Recombination of Hepatitis D Virus RNA Sequences and its Implications

Jaw-Ching Wu,*† Tzen-Yuh Chiang,‡ Wei-Kei Shue,§ Sheng-Yung Wang,* I-Jane Sheen,* Yi-Hsiang Huang,* and Wan-Jr Syu||

*Division of Gastroenterology, Department of Medicine, Veterans General Hospital, Taipei, Taiwan; †Institute of Clinical Medicine, National Yang-Ming University School of Medicine, Taipei, Taiwan; ‡Department of Biology, Cheng-Kung University, Tainan, Taiwan; §Department of Mathematics and Statistics, Southern Illinois University at Edwardsville; and ||Institute of Microbiology and Immunology, National Yang-Ming University School of Medicine, Taipei, Taiwan

Recombination between RNA sequences plays a role in the fast evolution of a few viruses. There has been no report on hepatitis D virus (HDV) recombination. In this study, we analyzed genetic recombination of HDV and its possible impact on evolution and clinical course. The aligned HDV sequences allowed us to construct a phylogenetic tree which supported the notion of distinct lineages of HDV. The tree was also used in the analysis of recombination using partial likelihoods assessed through optimization. Nine segments of the HDV genome with significant levels of genetic recombination were detected. Five segments were in the hypervariable region, and four were in the delta-antigen-coding region. None could be found in the well-conserved autocleavage region that is essential for replication. Recombination occurred both between and within types. The results of this study indicated that the remarkable variation in HDV genomic sequences, particularly in the hypervariable region, among different genotypes may at least partly result from multiple episodes of genetic recombination during evolution. Genetic recombination may play a significant role in increasing genetic diversity. Importantly, a genetic recombination (nt 1082–1093) occurred in one of the immunogenic domains of hepatitis delta virus antigen recognized by human and woodchuck antibodies (amino acids 174–195). Genetic recombination also occurred at another segment between nt 1517 and 1535, which was close to one of the predicted T-cell epitopes (amino acids 26–41). In longitudinal analysis of HDV genomes at different time points during chronic infection, novel dominant HDV strains with amino acid changes at these epitopes usually emerged after severe hepatitis attacks. In the comparison of HDV clones during or shortly after flare-up of liver disease, K/Ks ratios of >1 were frequently found, suggesting Darwinian positive selection. Therefore, recombination in these two segments may play an important role for HDV in the evasion of immunity.

Introduction

Hepatitis D virus (HDV) is a defective circular RNA virus (Wang et al. 1986; Makino et al. 1987; Lai 1995) and requires a supply of hepatitis B surface antigen (HBsAg) envelope from hepatitis B virus (HBV) for its assembly and transmission (Rizzetto et al. 1980; Wu et al. 1991). The HDV genome is about 1.7 kb in length. It is composed of a hypervariable region, a delta-antigen-coding region, and a more conserved autocleavage region (fig. 1) (Wang et al. 1986; Makino et al. 1987; Lai 1995). The HDV is a virus of negative polarity and encodes the only protein, hepatitis delta antigen, through the open reading frame on the antigenomic strand (for review, see Lai 1995). There are two forms of the delta antigen. Both forms share the same open reading frame, and the translation of the large delta antigen requires editing of the antigenomic HDV strand. The small form is essential for HDV replication, and the large form, with a 19-amino-acid extension at the carboxyl end, is crucial for the packing virions and for trans-dominant negative inhibition (for review, see Lai 1995).

Microheterogeneity is up to 1.43% in nucleotide sequences of different HDV clones from a single subject (Wu et al. 1995a). HDV was classified into three genotypes (Casey et al. 1993; Wu et al. 1995a, 1995c), and a novel subtype of genotype II was recently identified (Wu, Chiang, and Sheen 1998). The divergence in nucleotide sequences ranges from 5% to 14% among different isolates of the same genotype and from 22% to 38% among different genotypes (Casey et al. 1993; Wu et al. 1995a; Wu, Chiang, and Sheen 1998). It has been suggested that a faster rate of evolution in the RNA viral genome (Imazeki, Omata, and Ohto 1990; Lee et al. 1992) may be due to the reduced proofreading activity of RNA polymerase. Nevertheless, recombination may play another key role in the fast evolution of a few viruses (Lai 1992). Recently, Chao et al. reported the advantage of sex (recombination) and the cost of intra-host competition in RNA viruses (Chao 1997; Chao, Tran, and Tran 1997; Turner and Chao 1998). So far, there has been no report on HDV recombination.

