Journal of General Virology (1998), **79**, 3085–3089. Printed in Great Britain SHORT COMMUNICATION

Sequence variability in the putative coding region of TT virus: evidence for two rather than several major types

S. Viazov,[1,2] R. S. Ross,[1] C. Niel,[3] J. M. de Oliveira,[3] C. Varenholz,[4] G. Da Villa[5] and M. Roggendorf[1]

[1] Institute of Virology, University of Essen, Robert-Koch-Haus, Hufelandstr. 55, 45122 Essen, Germany

[2] Ivanovsky Institute of Virology, Gamaleya St 16, Moscow 123098, Russia

[3] Department of Virology, Oswaldo Cruz Institute, Avenida Brasil 4365, 21040-900 Rio de Janeiro, Brazil

[4] Institute of Hygiene and Occupational Medicine, University of Essen, Hufelandstr. 55, 45122 Essen, Germany

[5] Italian Institute for Prevention of Liver Disease, Via Orsini 42, 80132 Naples, Italy

Recently a new human virus, TT virus (TTV) was identified in the serum of a patient with post-transfusion hepatitis of unknown aetiology. Comparative sequence analysis of a 222 nt fragment of ORF1 of TTV was performed to assess the genomic variability of this virus. Phylogenetic analysis of the nucleotide sequences of 76 TTV isolates collected in 17 countries segregated them into two major groups: TTV 1 and TTV 2. The TTV 1 group comprised two distinct subgroups, which corresponded to previously described TTV subtypes 1a and 1b. The TTV 2 group was separated into four main branches, two of which included sequences previously provisionally attributed as TTV types 2 and 3. Bootstrap resampling, however, did not support the reliability of this grouping, suggesting that the isolates in the TTV 2 group should be considered as subtypes of a single type rather than different TTV types.

Recently, a novel human infectious agent was identified in a serum sample of a Japanese patient with post-transfusion hepatitis of unknown aetiology (Nishizawa *et al.*, 1997; Okamoto *et al.*, 1998). This agent was designated TT virus (TTV), after the name of the patient. The TTV genome is a single-stranded DNA of at least 3739 bases and contains two overlapping open reading frames (ORF1 and ORF2), encoding 770 and 202 amino acids, correspondingly. TTV is resistant to Tween 80 treatment and has a buoyant density of 1·26 g/ml in a sucrose gradient. Viral DNA can be detected in plasma but also in liver tissues of infected subjects, suggesting that TTV is hepatotropic. No serological tests for TTV infection markers are available and PCR is at present the only available diagnostic tool. Use of PCR has demonstrated that TTV causes both acute and persistent infections in humans but the epidemiology and clinical significance of these infections remain uncertain. TTV can be transmitted parenterally by blood and blood products (Nishizawa *et al.*, 1997; Okamoto *et al.*, 1998; Simmonds *et al.*, 1998) and probably also non-parenterally by a faecal–oral route (Miyakawa *et al.*, 1998). TTV DNA is present in a large proportion of patients with different forms of non-A–G hepatitis. For example, it was detected in 47% of patients with fulminant hepatitis and 46% of patients with chronic liver disease of unknown aetiology in Japan (Nishizawa *et al.*, 1997; Okamoto *et al.*, 1998), in 19% of patients with fulminant hepatic failure from Scotland (Simmonds *et al.*, 1998), and in 25% of patients with chronic liver disease from England (Naoumov *et al.*, 1998). Of importance is the fact that in the last study the majority of TTV-positive patients had no biochemical or histological evidence of significant liver damage. This finding, as well as the data on the presence of TTV DNA in a surprisingly high proportion (1·9–63%) of evidently healthy volunteer blood donors from several countries (Okamoto *et al.*, 1998; Simmonds *et al.*, 1998; Niel *et al.*, 1998), suggest the possibility of asymptomatic carriage of TTV. These results also indicate that TTV is probably not the causative agent of hepatitis, at least in the majority of patients with chronic TTV infection (Naoumov *et al.*, 1998; Cossart, 1998). Absence of strong evidence for an association between TTV infection and liver disease or any other clinical condition recalls the situation with the human virus GBV-C, and suggests that TTV may be another 'harmless' virus.

