Published in final edited form as:

J Med Virol. 2012 February ; 84(2): 242–252. doi:10.1002/jmv.22269.

Analysis of a non-structural gene reveals evidence of possible hepatitis C virus (HCV) compartmentalization

Jason T. Blackard^{1,*}, Gang Ma¹, Jeffrey A. Welge², Christina M. Martin¹, Kenneth E. Sherman¹, Lynn E. Taylor³, Kenneth H. Mayer³, and Denise J. Jamieson⁴

¹Department of Internal Medicine, Division of Digestive Diseases, University of Cincinnati College of Medicine

²Departments of Psychiatry and Environmental Health, University of Cincinnati College of Medicine

³Miriam Hospital and Department of Medicine, Brown University

⁴Division of Reproductive Health, Centers for Disease Control and Prevention

Abstract

Viral diversity is a hallmark of hepatitis C virus (HCV) infection; however, only limited data are available regarding HCV variability in extrahepatic sites, and none have systematically compared diversity in non-structural and structural genomic regions. Therefore, HCV diversity in the NS5B and envelope 1 (E1) hypervariable region 1 (HVR1) genes was evaluated in matched sera and peripheral blood mononuclear cells (PBMCs) obtained from 13 HCV-infected women. Multiple clonal sequences were compared to evaluate quasispecies diversity and viral compartmentalization in PBMCs.

Genetic distances were higher for E1/HVR1 compared to NS5B in both the sera and PBMCs ($p = 0.0511$ and $p = 0.0284$). Genetic distances were higher in serum NS5B compared to PBMC NS5B ($p = 0.0003$); however, they were not different when comparing E1/HVR1 in sera to PBMCs. By phylogenetic analysis of NS5B, evidence of possible PBMC compartmentalization was observed for 1 woman, while statistical methods were consistent with PBMC compartmentalization for 6 women. Evidence of compartmentalization within a non-structural genomic region may suggest that viral adaptation to a unique extracellular microenvironment(s) may be required for efficient replication and could contribute to HCV persistence.

Keywords

NS5B; HVR1; diversity; quasispecies; extrahepatic replication

Introduction

Globally, there are over 170 million people infected with the hepatitis C virus (HCV). HCV infection represents a major cause of chronic liver disease [Seeff, 1995]. While hepatocytes constitute the major site of viral replication, there is a broad clinical spectrum of disease and extrahepatic complications associated with HCV infection (reviewed in [Blackard et al., 2006]). One strategy that HCV may utilize to evade the host immune response in infected individuals is infection of extrahepatic cell types. A number of studies provide evidence for

*Address all correspondence and requests for reprints to: Jason Blackard, PhD, Division of Digestive Diseases, University of Cincinnati College of Medicine, ML 0595, Cincinnati, OH 45267; Tel: 513-558-4389; Fax: 513-558-1744; jason.blackard@uc.edu.

extrahepatic replication of HCV in peripheral blood mononuclear cells (PBMCs), including granulocytes, monocytes/macrophages, dendritic cells, B lymphocytes, as well as other extrahepatic tissues (reviewed in [Blackard et al., 2006]). Furthermore, extrahepatic replication has important implications for disease progression and treatment outcome. For instance, patients with detectable negative-strand HCV RNA in PBMCs have lower interferon sustained response rates compared to those without detectable negative-strand HCV RNA in PBMCs [Gong et al., 2003]. Moreover, HCV RNA, including replicative intermediate forms, can persist at very low levels in peripheral lymphoid cells for many years after spontaneous or treatment-induced resolution of chronic HCV [MacParland et al., 2009; Pham et al., 2004].

The RNA-dependent RNA polymerase (NS5B) lacks a proofreading mechanism thereby generating mutations at a rate of approximately one mutation per genome per replication cycle [Simmonds, 2004]. This results in a population of distinct but closely related viral variants, termed the *viral quasispecies*, within an infected individual. Importantly, these viral variants impact viral fitness, cell tropism, immunologic escape, and/or antiviral drug resistance. Several studies have demonstrated a non-random distribution of HCV diversity in hepatic and/or extrahepatic compartments. For instance, differences in the 5' untranslated region (UTR) and/or internal ribosome entry site (IRES) have been identified in PBMCs [Di Liberto et al., 2006; Vera-Otarola et al., 2009]. Functional studies demonstrate that these mutations also affect translational efficiency and imply that HCV has adapted to replication within PBMCs in a cell type-dependent manner [Forton et al., 2004; Laporte et al., 2003; Laporte et al., 2000; Lerat et al., 2000]. Furthermore, differences in the hypervariable region 1 (HVR1) have also been identified between serum/plasma, liver, and/or PBMC variants [Maggi et al., 1999; Maggi et al., 1997; Navas et al., 1998; Okuda et al., 1999; Ramirez et al., 2009; Shimizu et al., 1997; Zehender et al., 2005] due to variation in the viral envelope determinants of HCV entry and/or variability in the ability of the host immune response to interact with and remove particular viral variants. In contrast, while non-structural genomic regions, such as NS5B, may not directly participate in virus-host interactions at the cell surface, they can dramatically impact viral replication [Blackard et al., 2010; Lvque and Wang, 2002; Lohmann et al., 2000]. For instance, NS5A variability results in modified transcriptional activation suggesting that cellular compartmentalization of virus-host interactions may occur during chronic HCV infection [Pellerin et al., 2004]. Furthermore, genomic regions other than the viral envelope may regulate cell tropism of other viruses [Andino et al., 1999; Kauder and Racaniello, 2004; Oskarsson et al., 2007; Payne et al., 1999; Stieler et al., 2010]. Thus, it is likely that non-envelope determinants of HCV tropism and extrahepatic replication also exist. Therefore, non-structural gene (NS5B) variability in serum and PBMC samples from women with chronic HCV infection was assessed. These data were then compared to a structural gene (E1/HVR1) to assess if non-structural regions of the HCV genome may also influence cell tropism through adaptive mutation and/or cell-type specific interactions with cellular proteins.

