A Canadian isolate of hepatitis D (delta) virus

Véronique Poisson BSc[1], Daniel B Ménard MD[2], Éric Frost PhD[3], Jean-Pierre Perreault PhD[1]

V Poisson, DB Ménard, É Frost, J-P Perreault. A Canadian isolate of hepatitis D (delta) virus. Can J Gastroenterol 2000; 14(Suppl A):36B-40B. Hepatitis D (delta) virus (HDV) is an infectious agent that propagates in hepatocytes only in the presence of hepatitis B virus, causing fulminant or chronic hepatitis with liver cirrhosis. HDV is a 36 nm particle that includes a circular RNA genome of 1.7 kilobases with an extensive internal complementary that allows it to fold into a rod-like structure. The relationships among genotypes, sequence variability, geographical distribution and disease severity of HDV remain unknown. Consequently, in the present study, the complete nucleotide sequence of an HDV isolated from a Canadian patient was determined. The viral RNA from serum was amplified using reverse transcription coupled to polymerase chain reaction amplification. The resulting complementary DNA was cloned and sequenced. Sequence analysis revealed that this new isolate contained 1672 nucleotides corresponding to genotype 1, which has a worldwide distribution. Sequencing of four independent clones revealed 17 substitutions, corresponding to an overall sequence variability of 1%. Surprisingly, seven mutations were found in the 48-nucleotide region located between the two highly conserved self-catalytic motifs. This is the first demonstration that many substitutions are identified in this region of HDV, and prompts the present authors to define it as a hypervariable region.

Key Words: *Genotype; Hepatitis D (delta) virus; Reverse transcription-polymerase chain reaction; Sequence analysis; Sequence variability; Viral hepatitis*

Un isolat Canadien du virus de l'hépatite D (delta)

RÉSUMÉ : Le virus de l'hépatite delta (VHD) est un agent infectieux qui se propage dans les hépatocytes seulement en présence du virus de l'hépatite B, causant ainsi des hépatites fulminantes et chroniques avec des cirrhoses du foie. VHD est une particule de 36 nm contenant un génome d'ARN circulaire de 1,7 kilobases qui présente beaucoup de complémentarité interne le conduisant à se replier en une structure semblable à une tige. Les relations entre génotypes, degré de variabilité des séquences, distribution géographique, et sévérité de la maladie sont des aspects encore incompris. Par conséquent, la séquence nucléotidique complète d'un VHD isolé d'un individu canadien a éte déterminée. L'ARN viral d'un sérum fut amplifié en utilisant une transcription inverse couplée à une réaction de polymérisation en chaîne. L'ADNc résultant fut ensuite cloné et séquencé. L'analyse de la séquence a révélé que ce nouvel isolat contenant 1672 nucléotides correspond au génotype 1, lequel a une distribution mondiale. Le séquençage de quatre clones indépendants a révélé 17 substitutions, correspondant à une variabilité de séquence moyenne de 1%. Il est intéressant de noter que sept de ces mutations ont été retrouvées dans la petite région constituée de 48 nucléotides localisée entre les deux motifs autocatalytiques hautement conservés. C'est la première fois qu'autant de substitutions sont identifiées dans cette région de VHD, ce qui a conduit les auteurs a suggérer de la définir comme une région hypervariable.

Hepatitis D (delta) virus (HDV) is an infectious agent that infects and propagates in hepatocytes exclusively in the presence of hepatitis B virus (HBV), causing fulminant or chronic hepatitis with liver cirrhosis (1,2). The symptoms associated with both acute and chronic HDV infection are frequently more severe than those seen with HBV alone, and appear to differ in severity depending on the geographic location. HDV is a 36 nm particle and is unique among animal viruses in that it contains a single-stranded, circular RNA genome of 1.7 kilobases with an extensive internal complementary that allows it to assume a rod-like structure (1,2). These features are reminiscent of those seen

[1]*Département de Biochimie,* [2]*Département de Médecine,* [3]*Département de Microbiologie et Infectiologie, Faculté de médecine, Université de Sherbrooke, Sherbrooke, Québec*

Correspondence and reprints: Dr Jean-Pierre Perreault, Département de biochimie, Université de Sherbrooke, Sherbrooke, Québec J1H 5N4. Telephone 819-564-5310, fax 819-564-5340, email jperre01@courrier.usherb.ca

