Cloning and sequencing of TT virus genotype 6 and expression of antigenic open reading frame 2 proteins

Laura Kakkola, Klaus Hedman, Heidi Vanrobaeys, Lea Hedman and Maria Söderlund-Venermo

Department of Virology, Haartman Institute and Helsinki University Central Hospital, POB 21, FIN-00014 University of Helsinki, Finland

The near-full-length genome of a TT virus (TTV) (HEL32), closely related to the previously uncharacterized genotype 6, was cloned and sequenced. The genomic organization of HEL32 was compared to 41 published near-full-length TTV sequences representing 17 genotypes. In the majority of genomes, the open reading frame (ORF) 2 region was divided into two separate ORFs, 2a and 2b. The ORF2a sequence was conserved among all genotypes, while the ORF2b region showed more variability. The two corresponding putative proteins of HEL32 were expressed in prokaryotes and their antigenic potential was studied. IgM and IgG antibodies to the respective ORF2-encoded proteins, fp2a and fp2b, and the presence of TTV DNA were studied in the sera of 89 constitutionally healthy adults. By immunoblot using the small TTV proteins as antigens, strong IgM and IgG reactivities were found in 9 and 10% of subjects, respectively. Follow-up studies for 12–15 years of three subjects showed either a persistent coexistence of IgM and TTV DNA or the appearance of viral DNA regardless of pre-existing antibodies. The low prevalence of IgG could be due to the weak immunogenicity of these probably non-structural proteins or to a genotype-specific antibody response. By nested PCR of the conserved ORF2a region, the prevalence of TTV DNA was 85%. TTV genotype 6 sequences were found by specific PCR in 3 of 35 (8.6%) subjects.

The low prevalence of TTV IgG compared to the high TTV DNA prevalence, the coexistence of antibodies and viral DNA and the appearance of TTV DNA regardless of pre-existing antibodies suggest that the B-cell immunity against these minor TTV proteins would not be cross protective.

Introduction

TT virus (TTV) was discovered in 1997 from the serum of a Japanese patient with post-transfusion hepatitis of unknown aetiology (Nishizawa et al., 1997). TTV is a small, non-enveloped virus with a circular, single-stranded, negative-sense DNA genome of ~3.8 kb (Okamoto et al., 1998; Miyata et al., 1999; Mushahwar et al., 1999). TTV has similarities in genome organization to chicken anaemia virus (CAV), which is a member of the Circoviridae family, and thus TTV may represent the first circovirus to be found in humans (Miyata et al., 1999). Based on DNA sequence variation in the short N22 region, described by Nishizawa et al. (1997), TTV is classified into at least 27 different genotypes (Muljono et al., 2001; Okamoto et al., 2001). The evolutionary distance between the classified genotypes, as measured by nucleotide substitutions per site, is greater than 0.30 (Okamoto et al., 1999a). Besides the multitude of different genotypes, recombinants are also common (Worobey, 2000).

Full-length or near-full-length sequences from representatives of genotypes 1–3 and 10–23 have been deposited in GenBank. DNA sequence analysis of the identified TTVs indicate that most genotypes hitherto sequenced contain at least 2–3 open reading frames (ORFs) (Erker et al., 1999; reviewed by Bendinelli et al., 2001). The longest ORF, ORF1, encodes the putative capsid protein, whereas the shorter ORF2, often divided into ORFs 2a and 2b, possibly encode non-structural proteins. It has also been shown that mRNA splicing creates additional ORFs (Kamahora et al., 2000; Okamoto et al., 2000b).

As measured in serum or plasma by PCR, TTV is ubiquitous, with an age-dependent prevalence that approaches 94% among healthy adults (Hijikata et al., 1999a; Saback et al.,...
TTV DNA can persist for several months or years and multiple infections with several genotypes are frequently detected in the blood (Nishizawa et al., 1997; Ball et al., 1999; Biagini et al., 1999; Gallian et al., 1999; Okamoto et al., 1999a; Takayama et al., 1999; Niel et al., 2000) and in a diverse range of tissues (Okamoto et al., 2001). The prevalence of the various genotypes does not seem to differ geographically (Mushahwar et al., 1999; Gallian et al., 2000). Primate infection studies have shown TTV to be a transmissible agent (Mushahwar et al., 1999; Luo et al., 2000; Tawara et al., 2000) but it has not been shown to be the cause of any particular disease (reviewed by Bendinelli et al., 2001; Okamoto & Mayumi, 2001).