HDV superinfection is a common etiology of fulminant hepatic failure (Govindarajan et al. 1984; Wu et al. 1994). It is also an important cause of the progression of chronic hepatitis B to cirrhosis or liver failure (Govindarajan, De Cock, and Redeker 1986; Wu et al. 1990, 1995b). It has been reported that immunization with a recombinant hepatitis delta antigen fails to protect carrier woodchucks from HDV superinfection (Karayiannis et al. 1990). Thus, reinfection or mixed infection of HDV is highly likely to occur. Recently, mixed-geno-
type infection of HDV has been found in some subjects, particularly in those repeatedly exposed to risk behaviors (Wu et al. 1999). The HDV replicates through a double rolling circle mechanism (Kuo et al. 1988; Wu and Lai 1989; Lai 1995). The antigenomic strand, complementary to the genomic RNA, is produced using the genomic strand as a template, and the genomic RNA is, in turn, produced using the antigenomic strand as a template. The genomic or antigenomic HDV monomers are formed through autocleavage and autoligation of multimers at specific sites. It is reasonable to expect the occurrence of HDV recombination in such a situation. In this study, we show that recombination is more frequent in regions recognized as epitopes of human and woodchuck antibodies. This finding suggests that recombination may be an important evolutionary force in the evolution of HDV genomes during the progression of infection.

**Materials and Methods**

Serum samples were obtained from 3 patients with chronic HDV infection (Wu et al. 1995a, 1995b; Wu, Chiang, and Sheen 1998). They were all positive for serum HBsAg and antibody to HDV antigen (anti-HDV), and negative for immunoglobulin M antibody to hepatitis B core antigen (Austria II-125, anti-Delta, and CORAB-M; Abbott Laboratories, North Chicago, Ill.). Serum alanine aminotransferase (ALT) was measured with a sequential multi-autoanalyzer (Technicon SMAC; Technicon Instruments, Tarrytown, N.Y.).

For cloning of the whole HDV genomes, two sets of primers, (1) antisense #88 (complementary to nt 1663–1644) and sense #120 (homologous to nt 889–912) and (2) antisense #945 (complementary to nt 945–926) and sense #1622 (homologous to nt 1622–1641), were used in PCR as previously reported (Wu, Chiang, and Sheen 1998). Reverse transcription polymerase chain reaction (RT-PCR) was performed as previously reported, but with some modification according to the sequence of genotype I, IIa, or IIb (Chao et al. 1991; Wu, Chiang, and Sheen 1998; Wu et al. 1999). In brief, viral RNA was extracted from 50 μl of serum, dissolved in 10 μl DEPC-treated H₂O, and heated at 70°C for 10 min. Then, cDNA was generated in the presence of reverse transcriptase (SUPERSCRIPT II, GIBCOBRL, Life Technologies, Rockville, Md.) according to the manufacturer’s instructions. Each 100 μl of PCR reaction mixture contained 5 μl of cDNA, 0.5 μl (5 U/μl) of thermostable polymerase (TaKaRa Taq, Takara Shuzo Co., Ltd., Biomedical Group, Shiga, Japan), 10 μl of 10 × PCR buffer, 8 μl of dNTP mixture (2.5 mM each), 4 μl of primers (10 pmol/μl each), and 72.5 μl of H₂O. PCR was performed under the following conditions: denaturing at 95°C for 2 min, followed by 35 cycles of amplification (each cycle: 95°C for 20 s, 55°C for 40 s, 72°C for 1 min), followed by 72°C for 10 min. Strict
procedures were followed to avoid false positive results (Kwok and Higuchi 1989).

The amplified PCR products were ligated into the plasmid pCR2 vector (Original TA cloning kit, Invitrogen Corporation, Carlsbad, Calif.) according to the manufacturer’s instructions. The ligation mixture was used to transform the competent strain Escherichia coli DH5α (Gibco BRL, Life Technologies, Gaithersburg, Md.) (Wu et al. 1995a). Positive colonies were picked up and cultured in Luria-Bartani medium. Plasmid DNA was extracted and subjected to the dye terminator cycle sequencing reaction according to the standard protocol provided by the manufacturer (Dye Terminator Cycle Sequencing Core kit #402117, Perkin Elmer Cetus Corp., Norwalk, Conn.). The sequencing products were precipitated with alcohol and analyzed in an ABI 373A sequencer (Perkin Elmer Cetus Corp., Norwalk, Conn.).

