JOURNAL OF CLINICAL MICROBIOLOGY, Nov. 2010, p. 4256–4259 Vol. 48, No. 11
0095-1137/10/$12.00 doi:10.1128/JCM.01613-10

Variability of the Polymerase Gene (NS5B) in Hepatitis C Virus-Infected Women▽

Jason T. Blackard,[1]* Gang Ma,[1] Berkeley N. Limketkai,[1] Jeffrey A. Welge,[2] Peter D. Dryer,[1] Christina M. Martin,[1] Yoichi Hiasa,[3] Lynn E. Taylor,[4] Kenneth H. Mayer,[4] Denise J. Jamieson,[5] and Kenneth E. Sherman[1]

Department of Internal Medicine, Division of Digestive Diseases, University of Cincinnati College of Medicine, Cincinnati, Ohio[1]; Departments of Psychiatry and Environmental Health, University of Cincinnati College of Medicine, Cincinnati, Ohio[2]; Department of Gastroenterology and Metabology, Ehime University Graduate School of Medicine, Ehime, Japan[3]; Miriam Hospital and Department of Medicine, Brown University, Providence, Rhode Island[4]; and Division of Reproductive Health, Centers for Disease Control and Prevention, Atlanta, Georgia[5]

Received 10 August 2010/Returned for modification 20 August 2010/Accepted 20 August 2010

There are limited data on diversity within the hepatitis C virus polymerase (NS5B). In concordance with its key functional role during the life cycle, NS5B intrapatient variability was low. Moreover, differences between NS5B nonsynonymous (*dN*) and synonymous (*dS*) mutation rates (*dN* − *dS*) were positively correlated with CD4 cell count, while nonsynonymous mutations were strongly correlated with reduced replication *in vivo*.

The NS5B protein of hepatitis C virus (HCV) is an RNA-dependent RNA polymerase that lacks a proofreading mechanism, resulting in a population of distinct but closely related viral variants, termed viral quasispecies, within an infected individual (16). The advent of potent NS5B inhibitors mandated a focus on the identification and monitoring of putative NS5B resistance mutations. HCV quasispecies diversity is an important predictor of liver disease progression as well as HCV treatment outcome (7, 13–15, 18). Most of these studies have focused on structural genomic regions; thus, limited data are available regarding nonstructural regions. To address this issue, we assessed serum NS5B variability in chronic HCV infection. (These data were presented at the 4th International Workshop on HIV and Hepatitis Co-Infection held in Madrid, Spain, in June 2008.)

From 1993 to 2000, a prospective natural history study of HIV infection, the HER Study, was conducted with U.S. women (17). Women were included in the current study if they were (i) HCV antibody positive, (ii) from the Rhode Island study site, and (iii) had available serum samples for analysis. All participants provided informed consent prior to sample and clinical data collection. RNA extractions and amplifications of a 384-bp region of NS5B were performed as described previously (4). Gel-purified PCR products were ligated into the pGEM-T Easy vector, and a median of 10 plasmids per sample was sequenced. We and others have shown that analysis of ~10 sequences accurately and reproducibly represents intrapatient quasispecies diversity (1, 20). All alignments were performed using Clustal X and compared to database references to determine the HCV genotype. Intrapatient genetic distances, Shannon entropy, and differences between nonsynonymous (*dN*) and synonymous (*dS*) rates (*dN* − *dS*) were calculated as described elsewhere (1). Codons under positive or negative selection were detected via the fixed effects likelihood (FEL) and random effects likelihood (REL) methods as implemented in the Datamonkey program (12). *P* values of <0.05 were considered statistically significant.