Diagnoses of TTV infections have been based almost solely on PCR. Due to high sequence variation, however, it is difficult to find a universal PCR primer set for all TTV genotypes (Okamoto et al., 1999a). Only a few TTV antibody studies have been published. These studies have used either fragments of the ORF1-encoded protein or crude TTV particles as antigen. Robust immunoprecipitation experiments have suggested the existence of TTV-specific antibodies (Nishizawa et al., 1999; Tsuda et al., 1999) and an IgM-capture method combined with PCR has revealed IgM class antibodies (Tsuda et al., 2001). TTV antibodies in serum have been detected by immunoblot assays using bacterially expressed N- and C-terminal fragments of the ORF1-encoded protein as antigen (Handa et al., 2000; Ott et al., 2000). In addition, electron microscopy has revealed that TTV particles in faeces are in free form as opposed to those found in serum, which are complexed with IgG antibodies (Itoh et al., 2000).

The aims of this study were as follows: (1) to clone and sequence a TTV genome, to map its genomic organization and compare it with other TTV genomes; (2) to express viral proteins; and (3) to study the antigenicity of those proteins. We have cloned and sequenced the near-full-length genome of a TTV (HEL32), representing a previously uncharacterized genotype 6, and compared the genomic organization of HEL32 with those of 41 TTV strains of 17 genotypes. Three ORFs were cloned, two of which (ORFs 2a and 2b) were successfully expressed and used as antigen in immunoblot assays. Specific antibodies and TTV DNA were measured in the sera of 89 non-symptomatic individuals, three of whom were followed retrospectively for 12–15 years.

Methods

Serum samples. Sera were obtained, with informed consent, from 89 non-symptomatic adults, 80 of whom were ethnically Finnish. DNA was purified from 10 to 100 µl of serum by proteinase-K treatment, phenol-chloroform extraction and ethanol precipitation, followed by resuspending to the original volume in water.

To avoid contamination, the samples and PCR mixtures were prepared in separate rooms and under laminar flow hoods, using disposable racks and aerosol-resistant tips. Water was included as a negative control.

Cloning of the TTV genome. A 3381 bp region of a TTV DNA isolate, termed HEL32, was amplified and cloned in two overlapping pieces. A 3269 bp PCR product, corresponding to nt 113–3381 of the HEL32 genome, was amplified by nested PCR from the serum of a non-symptomatic Finnish female, using primers NG133, NG135, NG134 and NG136 (Okamoto et al., 1999a). Both PCR reactions (50 µl) contained 200 µM of each dNTP (Roche), 320 nM of each primer, 1.5 mM MgCl$_2$ and 2.6 U of Expand High Fidelity enzyme mix (Boehringer Mannheim). The sample volume was 25 µl for the outer PCR and 2 µl for the inner PCR. The PCR protocol consisted of annealing at 54 °C for 30 s and extension at 68 °C for 2:5 min for the first 10 cycles, with 5 s per cycle added to the extension time for the remaining 20 cycles. The amplicons were electrophoresed, stained with ethidium bromide (EtBr) and visualized under UV. The band of the expected size was excised and DNA was isolated with the QiAquick Gel Extraction kit (Qiagen). The 3269 bp PCR product was cloned into pSTBlue-1 (Novagen) and transformed into Escherichia coli DH5α cells.

Of the TTV genome, nt 1–222 were amplified with the semi-nested primers NG054, NG147 and NG132 (Okamoto et al., 1999a). The reaction mixture and PCR program were as described for universal PCR (see below), except that annealing occurred at 50 °C, the extension time was 45 s and the cycles were repeated 30 times, with a final extension of 3 min. Amplicons were isolated and cloned as described above.

Sequence analysis. All new TTV sequences reported in this paper were obtained using ABI PRISM (Perkin-Elmer) at the sequencing core facility of the Haartman Institute, University of Helsinki, Finland. The genomic organization in HEL32 was determined and compared to 41 other near-full-length TTV sequences (obtained from GenBank) using MALIGN (DNAstar) and DNASTRIDER, version 1.0 (Marc C., Commissariat a l‘Energie Atomique, France). Multiple sequence alignment and evolutionary distance analyses were obtained using the MALIGN, CLOUDBREW and DISTANCES programs, all of which are part of the Wisconsin package 10.1 (Genetics Computer Group) and are maintained by the Center for Scientific Computing (Espoo, Finland). For genotyping, the sequence of the N22 area of HEL32 was compared to those of 20 published TTV genotypes (Mulono et al., 2001) and the ORF1 sequence was compared to 17 genotypes reported by Okamoto et al. (2001). Phylogenetic trees (see Fig. 2) were constructed with TREEPZUZZLE using the maximum-likelihood approach (Strimmer & von Haesler, 1997); 10000 puzzling steps were applied using the HKY model of substitution (Hasegawa et al., 1985). For confirmation, the trees were also constructed using the neighbour-joining algorithm of the PHYLIP program package with 500 bootstrap replicates (Felsenstein, 1989). The program BLAST was used for database searches (NCBI).