Whole genomic sequences of HDV were obtained from GenBank and from previous reports (Wang et al. 1986; Makino et al. 1987; Lai 1995). Isolates Taiwan D2667 (TWD2667, Genbank accession number AF104263), TWD2476 (Genbank accession number AF104264), and TWD62 (Genbank accession number AF018077) were cloned from our laboratory and belonged to genotypes I, IIa, and IIb, respectively (Wu, Chiang, and Sheen 1998). HDV nucleotide sequences were aligned using the PILEUP program implemented in the Genetics Computer Group (GCG) package. Phylogeny of HDV based on the genomic sequences was estimated by the phylogenetic analysis using parsimony program (PAUP, version 3.1.1) (Swofford 1993) with the branch-and-bound algorithm, as well as by maximum likelihood using PHYLIP software (Felsenstein 1993). Characters were weighted equally (Casey et al. 1993; Waters and Wierling 1999). South American sequences, i.e., Peru-1, Peru-2, and Colombina (Casey et al. 1993), were chosen as outgroups based on previous studies (Casey et al. 1993; Wu, Chiang, and Sheen 1998) and biogeographical evidence. The confidence of clades was tested by bootstrap (Felsenstein 1985) with 1,000 replicates (Hedges 1992). The nodes with bootstrap values >0.70 are significantly supported, with >95% probability (Hillis and Bull 1993).

Genetic recombination among HDV nucleotide sequences was detected using partial likelihoods assessed through optimization (PLATO) (Grassly and Holmes 1997), a computer program developed for detecting gene regions that do not fit with a “global” phylogenetic topology based on Monte Carlo simulations. Using maximum-likelihood phylogeny, the likelihood for each site of a sequence can be calculated independently. The significance of an anomalous region could then be tested at a desired level, say, alpha = 0.05. Since the tests are carried out for each region from 5 nt up to half of the sequence length, the Bonferroni inequality is applied, whereby the level of alpha used for each test of significance is divided by the number of region size classes considered. For the sequence considered in table 1, the calculated Z value of >3.859463 indicates statistical significance.

The ratio of nonsynonymous to synonymous substitution rates was calculated to estimate the amount and types of selection affecting a gene (Messier and Stewart 1997; Sharp 1997). We estimated the number of synonymous substitutions per synonymous site ($K_s$) and the number of nonsynonymous substitutions per nonsynonymous site ($K_a$) based on Kimura’s (1968) two-parameter model using K-Estimator 4.5 (Comeron 1995).

### Results

**Phylogeny of HDV and Detection of Recombination**

A single tree (fig. 2) based on whole HDV genomic sequences was recovered by PAUP analysis which had an identical topology to that of the tree identified by PHYLIP. Four types of HDV were significantly supported by bootstrap analysis. Types IIa and IIb group together, and then form a group with type I that is distinct from type III.

Nine HDV genome segments with significant levels of genetic recombination were detected. Five segments were in the hypervariable region, and four were in the delta-antigen-coding region. None could be found in the autocleavage region. The corresponding positions of the nine segments with genetic recombination on the HDV
Recombination of Hepatitis D Virus

**Fig. 2.**—Phylogenetic analysis of HDV based on whole genomic sequences using PAUP. The confidence of the clades, i.e., the monophyly, was tested by bootstrapping with 1,000 replicates of heuristic searches. The number on each node indicates a bootstrap value (%). The nodes with bootstrap values $>70\%$ are significantly supported with $>95\%$ confidence. Peru1 and US2 are from Casey et al. (1993); Lebanon is from Lee et al. (1992); Nauru is from Chao et al. (1990); Italy is from Wang et al. (1986); Japan1 is from Imazeki, Omata, and Ohno (1990); T3 is from Lee et al. (1996); and the remaining TW (Taiwan) isolates are from our laboratory. The Roman numerals indicate genotypes.

