Positive Selection of Hepatitis Delta Antigen in Chronic Hepatitis D Patients

Shen-Yung Wang,1,5 Jaw-Ching Wu,1,3* Tzen-Yuh Chiang,6 Yi-Hsiang Huang,2,3 Chien-Wei Su,2,4 and I-Jane Sheen6

Department of Medical Research and Education1 and Division of Gastroenterology, Department of Medicine,2 Taipei Veterans General Hospital, Taipei, Institute of Clinical Medicine3 and Faculty of Internal Medicine,4 National Yang-Ming University School of Medicine, Taipei, Division of Gastroenterology, Department of Medicine, Mackay Memorial Hospital, Taipei,6 and Department of Life Sciences, National Cheng-Kung University, Tainan,6 Taiwan

Received 22 December 2006/Accepted 5 February 2007

Liver disease may become ameliorated in some patients with chronic hepatitis D virus (HDV) infection. We present here a study based on longitudinal sampling to investigate the viral dynamics in chronic HDV infection. We examined the HDV variants from different time points, especially those before and after the elevation of serum aminotransferase levels. The datasets from each patient were tested for positive selection by using maximum-likelihood methods with heterogeneous selective pressures along the nucleotide sequence. An average of 4.9%, ranging from 3.1 to 6.8%, of the entire delta antigen sites was regulated by a diversifying selection. Most of the positively selected sites were associated with immunogenic domains. Likelihood ratio tests revealed a significant fitness of positive selection over neutrality of the hepatitis delta antigen gene in all patients. We further adapted a neural network method to predict potential cytotoxic T ligand epitopes. Among the HLA-A*0201 cytotoxic T ligand epitopes, three consistent epitopes across all three genotypes were identified: amino acids (aa) 43 to 51, 50 to 58, and 114 to 122. Three patients (60%) had sites evolving under positive selection in the epitope from aa 43 to 51, and four patients (80%) had sites evolving under positive selection in the epitope from aa 114 to 122. The discovery of immunogenic epitopes, especially cytotoxic-T-lymphocyte ligands, associated with chronic HDV infection may be crucial for further development of novel treatments or designs in vaccine for HDV superinfection.

There are two modes of the HDV infection (19). Coinfection results from acute infection with both hepatitis B virus (HBV) and HDV, whereas superinfection results from HDV infection of patients with underlying chronic hepatitis B infection. The great majority of patients with HDV superinfection progress to chronicity, whereas few patients with coinfection do (27). The disease spectrum of HDV infection varies greatly from fulminant hepatitis, rapidly progressive disease, to a subclinical course (7, 9, 20, 26, 27). Persistent replication of HDV associated with continuous hepatic inflammation and elevated alanine aminotransferase (ALT) levels is a characteristic of chronic active hepatitis D, and infection of different HDV genotypes may influence the outcomes (27, 30).

The pathogenesis of HDV infection is still unclear. Cytotoxicity was hypothesized in some studies but was not supported by many others (6, 17). An immune mechanism is believed by most investigators to be responsible for hepatic inflammation and clearance of HDV (10, 16). B-cell and T-cell epitopes were mapped based on in vitro studies (16, 25). Recently, we reported two potential cytotoxic T-cell epitopes in animal and in vitro studies (10, 11). We reported that HDV variants in quasispecies might be changed during clinical course of chronic hepatitis D (29). However, the mechanism and selected sites for the changes of these HDV variants are still obscure.

Only a few statistical studies have ever analyzed selection forces in HDV (2, 12, 29). A positive selection can be inferred if more than synonymous substitutions occur, whereas negative selection removes most nonsynonymous changes. In our previous study, positive selection of the delta antigen sequence
was inferred with pairwise comparisons of synonymous and non synonymous substitution rates (Ks/Ka) in two patients with chronic hepatitis D (29). The substitution rate was averaged over the entire delta antigen sequence, whereas the positively selected sites were not specified. Anisimova and Yang (2) reported an analysis that allowed selective pressures to vary among codon sites on the small delta antigen (HDAg-S) gene of 33 HDV clones. About 11% sites, mostly located within the immunogenic domains of the HDAg-S gene, were predicted to be shaped by positive selection. Nevertheless, the clones used in that analysis did not display any geographical and temporal relationships. Furthermore, the C-terminal extension of the large-form HDAg has never been studied. Serving as an iso- prenylation signal required for packaging of HDV virion, the sequence of the C-terminal extension of large HDAg is conserved within each genotype but highly diverged between different genotypes. The divergence caused difficulties in evaluating the selective pressures on these sites especially since different genotypes were pooled together into analysis. Moreover, the relationship between those presumably selected sites of HDAg and the clinical course of chronic HDV infection remains unanswered.