Cloning of ORFs into expression plasmids. In-frame PCR primers specific for each of the three ORFs (Fig. 1) were designed. Primers were as follows: for ORF1, forward 5′-TCGCCAATTCATGGCCTGCT-3′ and reverse 5′-GCTTCCTGGGCAAGCAGCCG-3′; for ORF2a, forward 5′-CCTGGGATCC-3′ and reverse 5′-TCGAGGTTTTCCACGCCCGTC-3′; and for ORF2b, forward 5′-TCGAGGTTTTCCACGCCCGTC-3′ and reverse 5′-TCGAGGTTTTCCACGCCCGTC-3′. In each primer, restriction enzyme sites (underlined) were incorporated to allow cloning into the expression plasmid. ORFs 2a and 2b were amplified together as a single PCR amplicon (ORF2a/b). ORF2b was also amplified separately. The reverse primers were designed downstream of the stop codons in each ORF. However, to verify the presence of the ORF2a stop codon at nt 255, the same reverse primer that was used for ORF2b was employed.
ORF2 of TTV genotype 6

Fig. 1. Schematic representation of TTV HEL32 in comparison to TA278 (GenBank accession number AB017610). Two overlapping, cloned PCR products of HEL32, labelled a and b, are shown in grey. Three putative ORFs, ORFs 2a (nt 106–255), 2b (nt 239–709) and 1 (nt 583–2793), are shown in black. The N22 (nt 1843–2171) and universal (nt 90–232) PCR regions are also indicated (thin bars).

ORFs were amplified from the HEL32 plasmid clone containing the 3269 bp PCR insert. For ORF2a, nt 109–112, which were not present in the 3269 bp plasmid clone, were included in the ORF2a forward primer. The purified ORF restriction fragments were cloned into pGEX-4T-1 expression plasmids (Amersham Pharmacia) and the resulting plasmids were transformed into E. coli DH5α cells. The sequences of the entire cloned ORF inserts were verified by sequencing.

Expression of TTV proteins. The ORF amplicons in pGEX-4T-1 were expressed as GST fusion proteins in E. coli strain BL21 (Amersham Pharmacia). The GST protein alone was also expressed as a control. IPTG was used (0.4 mM) for 2-, 4-, 6- and 8-h inductions, both at room temperature and at 37 °C. Cells were lysed as described previously (Söderlund et al., 1992). Whole-cell lysates were studied by 10% SDS–PAGE and protein bands were visualized with Coomassie stain (Serva).

Immunoblotting. Proteins from SDS–PAGE were transferred to a Protran nitrocellulose membrane (Schleicher & Schuell) at 400 mA for 2 h (Söderlund et al., 1992). Membranes were blocked with 5% fat-free milk powder (Valio) and 0.2% Triton X-100 in PBS. For immunodetection of the GST fusion proteins, the primary antibody was goat anti-GST (1:1000; Amersham Pharmacia) and the secondary antibody was peroxidase-conjugated, affinity purified rabbit anti-goat IgG (1:100; Jackson Immunoresearch). Peroxidase-conjugated, affinity purified rabbit anti-goat IgG (1:100; Amersham Pharmacia) and the secondary antibody was the GST fusion proteins, the primary antibody was goat anti-GST (Jackson Immunoresearch).

The respective peroxidase conjugates were horseradish peroxidase-conjugated, rabbit anti-human IgG and IgM (Dako); bound antibodies were detected with hydrogen peroxide and DAB (Söderlund et al., 1992). Sera showing antibody reactivity with the GST fusion proteins fp2a and fp2b were re-examined with the expressed GST control protein and uninduced lysates of E. coli. The possibility of rheumatoid factor interference in IgM detection was eliminated by the prior removal of IgG using Gullsorb (Meridian Diagnostics).

Universal PCR. A conserved sequence (nt 90–232) (Fig. 1), immediately downstream of the untranslated region and partially overlapping ORF2a, was amplified using primers NG133, NG147, NG134 and NG132 (Okamoto et al., 1999a), resulting in a 110 bp product. PCR mixtures contained 200 μM of each dNTP (Roche), 600 nM of each primer, 1.5 mM MgCl₂ and 2.5 U of AmpliTaq Gold polymerase (Applied Biosystems/Roche) in 25 μl. The sample volume in the outer PCR was 8 μl and in the inner PCR was 2 μl. Amplification conditions were as described by Okamoto et al. (1999a), except that both PCRs comprised 35 cycles. Amplicons were detected by agarose gel electrophoresis followed by EtBr staining.