The genome are indicated in figure 1. The starting and ending nucleotides of each HDV segment with genetic recombination and the presumed provenance of the recombinants are shown in table 1. Three of the five segments in the hypervariable region were longer than 50 nt, while none of the four segments in the delta-antigen-coding region was longer than 20 nt. An example of recombination within the hypervariable region is shown in figure 3. Recombination occurred between different types: PLATO analysis identified sequence regions between nt 341 and 353 with recombination occurring between types I and IIb, as well as between types I and IIa, and regions from nt 495 to nt 638 with recombination occurring between types I and IIa. Recombination also occurred within types: PLATO analysis identified sequence regions from nt 1082 to nt 1093 with recombination occurring within type I HDV as well as between types I and IIb, and regions from nt 1517 to 1535 with recombination occurring within type I, as well as among types I, IIb, and IIa.

**Implications of Genetic Recombination in Immunogenic Domains**

Importantly, a genetic recombination (nt 1082–1093) occurred in one of the immunogenic domains of hepatitis delta virus antigen recognized by human and woodchuck antibodies (amino acids 174–195) (fig. 4) (Wang et al. 1990). Genetic recombination also occurred at another segment between nt 1517 and 1535, which was close to one of the predicted T-cell epitopes (amino acids 26–41) (Nisini et al. 1997).

To study the implication of genetic recombination in immunogenic domains, HDV genomes were isolated at different time points from patients D62 (AF165192–AF165210) and 2667 (AF165940–AF165966). At least five clones (5–11) were analyzed at each time point. A total of 46 sequences (645 nt in length for genotype I, 642 nt for genotype IIb) were analyzed. Patient D62 had acute genotype IIb HDV superinfection. Early time point a was at the acute stage of infection, while time points b and c were 75 and 86 months after time point a, respectively (fig. 5A). The HDV superinfection in patient D62 progressed to chronicity. There were multiple episodes of flare-ups of liver diseases accompanied by marked elevation of serum ALT levels between time points a and b, while there was no flare-up of liver disease, as reflected by mild elevation of ALT levels, between time points b and c. As shown in figure 5B, a novel dominant strain different from the original one at time point a was found at time points b and c. The most drastic change in amino acid sequences of HDV clones between early and late time points occurred in a B-cell

**Fig. 3.**—Recombination in the hypervariable region of HDV. The numbers on the aligned sequences indicate nucleotide numbers (Lai 1995). The boxed areas and the boxed, shaded areas indicate recombined sequences. Dots indicate identical nucleotides, whereas dashes indicate missing nucleotides. The sources of the HDV strains are indicated in figure 2. The Roman numerals indicate genotypes.
epitope spanning from amino acid 169 to amino acid 194 (the boxed, shaded area), overlapping with the above-mentioned area of genetic recombination. Furthermore, the substitution of alanine to valine (V) at residue 190 was a nonconservative amino acid change. Interestingly, the threonine (T) at residue 190 and the V at residue 187 could be found in the corresponding positions of some genotype I HDV strains (fig. 4). Many changes of nucleotide sequences in other regions of the delta-antigen-coding sequences at different time points were synonymous and did not induce amino acid changes.

As shown in figure 6A, patient 2667 had chronic hepatitis D with persistent hepatic inflammation and relapsing acute exacerbations. The early time point a was at the peak of ALT elevation, and time point b was just prior to a high peak of ALT elevation, 24 months from time point a. Time point c, 6 months after the time point b, was during a temporary remission stage with a normal ALT level. There were no significant changes in the B-cell epitope from amino acid 170 to amino acid 195 between the dominant strains at time points a and b, although there were some amino acid changes outside this epitope (fig. 6B). On the contrary, the dominant strains (clones a-1 to b-7) at time points a and b were mostly replaced by a novel dominant strain (clones c-1 to c-5) at time point c. These changes occurred within both T-cell and B-cell epitopes reported previously (Wang et al. 1990; Nisini et al. 1997) and were particularly striking in the latter. In the B-cell epitope shown in the boxed, shaded area in figure 6B, 5 of 26 amino acids were substituted by other amino acids. For these five substituted amino acids, the 170 changes from asparagine (N) to serine (S) and the 172 changes from leucine (L) to glutamine (Q) were nonconservative changes. Interestingly, these substitutions occurred in the recombined area reported above. Of note, the amino acids on the substitution sites could be found in the corresponding positions of other strains of the same genotype or in other genotypes (fig. 4). However, the arginine at residue 121 of the T-cell epitope of the novel dominant strain at time point c could not be found in other strains of the same or different genotypes. Although the dominant strains (clones a-1 to b-7) at time points a and b were mostly replaced by a novel dominant strain (clones c-1 to c-5), some clones still persisted as minor variants (c6–8) at time point c.