Phylogenetic analysis of a number of TTV isolates from Japan provided evidence for TTV genome heterogeneity and the existence of several viral types and subtypes (Okamoto *et al.*, 1998). Preliminary data on the sequence variability of TTV strains, confirming these observations, have been presented

Author for correspondence: Sergei Viazov (at University of Essen). Fax +49 201 723 5929. e-mail sergei.viazov@uni-essen.de

The GenBank accession numbers of the sequences reported in this paper are AF067973–AF067984, AF081078–AF081087 and AF084105–AF084137.

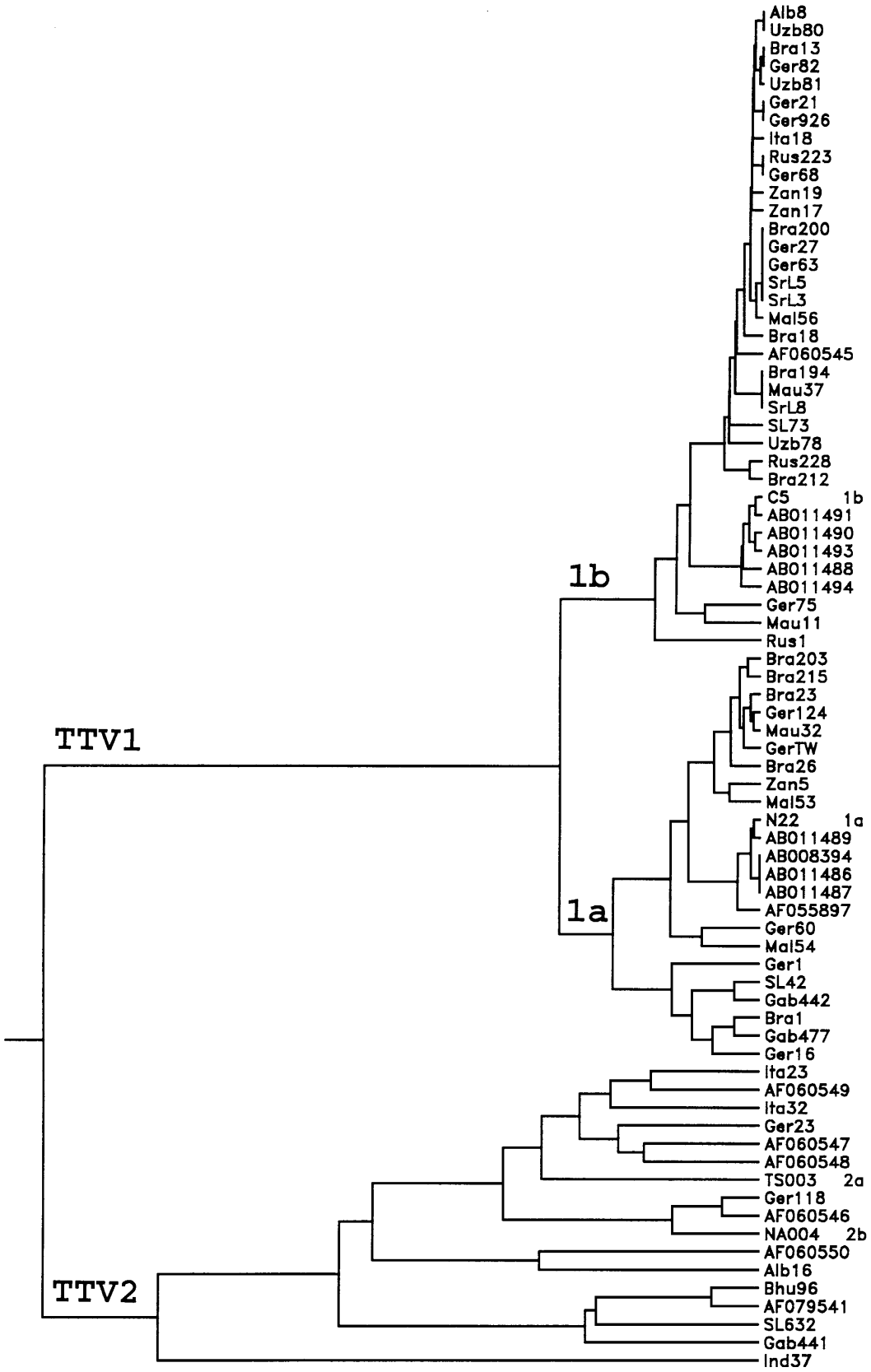

Fig. 1. For legend see facing page.

recently for another part of the world (Simmonds *et al.*, 1998; Naoumov *et al.*, 1998). Thus, at the moment we cannot exclude the possibility that different TTV genetic variants might possess distinct biological potential and that some of them might be more virulent. In the present paper, we report comparative sequence analysis of a DNA fragment from the putative ORF1 of 76 TTV isolates collected in 17 countries and provide some data suggesting even more pronounced genome heterogeneity for TTV than was reported previously.

DNA was purified from 200 µl of serum with the QIAamp blood kit (Qiagen), suspended in 50 µl of distilled water, and amplified by semi-nested PCR as described elsewhere (Okamoto *et al.*, 1998). Fifty-five TTV DNA-positive individuals were identified and enrolled in this study: two blood donors from Albania (Alb); one blood donor from Bhutan (Bhu); five blood donors and five hepatitis non-A–G patients from Brazil (Bra); three human immunodeficiency virus-infected individuals from Gabon (Gab); seven patients with end-stage liver disease, two patients with haematological disorders and five blood donors from Germany (Ger); one chronic hepatitis patient from Indonesia (Ind); three blood donors from Italy (Ita); three blood donors from the Maldive Islands (Mal); three blood donors from Mauritius (Mau); two chronic hepatitis patients and one blood donor from Russia (Rus); three human immunodeficiency virus-infected persons from Sierra Leone (SL); three blood donors from Sri Lanka (SrL); three patients with chronic hepatitis from Uzbekistan (Uzb); and three blood donors from Zanzibar (Zan). PCR products were purified from agarose gel with the QlAquick gel extraction kit (Qiagen) and subjected to direct sequencing in both directions using the ABI PRISM dye terminator cycle sequencing ready reaction kit (Perkin Elmer). Twenty-one TTV DNA sequences were taken from GenBank: six sequences from Germany, accession nos AF060545–AF060550; thirteen sequences from Japan, AB008394, AB011486–AB011494 and from Nishizawa *et al.* (1997) and Okamoto *et al.* (1998); one sequence from Scotland, UK, AF079541 (Simmonds *et al.*, 1998); and one TTV DNA sequence from China, AF055897. Phylogenetic analysis of sequences was accomplished with the programs Seqboot, Dnadist, Neighbor, Consense and Drawgram from the package PHYLIP, version 3.5c (Felsenstein, 1993).

A partial ORF1 sequence of 222 bp was determined for 55 TTV isolates. The resulting sequences were compared with each other, with five sequences of prototype TTV strains (Nishizawa *et al.*, 1997; Okamoto *et al.*, 1998; Simmonds *et al.*, 1998), and with TTV sequences available in GenBank. Overall, 76 partial DNA sequences from TTV isolates collected in 17 countries were subjected to a phylogenetic analysis. Two levels of sequence diversity may be seen in the consensus UPGMA tree depicted in Fig. 1. All TTV sequences were segregated into two major groups. Within the first group two clusters of more closely related variants were observed. These two clusters correspond to the previously described subtypes 1a and 1b (Okamoto *et al.*, 1998). The bootstrap analysis (1000 replicates) confirmed the reliability of these groupings; the corresponding values for two major groups and for subtypes 1a and 1b were higher than the arbitrary cut-off value of 75%.