Products of the universal PCR were confirmed by dot-blot hybridization. A digoxigenin-labelled 110 bp probe was prepared by incorporation of digoxigenin-11 dUTP (Boehringer Mannheim) during a separate universal PCR of cloned HEL32. Membranes were hybridized at 42 °C overnight. Non-labelled amplicons of the probe PCR, the PCR product of a TTV-negative control, pSTBlue-1 either with and without nt 1–222 of HEL32 and water were included as the controls.

Three universal PCR products that were longer than expected, as well as four randomly chosen amplicons, were cloned into pSTBlue-1 and the plasmid inserts were sequenced.

N22 PCR. The N22 region of ~270 bp, illustrated in Fig. 1 (Nishizawa et al., 1997), was amplified by nested PCR using primers TT6–9 (Höhne et al., 1998). PCR conditions were as described for universal PCR, except that annealing was at 42 °C and the elongation time was 45 s, plus 3 min at the end.

Genotype 6 PCR and hybridization. Nested primers within the N22 region were designed for genotype 6-specific detection of HEL32-like DNA. The primers used were as follows: for outer PCR, forward 5′-GCCCTGGAGCATACCAAG 3′ (nt 1805–1822) and reverse 5′-CATGACTCTAGCTGTTGGAAC 3′ (nt 2181–2202); for inner PCR, forward 5′-CAGTATACTCAGAAAAAGCAGAAG 3′ (nt 1898–1919) and reverse 5′-CATTTCCACCTCATTCTTATAGG 3′ (nt 2146–2168). Amplification conditions were as described for universal PCR, except that the annealing temperature was 55 °C for the outer PCR and 52 °C for the inner PCR.

The genotype 6 PCR products underwent dot-blot hybridization. A digoxigenin-labelled 273 bp probe was prepared as described above, but using the inner genotype 6 PCR primers and cloned HEL32 DNA as template. The genotype 6 PCR products obtained were also sequenced.

Results

HEL32 cloning and sequence analysis

The overlapping 3269 bp and 222 bp PCR products (Fig. 1) spanning nt 1–3381 were cloned and sequenced (GenBank accession number AY034068). Comparison of the entire sequences of HEL32 and the prototype TTV, TA278 (AB017610), revealed a divergence of > 40%. Phylogenetic comparison of the ORF1 sequence of HEL32 with those of 29 other TTV DNA isolates (Okamoto et al., 2001) revealed by high support values that HEL32 is unique (Fig. 2a). When comparing only the N22 region between HEL32 and 42 sequences representing 20 TTV genotypes (Muljono et al., 2001), HEL32 was most closely related to genotype 6 (Fig. 2b).

Sequence analysis of HEL32 DNA revealed three possible
Fig. 2. A phylogenetic tree based on (a) ORF1 and (b) N22 sequences of TTV DNA isolates found in GenBank and HEL32. Support values are given in numbers next to branches. Names of the TTV isolates are followed by the respective genotype (gX).
ORFs (Figs 1 and 3a): ORF1 in reading frame 1 (nt 583–2793), ORF2a in reading frame 1 (nt 106–255) and ORF2b in reading frame 2 (nt 239–709). Additional small ORFs were found in reading frames 1 (nt 2857–2961), 2 (nt 2486–2812) and 3 (nt 2805–2984). A TATA box (TATAA) was recognized upstream of ORF2a at nt 85 and a variant poly(A) sequence (ATTAAA) was situated 186 nt downstream of the ORF1 gene.

The genomic organization of HEL32 was compared to those of 41 TTV DNA isolates, representing 17 different genotypes. Out of these 41 DNA isolates, 34 had an ORF2 organization that was similar to HEL32 (Fig. 3a–c), whereas only seven (all of genotype 1) had an organization like that of TA278 with a single ORF2 (Fig. 3d). Of the 34 HEL32-like isolates with separate ORFs 2a and 2b, nine had an organization identical to HEL32 (Fig. 3a), ten contained ORFs 2a and 2b in the same reading frame, with a stop codon in between (Fig. 3b), and 15 contained ORF2a in a reading frame distinct from that of both ORFs 1 and 2b (Fig. 3c). ORF2b was defined as the ORF comprising the conserved motif WX,H,CX,H present in all TTVs and CAV (Hijikata et al., 1999b).