In the present study, we investigated the viral population in chronic HDV infection. We examined the genetic variation and selective modes of HDV variants from different time points especially those before and after elevation of the serum ALT levels, which are supposed to be selectively shaped by the host immunities, using the likelihood methodology of Yang et al. (33). Models with variable selective pressures among sites were implanted to reveal the positively selected sites.

MATERIALS AND METHODS

Patients. Serum samples were collected at several time points from each of five patients with chronic hepatitis D. These patients went into remission status during longitudinal follow-up. Chronic hepatitis was diagnosed by an elevated ALT level lasting for more than 6 months. Biochemical remission was defined as normal ALT levels in at least three separate samples lasting for at least 12 months without histological or clinical evidence of cirrhosis or hepatocellular carcinoma (HCC) (22, 27). Of the patients studied, two were infected with HDV genotype I, another two were infected with genotype IV (previously named Ib) (17), while another was infected with genotype III (previously named Ib). The patients were all positive for serum hepatitis B surface antigen and antibody to HDV antigen (anti-HDV) for more than 6 months and negative for immunoglobulin M (IgM) antibody to hepatitis B core antigen (anti-HBc), antibody to hepatitis C virus (anti-HCV), and IgM antibody to hepatitis A virus (anti-HAV). The HDV variants before and after elevation of the serum ALT levels were cloned from the sera of these patients.

Detection of viral nucleic acids. Because HDAg is the only protein encoded by the HDV and plays an essential role in viral replication and assembly (14), this study focused on changes of sequences of the HDAg-coding region. Reverse transcription PCR using the primers 120 (homologous to a sequence from nucleotides 889 to 912) and 88 (complementary to a sequence from nucleotides 1663 to 1642) according to the HDAg-coding region in the antigenic strand of HDV RNA was performed as reported previously (4, 28). Pfu DNA polymerase (Promega, Madison, WI) with the highest fidelity was used to minimize the PCR error rate. All of the experiments were independently repeated, and the results were reproducible. Strict procedures were followed to avoid false-positive results (13). Further PCR were done as previously reported (27).

PCR cloning and sequencing. The amplified PCR products were ligated to a plasmid pCR2 vector (Original TA cloning kit; Invitrogen Corp., Carlsbad, CA). The ligation mixture was used to transform the competent Escherichia coli strain DH5α (Gibco-BRL/Life Technologies, Gaithersburg, MD) as reported elsewhere (5). Multiple colonies from each time point of each case were randomly selected and cultured in the Luria-Bertani medium. Plasmid DNA was extracted and subjected to the dye terminator cycle sequencing reaction (dye terminator cycle sequencing core kit 020117; Perkin-Elmer Cetus Corp., Norwalk, CT). The sequencing products were precipitated with alcohol and analyzed in an ABI 373A sequencer (Perkin-Elmer Cetus Corp.). The sequences were submitted to GenBank; the corresponding GenBank accession numbers are EF187041 to EF187227.

HDAg-DNA typing and serological viral markers. HLA (phenotypes were classified by using Micro-SSP HLA DNA typing trays (One Lambda, Inc., Canoga Park, CA). Assays for viral hepatitis markers, including hepatitis B surface antigen, hepatitis B e antigen (HBeAg), IgM anti-HBc, anti-HDV, and IgM anti-HAV were tested in each case. All of the markers were detected by radioimmunoassay kits (Austria II-125, HBeAg-RIA, CORAB-M, anti-Delta, and HAVAB-M; Abbott Laboratories, Chicago, IL). Anti-HCV was measured by a second generation enzyme immunoassay (Abbott Laboratories). Serum ALT levels were measured by a sequential multi-autoanalyzer (Technicon SMAC; Technicon, Tarrytown, NY).

Analysis of quasispecies. Several immunogenic domains over delta antigen have been demonstrated (16, 25). The composition of quasispecies may be altered during the disease course (29). The present study focused on the relationships between immunogenic domains and quasispecies. A dominant variant is defined as a viral strain represented for more than half of the viral population. The datasets of each patient were compiled for the model implantation of positive selection. The predicted amino acid sequences of dominant variants from each data set were processed into the prediction of cytotoxic T ligands.