Materials and Methods

Study population

From 1993 until 2000, a prospective natural history study of HIV infection - the HIV Epidemiologic Research (HER) Study - was conducted in US women [Smith et al., 1997]. By study design, one-half reported injection drug use (IDU), while the other half reported only sexual risk behavior for HIV. As described elsewhere, HCV serostatus was determined by Abbott HCV EIA 2.0 or Ortho HCV version 3.0 ELISA [Stover et al., 2003]. Overall, the seroprevalence of HCV was 56.5%, with rates of 48.0% and 60.8% among HIV-uninfected and HIV-infected women, respectively. In a subsequent study, HCV RNA was quantified for all HCV seropositive women [Thomas et al., 2001]. Of the 676 women tested, 520 had

detectable HCV RNA. For the current study, women were included if they were 1) HCV RNA positive, 2) from the Rhode Island site, and 3) had serum and PBMC samples available for additional analysis. All participants provided informed consent prior to sample and clinical data collection.

PCR amplifications

Serum and PBMC samples were processed and stored in accordance with standard laboratory procedures to ensure viral preservation. Briefly, RNA was extracted from 140 μ L of serum using the QIAamp Viral RNA Kit (Qiagen; Valencia, CA), and eluted in 60 μ L of DEPC-treated dH₂O. For PBMC samples, the number of cells available varied from $\sim 1\text{--}8 \times 10^6$ cells/mL. 500 μ L of cells were washed with DEPC-treated dH₂O and cellular RNA was extracted with TRIzol. The resulting RNA was resuspended in 40 μ L of DEPC-treated dH₂O and treated two times with DNase I (Ambion; Austin, TX).

A 384 base pair region of NS5B was amplified by hemi-nested reverse transcriptase polymerase chain reaction (RT-PCR) as described elsewhere [Laperche et al., 2005]. First round primers were P4 (5' – GCN GAR TAY CTV GTC ATA GCC TC – 3'; nucleotides 8644 – 8622 of H77) and P3 (5' – TAT GAY ACC CGC TGY TTT GAC TC – 3'; nucleotides 8256 – 8278); second round primers were P3 and P5 (5' – GCT AGT CAT AGC CTC CGT – 3'; nucleotides 8636 – 8619). cDNA corresponding to E1/HVR1 was generated using the antisense primer 5'-AGG CTT TCA TTG CAG TTC AAG GCC TTG CTA TTG ATG TGC C-3' (nucleotides 1639-1600) as described elsewhere [Blackard et al., 2004]. First round PCR amplifications using the sense primer 5'-GCG TCC GGG TTC TGG AAG ACG GCG TGA ACT ATG CAA CAG G-3' (nucleotides 802–841) were performed as follows: 94°C for 2 minutes, followed by 35 cycles at 94°C for 1 minute, 65°C for 2 minutes, and 72°C for 2 minutes, with a final elongation step at 72°C for 8 minutes. Second round PCR conditions were the same as first round conditions and utilized the antisense primer 5'-AGT TCA AGG CCG TGC TAT TGA TGT GCC AAC TGC CGT TGG T-3' (nucleotides 1626-1587) and the sense primer 5'-GGC ATG GGA TAT GAT GAT GAA CTG GTC CCC TAC-3' (nucleotides 1295–1327). These PCR primers have been utilized extensively to amplify HCV genotypes 1–4 and do not display appreciable mismatches with reference sequences representing these genotypes. To avoid potential cross-contamination, serum and PBMC samples from the same individual were handled separately. Additionally, all RT-PCR amplifications included one reaction containing no reverse transcriptase and another containing no template as negative controls. PCR products were gel purified and ligated into the pGEM-T Easy Vector (Promega). Plasmids were propagated and purified prior to sequencing using dye terminator chemistry. A median of 10 plasmids per sample type per woman was sequenced and edited using CodonCode Aligner (CodonCode Corporation, