Sequence Heterogeneity of TTV Virus

1Nagwa A. El-Esnawy and 2Wahiba A. Zarouk

The aim of this study was to detect blood Transmitted Transfusion Virus (TTV) among Hepatocellular Carcinoma (HCC) patients and investigate the genetic diversity of an isolated TTV strain. Fifty blood samples of patients with HCC non-A through non-G hepatitis were collected and 25 blood donors as controls. TTV DNA was detected by the second generation of TTV PCR primers (NG059, NG061, NG062 and NG063) after extraction of DNA from 50 µL serum. TTV DNA was detected in 90% (45/50) of HCC patients and the prevalence in control group was 40% (10/25). To assess genetic heterogeneity of the TTV genome in more detail, a sequence analysis of PCR fragments, amplified from open reading frame 1 (ORF1) was performed. A total of 137 bp PCR fragments were sequenced and compared to sequences derived from the corresponding TTV genome region deposited in Gen Bank. Present sequencing was closely related to TTV-like mini virus, complete genome (Accession NC_002195), at nucleotide number 855 to 1019. The phylogenetic analysis suggested that present strain might be sub-strain or mutation from the parent gene 2195. These observations suggest a high degree of genetic complexity within the TTV population.

Key words: TTV, sequence, PCR, hepatocellular carcinoma, phylogenetic analysis, heterogeneity

1Virology Laboratory, Department of Water Pollution Research, Environmental Research Division, 2Department of Immunogenetics, Human Genetics and Genome Research Division, National Research Centre, Cairo-Dokki, Egypt
INTRODUCTION

In 1997, a group of Japanese scientists discovered a novel virus in a patient hospitalized for an acute illness that was treated as hepatitis. But because the patient’s disease was not caused by any of the known hepatitis viruses, the researchers speculated that they had discovered a new kind of hepatitis virus, which they named Transmitted Transfusion virus (TTV)\(^1\). TTV-DNA infection more frequently occurs in patients with liver cirrhosis and Hepatocellular Carcinoma (HCC) than in those with chronic hepatitis\(^2\) and their was a link between TTV and hepatocellular carcinoma in patients positive or negative for HBV and HCV\(^3,4\). By cloning and sequencing over 90% of the prototype Japanese isolate TA278; the TTV genome was proposed to be linear single-stranded DNA, similar to that of paroviruses\(^5,6\). Subsequently, however, when sequencing of the viral genome was further extended toward its 5' and 3' ends, it was recognized that the two extremities were connected by a GC-rich stretch of about 100 nucleotides (nt) to form a closed circular molecule, resembling that of circoviruses\(^7\).

At the end of 1999, gene banks contained the entire nucleotide sequences of 10 TTV isolates, which varied in length between 3,808 nt (SANBAN isolate) and 3,833 nt (isolates TA278 and JA20). Judging from the presence of well-conserved ORFs among the sequenced isolates, the genome of TTV is divided into a potentially coding region of ~ 2.6 kb and an Untranslated Region (UTR) of ~ 1.2 kb. The former consist of two major potential protein-coding genes (ORF1 and ORF2), which are present in the plus strand complementary to the genomic DNA. In the prototype TA278 sequence, ORF1 spans nt 589 to 2898 and ORF2 spans nt 107 to 712, which correspond to a coding capacity of 770 and 150 amino acids (aa), respectively. However, the length of the two ORFs may vary somewhat in individual isolates. Analysis of the sequence of 11 isolates, revealed the presence of an additional small ORF (designated ORF3) with coding capacity for 57 aa, located immediately downstream of ORF1 (nt 2904 to 3074 in the TA278 isolate) and well conserved in most isolates (70 to 100% identity). Other ORFs have been recognized in some isolates, but they might be nonfunctional, as suggested by their absence in most other isolates\(^8\). More extensive investigation of genetic diversity and phylogenetic relationships of TTV isolates have been carried out by sequencing amplification obtained from the N22 segment\(^9\), which has been extensively used for molecular epidemiology surveys and genotyping\(^9\). With the growing of the characterized isolates, appreciation of the genetic diversity of TTV has rapidly increased\(^10\). So far, phylogenetic trees constructed on the N22 region have led investigators to subdivide TTV into up to 16 genotypes and several subtypes or to refrain from designating new types\(^11\).