Determination of $K_{s}/K_{a}$ and Its Implications

To determine $K_{a}$ and $K_{s}$ values, representative HDV sequences of patients D62 and 2667 were chosen and compared. To simplify the table, only one representative sequence was selected from several identical sequences. Based on pairwise comparisons, the $K_{a}$ and $K_{s}$ values between clones at the same time point (quasispecies) or clones of TWD62 at different time points are shown in table 2. In the comparisons of HDV quasispecies at the time point a (acute stage of infection) for the patient D62, only 1 of 10 $K_{a}/K_{s}$ ratios was >1. None of the $K_{a}/K_{s}$ ratios was >1 at time points b or c, when the hepatic inflammation was very mild (table 3). In the comparisons of HDV clones between time points separated by flare-ups of liver disease, 7 of 20 $K_{a}/K_{s}$ ratios between time points a and b and 3 of 15 $K_{a}/K_{s}$ ratios between time points a and c were >1. $K_{a}/K_{s}$ ratios were also determined in various HDV clones of patient 2667 and various HDV strains in the world shown in the phylogeny of figure 2 (data not shown). As shown in table 3, a large number of $K_{a}/K_{s}$ ratios were >1 at time points a (4/10) and b (8/10) when flare-up of liver disease occurred during chronic course of infection. In contrast, only one of six $K_{a}/K_{s}$ ratios was >1 at time point c, when patient 2667 was in a temporary remission state. In the comparisons of HDV clones at time points separated by flare-ups of liver disease, a large number of $K_{a}/K_{s}$ ratios were >1 (a vs. b: 14/25, a vs. c: 4/20; b vs. c: 7/20). In the comparisons among different strains of the world (fig. 2), none of the 45 $K_{a}/K_{s}$ ratios was >1.

Discussion

There has been no report of genetic recombination in the small 1.7-kb HDV, whereas it has been shown in a few RNA viruses (Lai 1992; Chao 1997; Chao, Tran, and Tran 1997; Turner and Chao 1998). In this study, recombination events between different genotypes as well as within the same genotype of HDV were detect-
Recombination of Hepatitis D Virus

**Fig. 5.**—Emergence of novel dominant HDV strains with amino acid changes clustering within T-cell and B-cell epitopes after acute exacerbations of an acute hepatitis D patient (patient D62) infected with genotype Ib HDV (Wu, Chiang, and Sheen 1998) with progression to chronicity. A. Clinical course. Shaded area indicates normal range of ALT level; a, b, and c indicate time points of blood sampling. B. Partial amino acid sequences of the delta antigen of randomly selected HDV clones at different time points. Boxed areas and boxed, shaded areas labeled T and B are T-cell and B-cell epitopes reported previously (Wang et al. 1990; Nisini et al. 1997). The numbers above amino acid sequences are the numbers of amino acids starting from the initiation of the delta antigen. Dots indicate identical amino acids. ALT = alanine aminotransferase.
Fig. 6.—Emergence of novel dominant HDV strains with amino acid changes clustering within T-cell and B-cell epitopes after an acute exacerbation of a chronic hepatitis D patient (patient 2667) infected with genotype I HDV. A, Clinical course. Shaded area indicates normal range of ALT level; a, b, and c indicate time points of blood sampling. B, partial amino acid sequences of the delta antigen of randomly selected HDV clones at different time points. Boxed areas and boxed, shaded areas labeled T and B are T-cell and B-cell epitopes reported previously (Wang et al. 1990; Nisini et al. 1997). The numbers above amino acid sequences are the numbers of amino acids starting from the initiation of the delta antigen. Dots indicate identical amino acids. ALT = alanine aminotransferase.
ed. The accuracy of detection of recombination is proportional to the rates of evolution along the tree (Grassly and Holmes 1997), simply because more changes provide more information along branches. In this case, the recombination detection appeared reliable according to the Z values obtained (table 1).

