Kit. Plasmid DNA containing the inserted PCR products was purified using a QIAfilter Plasmid Midi Kit (Qiagen) and sequencing was carried out by Cambridge Bioscience. The resulting three fragments were assembled using programs EditSeq and MegAlign in the Lasergene Navigator package (Dnastar) to give the complete genome sequence of 3736 nucleotides.

■ **Sequence analysis.** Sequences contained in the GenBank/EMBL and SWISS-PROT databases were searched using BLASTN, BLASTP and FASTA (Altschul *et al.*, 1997; Pearson & Lipman, 1988). DNA and protein alignments were carried out using the ClustalW algorithm (Thompson *et al.*, 1994). Pairwise distances were calculated and phylogenetic analysis performed using programs in the Phylip package (Seqboot, DNAdist and Fitch, Neighbor and DNAml) (Felsenstein, 1993). Tree diagrams were generated using Treeview (Page, 1996). Secondary structure predictions of the non-coding regions in the PMV genome were made using the GCG program Mulfold (Jaeger *et al.*, 1989). By using the GCG program Plotsimilarity, with a sliding window of 50 nucleotides, sequence similarity was plotted across a multiple alignment of PMV and six full or near full-length TTV genomes representing the most divergent genetic types. This analysis was also carried out on the amino acid sequences of ORFs 1 and 2 using the same genomes and a sliding window of 20 amino acids.

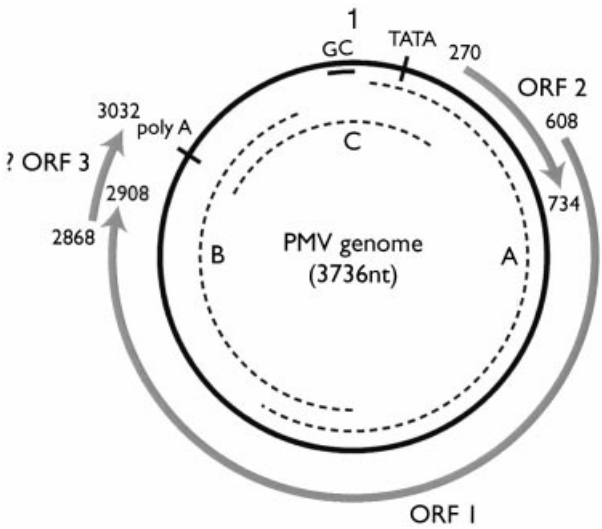


Fig. 1. Putative structure of the PMV genome. Dotted lines (A, B and C) show the positions of the three PCR products that were assembled to provide the complete genome sequence. Shaded arrows indicate the conserved ORFs, with the numbers showing their nucleotide positions. Other conserved characteristics are a TATA box, polyadenylation signal (polyA) and GC-rich region (GC).

Results

Initial detection and cloning

The initial TTV PCR, carried out on serum from a 4-year old girl (PM), revealed a 222 bp fragment of DNA from ORF-1 that was highly divergent from all other equivalent TTV sequences deposited in GenBank. This sequence was recovered from eight of ten colonies of cloned PCR products (the remaining two colonies contained one point mutation in each), and also from an archived liver sample from the same patient. This UK patient had acute non-A–E hepatitis and coagulopathy. She had no risk factors for bloodborne infection or travel outside the country.

Genome analysis

The entire genome of PMV was sequenced by producing three overlapping PCR products, as shown in Fig. 1. PCR products A and B were generated with standard long-range amplification conditions. However, PCR product C was not amplifiable this way due to the presence of a GC-rich region in the template DNA, and a PCR under conditions specific for GC templates was necessary to amplify this fragment (see Methods). A database homology search was carried out using the nucleotide sequence of the PMV genome. It did not reveal any significant nucleotide similarity between the PMV sequence and other sequences, other than TTV sequences.