Tests of positive selection. Analysis of selective pressures was conducted with a maximum-likelihood approach, which implements different models in evaluating variable ω, equivalent to nonsynonymous/synonymous substitution ratios, among codon sites (33). Model M0 assumes a single ω for all sites. Model M1, for evaluation of neutral evolution, applies two classes: a portion of conserved sites with ω = 0 and the neutral sites with ω = 1. Model M2, denoted as a selection model, adds an additional class of sites with ω2 to be estimated from the data. Model M3 uses a discrete distribution to estimate the ω value in three classes and provides a sensitive test for positive selection. Likelihood ratio tests were performed to analyze the statistical significance. Sequential to model implantation, the Bayes theorem was applied to identify positively selected sites through maximum-likelihood approach (32). All methods described here were applied on the sequence sets by using CODEML programs of PAML package v3.13 (31).

Prediction of MHC class I-restricted Ligands. In addition to the known immunogenic domains over delta antigen, the domains of major histocompatibility complex (MHC) class I-restricted T-cell epitopes remain to be elucidated. The present study estimated MHC class I-restricted ligands within delta antigen with the use of NetMHC, which adapts a neural network method to predict binding affinity of peptides being presented by MHC class I molecules (15). Based on the assumption that host immunity interacts mostly with the majority in the viral population, the dominant variants of the first time point from each patient were selected for analysis. Among the high-affinity list of the predicted ligands, the cutoff binding level was determined by sites tested in a HLA-A*0201 transgenic mouse model (10). The analyses were done by the NetMHC program v2.1 (15).

RESULTS

Test of positive selection. The results of likelihood ratio tests and the maximum-likelihood estimates of different models’ implantation are listed in Table 1. Model M2 suggested that 3.1 to 6.8% (mean 4.9%) of the delta antigen sites evolved under a positive selection with a ω2 value ranging from 7.582 to 23.958. Model M2 assumes a class of constrained sites with ω0 = 0. More than half of the sites in delta antigen gene ranging from 62.7 to 79.7% (mean, 68.1%) were conserved. Model M2 all had a significantly better log-likelihood score in five patients, indicating a greater fitness of positive selection over neutrality.

Anisimova and Yang (2) reported that model M2, by adding an additional class of sites to neutral model M1, may identify purifying rather than positive selected sites (2). In the present study, the optimal peaks of model M2 all favored the hypothesis of positive selection. The likelihood ratio tests also showed strong evidences of positive selection in all five patients. Positively selected sites predicted by the Bayes theorem under model M3 with a possibility of >90% are presented in Table 2.

Downloaded from http://jvi.asm.org/ on November 12, 2017 by guest
Most positively selected sites (ranging from 50 to 83% in each patient) were located within the immunogenic domains. Moreover, patient A had only four sites (9, 116, 121, and 180), patient B had three sites (9, 121, and 180), patient C had seven sites (9, 17, 28, 149, 158, 180, and 191), patient D had four sites (6, 16, 171, and 187), and patient E had five sites (6, 16, 171, 187, and 190) that were homologous to those reported by Anisimova and Yang (2). By analyzing HDV variants before and after intense selection pressures, additional sites under positive selection were disclosed. Indeed, these sites demonstrated a stronger association with the immunogenic domains (ranging from 73.6 to 100%). In the present study, we also observed a higher percentage of conserved sites than previously reported (2). This suggests that most sites essential for viral structure or replication are regulated by a purifying selection.

**CTL epitope prediction.** Of the five patients examined in the present study, four had HLA typing of HLA-A*0201, and three of them had a mixture of another allele (two HLA-A*1101 and the other HLA-A*2403). The only non-HLA-A2 patient was HLA-A*1101 (Table 3). Huang et al. (10) identified two epitopes HDV 26-34 and HDV 43-51 in four HLA-A*0201 patients. By analyzing the published peptides’ binding affinity along with those predicted by the neural network algo-

### TABLE 2. Positively selected sites predicted in the delta antigen gene

<table>
<thead>
<tr>
<th>Patient</th>
<th>Positively selected sites*</th>
</tr>
</thead>
<tbody>
<tr>
<td>A</td>
<td>9, 34*, 47, 48, 49, 112, 113, 116, 121, 180, 188</td>
</tr>
<tr>
<td>B</td>
<td>8, 9, 43, 48, 78*, 81, 121, 170, 172, 173*, 174, 180</td>
</tr>
<tr>
<td>C</td>
<td>7, 9, 16, 17, 26, 27, 28, 29, 37, 43, 44, 48, 78, 112, 115, 124, 131, 141, 149, 154, 155, 158, 189, 180, 184, 191</td>
</tr>
<tr>
<td>D</td>
<td>6*, 16*, 171, 172, 181, 187</td>
</tr>
</tbody>
</table>

*That is, the sites under selection inferred under model M3. Sites predicted to be under positive selection with a probability of >0.9 were included. Those with a probability of >0.95 are indicated with an asterisk; those with a probability of >0.99 are in boldface.