Sequences of the second major group were also further separated into several branches. Two of them included previously described prototype sequences of subtypes 2a and 2b (Okamoto *et al.*, 1998); the third branch included sequence AF079541, which has been recently attributed to type 3 (Simmonds *et al.*, 1998). One sequence from Indonesia (Ind37) and two sequences from Germany (AF060550) and from Albania (Alb16) formed two additional branches. The separation of these sequences into the four main branches was not supported, however, by bootstrap resampling. Thus, the virus isolates in this second major group most probably should be considered as subtypes of a single type (TTV type 2) rather than different TTV types, as has been suggested recently on the basis of analysis of the much smaller number of sequences from a single geographical region (Simmonds *et al.*, 1998).

Predicted amino acid sequences for all TTV isolates were subjected to a phylogenetic analysis using the programs Protdist and Neighbor from the PHYLIP package (Felsenstein, 1993). The resulting phylogenetic tree (not shown) was an almost exact replica of the tree constructed by analysis of the corresponding nucleotide sequences (Fig. 1). All sequences were segregated into two groups corresponding to TTV types 1 and 2. Type 1 sequences segregated into two subgroups (1a and 1b). The type 2 group was represented by a single set of very heterogeneous sequences.

Fig. 2 is an alignment of nucleotide and amino acid sequences of several TTV isolates belonging to different proposed types and subtypes. Analysis of the alignment data of all TTV sequences was used to calculate partial nucleotide identity for different isolates. Sequence identity between isolates of two major TTV types varies within the range 59–71% (59–68% at the amino acid level). The similarity level of TTV type 1 and 2 nucleotide sequences is somewhat closer to the range noted for autonomous parvoviruses infecting different species (63–66%) than to the level of nucleotide identity observed for different rodent parvoviruses (72–90%; Ball-Goodrich *et al.*, 1998; L. Ball-Goodrich, personal communication).

Of note is the difference in topology of branches forming type 1 and 2 groups (Fig. 1). The first group (type 1 sequences)

Fig. 1. Phylogenetic tree drawn from a phylogenetic analysis [by unweighted pair group method with arithmetic mean (UPGMA)] of 76 sequences of the TTV ORF1 (positions 1939–2160, according to the sequence of the TTV prototype isolate, GenBank accession no. AB008394). Sequences of the N22, C5, TS003 and NA004 prototype isolates of TTV were taken from Nishizawa *et al.* (1997) and Okamoto *et al.* (1998). Sequences taken from GenBank are indicated by their accession number.

(*a*)

```
N22     (1a)   CTAAGCAAAAAAAACATGAACTATGACAAACTACAAAGTAAATGCTTAATATCAGACCTA
C5      (1b)   ....CT.............A.........G.G........G...C..G..G........
Zan5    (1a)   ..............T..............G.............T................
Ger926  (1b)   ..T..T............G.A.........G.G.....C..G...C..G..G........
Ita32   (2)    ....CT...G.TG..TCAGC....TCA...AC.AGC.........TC.C...GAGA....G
Ger118  (2)    ....CT...G.T.C.TCAGTA.........GC...G.........TC.T...CA....A.G
Alb16   (2)    ..T..T.........T..G.A...TC....GCG..G..C..G..TC.....G..A....G
Bhu96   (2)    G..TCT...T.TG..TCTGTG..CTCA...AC......C............GA.A.....
Ind37   (2)    ...TTA....C.G..TCT.GA........GAC..GC..C......C.T...GA.A.A..T

N22     (1a)   CCTCTATGGGCAGCAGCATATGGATATGTAGAATTTTGTGCAAAAAGTACAGGAGACCAA
C5      (1b)   ..A..G..................T...T.......C..CT.T.....C.........AC.
Zan5    (1a)   ....................C.....CT..............................
Ger926  (1b)   ..A..G.................G...T.......C..CT.T.....C.........AC.
Ita32   (2)    ..C..G.....CT...T...C.....CAA...G..C...T.C...GTA.........AC.
Ger118  (2)    ..CT.G.....CT.T.T...C....TCTCC..G.AC..CAGT...GTA.........AC.
Alb16   (2)    .....G.....C....T......G...T.......C..CTGT...GCA.........AC.
Bhu96   (2)    .....G.....C.....T.T........C.....AC..CAGC...GCC.........AC.
Ind37   (2)    ...T.G.....TT...T...C..G..C.C...G.AC..C..C..GGCC.........TCT

N22     (1a)   AACATACACATGAATGCCAGGCTACTAATAAGAAGTCCCTTTACAGACCCACAACTACTA
C5      (1b)   ....................A.................T...........C..G...A..
Zan5    (1a)   ....................A.............................C.........
Ger926  (1b)   ....................A..GT..........................C..G...A..
Ita32   (2)    ......G.ACAC..CTG...ATGTG.T........C....AC....TA........GT..
Ger118  (2)    ......G.ACAC..CTG...ATGTG.T..T.....C....AC....TG........G...
Alb16   (2)    ........TC....ATG...ATGTG.T........C.....C........C..G..GT..
Bhu96   (2)    ......G.ACAA..CTG...AG..G.T..T.....C.....C..TA....T..G......
Ind37   (2)    ......G.......C.....AG..G.T..C..G.......A...TACA..C...A.GA..

N22     (1a)   GTACACACAGACCCCACAAAAGGCTTTGTTCCTTACTCTGTA
C5      (1b)   .........A.......T...........A..C..T...T..
Zan5    (1a)   ................................C......T..
Ger926  (1b)   .........A......AT...........A.....T...T..
Ita32   (2)    .AC....ACA.T...CTC.G.....AC..G.....TAGCC..
Ger118  (2)    .AT....ACA.....CTC.GG..A.AC..G..C...AG.T.T
Alb16   (2)    .AC.....CA.....CTT...........A.....T...T..
Bhu96   (2)    .AC..T.ACA....TCT..G...G.AC..C..C.....AA..
Ind37   (2)    .ACAC..AT......CTC.GG.....CA.AGTA...AGCT..
```