Hepatocellular carcinoma is developing world, especially in the Africa, Hepatocellular Carcinoma (HCC) is one of ten most types of cancer, occurring predominantly in 40-60 years olds. In epidemic areas of the world, occurs mainly in 20-40 age group. Worldwide, the male to female patients is 3:2, but this male predominance is more marked (11:1) in cases where liver cirrhosis is the pathogenic mechanism. There are two major etiological agents are microorganisms. HCC is now a common malignancy in Egypt, which usually develops on top of cirrhosis of viral origin in 82%. Hepatic resection is the only method of treatment with a low respectability rate. Using serum AFP (alphafetoproteins) levels may allow early detection of tumor development via rising AFP levels, permitting early intervention\(^12\).

The aim of the present study was to screening the presence of TTV between HCC patients and to investigate the genetic diversity of the isolated TTV strain.

MATERIALS AND METHODS

Samples: During 2004, sera were obtained from 50 patients with hepatocellular carcinoma non-A through non-0 hepatitis, from Oncology Diagnostic Unit, Biochemistry Department, Faculty of Medicine, Ain Shams University. Also, 25 samples from healthy blood donors were collected as control.

Oligonucleotide primers: Two pairs of primers were used for first and second nested PCR. The first pair NGO59 (antisense) (5' -ACA GAC AGA GGA GAC ATG 3') and (sense) NGO63 (5' -CTG GCA TTT TAC CAT TTC CAA AGT 3'). The second pair (antisense) NGO61 (5' -GGC AAC ATG TTA TGG ATA GAC TGG 3') and (sense) NGO62 (5' -TAG TAG CCT GGC ATT CAT 3').

PCR amplification was performed in the ORF1 region of the TTV genome using degenerate oligonucleotide primers NGO59 and NGO63\(^k\). These primers are located in a quite variable region and permit the amplification of type 1, 2 and 3 isolates.

Extraction of TTV DNA and PCR: DNA was extracted from 50 µL of serum by the standard method, which included pretreatment of samples with proteinase K-sodium dodecyl sulfate and extraction with phenol and chloroform. DNA samples were dissolved in 20 µL of Tris-HCl buffer (pH 8.0), denatured at 95°C for 5 min and then chilled on ice. Two microliter of the denatured DNA was used for 30 cycles of the first-round PCR amplification with external primers at 94°C for 1 min, 50°C for 40 sec and 72°C for 1 min. Two microliter of the first-round PCR product was used to initiate the second-round PCR with internal primers at 94°C for 40 sec, 55°C for 40 sec and 72°C for 45 sec for 25 cycles\(^f\).
First-round PCR primers were used in a final volume of 25 µL at a concentration of 0.1 mM in a reaction mixture consisting of 10 mM Tris-HCl (pH 8.3), 50 mM KCl, 1.8 mM MgCl₂, 0.2 mM of each deoxynucleoside triphosphatase (dNTP), 5 U of Taq DNA polymerase (Promega, Madison, Wis.) and 5 µL of DNA. Using Perkin-Elmer, Gene Amp., PCR system, 9600 DNA thermal cycler. PCR products were analyzed in a 2% low melting point agarose gel. Two microliter of the first round PCR product was submitted to a second round PCR using 1 µL of each nested primer (NG061 and NG062). A final volume of 25 µL was prepared by add 21 µL DEPC-H₂O. PCR products were analyzed in 2% agarose gel. Electrophoresis was performed in TAE (50 mM Tris Base, 0.08 M Glacial acetic acid and 0.5M EDTA). Eight samples TTV-DNA positive by PCR from studied groups were selected for sequencing.

**Sequencing:** PCR products of positive TTV DNA signal samples were purified for sequencing using Wizard PCR Prep DNA purification system (Promega, Madison, Wis.). The primary structure of purified PCR fragments was determined by using an automated DNA sequencing (ABI PRISM™ 310 Genetic Analyzer, Applied Biosystems Division, Foster City, Calif. and USA). The amplified DNA fragment was sequenced directly in both directions with each second-round PCR primer⁴. Gene Sequence fragment was (annttgccc g c tt tga at caa arca tatactgctca ccatagaa g tagc tga ttagaaca tt tac ttc tac ttggc atag ttc atat tttttg tta gca ta acaat g g tggc a).