Received for publication May 18, 1999. Accepted June 4, 1999

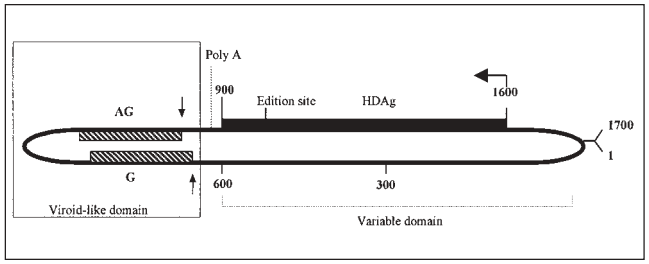

Figure 1) *Schematic diagram of the hepatitis D (delta) virus (HDV) genome. The RNA genome of HDV folds into a rod-like structure. The left-hand domain (boxed) is reminiscent of plant viroids and includes self-cleaving sequences of both genomic and antigenomic polarities (striped rectangles). The self-cleavage sites are indicated by arrows. The open reading frame for both forms of delta antigen (HDAg) (black box) is on the upper strand of the right-hand domain, while the lower strand is the variable domain. AG Antigenomic ribozymes; G Genomic ribozymes; Poly A Polyadenylation site*

TABLE 1
Synthetic oligonucleotide primers*

Name	Sequence	Position†
	Sense primers	
S200	5 CGATCCGAGGGGCCCAAC3	202-219
S275	5 CACTCCGGCCCGAAGGGTTG3	236-255
S670	5 TGGCCGGCATGGTCCCAGCC3	685-704
S1100	5 GCCCGCCCCGGGCGCTCC3	1101-1118
	Antisense primers	
AS275	5 CAACCCTTCGGGCCGGAGTG3	255-236
AS300	5 GCTGAAGGGGTCCTCTGG3	324-304
AS800	5 CGCACGTCCACTCGGATG3	844-827
AS1100	5 GGAAAAGAAGAGTAGCCG3	1180-1163
AS1200	5 GAAGGAAGGCCCTCGGAACAAG3	1288-1266
	Sequencing primers	
F606	5 GAAGGCGGATCGAGGGGAGC3	1133-1114
F606N	5 CTTTCCTCTTCGGGTCGG3	911-894
F836N	5 GCTTCTTGGGAGTAGTTTC3	106-88
F836Z	5 GAGGACGAAAATCCCTGG3	1466-1449

**Primers were designed using a hepatitis D (delta) virus sequence alignment described in the 'Results and Discussion' section; †The position of the primers is as outlined by Makino et al (13)*

in the plant RNA pathogens known as viroids. Moreover, both HDV and viroids have RNA genomes that undergo RNA-dependent RNA replication via a rolling circle mechanism that is DNA-independent (ie, involving only RNA intermediates) (2).

The HDV genome is composed of two domains known as the left- and right-hand domains (Figure 1). The left-hand domain comprises the viroid-like region and includes self-cleaving sequences for RNAs of both genomic and antigenomic polarities. These two self-cleaving RNA motifs are crucial for HDV replication. The right-hand domain includes two regions: the upper strand encodes the messenger RNA (mRNA) for a single protein, the delta antigen (HDAg); and the lower strand is a variable domain to which no function has been attributed. HDAg, which is encoded by a mRNA of sequence corresponding to the antigenomic polarity, exists as two distinct species: a large HDAg (L-HDAg, 27 kDa, 214 amino acids) and a small HDAg (S-HDAg, 24 kDa, 195 amino acids). The L-HDAg is produced following an editing event that alters the first stop codon to a tryptophan codon, thereby extending the translation to the next stop codon. S-HDAg acts as a transcriptional factor during HDV RNA replication, whereas L-HDAg inhibits replication and is required for virus assembly.