A randomly selected subset of genotype 1-infected women that included 25 with HIV/HCV coinfection and 4 with HCV monoinfection were analyzed. Mean aspartate transaminase (AST) and alanine aminotransferase (ALT) levels were 58.5 U/ml and 51.9 U/ml, respectively. For the HIV/HCV-coinfected women, the median CD4 cell count was 358.7 cells/ml, and 23 (92%) had detectable plasma HIV viral loads. Of 22 HIV/HCV-coinfected women with available data, 13 (59%) were receiving antiretroviral therapy at the time of serum collection. Twenty-six women were infected with HCV genotype 1a, and 3 were infected with genotype 1b. Among the partial NS5B consensus sequences generated, 21 of 114 (18.4%) amino acid positions analyzed were variable (Fig. 1). The NS5B motifs A to D were conserved at nearly all amino acid positions, including the GDD active site. S282T and C316Y resistance mutations were not observed in any serum samples, although we have previously reported the persistence of the C316Y mutation in a treatment-naïve HCV-monoinfected woman from the same cohort (4). We further evaluated multiple measures of NS5B intrapatient variability. The median intrapatient genetic distance was 0.90% (range, 0.30% to 1.50%). These values are lower than we reported previously for hypervariable region 1 (HVR1) but higher than genetic distance measures for the 5′-untranslated region (1). Median intrapatient entropy was 0.83 (range, 0.53 to 1.00). The median intrapatient *dN* − *dS* value for NS5B was −0.011 (range, −0.035 to 0.005), indicating negative or purifying immune selection pressures have preserved this functionally constrained genomic region. While it was not our intent to compare NS5B diversity in the presence/absence of HIV coinfection, we did observe a higher median genetic distance for the HIV/HCV-

* Corresponding author. Mailing address: Division of Digestive Diseases, University of Cincinnati College of Medicine, ML 0595, Cincinnati, OH 45267. Phone: (513) 558-4389. Fax: (513) 558-1744. E-mail: jason.blackard@uc.edu.
▽ Published ahead of print on 1 September 2010.

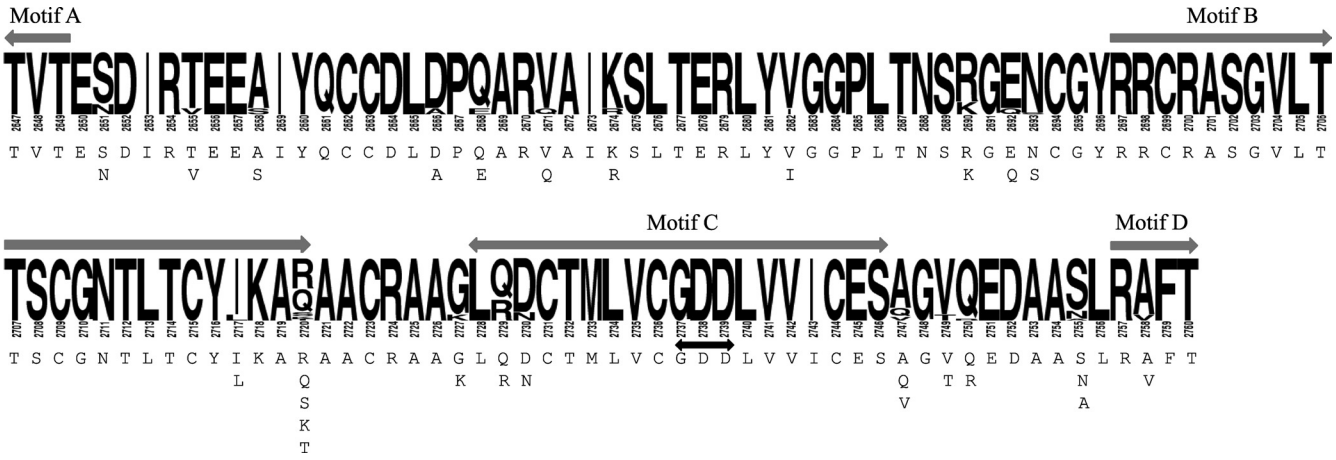

FIG. 1. Sequence logo diagram (http://weblogo.berkeley.edu) showing amino acids 227 through 340 of the HCV NS5B protein from 26 genotype 1a consensus sequences. Numbers shown below the sequence represent the amino acid position in the complete HCV polyprotein. For polymorphic sites (i.e., those with more than one letter shown per position), the height of each letter represents the relative proportion of each amino acid at that position.