**Computer-assisted sequence analysis:** The TTV nucleotide sequences were aligned with the Clustal X program¹⁰ and edited with the GeneDoc program version 2.5.000 (Nicholas KB, Nicholas HB Jr., GeneDoc: a tool for editing and annotating multiple sequence alignments, 1997, distributed by the authors). The nucleotide sequences of the TTV-DNA amplified from our patients were compared with the TTV-DNA sequences retrieved from Gen Bank, corresponding to all of the published TTV genotypes. Phylogenetic analysis was performed using the DNA STAR¹⁰. The genetic distances were estimated by the Kimura two-parameters method and the unrooted phylogenetic tree was constructed by the neighbor-joining method. The final output of the tree was obtained using the tree view program version 1.5.2²⁶.

### RESULTS AND DISCUSSION

The Circoviridae family consists of nonenveloped, single-stranded circular DNA viruses¹⁰; there have been only three animal circoviruses known, i.e. chicken anemia virus, beak and feather disease virus of parrots and porcine circovirus¹⁷. Okamoto et al.¹⁸ and Mushahwar et al.¹⁹ discussed that TTV due to its circular genomic structure, may qualify as a fourth animal circovirus, although its genomic size (3,818 to 3,853 nt, depending on genotype) is much larger than those of the other three (1,758 to 2,319 nt).

The prevalence rate of TTV-DNA infection among hepatocellular carcinoma patients was 50% (45/90) with absence of other hepatitis, while in control group was 40% (10/25) (Table 1).

<table>
<thead>
<tr>
<th></th>
<th>HCC</th>
<th>Control</th>
</tr>
</thead>
<tbody>
<tr>
<td>Positive</td>
<td>95% (45/50)</td>
<td>50% (10/25)</td>
</tr>
<tr>
<td>Negative</td>
<td>10% (5/50)</td>
<td>60% (15/25)</td>
</tr>
</tbody>
</table>

**Fig. 1:** Neighbor-joining tree constructed with the nucleotide sequences of the TTV isolates from serum hepatocellular carcinoma. The Gen Bank TTV sequences included in the tree are: strain egt205 ORF1 gene; strain egt238 ORF1 gene; isolate patient 30 variant 5; TTV-like mini virus (NC-002195); isolate 1L2 ORF1 gene; clone 12.S1 ORF1 gene; clone 12.B1 ORF1 gene; isolate patient 30 variant 3; isolate BSH1109; isolate 8195 ORF1 gene; TTV virus complete genome (NC-002076); isolate p30e23; isolate p30e65; isolate p30e60; strain egt216 ORF1 gene; isolate KSA20-Cr ORF1 gene; strain egt777 ORF1 gene; strain egt236 ORF1 gene; isolate ist2 unknown gene.
Fig. 2: Average homology between our sequence and TTV-like mini virus

Fig. 3: Amino acid sequences of our sequence and TTV-like mini virus

Similar observations was found in a series of 26 patients with hepatocellular carcinoma TTV-DNA was 100%.