Comparison of HDV sequences indicated the existence of three distinct genotypes (3,4). It has been observed that genotypes 2 and 3 are 80% and 65% homologous, respectively, with genotype 1. Genotype 1 is widespread in Italy, the United States, Taiwan, Nauru, France, Lebanon and China. The disease pattern associated with infection by HDV genotype 1 appears to be highly variable, ranging from severe to mild. Genotype 2 has been isolated only in Japan and Taiwan, and is associated with a milder disease than that caused by genotype 1 infection. Genotype 3 has been isolated only in northern parts of South America where the disease is the most severe. Moreover, a sequence microheterogeneity was observed in patients. The role of HDV genetic variations in the severity of HDV is not yet clear, but some studies suggest that these factors are linked (3,5). Conversely, the relationship between HDV genotypes and the degree of HDV sequence variability in relation to geographical distribution and disease severity remains unclear. To address this question we determined the nucleotide sequence of a Canadian HDV isolate and investigated the sequence variability throughout the whole genome (compared with most studies that are limited to the protein-coding region).

PATIENTS AND METHODS

This study was approved by the ethical review board of the Université de Sherbrooke, Centre de Recherche Clinique. The serum sample from which the HDV RNA was isolated was collected from a 50-year-old Canadian male resident of the province of Quebec. This patient suffered from severe chronic hepatitis that was rapidly evolving into cirrhosis and liver failure. The serum was previously analyzed for hepatitis B surface antigen, immunoglobulin M antibody to hepatitis B core antigen, total antigen to hepatitis delta, immunoglobulin M antibody to hepatitis delta and antibody to hepatitis C virus (as determined by enzyme-linked immunosorbent assay [ELISA] II). The diagnosis of hepatitis C was also confirmed by polymerase chain reaction (PCR).

The serum (50 µmL) was subjected to proteinase K digestion and then extracted using phenol/chloroform (one to one volume) extracted. The mixture of cellular and viral RNA was precipitated with isopropanol, washed with 70% ethanol and dissolved in 50 µmL diethylpyrocarbonate-treated water. A strategy of reverse transcription (RT) coupled to PCR was used to amplify four fragments of the