Of note, genetic recombination occurred in the hypervariable and the delta-antigen-coding regions, but not in the more conserved autocleavage region. The well-conserved autocleavage region is responsible for autocleavage and autoproteolysis during double rolling circle replication of HDV (Kuo et al. 1988; Wu and Lai 1989; Lai 1995). Recombinants of this region, if they ever do occur, may cause a failure in replication (Kuo et al. 1988; Wu and Lai 1989; Lai 1995). The autocleavage region downstream of the hypervariable region may result from multiple genetic recombinations during evolution. In recent reviews of RNA recombination (Lai 1992; Nagay and Simon 1997), a copy choice mechanism predicts that recombination would occur more frequently at RNA sites with strong secondary structures, which promote transcription pausing. The autocleavage region downstream of the hypervariable region happens to have a strong secondary structure (Lai 1992; Wu et al. 1992) which may facilitate transcription pausing and recombination. Further studies are needed to prove this hypothesis.

It is generally true that an RNA virus usually has a more rapid rate of evolution than a DNA virus does. The rapid evolutionary rate may account for the extremely heterogeneous viral genome populations within a single subject and the high variation of HDV genomic sequences among viral genomes of different subjects or different genotypes (Casey et al. 1993; Wu et al. 1995a), but how rapid the rate of evolution would be still remains controversial (Imazeki, Omata, and Ohto 1990; Lee et al. 1992; Chao, Tang, and Hsu 1994; Netter et al. 1995). For example, Netter et al. (1995) found that HDV RNA stability (no sequence change) was observed in a serial passage six times in woodchucks over a period of 265 days. What is the cause for the discrepancy for the estimation of the evolutionary rate of HDV? The estimation of the evolutionary rate of HDV was based on the comparison of viral sequences cloned at different time points in earlier reports (Imazeki, Omata, and Ohto