Downloaded from www.sgmjournals.org by
IP: 93.91.26.97
On: Tue, 28 Jul 2015 19:38:15

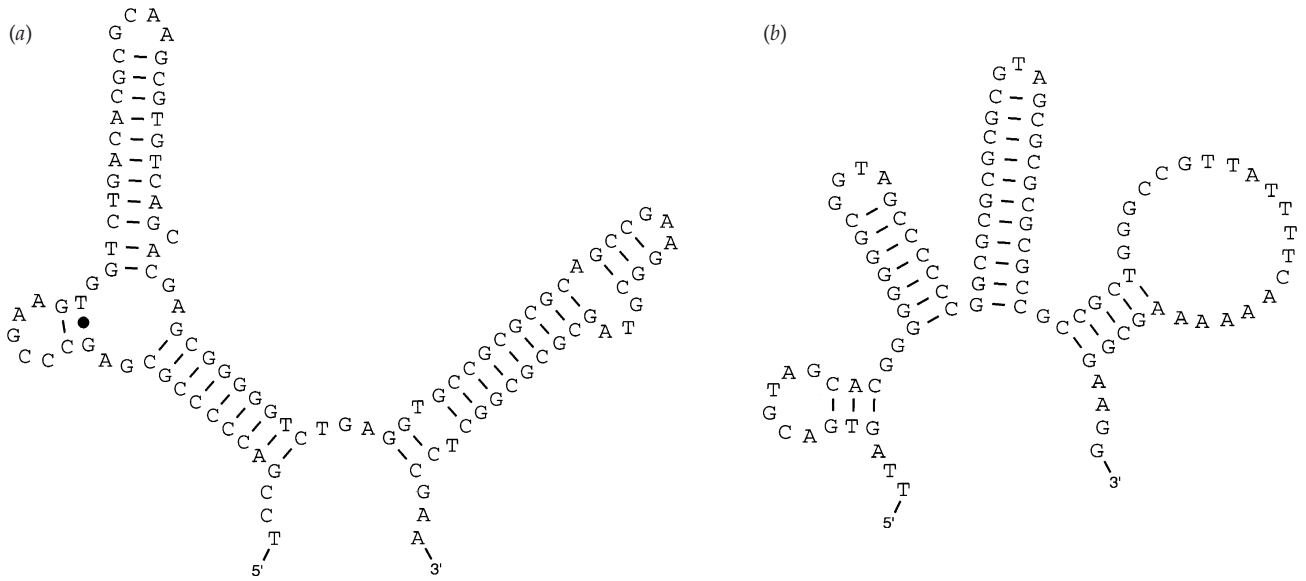


Fig. 2. Predicted secondary structure of two areas of the PMV non-coding region. The genomic strand (negative sense) was used in the analysis: (*a*) region immediately downstream of the polyadenylation signal (nucleotides 3133–3237); (*b*) truncated GC-rich region (nucleotides 3656–13).

In a comparison with full or near full-length TTV genomes it was found that PMV shared characteristics, including the positions of ORFs 1 and 2 and the non-coding regions, a GC-rich string of bases and a TATA box (Fig. 1). A polyadenylation signal (AATAAA), starting at position 3028 in PMV, was found in all full and near full-length genomes with the exception of SANBAN (Hijikata *et al*., 1999).

Open reading frames

The longest ORF, ORF-1, of the PMV genome encodes 767 amino acids and has an average amino acid identity of only 31·1% with the corresponding ORFs of other TTVs. It has an N-terminal arginine-rich region, noted previously in other TTVs (Takahashi *et al*., 1998*a*). When the amino acid sequence of ORF-1 was used to BLAST search the SWISS-PROT database, 37% identity to a cuttlefish protamine protein (PRT2 SEPOF) was revealed.

A shorter ORF, ORF-2, was found in PMV, but again there was a high degree of sequence divergence from other TTVs, the average amino acid identity being 32·7%. The sequence motif (ACCATGG) containing the initiation codon of ORF-1 in PMV had a single nucleotide difference from that described by Kozak (1986) as the optimal sequence for translation initiation. The ORF-2 initiation motif had two suboptimal nucleotides. Conserved Sp1 (GGCGGG) and Cap (CAATTC) sites, immediately upstream of the ORF-2 initiation codon, were also present in the PMV genome (Hijikata *et al*., 1999).

The putative ORF-3 proposed by Erker *et al*. (1999) was 165 nucleotides long in PMV. However, it overlapped with the C-terminal end of ORF-1 by 40 nucleotides, a feature not seen in TTV genome sequences so far described. Erker *et al*. (1999) reported ORF-3 to be highly conserved, but when the full or near full-length TTV sequences from their study were analysed with sequences TUS01, SANBAN and PMV, ORF-3 was revealed to be variable. The average amino acid sequence identity between PMV and these ORF-3 sequences was 46·3%. The Kozak initiation motif was not identified at the start codon in PMV.