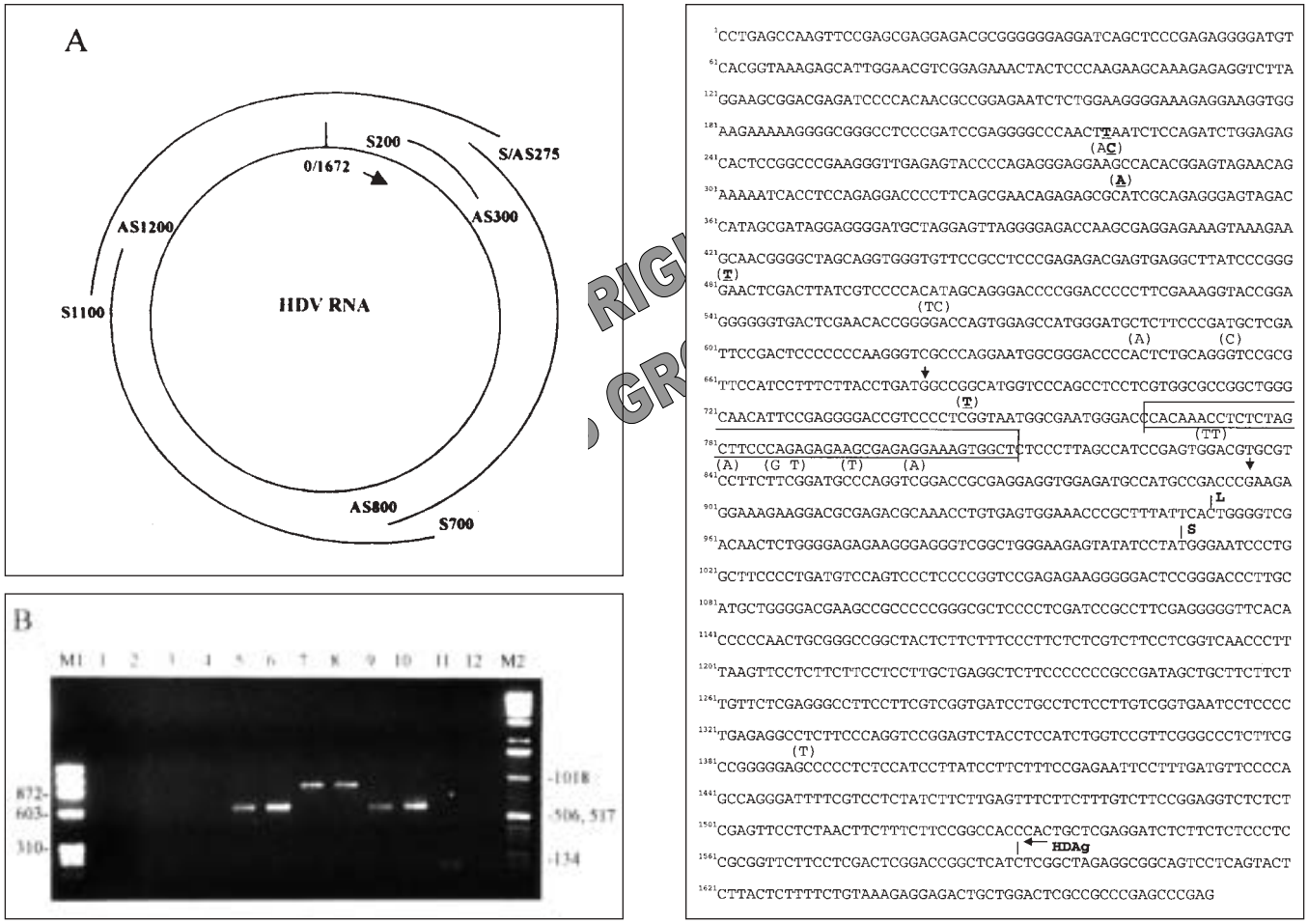

Figure 2) *Molecular cloning and sequencing of a hepatitis D (delta) virus (HDV) isolate.* **A** *Schematic representation of the fragments amplified by polymerase chain reaction (PCR). The arrow indicates the orientation of the genomic RNA.* **B** *PCR products were purified on 1.5% agarose gels. Lanes 1 to 4 are negative controls performed using the set of primers S700-AS1200: lane 1, reverse transcription (RT)-PCR performed in the absence of RNA as template; lane 2, RT reaction performed in the absence of the antisense primer, which was only added for the subsequent PCR reaction; lane 3, RT-PCR using the sense primer to perform the RT reaction; and, lane 4, omission of the RT reaction. Similar controls were performed with all four sets of primers (data not shown). Lanes 5, 7, 9, and 11 are positive controls performed using a cloned HDV dimer sequence (pSVL [D3], generously provided by Dr John Taylor, as per reference 14), while lanes 6, 8, 10 and 12 are the RT-PCR amplifications performed using the RNA extracted from the Canadian patient as template. The primer sets used for the PCR amplifications are lanes 5 and 6, S700/AS1200; lanes 7 and 8, S1100/AS275; lanes 9 and 10, S275/AS800; and lanes 11 and 12, S200/AS300. Lanes M1 and M2 are the DNA size markers, HaeIII digested ΦX174 (Promega, Madison, Wisconsin) and 1 kilobase DNA ladder (Gibco BRL, Grand Island, New York), respectively.* **C** *RNA sequence of the HDV isolate. The genomic site corresponding to the start codon and the orientation of the delta antigen (HDAg) are indicated. The vertical lines indicate the termination sites for both the small HDAg and the large HDAg. The arrows at positions 681 and 895 indicate the self-cleaving sites of the genomic and antigenomic strands, respectively. The nucleotide variations observed are indicated in parentheses. The underlined nucleotides correspond to positions that are shown to vary for the first time. The sequence in the box corresponds to the small hypervariable region, located between the self-cleaving sequences, that the authors identified*

HDV genome. The sequences of the primers used for RT-PCR are shown in Table 1. Only one oligonucleotide (AS1100) was required for the RT reaction. Briefly, one-tenth of the extracted RNA and 100 pmol of primer AS1100 were annealed in a final volume of 18 L by incubating at 70 C for 5 mins. The reaction mixture was then adjusted to 30 L such that it contained 50 mM Tris-hydrogen chloride, pH 8.3, 75 mM potassium chloride, 3 mM magnesium chloride, 10 mM dithiotreitol, 0.67 mM of each dNTP, 30 units of the ribonuclease inhibitor RNA Guard (Pharmacia Biotech, Piscataway, New Jersey) and 200 units of SuperScript TM II reverse transcriptase (Gibco BRL, Grand Island, New York), and was incubated at 42 C for 50 mins. The pairs of oligonucleotides shown in Figure 2A were used in the subsequent PCR steps. One-fifth of the complementary DNA (cDNA) pool (6 L) was added to a PCR mixture containing 10 mM Tris-hydrogen chloride, pH 8.3, 50 mM potassium chloride, 2.5 mM magnesium chloride, 0.2 mM of each dNTP, 100 pmol of both sense and antisense primers and 1 unit Taq DNA polymerase in a final volume of 100 L. Thirty amplification cycles (1 min at 94 C, 1.5 mins at a temperature between 52 C and 65 C depending of the primer set used, and 1 min at 72 C) were performed.