---

### Table 2

<table>
<thead>
<tr>
<th>Pairwise Comparisons of Synonymous (Ks) and Nonsynonymous (Ka) Substitution Rates in the Delta-Antigen-Coding Region of Various HDV Clones Isolated from Serum Samples of Patient D62 at Different Time Points</th>
<th>Ks/Ka</th>
<th>a-1</th>
<th>a-3</th>
<th>a-4</th>
<th>a-6</th>
<th>a-7</th>
<th>b-1</th>
<th>b-2</th>
<th>b-3</th>
<th>b-4</th>
<th>c-3</th>
<th>c-5</th>
<th>c-6</th>
</tr>
</thead>
<tbody>
<tr>
<td>a-1 . . .</td>
<td>0.0000</td>
<td>0.00000</td>
<td>0.01090</td>
<td>0.00232</td>
<td>0.00934</td>
<td>0.01171</td>
<td>0.01127</td>
<td>0.01172</td>
<td>0.00934</td>
<td>0.01171</td>
<td>0.01171</td>
<td></td>
<td></td>
</tr>
<tr>
<td>a-3 . . .</td>
<td>0.00487</td>
<td>0.00846</td>
<td>0.01131</td>
<td>0.00232</td>
<td>0.00934</td>
<td>0.01171</td>
<td>0.01127</td>
<td>0.01172</td>
<td>0.00934</td>
<td>0.01171</td>
<td>0.01171</td>
<td></td>
<td></td>
</tr>
<tr>
<td>a-4 . . .</td>
<td>0.00846</td>
<td>0.01134</td>
<td>0.01090</td>
<td>0.00232</td>
<td>0.00934</td>
<td>0.01171</td>
<td>0.01127</td>
<td>0.01172</td>
<td>0.00934</td>
<td>0.01171</td>
<td>0.01171</td>
<td></td>
<td></td>
</tr>
<tr>
<td>a-6 . . .</td>
<td>0.01331</td>
<td>0.01829</td>
<td>0.02188</td>
<td>0.01137</td>
<td>0.01563</td>
<td>0.01803</td>
<td>0.01758</td>
<td>0.01804</td>
<td>0.01563</td>
<td>0.01803</td>
<td>0.01803</td>
<td></td>
<td></td>
</tr>
<tr>
<td>a-7 . . .</td>
<td>0.00000</td>
<td>0.00487</td>
<td>0.00846</td>
<td>0.01131</td>
<td>0.01171</td>
<td>0.01171</td>
<td>0.01409</td>
<td>0.01364</td>
<td>0.01410</td>
<td>0.01171</td>
<td>0.01409</td>
<td></td>
<td></td>
</tr>
<tr>
<td>b-1 . . .</td>
<td>0.00489</td>
<td>0.00977</td>
<td>0.01342</td>
<td>0.01826</td>
<td>0.00489</td>
<td>0.00000</td>
<td>0.00232</td>
<td>0.00192</td>
<td>0.00232</td>
<td>0.00000</td>
<td>0.00232</td>
<td></td>
<td></td>
</tr>
<tr>
<td>b-2 . . .</td>
<td>0.00981</td>
<td>0.01473</td>
<td>0.01836</td>
<td>0.02331</td>
<td>0.00981</td>
<td>0.00490</td>
<td>0.00232</td>
<td>0.00192</td>
<td>0.00232</td>
<td>0.00000</td>
<td>0.00232</td>
<td></td>
<td></td>
</tr>
<tr>
<td>b-3 . . .</td>
<td>0.00981</td>
<td>0.01489</td>
<td>0.01832</td>
<td>0.02322</td>
<td>0.00981</td>
<td>0.00487</td>
<td>0.00232</td>
<td>0.00192</td>
<td>0.00232</td>
<td>0.00000</td>
<td>0.00232</td>
<td></td>
<td></td>
</tr>
<tr>
<td>b-4 . . .</td>
<td>0.01340</td>
<td>0.01340</td>
<td>0.02206</td>
<td>0.02688</td>
<td>0.011340</td>
<td>0.00851</td>
<td>0.01277</td>
<td>0.00424</td>
<td>0.00192</td>
<td>0.00424</td>
<td>0.00192</td>
<td></td>
<td></td>
</tr>
<tr>
<td>c-3 . . .</td>
<td>0.01834</td>
<td>0.02325</td>
<td>0.02701</td>
<td>0.03194</td>
<td>0.01834</td>
<td>0.01344</td>
<td>0.00851</td>
<td>0.01834</td>
<td>0.01717</td>
<td>0.00232</td>
<td>0.00465</td>
<td></td>
<td></td>
</tr>
<tr>
<td>c-5 . . .</td>
<td>0.01342</td>
<td>0.01830</td>
<td>0.02206</td>
<td>0.02687</td>
<td>0.01342</td>
<td>0.00851</td>
<td>0.01345</td>
<td>0.01336</td>
<td>0.01717</td>
<td>0.02211</td>
<td>0.00232</td>
<td></td>
<td></td>
</tr>
<tr>
<td>c-6 . . .</td>
<td>0.00489</td>
<td>0.00977</td>
<td>0.01342</td>
<td>0.01826</td>
<td>0.00489</td>
<td>0.00000</td>
<td>0.00490</td>
<td>0.00487</td>
<td>0.00851</td>
<td>0.01344</td>
<td>0.00851</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

*K/K > 1.
A single clone or only a few clones of HDV were obtained at each time point, and the estimation appeared to be based on the assumption that variations of HDV genomic sequences were accumulating. The concept of "quasispecies" was not taken into consideration during the estimation (Van Regenmortel 1997). It is likely that some of the "mutations" or "substitutions" observed in late periods may be some minor variants that had been presented in the original viral populations but were not detected. In this study, the analysis of multiple clones (quasispecies) at different time points clearly indicates that the comparisons of a dominant strain at each time point may overestimate the evolutionary rate (patient 2667 in fig. 6). The observation of stability of HDV RNA in woodchucks also supports this assumption, because the inoculum of woodchucks was obtained from particles released into medium by culture cells cotransfected with a single clone of HDV cDNA and a woodchuck hepatitis virus cDNA clone. The lack of quasispecies of HDV genome in an inoculum makes the estimation of evolutionary rate slower. Therefore, a faster RNA virus evolutionary rate may not be explained by a higher mutation rate due to lack of proofreading activity of RNA polymerase alone. This study indicates a higher mutation rate due to lack of proofreading activity of RNA polymerase. The lack of quasi-species may at least partly result from multiple episodes of genetic recombination (either between genotypes, within a genotype, or within a quasispecies in an infected subject) during evolution.