When the ORF2a sequences of all the 42 TTV DNA isolates were compared to each other, this region was found to
Table 1. TTV DNA prevalence in 87 non-symptomatic subjects correlated with IgM reactivities for small TTV proteins

<table>
<thead>
<tr>
<th></th>
<th>IgM +</th>
<th>IgM ±</th>
<th>IgM −</th>
<th>Total (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>IgM for fp2a PCR-positive</td>
<td>2 (2)</td>
<td>2 (2)</td>
<td>71 (82)</td>
<td>75 (86)</td>
</tr>
<tr>
<td>PCR-negative</td>
<td>0</td>
<td>1 (1)</td>
<td>11 (13)</td>
<td>12 (14)</td>
</tr>
<tr>
<td>Total</td>
<td>2 (2)</td>
<td>3 (3-4)</td>
<td>82 (94.3)</td>
<td>87 (100)</td>
</tr>
</tbody>
</table>

|                        | IgM for fp2b PCR-positive | 4 (5) | 5 (6) | 66 (76) | 75 (86) |
| PCR-negative           | 2 (2) | 1 (1) | 9 (10) | 12 (14) |
| Total                  | 6 (7) | 6 (7) | 75 (86) | 87 (100) |

Table 2. TTV DNA prevalence in 87 non-symptomatic subjects correlated with IgG reactivities for small TTV proteins

<table>
<thead>
<tr>
<th></th>
<th>IgG +</th>
<th>IgG ±</th>
<th>IgG −</th>
<th>Total (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>IgG for fp2a PCR-positive</td>
<td>0</td>
<td>2 (2)</td>
<td>73 (84)</td>
<td>75 (86)</td>
</tr>
<tr>
<td>PCR-negative</td>
<td>0</td>
<td>0</td>
<td>12 (14)</td>
<td>12 (14)</td>
</tr>
<tr>
<td>Total</td>
<td>0</td>
<td>2 (2)</td>
<td>85 (98)</td>
<td>87 (100)</td>
</tr>
</tbody>
</table>

|                        | IgG for fp2b PCR-positive | 9 (10) | 1 (1) | 65 (75) | 75 (86) |
| PCR-negative           | 0     | 0     | 12 (14) | 12 (14) |
| Total                  | 9 (10) | 1 (1) | 77 (89) | 87 (100) |

be highly conserved both at the nucleotide level and at the amino acid level. The ORF2a nucleotide and amino acid divergences between TA278 and HEL32 were 8 and 12%, respectively. Also, the length (49 aa) of the predicted ORF2a product was conserved in 26 of 34 ORF2a-containing isolates. The ORF2b region varied both in sequence (divergence of > 50%) and in length. The lengths of the predicted ORF1 products were found to vary by 646–816 aa. Interesting exceptions were isolates US32 (detected by Erker et al., 1999) and SENV-B, which had stop codons within the ORF1 regions that disrupted the potential proteins. The ORF1-encoded protein of HEL32 was predicted to be 736 aa in length and to contain the conserved arginine-rich region described by Okamoto et al. (1998) and Mushahwar et al. (1999).

Expression of TTV proteins

The three putative ORFs (ORFs 1, 2a and 2b) were amplified by PCR and cloned into a GST fusion expression plasmid. Both fp2a and fp2b were successfully expressed at 4, 6 and 8 h post-induction, both at room temperature and at 37 °C (Fig. 4a), with increasing yields over time. However, no protein product was obtained from the ORF1–GST construct. The correct molecular masses of fp2a (32 kDa), fp2b (43 kDa) and the GST moiety (26 kDa) were confirmed by immunoblotting with an anti-GST antibody (Fig. 4b). In all future assays, an 8 h induction at room temperature was used.

Antibody detection

For the detection of specific IgM and IgG antibodies, the recombinant fp2a and fp2b proteins were used as antigen in immunoblot experiments with serum samples from 89 non-symptomatic adults (the TTV DNA or antibody status of 2 of 89 subjects changed during follow up and are therefore not included in Tables 1 and 2). We classified as positive both very strong and moderately strong antibody reactivities and as
borderline, lighter bands (example shown in Fig. 5b). We classified the result as negative if no, or extremely faint, antibody reactivity was detectable or if repeated tests showed poorly reproducible borderline results.

Of 87 subjects, 2 (2%) showed strong fp2a IgM reactivity (Table 1). Both were fp2a and fp2b IgG-negative. Strong fp2b IgM reactivity was observed in six (7%) subjects (Table 1), one of which was also strongly IgG-positive for the same TTV protein. By inclusion of borderline results, these IgM prevalences would rise to 6 and 14%, respectively. By universal PCR, 6 of 8 strongly fp2 IgM-positive subjects were also TTV DNA-positive. Among the 87 subjects, strong fp2a IgG reactivity was not seen, whereas borderline reactivity was seen in two (2%) subjects (Table 2). Strong fp2b IgG reactivity was found in 9 of 87 (10%) subjects, in addition to one borderline result (Table 2). All subjects with strong or borderline fp2a or fp2b IgG results had circulating TTV DNA, i.e. all TTV DNA-negative subjects were also IgG-negative for both antigens. On the other hand, among the 63 seronegative subjects, 55 (87%) contained TTV DNA in their serum.