After amplification, the mixtures were extracted with phenol/chloroform and 20 L aliquots were analyzed by 1.5% agarose gel electrophoresis (Figure 2B). Gel slices con-

taining the four PCR products (838 base pairs [bp], 610 bp, 606 bp and 125 bp in length) were isolated and the DNA extracted, precipitated and washed with ethanol. The resulting DNA fragments were cloned by taking advantage of the property of Taq DNA polymerase of adding an adenosine at the 3 ends of the PCR products. The PCR amplified fragments were ligated in a 'sticky end' fashion to linearized PCR^R 2.1 vector, which possesses an extra thymidine residue at each 5 end, as recommended by the manufacturer (TA cloning kit; Invitrogen, San Diego, California) (6).

Four white colonies for each PCR product were chosen after plating on Petri dishes containing ampicillin, isopropylthiogalactoside (IPTG) and X-galactose (X-Gal), and were sequenced by the dideoxyribonucleotide chain termination method using the T7 sequencing kit (Pharmacia Biotech). Sequencing reactions were carried out using the M13 universal primer and custom oligonucleotides derived from the HDV genome as shown in Table 1.

RESULTS AND DISCUSSION

In order to sequence the entire genome of an HDV isolate from a Canadian patient, the authors used a strategy of RT for the production of a cDNA molecule coupled to amplification by PCR. Although HDV is a highly structured RNA molecule, the oligonucleotide AS1100 was sufficient to produce a full length cDNA molecule. Oligonucleotides complementary to various HDV regions were tested, but they failed to produce full length cDNAs. Oligonucleotide AS1100 annealed to a region of the HDV genome that is highly conserved in all genotypes and may, therefore, be considered as an original and universal primer for the production of full length HDV cDNA.

PCR amplification of the cDNA was performed using four distinct pairs of primers that produced overlapping fragments spanning the entire genome (Table 1, Figure 2A). In most of the previous studies, five or more sets of primers were required to amplify the complete genome (7-10). The primers used in this work were selected from highly conserved regions of the HDV genome as determined by a sequence alignment of all known isolates (11) and hence resulted in optimal RT-PCR reactions. Primers located in highly conserved regions should be more reliable and, therefore, more sensitive than others in detecting HDV sequences. Furthermore, the approach of using several regions, as used here, facilitated the subsequent sequencing step, as well as reducing the probability of PCR-derived sequence artefacts. The four primer pairs used here produced the predicted fragments (125 to 838 bp) as revealed by agarose gel electrophoresis (Figure 2B).

Each RT-PCR product was extracted from the agarose gel, cloned and sequenced. The complete sequence of this HDV isolate is shown in Figure 2C (GenBank accession number: AF098261). The sequence data reveal it to be an RNA molecule 1672 bp in length with characteristics similar to other HDV species. For example, one open reading frame encodes a functional protein with all of the features required for its activity and possesses the appropriate S-HDAg stop codon of UAG, which can be edited to produce the L-HDAg form. The alignment of this sequence with the other 22 complete sequences reported to date indicates that this HDV isolate is of genotype 1. The sequences of the four chosen clones showed the existence of some variability in the patient (Figure 2C), ie, for any given fragment, none of the clones were completely identical.

Sequence divergence was detected over the entire genome, reflecting selective pressures. For example, the protein-coding region (HDAg) was highly conserved, while the variable region had several mutations. In total, mutations were detected at 17 positions, corresponding to 1% variation. Among the observed mutations, five appeared for the first time (these are underlined in Figure 2C).