Non-coding region

A conserved non-coding region was found to stretch from the polyadenylation signal to the initiation codon of ORF-2, encompassing the GC-rich region. Comparisons were carried out between PMV and five full-length TTV sequences (TA278, JA20, US35, TUS01 and SANBAN). For four of these, the non-coding region was about 1040 nucleotides long; in TUS01 it was shorter at 1014 nucleotides. The equivalent region in PMV was only 972 nucleotides long, the most noticeable truncation being in the GC-rich region, which was 58 nucleotides in length compared to 117 in the prototype TTV genome TA278 (Miyata *et al*., 1999). The average nucleotide identity between PMV and the five analysed sequences in this region was 69·8%.

Hijikata *et al*. (1999) proposed that the domain immediately downstream of the polyadenylation signal in the non-coding region of TTV forms a stem–loop structure. The corresponding region in PMV, analysed in this way, predicted a secondary structure that was very similar to the TTV structure (Fig. 2*a*). By contrast, the secondary structure of the GC-rich region of the PMV genome is likely to be affected by deletions: Fig. 2(*b*) shows how the GC-rich region of PMV might be different from the GC-rich stretch in isolates TA278 and TUS01 (Okamoto *et al*., 1999*b*).

Downloaded from www.sgmjournals.org by
IP: 93.91.26.97
On: Tue, 28 Jul 2015 19:38:15

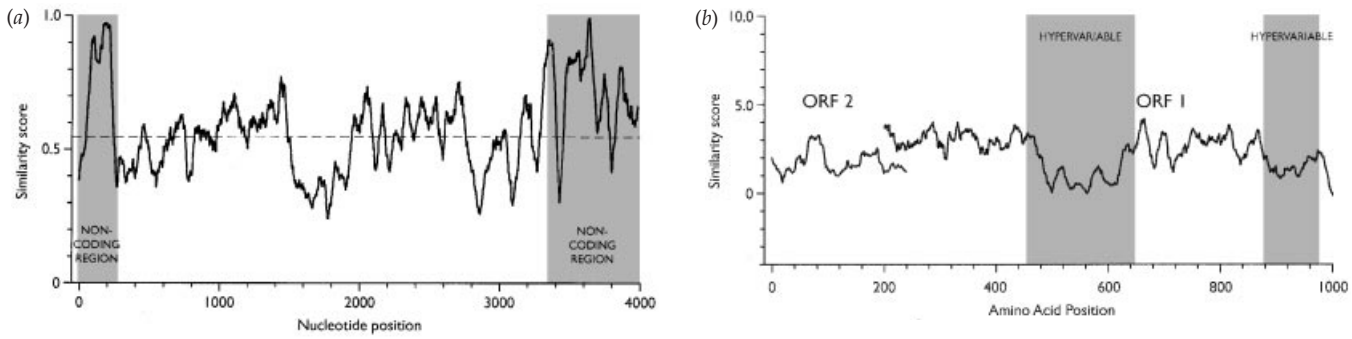
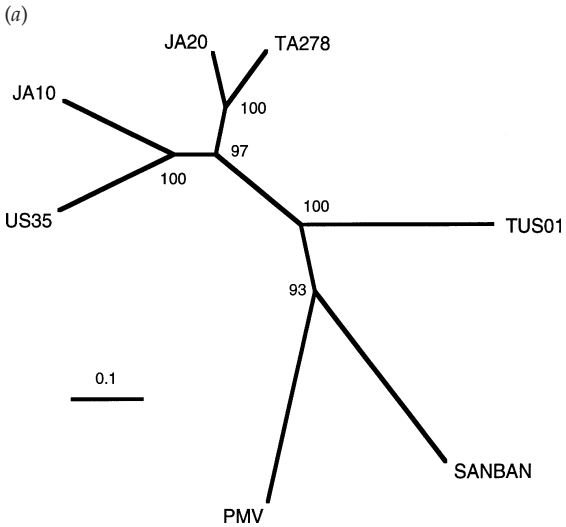


Fig. 3. (*a*) Similarity plot of an alignment of the PMV genome sequence with six representative TTV genomes. The average nucleotide sequence identity is indicated by the dashed line. (*b*) Amino acid similarity plot of an alignment of ORFs 1 and 2 from the same seven sequences. The bars denote hypervariable regions within ORF-1. The six genomes used are TA278, JA20, US35, JA10, TUS01 and SANBAN (Miyata *et al*., 1999; Erker *et al*., 1999; Okamoto *et al*., 1999*b*; Hijikata *et al*., 1999).