coinfected women than for the HCV-monoinfected women (0.8% versus 0.4%; $P = 0.002$), even with the small population size. Site-specific codon selection was examined in each patient as summarized in Table 1. Using the FEL method, there were no positively selected amino acid sites in NS5B among the 29 women examined; however, there were 14 negatively selected amino acid sites identified in 9 women. Three of these amino acids—2730, 2747, and 2769 of the region analyzed—were identified in more than one woman. Using the REL method, 2 amino acids—2807 and 2818—were found to be under positive

TABLE 1. Codons under positive or negative selection pressure among 29 HCV-seropositive women with genotype 1 based on the FEL and REL methods

Patient no.	Codon no. based on FEL with[a]:		Codon no. based on REL with[a]:	
	Positive selection	Negative selection	Positive selection	Negative selection
62	None	None	None	2785, 2788
90	None	None	None	None
106	None	None	None	None
116	None	2771	None	None
119	None	None	None	None
124	None	None	None	None
135	None	2733, 2769	None	2718, 2722, 2733, 2737, 2764, 2769, 2790, 2792, 2813, 2817
149	None	None	None	None
184	None	None	None	None
186	None	None	None	None
199	None	None	None	None
211	None	2761	None	2719, 2743, 2750, 2754, 2761, 2769, 2776, 2781, 2786, 2795, 2798, 2810, 2815
217	None	None	None	None
220	None	None	2818	None
230	None	2715	None	None
249	None	None	None	None
269	None	None	2807	None
282	None	None	None	None
285	None	2747	None	2742, 2747, 2780, 2784
287	None	None	None	None
304	None	2717, 2738	None	None
340	None	2809	None	None
383	None	None	None	2708, 2817, 2818
389	None	None	None	None
392	None	None	None	2712, 2715, 2766, 2797, 2800, 2804, 2805, 2811, 2816
394	None	2733	None	None
504	None	2728, 2747, 2761, 2769	None	2728, 2747, 2756, 2769, 2775, 2780, 2800
514	None	None	None	None
520	None	None	None	None

[a] Codons under positive or negative selection pressure among 29 HCV-seropositive women with genotype 1. Shown are codons with a significance of $P \leq 0.05$ (FEL method) or posterior probability of ≥ 0.99 (REL method). Codons detected in more than one individual are underlined.

TABLE 2. Correlation of NS5B intrapatient diversity with clinical variables among 29 HCV-infected women

Comparison	Spearman correlation coefficient	*P* value[a]
NS5B genetic distance vs HCV RNA	0.286	NS
NS5B entropy vs HCV RNA	0.163	NS
NS5B *dN* − *dS* vs HCV RNA	−0.492	0.007
NS5B genetic distance vs CD4 cell count	−0.033	NS
NS5B entropy vs CD4 cell count	0.186	NS
NS5B *dN* − *dS* vs CD4 cell count	0.387	0.056
NS5B genetic distance vs ALT	−0.222	NS
NS5B entropy vs ALT	−0.002	NS
NS5B *dN* − *dS* vs ALT	0.199	NS
NS5B genetic distance vs AST	−0.178	NS
NS5B entropy vs AST	−0.144	NS
NS5B *dN* − *dS* vs AST	0.165	NS
NS5B genetic distance vs HIV viral load	0.122	NS
NS5B entropy vs HIV viral load	−0.079	NS
NS5B *dN* − *dS* vs HIV viral load	−0.142	NS

[a] *P* values of <0.10 are shown. NS, not significant.

selection pressure, while 48 amino acids in seven women (2 to 13 sites per individual) were under negative selection pressure. Four of these amino acids—2747, 2789, 2800, and 2817—were identified in at least two women, while amino acid 2769 was identified in three women. We also analyzed potential correlations among various clinical parameters and measurements of quasispecies diversity (Table 2). Median intrapatient NS5B *dN* − *dS* values trended toward a positive correlation with CD4 cell count (Spearman *r*, 0.387; $P = 0.056$) but were negatively correlated with HCV RNA levels (r, −0.492; $P = 0.007$). NS5B genetic distance, entropy, and *dN* − *dS* were not correlated with ALT, AST, or HIV viral load. Similar results were obtained when the study population was restricted to those with HIV/HCV coinfection.