None of the TTV antibody-positive sera showed comparable reactivity with lysates of uninduced E. coli or with the GST protein moiety (Fig. 5). The IgM result of the only sample that was both IgM- and IgG-positive for the same antigen (fp2b), was retained after Gullsorb treatment, i.e. it was not abrogated by the removal of IgG.

**Follow-up studies**

Sequential serum samples covering 12–15 years were available from three subjects and a fourth subject was followed for 1–5 years (Fig. 6). The DNA and antibody status remained unaltered in two subjects (#26 and #55), who maintained very strong fp2b-specific IgM reactivity but no IgG reactivity in all samples. Both subjects were TTV DNA-positive for at least 14 and 15 years, respectively. The DNA and antibody status of two subjects changed with time. Subject #1 showed strong fp2b IgM reactivity for three years, after which it gradually declined below levels of detection. This subject was the only one in the study presenting with strong fp2a IgG immunoreactivity, which, furthermore, persisted throughout the 12 year follow-up period. Only the last sample was TTV DNA-positive. Subject #97 showed fluctuating borderline fp2b IgM and IgG reactivity for 1–5 years and, in the last sample, converted to TTV DNA-positivity. The TTV strains of subjects #1 and #97, detected by universal PCR, were (as described below) closely related to the TTV-like miniviruses (TLMV; Okamoto et al., 2000a; Takahashi et al., 2000).

**Prevalence of TTV DNA**

Serum samples of all 89 subjects were studied for TTV DNA by universal PCR. The sera of most (80) subjects were also analysed by N22 PCR. Each PCR was repeated at least once and the universal PCR results were confirmed by dot-blotting and sequencing. Six samples showed poorly reproducible results in universal PCR, probably due to low DNA concentrations, but were considered to be TTV DNA-positive. All water controls were always negative.

TTV DNA in serum was detected by universal PCR in 85% (74 of 87) of the subjects (Tables 1 and 2). Two additional subjects became positive during follow-up (see above and Fig. 6). The less sensitive, or more specific (detecting fewer genotypes), N22 PCR was positive in 18% (14 of 80) of subjects, one of whom was negative by universal PCR. Thus, the overall prevalence of TTV DNA among these 87 non-symptomatic subjects was 86% (75 of 87).

By universal PCR, three samples (subjects #1, 52 and 97) gave PCR products that were longer than expected (130–160 bp versus 110 bp) and did not hybridize with the TTV probe. To avoid potential interference of multiple genotypes in sequencing, these three amplicons were cloned into plasmids and sequenced. Sequence comparison, using the BLAST and CROWTREE programs, showed that all three sequences grouped
Fig. 6. Prevalence of TTV DNA and antibodies in three patients followed up over a 12–15 year period and one patient followed up over a 1.5 year period. ND, Not done.

with those of TLMVs (Okamoto et al., 2000a; Takahashi et al., 2000) (data not shown).

Genotype 6 DNA detection

For detection of TTVs of the HEL32 genotype, samples of all 17 IgM- and/or IgG-positive subjects (15 of whom were, or became, TTV DNA-positive) and 18 antibody negative subjects (16 of whom were TTV DNA-positive), including all samples of three subjects followed up (#1, #26 and #55), were studied by genotype 6 PCR. Serum of the HEL32 donor and the plasmid clone were included as positive controls. PCR results were confirmed both by sequencing and by hybridization with a genotype 6-specific probe.

By genotype 6 PCR, only 2 (#23 and #86) of 35 subjects, in addition to the HEL32 donor, had TTV of genotype 6, revealing a genotype 6 DNA prevalence of 8.6% in total and 9.7% (3 of 31) in the TTV DNA-positive group. Comparison of the amplified DNA sequences of subjects #23 and #86 to that of HEL32 revealed 2 and 8 of 213 nucleotide differences, respectively. Subject #23 was also universal, but not N22 PCR-positive, whereas subject #86 was both universal and N22 PCR-positive. Both subjects had IgG antibodies forfp2b. The donor of our HEL32 clone, on the other hand, lacked antibodies forfp2a andfp2b, but was both universal and N22 PCR-positive and, based on the sequenced PCR products, carried at least two other TTV genotypes (data not shown).