(*b*)

```
N22     (1a)   LSKKNMNYDKLQSKCLISDLPLWAAAYGYVEFCAKSTGDQNIHMNARLLIRSPFTDPQLL
C5      (1b)   .T....K...V.....VA...........L...S.....T...................I
Zan5    (1a)   ..........V..................L.............................
Ger926  (1B)   ......E...V.....VA...........L...S.....T...................I
Ita32   (2)    .T.DDSA.S.TS.....EN.....SV...K...S.V...T..EH.C.CV....Y.V....
Ger118  (2)    .T.DTSV...A......Q.M....SV..FS.Y.S.V...T..EH.C.CV....Y.V....
Alb16   (2)    .....LE.S.A......AN......V...L...C.A...T...LKC.CV...........
Bhu86   (2)    V..YDSV.S.T......EN.......F..A.Y.S.A...T..EQ.C.VV......N....
Ind37   (2)    .L.TDSR...TR.....EK.....SV...A.Y...A...S..D....VV....Y.T..MI

N22     (1a)   VHTDPTKGFVPYSV
C5      (1b)   ...N.........L
Zan5    (1a)   ............L
Ger926  (1b)   ...N.N.......L
Ita32   (2)    D.NN.LR.Y....L
Ger118  (2)    D.NN.LR.Y....F
Alb16   (2)    D..N.L.......L
Bhu86   (2)    D.NN.LR.Y....I
Ind37   (2)    DTN..LR..IV..L
```

Fig. 2. Comparison of (*a*) nucleotide and (*b*) amino acid sequences in the putative ORF1 of nine TTV isolates. Nucleotide positions 1939–2160, according to the sequence of the TTV prototype isolate, GenBank accession no. AB008394. Attribution of TTV type and subtype is given in parentheses.

forms only two clusters, 1a and 1b, consisting of numerous very closely related isolates. The sequence similarity between subtype 1a and 1b isolates (84–92%) is also high. In contrast, the second group (TTV type 2) demonstrates an asymmetric branching pattern. The nucleotide similarity between sequences forming four major branches varied from 65–81%. To some extent the differences in the branching structure of the TTV phylogenetic tree recall those observed in the phylogenetic tree of flaviviruses, in which evolutionary lineages of viruses transmitted by mosquitoes and ticks are characterized by different branching topology (De Zanotto *et al*., 1996). One might suggest that the topological distinction between TTV 1 and 2 variants also reflects the differences in their biological properties or in the epidemiology of their infections.