Few data are available for nonstructural genomic regions, despite their potential as therapeutic targets to reduce viral replication. Our data support the concept that viral load is influenced by variability within the polymerase gene. This observation is supported by a number of *in vitro* studies that have identified key positions within NS5B that can dramatically affect HCV replication (9, 10). Moreover, a correlation has been demonstrated between the number of amino acid substitutions in NS5B and replicative capacity (6, 8), as well as HCV treatment outcome (5). Thus, increased mutations within NS5B may have a deleterious effect on viral replication both *in vitro* and *in vivo*. In the current study, intrapatient NS5B variability was detected, although variability was not as high as that observed in HVR1 from the same individuals (unpublished observations). The existence of multiple unique NS5B variants within an infected individual permits rapid viral adaptation to immunologic and antiviral selection pressures and the cellular microenvironment. We identified a number of amino acid positions that were under negative selection, but only two amino acids were found to be under positive selection. Moreover, among genotype 1-infected women, the median *dN* − *dS* value for NS5B was <0, suggesting that negative immune selection pressures and/or strong functional constraints on enzymatic activity limit NS5B variability compared to other more variable genomic regions. Nonetheless, intrapatient *dN* − *dS* values were positively correlated with CD4 cell count but negatively correlated with HCV RNA levels, suggesting that nonsynonymous mutations within NS5B could impact viral replication. In contrast, NS5B variability was not associated with HIV viral load, although the median NS5B genetic distance was higher for HIV/HCV-coinfected women than HCV-monoinfected women. This finding is particularly interesting given that HCV RNA levels are significantly elevated during HIV/HCV coinfection (19, 21) and HIV coinfection is associated with lower HCV treatment response (2, 11). Thus, HIV may influence NS5B variability, although this intriguing possibility requires a longitudinal analysis of NS5B diversity.

Our study has several potential limitations, including the modest population size, analysis of partial NS5B sequences, and the inclusion of relatively few HCV-monoinfected controls. The analysis included only HCV genotype 1, and the relevance of our findings to non-1 genotypes is unknown. While our data may be influenced by PCR-generated mutations, the various measures of quasispecies diversity reported here suggest that NS5B diversity was in excess of the *Taq* polymerase error rate (3). Future studies aimed at functional characterization of unique NS5B sequences identified *in vivo* will also provide data on how NS5B variability impacts replicative fitness. In summary, our data demonstrated a link between NS5B viral diversity and viral load that could impact targeted drug therapy. Longitudinal analysis of the cohort with relation to disease and treatment outcome is warranted.

Nucleotide sequence accession numbers. Consensus NS5B sequences have been submitted to GenBank under accession numbers GU131368 to GU131396.

We thank the HER Study staff and participants. The HER Study group consists of Robert S. Klein, Ellie Schoenbaum, Julia Arnsten, Robert D. Burk, Penelope Demas, and Andrea Howard from Montefiore Medical Center and the Albert Einstein College of Medicine; Paula Schuman, Jack Sobel, Suzanne Ohmit, William Brown, Michael Long, Wayne Lancaster, and Jose Vazquez from the Wayne State University School of Medicine; Anne Rompalo, David Vlahov, and David Celentano from the Johns Hopkins University School of Medicine; Charles Carpenter, Kenneth Mayer, Susan Cu-Uvin, Timothy Flanigan, Joseph Hogan, Valerie Stone, Karen Tashima, and Josiah Rich from the Brown University School of Medicine; Ann Duerr, Lytt I. Gardner, Chad Heilig, Scott D. Holmberg, Denise J. Jamieson, Janet S. Moore, Ruby M. Phelps, Dawn K. Smith, and Dora Warren from the Centers for Disease Control and Prevention; and Katherine Davenny from the National Institute on Drug Abuse.