Discussion

We have cloned the near-full-length genome of a TTV (HEL32), which, based on sequence comparison, is closely related to genotype 6. This is, to our knowledge, the first description of a genotype 6 TTV genome. By phylogenetic analysis of the ORF1 gene, our HEL32 DNA isolate was distant from all other TTV DNA isolates. A conserved TATA box was, however, found to begin at nt 85, as in other TTV genotypes (Miyata et al., 1999; Okamoto et al., 1999b; Kamahora et al., 2000; Tanaka et al., 2000), whereas a conserved poly(A) signal, located 177 nt downstream of ORF1 (Erker et al., 1999), was not found in HEL32. Instead, a variant poly(A) signal (ATTTAA) was identified 186 nt downstream
of ORF1. Okamoto et al. (1998) reported that strain TA278 contained such a variant motif but at a different location.

In CAV, the longest ORF encodes the capsid protein, whereas the other ORFs encode non-structural proteins (Todd et al., 1990; Noteborn et al., 1992; Douglas et al., 1995; Kato et al., 1995). In light of the similarity of genome organization, TTV ORF1 could encode the capsid protein and ORFs 2a and 2b the non-structural proteins. We compared the ORF2 sequences of 18 genotypes (altogether 41 sequences obtained from GenBank, plus our HEL32 strain) and found ORF2a to be highly conserved in all TTV strains. In contrast, the ORF2b sequences varied extensively. This has been reported previously, albeit with fewer genotypes (Erker et al., 1999; Tanaka et al., 2000).

In HEL32, ORFs 1 and 2a are in the same reading frame and ORF2b is in a separate one, partly overlapping ORF2a. Similar genomic organizations were found in several genotypes. Starting at ATG in frame 1, the ORF2a of HEL32 encodes a putative protein of 49 aa. ORF2b, in frame 2, contains three possible transcription initiation sites at nt 239, 272 and 356, and the longest transcript encodes a putative protein of 156 aa.

In recent TTV mRNA studies, due to late transcription start sites and splicing (removing the first two initiation sites), only the third ATG at nt 353 or 338 was suggested to be used for translation of the ORF2 gene (Kamahora et al., 2000; Okamoto et al., 2000b). Consequently, if this would apply for HEL32, the ORF2a-encoded protein would not be produced at all and the ORF2b-encoded protein would be shortened considerably. However, by Kozak’s criteria (Kozak, 1996), this third A^{356}TG of ORF2b in HEL32 is in a weak context for initiation and is therefore not likely to be a start site, whereas the context of the first A^{239}TG is considered to be strong. Alternative transcription initiation (Ellis et al., 1987; Suzuki et al., 2001) and splicing may thereby occur in HEL32-like TTV genotypes, allowing the use of these first initiator codons. Interestingly, we found in every single TTV sequence, at nt ~ 107, an ATG initiation site that adequately satisfied Kozak’s criteria (Kozak, 1996). The short 5’ leader upstream of this ATG would not necessarily hinder translation, as shown for even shorter leaders (Ellis et al., 1987; Maicas et al., 1990). Furthermore, based on sequence comparison, the ORF2a region is highly conserved among all genotypes (also at the protein level). Obviously, this highly conserved area has some important function, either at the protein level or at the nucleotide level. Finally, by in vitro transcription and translation, the full-length 202 aa protein of genotype 1 ORF2 has been produced, beginning from A^{197}TG (Tanaka et al., 2000). However, in a genotype 3-related isolate with a divided ORF2, ORF2a, starting at A^{197}TG was, for some reason, not translated. The translation of ORF2b began at the first A^{356}TG initiation site. The more important in vivo effects of transcription and translation initiation sites and alternative splicing of ORFs 2a and 2b of HEL32 genotype 6 (and other similar genotypes) remain to be studied.

We have in E. coli productively expressed both of the HEL32 ORF2 putative protein-coding regions, ORFs 2a (nt 106–255) and 2b (nt 239–709). Interestingly, ORF1 resisted expression, possibly due to the cytotoxicity of the product. The suitability of the corresponding recombinant proteins fp2a and fp2b as antigens for the detection of specific IgM and IgG antibodies was investigated in constitutionally healthy adults.

Fp2a- and fp2b-specific IgM antibodies were found in an unexpectedly high (~ 9%) prevalence. By using TTV particles of faecal origin, short-lived IgM responses have been previously detected in three liver patients but not in healthy controls (Tsuda et al., 2001). Another study found no IgM reactivity against the C terminus of the ORF1 protein among patients with hepatitis of unknown aetiology or among blood donors or healthy children (Ott et al., 2000). In most virus infections, the typical IgM responses are early and short-lived. However, we observed strong fp2b IgM responses that persisted for years. Most of our fp2 IgM-positive individuals were fp2 IgG-negative but carried circulating TTV DNA, a pattern that could possibly infer an acute stage preceding virus clearance. However, such a conclusion lends no support to our follow-up study, where persistent IgM and DNA co-existed. The reason for the absence of an immunoglobulin class shift (IgM → IgG) in our long-term follow-up subjects is at present unknown. One possible reason could be impaired helper T-cell activity, which in turn might be due to T-cell cross-tolerance against multiple TTV genotypes.