### TABLE 3. Characteristics of patients and HDV quasispecies

<table>
<thead>
<tr>
<th>Patient</th>
<th>Genotype</th>
<th>HLA typing</th>
<th>Selected blood samples*</th>
<th>Total no. of clones</th>
</tr>
</thead>
<tbody>
<tr>
<td>A</td>
<td>I</td>
<td>A2, A11</td>
<td>1435, 5132</td>
<td>22</td>
</tr>
<tr>
<td>B</td>
<td>I</td>
<td>A2</td>
<td>1573, 2643, 2667, 2688, 2748, 2871, 3123, 3678</td>
<td>67</td>
</tr>
<tr>
<td>C</td>
<td>II</td>
<td>A2, A24</td>
<td>1385, 4077</td>
<td>14</td>
</tr>
<tr>
<td>D</td>
<td>IV</td>
<td>A2, A11</td>
<td>1025, 3038</td>
<td>20</td>
</tr>
<tr>
<td>E</td>
<td>IV</td>
<td>A11</td>
<td>62, 78, 338, 962, 2199, 2621</td>
<td>64</td>
</tr>
</tbody>
</table>

*The numbers indicate serial numbers of blood samples in our serum bank that have been collected and labeled consecutively.
In a cutoff level of 10^{-5} M was determined. Of the potential epitopes, only those predicted binding affinity stronger than 10^{-5} M were included. Table 4 shows the potential cytotoxic T-cell epitope from aa 30 to 47, and in five patients (100%) in a B-cell epitope from aa 2 to 7. The predicted affinity in the cytotoxic T-cell epitope from aa 43 to 51, and four (80%) patients had sites under positive selection in an epitope from aa 114 to 122. The C-terminal extension of the large HDAg gene had only one potential epitope of aa 198 to 206 predicted within genotype I. Of the HLA-A*0201 patients, all genotypes had sites within normal limits for 70 months at the last follow-up. Figure 1 demonstrates the emergence of a novel HDV dominant variant after acute exacerbation in patient A. Although this patient had active cirrhosis before the last episode of acute exacerbation, the disease course went into biochemical remission. The novel dominant HDV variant of sample 3038 had an alteration of amino acid residues at the CD4 epitopes and cryptic RNA-binding domain (aa 2 to 27), coiled-coil structure (aa 31 to 52), and helix-loop-helix (aa 108 to 135) of the RNA-binding domain (aa 97 to 146). The serum ALT level gradually became normalized and remained within normal limits for 70 months at the last follow-up. Figure 2 demonstrates the emergence of a novel HDV dominant variant after acute exacerbation of chronic hepatitis D. The novel dominant variant of sample 3038 had amino acid residues changes at B-cell epitopes and cryptic RNA-binding domain (aa 2 to 27).

**DISCUSSION**

We examined the population of HDV during chronic infection. Despite the majorities of delta antigen sites being functionally constrained and thereby under purifying selection, the analysis revealed that positive selection predominates the modeling of HDV population during intense immune responses from host. An average of 4.9% HDAg sites were shaped by a positive selection in our study. Crucially, most of these positively selected sites were located within immunogenic domains. Various immunologic mechanisms contributed to the modeling of HDV population. In our previous study, DNA-based immunization could lead the endogenous small
HDAg to generate a significantly higher anti-HDV response than the large HDAg (11). Epitope mapping further showed that the peptide from aa 174 to 195 could bind more strongly with antibodies induced by the DNA vaccine expressing the small HDAg than those induced by the DNA vaccine expressing the large HDAg. In another study, we also observed that HDV variants with residue changes within the epitope from aa 174 to 195 often emerged after severe hepatitis attacks in chronic HDV-infected patients (29). Interestingly, all patients in the present study had a positive selection occurred within the B-cell epitope. The humoral immunity associated with this epitope may have an important role in HDV chronicity.