This work was supported by the Partners/Fenway/Shattuck Center for AIDS Research (NIH P30-AI42851); the National Institute on Drug Abuse (R21 DA022148-01); the Life Span/Tufts/Brown Center for AIDS Research (NIH P30 AI42853); the Japanese Ministry of Education, Culture, Sports, Science, and Technology (JSPS KAKENHI 21590848); and the Centers for Disease Control and Prevention (U64/CCU106795 to fund data collection at Brown University).

REFERENCES

1. **Blackard, J., Y. Yang, P. Bordoni, K. Sherman, and R. Chung for the AIDS Clinical Trials Group 383 Study Team.** 2004. Hepatitis C virus (HCV) diversity in HIV-HCV coinfected subjects initiating highly active antiretroviral therapy. J. Infect. Dis. **189:**1472–1481.
2. **Chung, R., J. Andersen, P. Volberding, G. Robbins, T. Liu, K. Sherman, M. Peters, M. Koziel, A. Bhan, B. Alston, D. Colquhoun, T. Nevin, G. Harb, and C. van der Horst for the AIDS Clinical Trials Group A5071 Study Team.** 2004. Peginterferon alpha-2a plus ribavirin versus interferon alpha-2a plus ribavirin for chronic hepatitis C in HIV-coinfected persons. N. Engl. J. Med. **351:**451–459.
3. **Cline, J., J. Braman, and H. Hogrefe.** 1996. PCR fidelity of Pfu DNA