In the present study, fp2a- and fp2b-specific IgG antibodies were also detected but, relative to IgM, in a surprisingly low (10%) prevalence. Interestingly, all the IgG-positive subjects had TTV DNA in their sera. Two other immunoblot studies, using prokaryotically expressed proteins of TTV genotype 1 as antigen, have reported considerably higher TTV IgG prevalence values of 38% for the N-terminal ORF1 product (Handa et al., 2000) and 98%6% for the C-terminal part (Ott et al., 2000). In both studies, the ORF1-specific antibodies and TTV DNA coexisted. In light of their high prevalence (Ott et al., 2000), the ORF1-specific antibodies could be highly cross-reactive and non-protective.

Using two nested PCRs, we found the prevalence of TTV DNA to be 86%. This result is in line with a previous study that reported a TTV DNA prevalence of 73% in Finnish blood donors (Simmonds et al., 1999). Other studies worldwide have yielded comparable values of TTV DNA prevalence; e.g. for Italian blood donors, the prevalence of TTV DNA was 87.5% (Zehender et al., 2001) and for healthy Japanese subjects, the prevalence of TTV DNA was 94% (Hijikata et al., 1999a; reviewed Bendinelli et al., 2001).

By genotype 6-specific PCR, 3 of 35 (8.6%) subjects had TTV of genotype 6. However, because HEL32 is the only available (long) sequence of TTV genotype 6, the extent of intra-genotype sequence variation in the amplified regions that could affect the PCR result is not known. Of the three subjects with genotype 6 DNA, two had coexistent fp2b IgG but none
had IgM. The donor of our HEL32 clone carried at least two other TTV genotypes. Superinfections with several TTV genotypes are well documented, inferring lack of inter-genotype protection (Okamoto et al., 1999a; Takayama et al., 1999; Niel et al., 2000; Romeo et al., 2000).

Altogether, our results suggest that the ORF2 protein-specific antibodies are not cross-protective, because TTV DNA was found to coexist with the fp2 antibodies. The fp2b-specific antibodies could, theoretically, also be genotype-specific, because (1) the ORF2b sequence varies considerably among different TTV genotypes and (2) the fp2b antibodies were detected only in a minority of our subjects. On the other hand, the ORF2a sequences are very similar in all genotypes, also at the putative protein level. Consequently, antibodies for this protein are not likely to be genotype-specific. The very low prevalence of fp2a-specific antibodies could result from inefficient protein production (Kamahora et al., 2000; Okamoto et al., 2000b; Ellis et al., 1987), as discussed above. The low prevalence of fp2 antibody and the apparent lack of cross-protectiveness of these antibodies could also be easily understandable assuming that these proteins are non-structural and intracellular. Furthermore, it should be kept in mind that immunoblotting tends to favour antibodies against linear epitopes, the relative proportion of which may change during the course of virus infection (Sundsrud et al., 1995; Kaikkonen et al., 1999). Furthermore, prokaryotic expression does not create the post-translational modifications that may be required for optimal antigenicity. Finally, an interesting possibility is that the ORF2b antibodies are produced at very low levels in healthy individuals but more abundantly in some diseases yet to be associated with this virus, as has been suggested (von Poblotzki et al., 1995; Hemauer et al., 2000) for the non-structural protein of parvovirus B19, another human single-stranded DNA virus.

The authors would like to thank Tarja Sironen for help in the construction of the phylogenetic trees and Anna Harju and Sanna Ilvonen for excellent technical assistance. This research was financially supported by the Helsinki Biomedical Graduate School, The Sohberg Foundation, the ERASMUS (Socrates) student exchange program, the Finnish Foundation for Research on Viral Diseases, The Finnish Technology Advancement Fund, the Helsinki University Central Hospital Research and Education Fund and the Finnish Academy (project #76132).

References


Hijikata, M., Iwata, K., Ohta, Y., Nakao, K., Matsumoto, M., Matsumoto, H., Kanai, K., Baba, K., Samokhvalov, E. I. & Mishiro, S. (1999a). Genotypes of TT virus (TTV) compared between liver disease patients and healthy individuals using a new PCR system capable of differentiating 1a and 1b types from others. Archives of Virology 144, 2345–2354.


Received 16 May 2001; Accepted 7 January 2002