Furthermore, several studies have suggested the importance of cellular immunity in HDV infection. High HDV viral loads were detected in human immunodeficiency virus-infected chronic hepatitis D patients whose circulating T cells numbers were suppressed (21), and the activity of HDV-induced liver disease correlated with CD4⁺ T-cell response to HDAg (16). Nisini et al. (16) also reported four epitopes from HDAg-specific CD4⁺ T-cell clones. Three of these four epitopes are positively selected in at least 60% of our patients. The epitope from aa 106 to 121 has an occurrence of positive selection in 80% of our patients. This MHC class II-restricted T-cell epitope is reported to generate via extracellular processing of HDAg by serum protease (1). This peptide may function both as a “blocking” peptide, inhibiting MHC binding and the presentation of “endogenous” HDAg epitopes, and as a kind of decoy inducing the exhaustion of HDAg106-121-specific T cells and consequently facilitating the persistence of virus-infected cells. The alterations in this epitope may confer a greater

FIG. 1. Consequence of selection of novel dominant HDV variants after flare-up of ALT levels in patient A. (A) Clinical course. The time points of blood sampling are marked (1435 and 5132). (B) Partial amino acid sequences of delta antigen obtained from different time points. The initially dominant HDV variant is shown at top of the panel. Dots indicate identical amino acid residues. The labeled CD4⁺ T is a T-cell epitope reported previously (16). The labeled CD8⁺ CTL are predicted to be MHC-I-restricted epitopes.
fitness to the viral population for escaping attacks from the host's immune system. Importantly, CD8+ T lymphocyte-related cytotoxic immunity has a pivotal role in the elimination of intracellular pathogens. In HBV and hepatitis C virus infections, multispecific CD8+ T-cell responses to HBV and hepatitis C virus are closely associated with viral clearance (8, 24). Our previous study revealed that aa 26 to 34 and aa 43 to 51 of HDAg are novel HLA-A*0201-restricted CTL epitopes in genotype I HDV (10). Using a neural network method, our analysis suggested more potential CTL epitopes in the genotype I. Furthermore, novel epitopes were proposed in genotypes II and IV. All genotypes of HLA-A*0201 HDV-infected patients had three consistent epitopes: aa 43 to 51, aa 50 to 58, and aa 114 to 122. In our previous studies, the CD8 cells showing positive responses to HDV 43-51 were also detectable in two HLA-A*0201 chronic delta hepatitis patients (patients A and B), whose serum HDV RNA levels gradually decreased and progressed into disease remission during follow-up (10). In the present study, the majorities of the HDV population in our patients were shaped by positive selection related to aa 43 to 51 and aa 114 to 122. Therefore, cytotoxic immune response is closely associated with viral clearance, as shown by animal and human studies. These potential epitopes can be confirmed in further studies and could lead to the development of therapies.

It is worth noting that the predicted ligand 29-37 (ELERDLRKV) in patient A is located at the same loci of HDV 29-37 (DLERDLRKI) tested by Huang et al. (10). We tested these two peptides using the NetMHC algorithm. The predicted binding level of Huang et al.'s HDV 29-37 was 1.8096 × 10^-5 M, which was higher than the 0.7419 × 10^-5 M of ligand 29-37 or our designated 10^-5 M. Thus, the peptide in the Huang et al. study was suggested to have a weaker binding affinity. The
predicted ligand 29-37 (ELERDLRKV) may still be a potential ligand.

A recent study inspected the nucleotide changes after 1 year of in vitro replication in cell culture (3). These authors concluded that the observed nucleotide changes were essentially neutral. By supplying a consistent HDAg-S from the integrated cDNA sequence, the nucleotide changes were diverse and did not compromise replication competence. However, HDV may encounter a much more complicated environment in the human host than in cell cultures. A heterogeneous HDV population may supply the HDAg-S. Besides, the HDAg-S may be under intensive selective pressures from the host immunities. A chronic HDV infection could lead to a more complicated relationship between host and viral genomes. Our analysis revealed that, on the contrary, the nucleotide changes in chronic HDV infection were positively selected.

Positive selection is commonplace in chronic hepatitis D during intense selection forces from host immunity. Positively selected sites identified in a longitudinal study differ from those observed in a cross-sectional one. The majority of sites under diversifying selection associate with known immunogenic domains. More cytotoxic T-cell epitopes in HDAg were predicted. aa 43 to 51, 50 to 58, and 114 to 122 of HD-Ag are conserved across different HDV genotypes. The discovery of immune epitopes associated with selection and clearance of HDV variants may help to understand the development of novel treatments or vaccines for HDV superinfection.

ACKNOWLEDGMENTS

This study was supported by grants NSC94-2314-B-010-041, NSC95-2314-B-010-011, and NSC95-2314-B-010-021-MY3 from the National Science Council of the Executive Yuan, Taiwan.

REFERENCES


Downloaded from http://jvi.asm.org on November 12, 2017 by guest.