(b)

TA278	TA278					
JA20	0.17	JA20				
US35	0.35	0.33	US35			
JA10	0.35	0.34	0.34	JA10		
TUS01	0.54	0.53	0.58	0.57	TUS01	
SANBAN	0.64	0.62	0.68	0.63	0.58	SANBAN
PMV	0.57	0.57	0.65	0.60	0.63	0.58	PMV

Fig. 4. (*a*) A neighbour-joining phylogenetic tree demonstrating the genetic relatedness between PMV and six representative full or near full-length TTV genomes, with (*b*) the corresponding genetic distances. Bootstrap percentages from 1000 replicates are shown on the branches. The scale bar indicates the proportion of nucleotides substituted for a given branch length.

Relationship to other TTV-like genomes

The similarity of PMV to the overlapping part of the six representative sequences is shown in Fig. 3(*a*), which reveals regions of conservation and variability across the 3·7 kb studied. The average nucleotide sequence identity was 55% (by Plotsimilarity). The amino acid similarity plot for ORF-1 revealed hypervariable regions between amino acid positions 300 and 440, and between 710 and 760 in the multiple alignment. For ORF-2, the amino acid plot was smoother, indicating that a high level of variability occurs across the whole ORF, rather than being concentrated into hypervariable zones.

The phylogenetic tree and corresponding genetic distances produced from a multiple alignment of PMV with the six representative TTV sequences are shown in Fig. 4. It is clear that the sequences of TUS01, SANBAN and PMV are highly divergent from the prototype TTV sequence, TA278, and from each other. Six major groupings were observed, consisting of TA278 and JA20 as one group, with the other five sequences each representing a separate group. The average genetic distance between each group was 0·48. Strong support for these groupings was obtained from bootstrap resampling.

Phylogenetic sequence comparisons were also made between PMV, TTV, and complete CAV and TLMV sequences (CAE26P4 and CBD279 respectively), by first carrying out individual pairwise alignments to identify homologous regions, and then performing a multiple alignment using gap-stripped 2·5 kb sequences. Fig. 5 shows that there is one cluster of sequences that includes the TTVs (TA278, JA20, JA10 and US35), and five other sequences (PMV, TUS01, SANBAN, TLMV and CAV), each of which is divergent from the others.

Downloaded from www.sgmjournals.org by
IP: 93.91.26.97
On: Tue, 28 Jul 2015 19:38:15

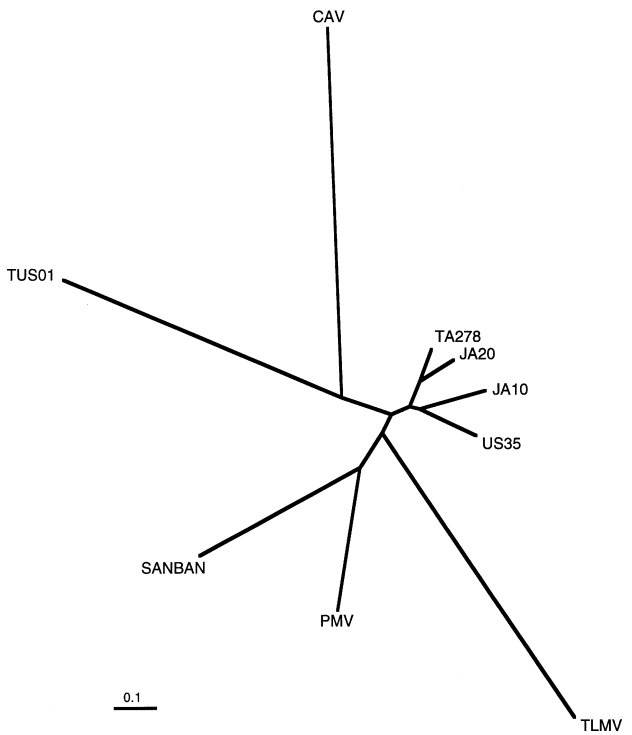


Fig. 5. A neighbour-joining phylogenetic tree showing the relationship between 2·5 kb homologous regions of TTV genomes, a CAV genome (CAE26P4) and a TLMV genome (CBD279). A multiple alignment was made of the homologous regions to the CAV genome, which were identified by individual pairwise alignments with CAV. The alignment was used in DNAdist and Neighbor with a transition/transversion ratio of 1·1. The scale bar indicates the proportion of nucleotides substituted for a given branch length.