polymerase and other thermostable DNA polymerases. Nucleic Acids Res. **24:**3546–3551.
4. **Dryer, P. D., B. N. Limketkai, C. M. Martin, G. Ma, K. E. Sherman, L. E. Taylor, K. H. Mayer, D. J. Jamieson, and J. Blackard.** 2009. Screening for hepatitis C virus non-nucleotide resistance mutations in treatment naïve women. J. Antimicrob. Chemother. **64:**945–948.
5. **Hamano, K., N. Sakamoto, N. Enomoto, N. Izumi, Y. Asahina, M. Kurosaki, E. Ueda, Y. Tanabe, S. Maekawa, J. Itakura, H. Watanabe, S. Kakinuma, and M. Watanabe.** 2005. Mutations in the NS5B region of the hepatitis C virus genome correlate with clinical outcomes of interferon-alpha plus ribavirin combination therapy. J. Gastroenterol. Hepatol. **20:**1401–1409.
6. **Itakura, J., K. Nagayama, N. Enomoto, K. Hamano, N. Sakamoto, L. J. Fanning, E. Kenny-Walsh, F. Shanahan, and M. Watanabe.** 2005. Viral load change and sequential evolution of entire hepatitis C virus genome in Irish recipients of single source-contaminated anti-D immunoglobulin. J. Viral Hepat. **12:**594–603.
7. **Layden-Almer, J. E., C. Kuiken, R. M. Ribeiro, K. J. Kunstman, A. S. Perelson, T. J. Layden, and S. M. Wolinsky.** 2005. Hepatitis C virus genotype 1a NS5A pretreatment sequence variation and viral kinetics in African American and white patients. J. Infect. Dis. **192:**1078–1087.
8. **Le Pogam, S., A. Seshaadri, A. Kosaka, S. Chiu, H. Kang, S. Hu, S. Rajyaguru, J. Symons, N. Cammack, and I. Nájera.** 2008. Existence of hepatitis C virus NS5B variants naturally resistant to non-nucleoside, but not to nucleoside, polymerase inhibitors among untreated patients. J. Antimicrob. Chemother. **61:**1205–1216.
9. **Lévêque, V. J., and Q. Wang.** 2002. RNA-dependent RNA polymerase encoded by hepatitis C virus: biomedical applications. Cell. Mol. Life Sci. **59:**909–919.
10. **Lohmann, V., A. Roos, F. Körner, J. O. Koch, and R. Bartenschlager.** 2000. Biochemical and structural analysis of the NS5B RNA-dependent RNA polymerase of the hepatitis C virus. J. Viral Hepat. **7:**167–174.
11. **Perez-Olmeda, M., M. Nunez, M. Romero, J. Gonzalez, A. Castro, J. Arribas, J. Pedreira, P. Barreiro, J. Garcia-Samaniego, L. Martin-Carbonero, I. Jimenez-Nacher, and V. Soriano.** 2003. Pegylated IFN-α2b plus ribavirin as therapy for chronic hepatitis C in HIV-infected patients. AIDS **17:**1023–1028.
12. **Pond, S. L., and S. Frost.** 2005. Datamonkey: rapid detection of selective pressure on individual sites of codon alignments. Bioinformatics **21:**2531–2533.
13. **Qin, H., N. J. Shire, E. D. Keenan, S. D. Rouster, M. E. Eyster, J. J. Goedert, M. J. Koziel, and K. E. Sherman.** 2005. HCV quasispecies evolution: association with progression to end-stage liver disease in hemophiliacs infected with HCV or HCV/HIV. Blood **105:**533–541.
14. **Sherman, K. E., S. D. Rouster, S. Stanford, J. T. Blackard, N. Shire, M. Koziel, M. Peters, and R. T. Chung for the ACTG 5071 Study Team.** 2010. Hepatitis C virus (HCV) quasispecies complexity and selection in HCV/HIV coinfected subjects treated with interferon-based regimens. J. Infect. Dis. **201:**712–719.
15. **Shire, N. J., P. S. Horn, S. D. Rouster, S. Stanford, M. E. Eyster, and K. E. Sherman.** 2006. HCV kinetics, quasispecies, and clearance in treated HCV-infected and HCV/HIV-1-coinfected patients with hemophilia. Hepatology **44:**1146–1157.
16. **Simmonds, P.** 2004. Genetic diversity and evolution of hepatitis C virus: 15 years on. J. Gen. Virol. **85:**3173–3188.
17. **Smith, D., D. Warrne, D. Vlahov, P. Schuman, M. Stein, B. Greenberg, and S. Holmberg for the Human Immunodeficiency Virus Epidemiology Research Study Group.** 1997. Design and baseline participant characteristics of the Human Immunodeficiency Virus Epidemiology Research (HER) study: a prospective cohort of human immunodeficiency virus infection in US women. Am. J. Epidemiol. **146:**459–469.
18. **Sullivan, D. G., D. Bruden, H. Deubner, S. McArdle, M. Chung, C. Christensen, T. Hennessy, C. Homan, J. Williams, B. J. McMahon, and D. Gretch.** 2007. Hepatitis C virus dynamics during natural infection are associated with long-term histological outcome of chronic hepatitis C disease. J. Infect. Dis. **196:**239–248.
19. **Tedaldi, E., R. Baker, A. Moorman, C. Alzola, J. Furhrer, R. McCabe, K. Wood, and S. Holmberg for the HIV Outpatient Study (HOPS) Investigators.** 2003. Influence of coinfection with hepatitis C virus on morbidity and mortality due to human immunodeficiency virus infection in the era of highly active antiretroviral therapy. Clin. Infect. Dis. **36:**363–367.
20. **Torres-Puente, M., M. Bracho, N. Jiminez, I. Garcia-Robles, A. Moya, and F. Gonzalez-Candelas.** 2003. Sampling and repeatability in the evaluation of hepatitis C virus genetic variability. J. Gen. Virol. **84:**2343–2350.
21. **Yokozaki, S., J. Takamatsu, I. Nakano, Y. Katano, H. Toyoda, K. Hayashi, T. Hayakawa, and Y. Fukuda.** 2000. Immunological dynamics in hemophiliac patients infected with hepatitis C and human immunodeficiency virus: influence of antiretroviral therapy. Blood **96:**4293–4299.