It should be noted that final conclusions on TTV genome heterogeneity can be drawn only when the full genome sequences of different TTV isolates are available. Comparison of complete sequences should also answer the question whether comparison of the short genome sequences of TTV provide adequate estimates of overall degree of sequence dissimilarity of different TTV types. The rather high level of sequence dissimilarity between TTV isolates recorded in this study, based on analysis of a 222 bp fragment of ORF1, suggests, however, the validity of such an approach for TTV research.

In the current study, TTV DNA was detected in blood samples from 17 countries situated in different parts of the world. This indicates the ubiquitous distribution of TTV. Of interest is the fact that the most frequently detected type 1 isolates were identified in all geographical regions. It should be pointed out that at the moment we cannot exclude the possibility that the PCR primers used in this study are suboptimal and are able to amplify only two TTV types. Our numerous attempts to introduce degeneracy in the existing primers or to use new primers derived from other fragments of the TTV genome have not so far resulted in the identification of any additional TTV types.

The practical implications of the possible existence of TTV types (genotypes) are not yet clear. One cannot exclude the possibility that different TTV types posses different pathological potential or that infections caused by TTV type 1 or 2 have distinct epidemiological characteristics. The degree of amino acid dissimilarity observed between TTV types 1 and 2 would be expected to affect the antigenicity of the putative structural protein encoded by ORF1 with evident consequences for viral diagnostics.

We are grateful to Mrs Tanja Gross for excellent technical assistance, to Mrs Susanne Standar for help in DNA sequencing, and to Dr Lisa Ball-Goodrich and Dr Linda Prescott for providing unpublished data.

References

Ball-Goodrich, L. J., Leland, S. E., Johnson, E. A., Paturzo, F. X. & Jacoby, R. O. (1998). Rat parvovirus type 1: the prototype for a new rodent parvovirus serogroup. *Journal of Virology* **72**, 3289–3299.

Cossart, Y. (1998). TTV a common virus, but pathogenic? *Lancet* **352**, 164.

Felsenstein, J. (1993). PHYLIP (Phylogeny Inference Package) version 3.5c. Department of Genetics, University of Washington, Seattle, USA.

Miyakawa, Y., Okamoto, H. & Mayumi, M. (1998). TT virus (TTV): an unenveloped DNA virus associated with acute and chronic hepatitis of non-A to G etiology. Abstract L1, *5th International Meeting on Hepatitis C Virus and Related Viruses. Molecular Virology and Pathogenesis*. Venezia, Italy.

Naoumov, N. V., Petrova, E. P., Thomas, M. G. & Williams, R. (1998). Presence of a newly described human DNA virus (TTV) in patients with liver disease. *Lancet* **352**, 195–197.

Niel, C., De Oliveira, J. M., Ross, R. S., Gomes, S. A., Roggendorf, M. & Viazov, S. (1998). High prevalence of TT virus infection in Brazilian blood donors. *Journal of Medical Virology* (in press).

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.

Okamoto, H., Nishizawa, T., Kato, N., Ukita, M., Ikeda, H., Iizuki, H., Miyakawa, Y. & Mayumi, M. (1998). Molecular cloning and characterisation of a novel DNA virus (TTV) associated with posttransfusion hepatitis of unknown etiology. *Hepatology Research* **10**, 1–16.

Simmonds, P., Davidson, F., Lycett, C., Prescott, L., MacDonald, D. M., Ellender, J., Yap, P. L., Ludlam, C. A., Haydon, G. H., Gillon, J. & Jarvis, L. M. (1998). Detection of a novel DNA virus (TTV) in blood donors and blood products. *Lancet* **352**, 191–195.

De Zanotto, P. M., Gould, E. A., Gao, G. F., Harvey, P. H. & Holmes, E. C. (1996). Population dynamics of flaviviruses revealed by molecular phylogenies. *Proceedings of the National Academy of Sciences, USA* **93**, 548–553.

Received 10 July 1998; Accepted 27 August 1998