This microheterogeneity is likely real and not the result of PCR artefacts because this divergence rate is much higher than the reported error frequency of Taq polymerase (12). Rather, it likely reflects the actual microheterogeneity of HDV RNA within an individual. This phenomenon may contribute to the diversity seen in the clinical picture of HDV infection. More interestingly, seven of the 17 mutations were detected in the small region between both the genomic and antigenomic self-cleaving domains. Although no function was attributed to this region, this is the first time that numerous mutations have been observed in this small region of 48 bases. This may have significant consequences on the molecular and cellular biology of HDV, although it remains to be addressed. These results suggest that this region may be considered as being hypervariable.

CONCLUSIONS

Sequence analysis of an HDV species isolated from the serum of a Canadian patient indicated that it is genotype 1. The sequence showed a divergence, evenly distributed over the entire genome, from the previously known HDV isolates. These results confirm that HDV species isolated from different geographical areas are heterogeneous. The sequence heterogeneity of HDV RNA may, to some extent, explain the variations observed in the degree of HDV severity and in the virus transmissibility. However, the role of HDV genetic variations in determining its severity remains unclear. Other factors contributing to the clinical manifestations of HDV infections could be the nature and condition of the associated HBV infections. Further analysis of the worldwide distribution of different HDV sequences, and of sequence variations within HDV isolates, will be important for the further evaluation of the pathogenic and epidemiological aspects of this infectious agent.

ACKNOWLEDGEMENTS: We thank Michael Dionne for technical assistance and Dr Sirinart Ananvoranich for critical reading of the manuscript. This work was supported by a grant from the Medical Research Council (MRC) of Canada to Dr Jean-Pierre Perreault. Ms Véronique Poisson is a recipient of a studentship from Fonds

pour la Formation de Chercheurs et l'Aide à la Recherche (FCAR). Dr Jean-Pierre Perreault is an MRC scholar.

REFERENCES

1. Lai MMC. The molecular biology of hepatitis delta virus. Annu Rev Biochem 1995;64:259-86.
2. Poisson F, Roingeard P, Goudeau A. Le virus de l'hépatite delta: un mode de réplication bien singulier. Médecine/Sciences 1995;10:1379-87.
3. Casey JL, Brown TL, Colan EJ, Wignall FS, Gerin JL. A genotype of hepatitis D virus that occurs in Nothern South-America. Proc Natl Acad Sci USA 1993;90:9016-20.
4. Wu JC, Chiang TZ, Sheen IJ. Characterization and phylogenetic analysis of a novel hepatitis D strain discovered by restriction fragment length polymorphism analysis. J Gen Virol 1998;79:1105-13.
5. Wu JC, Choo KB, Chen CM, Chen TZ, Huo TI, Lee SD. Genotyping of hepatitis D virus by restriction-fragment length polymorphism and relation to outcome of hepatitis D. Lancet 1995;346:939-41.
6. Lafontaine DA, Mercure S, Perreault JP. Identification of a Crohn's disease specific transcript with potential as a diagnostic marker. Gut 1998;42:878-82.
7. Chao YC, Chang MF, Gust I, Lai MMC. Sequence conservation and divergence of hepatitis virus RNA. Virology 1990;178:384-92.
8. Chao YC, Lee CM, Pang HS, Govindarajan S, Lai MMC. Molecular cloning and characterization of an isolate of hepatitis delta virus from Taiwan. Hepatology 1991;13:345-52.
9. Lee CM, Bih FY, Chao YC, Govindarajan S, Lai MMC. Evolution of hepatitis delta virus RNA during chronic infection. Virology 1992;188:265-73.
10. Lee CM, Changchien CS, Chung JC, Liaw YF. Characterization of a new genotype II hepatitis delta virus from Taiwan. J Med Virol 1996;49:145-54.
11. Lafontaine AD, Mercure S, Perreault J-P. Update of the viroid and

viroid-like RNA sequence database: addition of the hepatitis delta RNA section. Nucleic Acids Res 1997;25:123-5.
12. Shakil AO, Hadziyannis S, Hoofnagle JH, Di Bisceglie AM, Gerin JL, Casey JL. Geographic distribution and genetic variability of hepatitis delta virus genotype 1. Virology 1997;234:160-7.
13. Makino S, Chang MF, Shih CK, et al. Molecular cloning and sequencing of human hepatitis virus RNA. Nature 1987;329:343-6.
14. Kuo MY, Chao M, Taylor J. Initiation of replication of the human hepatitis delta virus genome from cloned DNA: role of delta antigen. J Virol 1989;63:1945-50.