The nucleotide sequence of the TTV ORF-1 isolated from serum (PCR fragments), obtained in this study were directly sequenced. These sequences were compared to the corresponding sequences deposited in Gen Bank (Fig. 1). A significant degree of heterogeneity was observed between the compared sequences. The average homology between all TTV nucleotide sequences was 90% with TT virus strain egt203 ORF1 gene (Accession AF173107) partial cDNA, 217 bp DNA, Version AF173107.1 GI: 65959347; TT virus strain egt238 ORF1 gene 217 bp DNA (Accession AF173111), Version AF173111.1 GI: 65959354; TT virus isolates patient 30 variant 4 putative core/capsid protein gene, partial cDNA. 169 bp DNA (Accession AF126512) Version AF126512.1 GI: 4927780; TT virus isolates patient 30 variant 5 putative core/capsid protein gene, partial cDNA. 169 bp DNA (Accession AF126513) Version AF126513.1 GI: 4927781; TTV-like mini virus, complete genome. Accession (NC_002195) Version NC_002195.1 GI: 963495; TT virus isolate 1L2 ORF1 gene, partial cDNA. 222 bp DNA (Accession AF229328) Version AF229328.1 GI: 9082597; TT virus clone 12.S1 ORF1 gene, partial cDNA. 222 bp DNA (Accession AF216418) Version AF216418.1 GI: 8777804; TT virus clone 12.B1 ORF1 gene, partial cDNA. 222 bp DNA (Accession AF216419) Version AF216419.1 GI: 8777806; TT virus isolates patient 30 variant 3 putative core/capsid proteins Gene, partial cDNA (Accession AF126511) Version AF126511.1 GI: 4027779; TT virus isolate BeHCC109 unknown gene. 222 bp DNA (Accession AF136909) Version AF136909.1 GI: 4877854; TT virus isolate S7195 ORF1 gene, partial cDNA. 247 bp DNA (Accession AF289360) Version AF289360.1 GI: 10717119; TT virus, complete genome. 3852 bp ss-DNA (Accession NC_002076) Version NC_002076.2 GI: 29502191; TT virus isolate p30e23 putative core/capsid protein gene, partial cDNA. 199 bp DNA (Accession AF144998) Version AF144998.1 GI: 5354087; TT virus isolate p30e65 putative core/capsid protein gene, partial cDNA. (Accession AF144994) Version AF144994.1 GI: 5354079; TT virus isolate p30a30 putative core/capsid protein gene, partial cDNA. 199 bp DNA (ACCESSION AF144992) Version AF144992.1 GI: 5354075; TT virus strain egt216 ORF1 gene, partial cDNA. 217 bp DNA (Accession AF173109) Version AF173109.1 GI: 6595930; TT virus isolate KSA20-Cr ORF1 gene, partial cDNA. 222 bp DNA (Accession AF256665) Version AF256665.1 GI: 30027178.

The phylogenetic analysis performed with the nucleotide sequences of the isolates from our patients and all of the TTV genotypes deposited in Gen Bank (Fig. 1 and 2) showed that our sequence can be considered a genetic link, or closely related to TTV-like mini virus, complete genome (Accession NC_002195), at nucleotide number 855 to 1019. So the phylogenetic analysis suggested that our strain might be sub-strain from the parent gene 2195 or mutation. These observations suggest a high degree of genetic complexity within the TTV population.

Nucleotide sequences were translated into amino acid sequences (Fig. 3). A comparison of the derived protein sequences showed that the average homology was 71.8%. The minimal homology between protein sequences was found to be 44.1%.
TTV in spite of being a DNA virus, also has an extremely high level of sequence divergence, ranging to 60.5% for the entire genome[14] and at least 16 genotypes that differ by >30% from one another have been distinguished[15]. The reason for the outstanding genetic heterogeneity of TTV is not clear. The replication of TTV might involve reverse transcription, making for an accelerated mutation rate as is the case for hepatitis B virus[23].

A double-stranded DNA virus encoding a reverse transcriptase, a sequence motif for reverse transcriptase has not been identified in the TTV genome, however nucleotides sequences vary considerably even among TTV isolates of the same genotype. Sequence divergence resulting in amino acid conversions is the highest in a central portion of ORF1, the translation product of which has an arginine-rich sequence in the N terminus[26]. By analogy with VP1 of chicken anemia virus, this also has this sequence[27]. The ORF1 in TTV may encode a capsid protein. The nine TTV isolates of genotypes 1a were 92 to 96% similar in the amino acid sequence of the ORF1 product. Most amino acid substitutions in ORF1 (47 to 86%) clustered in HVRs that spanned 100 aa altogether; regions outside HVRs in ORF1 were well conserved, with a similarity of 98%. HVRs in TTV were comparable to the HVR in chicken anemia virus, which stretches for 13 aa (position, 139 to 151) in the 450 aa VP1[28].

ACKNOWLEDGMENTS

The authors are deeply indebted to Prof. Dr. Ali Khalifa soul's, Oncology Diagnostic Unit, Biochemistry Department, Faculty of Medicine, Ain Shams University for providing the serum samples.

REFERENCES