The delta antigen is the only protein encoded by HDV (Wang et al. 1986; Makino et al. 1987; Lai 1995). In this study, genetic recombination was also found within the delta-antigen-coding region of HDV. Importantly, genetic recombination events occurred in B- and T-cell epitopes reported previously (Wang et al. 1990; Nisini et al. 1997). Moreover, both of the chronic hepatitis D patients with acute exacerbations (a severe liver injury presumably caused by immune attack) who had been followed up longitudinally in this study showed drastic amino acid changes within the recombined region in the novel dominant HDV strains after acute exacerbation. As many as 4 to 5 of 26 amino acids in the B-cell epitope of the hepatitis delta antigen changed, and at least one of these changes was nonconservative. It is reasonable to expect that these changes may alter the immune-recognition epitope of the delta antigen and result in the emergence of escape mutants from strong immune selection associated with acute exacerbation. In our recent study, more delta-antigen-coding sequences isolated from more chronic hepatitis D patients with acute exacerbations were analyzed, and they all showed drastic amino acid changes in this region in the novel dominant HDV strains after acute exacerbation. In addition, another chronic hepatitis D patient showed drastic amino acid changes in a region from amino acid 26 to amino acid 41, one of the predicted T-cell epitopes on the delta antigen, in the novel dominant HDV strain after an episode of hepatic inflammation (unpublished data). This T-cell epitope is also one of the recombined regions in this study. Our findings suggest that these two epitopes are critical in immune selection during acute exacerbation. Therefore, point mutation and selection in these two segments may play an important role for HDV in the evasion of immunity. It has been reported that duplication of genes with inactivation of old ones, introduction of point mutations, and intralocus and interlocus recombination seem to be the major mechanisms generating amino acid polymorphism in the antigen recognition region of the major histocompatibility complex (MHC), which is important in coping with diverse pathogens (Hughes and Nei 1988; Vogel et al. 1999; Yeager and Hughes 1999). However, pathogens appear to co-evolve with the MHC class I alleles present in a population and escape from immune recognition (Hughes and Nei 1988; Vogel et al. 1999; Yeager and Hughes 1999).

Of note, the substituted amino acids in the B-cell epitope of the emerging dominant HDV strains could usually be found in other strains of the same genotypes or different genotypes, suggesting possible recombination. However, parallel virus evolution is another possibility for these changes (Crandall et al. 1999). The finding of the substitution of 5 of 26 amino acids in the B-cell epitope of the delta antigen in a short period of time (6 months between time points b and c in patient 2667) does not favor parallel evolution, because the estimated evolutionary rate based on comparisons of the sequences of dominant strains in our patients before and after exacerbation far exceeds the estimated evolutionary rate of HDV reported previously (Imazeki, Omata, and Ohno 1990; Lee et al. 1992; Chao, Tang, and Hsu 1994), which had already been overestimated (Netter et al. 1995).

A $K_a/K_s$ ratio of >1 was reported to be an indicator of Darwinian positive selection (Messier and Stewart 1997; Sharp 1997). However, Crandall et al. (1999) failed to detect selection using this ratio. This study showed that this ratio may be of some value in the detection of positive selection during or shortly after flare-ups of liver disease which result in the replacement of original dominant strains by novel ones. However, this ratio is of little value in the detection of remote positive selection long after acute exacerbation, when liver disease is in a quiescent state. As suggested by Crandall et al. (1999), pairwise comparisons, which count synonymous substitutions on internal branches multiple times, may give rise to an underestimate of $K_a/K_s$. In addition, "the selection is proportionally on a few sites, such that even if the non-synonymous rate is estimated correctly, it will still be swamped by synonymous change at other positions" (Crandall et al. 1999).

In summary, the analyses of our data indicate that recombination has occurred in HDV. Via changes in nucleotide and amino acid sequences, recombination may play a role in genetic diversity and in the escape from immune responses of the host.

**Acknowledgment**

This study was supported by grants NSC87-2314-B-010-062 and NSC88-2315-B-010-013 from the National Science Council, Taiwan.
LITERATURE CITED


DAVID RAND, reviewing editor

Accepted August 10, 1999