PMV was about equally closely related to TA278 and JA20, and to SANBAN.

Prevalence of PMV

PMV-like sequences were searched for in 197 sera belonging to two patient groups, UK blood donors ($n = 99$) and acute non-A–E hepatitis patients ($n = 98$). A PCR assay with set B primers situated in the non-coding region was used to amplify TTV DNA (Leary *et al.*, 1999). Cloned PMV DNA (fragment B, Fig. 1) was used as a positive control. A total of 49 (17 blood donors and 32 hepatitis patients) were positive and, of these, 36 gave readable DNA sequence data. Similar to TTV, but not PMV. On alignment with the set B PCR product from PMV it was seen that PMV remained phylogenetically unique, with the most closely related sequences showing around 10% divergence.

Discussion

The circular genome of a novel TTV-like virus, PMV, has been characterized. PMV shares its genome organization and a moderately conserved non-coding region with other TTV sequences. Its genome also contains the typical motifs identified in other TTVs. It is the shortest TTV-like genome reported so far, the largest deletions being in the GC-rich region. The similarity plot of an alignment of PMV and the six most divergent TTV sequences shows a high degree of sequence heterogeneity, with genetic distances between them varying from 0·17 to 0·68. Sequence variation is more common in the coding regions, which may lead to major antigenic differences between the viruses. The protein encoded by the ORF-1 of PMV was homologous to a cuttlefish protamine protein that is involved in the folding and packaging of sperm DNA in the nucleus. This is consistent with the ORF-1 product being a virion capsid protein that enwraps the viral DNA genome. Nishizawa *et al.* (1999) and Erker *et al.* (1999) have previously proposed that ORF-1 of TTV encodes the viral capsid protein on the basis of its N-terminal arginine-rich region, which is shared with the circovirus capsid proteins. The functions of the proteins encoded by ORF-2 and ORF-3 are not known.

The non-coding region is relatively conserved, suggesting that it plays an important regulatory role in virus replication. The effect that the truncated GC-rich stretch may have on the replication of PMV is yet to be determined.

Representative full-length and near full-length genomes of TTV-like viruses fall into six major groups on phylogenetic analysis, with an average genetic distance of 0·48 between them (Fig. 4). The groupings observed on analysis of the overlapping nucleotide sequences are also seen when the amino acid sequences of ORF-1 are aligned (data not shown). The six groups correspond to the previously defined groups 1a (the prototype TA278) and 1b (JA20), 2 (US35) and 3 (JA10) (Erker *et al.*, 1999), with TUS01, SANBAN and PMV forming three new branches. Following the nomenclature of Erker *et al.* (1999), TUS01, SANBAN and PMV could be considered new 'genotypes'. However, the level of genetic divergence within the TTV cluster is higher than would be expected if all these sequences were from viruses of the same species. Several groups have postulated a genotyping system based on the phylogenetic analysis of short PCR products (Tanaka *et al.*, 1998; Okamoto *et al.*, 1999*a*; Mulyanto *et al.*, 2000). The analysis presented here, using full and near full-length genomes, suggests that until the full extent of divergence within the TTVs is known, it is premature to classify short TTV-like sequences in this way. The possibility that further divergent genomes remain to be discovered cannot be excluded.

Several groups have reported short stretches of sequence similarity between TTV and CAV (Miyata *et al.*, 1999; Hijikata *et al.*, 1999). Furthermore, TLMV was recently discovered in human serum and reported as an intermediate relative of TTV and CAV (Takahashi *et al.*, 2000). In order to further assess the relationships between these viruses, we compared representative TTV sequences, a CAV genome and a TLMV genome (Fig. 5). Phylogenetic analysis shows that six lineages can be

Downloaded from www.sgmjournals.org by
IP: 93.91.26.97
On: Tue, 28 Jul 2015 19:38:15

discerned: the lineage to which the TTV cluster of groups 1, 2 and 3 belongs, and five others (PMV, SANBAN, TLMV, TUS01 and CAV).

The *Circoviridae* has recently been divided into two genera: *Gyrovirus*, containing CAV; and *Circovirus*, with porcine circovirus (PCV) and beak and feather disease virus (BFDV) as members (Pringle, 1999). Although there is very little sequence similarity between them, individual isolates of each virus are moderately conserved; for example, the maximum divergence shown between six full-length CAV genomes, available in GenBank, was 5% (not shown). By this criterion, TUS01, SANBAN, PMV and TLMV individually, and TA278, JA20, US35 and JA10 as a cluster, should have equal taxonomic rank with each other and with the *Gyrovirus* and *Circovirus* genera of *Circoviridae*.

Although TTV was initially considered a hepatitis virus, this is not yet substantiated. PMV may be hepatotropic as its sequence was also detected in liver, but since a pre-hepatitis serum specimen from the patient was not available we were unable to determine if PMV infection was associated with onset of disease. It is noted that circoviruses showing less sequence divergence than the TTV group can cause different clinical syndromes. For example, a strain of PCV, reported to be associated with a disease in pigs known as postweaning multisystemic wasting syndrome, shows only 26% divergence from the genome of the prototype PCV, which is common in domestic herds and apparently does not cause disease (Morozov *et al.*, 1998).

Further study is required to unravel the precise evolutionary and taxonomic connections between the circular single-stranded genomes examined here. While PMV may be unique genetically and might be hepatotropic, it appears to be a rare virus in humans. The emerging pattern of infection with TTV-like viruses is that of a collection of related but different viruses, with varying prevalences. Their pathogenic potential has yet to be elucidated.

We thank Dr Jim Gray for provision of serum and archived liver specimens.

References

Altschul, S. F., Madden, T. L., Schaffer, A. A., Zhang, J., Zhang, Z., Miller, W. & Lipman, D. J. (1997). Gapped BLAST and PSI-BLAST: a new generation of protein database search programs. *Nucleic Acids Research* **25**, 3389–3402.

Boom, R., Sol, C. J. A., Salimans, M. M. M., Jansen, C. L., Wertheim-van Dillen, P. M. E. & van der Noordaa, J. (1990). Rapid and simple method for purification of nucleic acids. *Journal of Clinical Microbiology* **28**, 495–503.

Coombs, N. J., Gough, A. C. & Primrose, J. N. (1999). Optimization of DNA and RNA extraction from archival formalin-fixed tissue. *Nucleic Acids Research* **27**, 12–14.

Erker, J. C., Leary, T. P., Desai, S. M., Chalmers, M. L. & Mushahwar, I. K. (1999). Analyses of TT virus full-length genomic sequences. *Journal of General Virology* **80**, 1743–1750.

Felsenstein, J. (1993). PHYLIP (phylogeny inference package) v3.5c. Department of Genetics, University of Washington, Seattle, USA.

Hijikata, M., Takahashi, K. & Mishiro, S. (1999). Complete circular DNA genome of a TT virus variant (isolate name SANBAN) and 44 partial ORF2 sequences implicating a great degree of diversity beyond genotypes. *Virology* **260**, 17–22.

Jaeger, J. A., Turner, D. H. & Zuker, M. (1989). Improved predictions of secondary structures for RNA. *Proceedings of the National Academy of Sciences, USA* **86**, 7706–7710.

Kozak, M. (1986). Point mutations define a sequence flanking the AUG initiator codon that modulates translation by eukaryotic ribosomes. *Cell* **44**, 283–292.

Leary, T. P., Erker, J. C., Chalmers, M. L., Desai, S. M. & Mushahwar, I. K. (1999). Improved detection systems for TT virus reveal high prevalence in humans, non-human primates and farm animals. *Journal of General Virology* **80**, 2115–2120.

Miyata, H., Tsunoda, H., Kazi, A., Yamada, A., Khan, M. A., Murakami, J., Kamahora, T., Shiraki, K. & Hino, S. (1999). Identification of a novel GC-rich 113-nucleotide region to complete the circular, single-stranded DNA genome of TT virus, the first human circovirus. *Journal of Virology* **73**, 3582–3586.

Morozov, I., Sirinarumitr, T., Sorden, S. D., Halbur, P. G., Morgan, M. K., Yoon, K.-J. & Paul, P. S. (1998). Detection of a novel strain of porcine circovirus in pigs with postweaning multisystemic wasting syndrome. *Journal of Clinical Microbiology* **36**, 2535–2541.

Mulyanto, M., Hijikata, M., Matsushita, M., Ingkokusmo, G., Widjaya, A., Sumarsidi, D., Kanai, K., Ohta, Y. & Mishiro, S. (2000). TT virus (TTV) genotypes in native and non-native prostitutes of Irian Jaya, Indonesia: implication for non-occupational transmission. *Archives of Virology* **145**, 63–72.

Mushahwar, I. K., Erker, J. C., Muerhoff, A. S., Leary, T. P., Simons, J. N., Birkenmeyer, L. G., Chalmers, M. L., Pilot-Matias, T. J. & Desai, S. M. (1999). Molecular and biophysical characterization of TT virus: evidence for a new virus family infecting humans. *Proceedings of the National Academy of Sciences, USA* **96**, 3177–3182.

Nishizawa, T., Okamoto, H., Konishi, K., Yoshizawa, H., Miyakawa, Y. & Mayumi, M. (1997). A novel DNA virus (TTV) associated with elevated transaminase levels in posttransfusion hepatitis of unknown etiology. *Biochemical and Biophysical Research Communications* **241**, 92–97.

Nishizawa, T., Okamoto, H., Tsuda, F., Aikawa, T., Sugai, Y., Konishi, K., Akahane, Y., Ukita, M., Tanaka, T., Miyakawa, Y. & Mayumi, M. (1999). Quasispecies of TT virus (TTV) with sequence divergence in hypervariable regions of the capsid protein in chronic TTV infection. *Journal of Virology* **73**, 9604–9608.

Okamoto, H., Akahane, Y., Ukita, M., Fukuda, M., Tsuda, F., Miyakawa, Y. & Mayumi, M. (1998). Fecal excretion of a nonenveloped DNA virus (TTV) associated with posttransfusion non A-G hepatitis. *Journal of Medical Virology* **56**, 128–132.

Okamoto, H., Takahashi, M., Nishizawa, T., Ukita, M., Fukuda, M., Tsuda, F., Miyakawa, Y. & Mayumi, M. (1999*a*). Marked heterogeneity and frequent mixed infection of TT virus demonstrated by PCR with primers from coding and noncoding regions. *Virology* **259**, 428–436.

Okamoto, H., Nishizawa, T., Ukita, M., Takahashi, M., Fukuda, M., Iizuka, H., Miyakawa, Y. & Mayumi, M. (1999*b*). The entire nucleotide sequence of a TT virus isolate from the United States (TUS01): comparison with reported isolates and phylogenetic analysis. *Virology* **259**, 437–448.

Page, R. D. M. (1996). Treeview: an application to display phylogenetic trees on personal computers. *Computer Applications in the Biosciences* **12**, 357–358.

Downloaded from www.sgmjournals.org by
IP: 93.91.26.97
On: Tue, 28 Jul 2015 19:38:15

Pearson, W. R. & Lipman, D. J. (1988). Improved tools for biological sequence comparison. *Proceedings of the National Academy of Sciences, USA* **85**, 2444–2448.

Pringle, C. L. (1999). Virus taxonomy at the XIth International Congress of Virology, Sydney, Australia, 1999. *Archives of Virology* **144**, 2065–2068.

Takahashi, K., Ohta, Y. & Mishiro, S. (1998*a*). Partial ~2·4-kb sequences of TT virus (TTV) genome from eight Japanese isolates: diagnostic and phylogenetic implications. *Hepatology Research* **12**, 111–120.

Takahashi, K., Hoshino, H., Ohta, Y., Yoshida, N. & Mishiro, S. (1998*b*). Very high prevalence of TT virus (TTV) infection in general population of Japan revealed by a new set of PCR primers. *Hepatology Research* **12**, 233–239.

Takahashi, K., Iwasa, Y., Hijikata, M. & Mishiro, S. (2000). Identification of a new human DNA virus (TTV-like mini virus: TLMV) intermediately related to TT virus and chicken anaemia virus. *Archives of Virology* (in press).

Tanaka, Y., Mizokami, M., Orito, E., Ohno, T., Nakano, T., Kato, T., Kato, H., Mukaide, M., Park, Y.-M., Kim, B.-S. & Ueda, R. (1998). New genotypes of TT virus (TTV) and a genotyping assay based on restriction fragment length polymorphism. *FEBS Letters* **437**, 201–206.

Thompson, J. D., Higgins, D. G. & Gibson, T. J. (1994). CLUSTAL W: improving the sensitivity of progressive multiple sequence alignment through sequence weighting, position-specific gap penalties and weight matrix choice. *Nucleic Acids Research* **22**, 4673–4680.

Received 15 May 2000; Accepted 19 June 2000

Downloaded from www.sgmjournals.org by
IP: 93.91.26.97
On: Tue, 28 Jul 2015 19:38:15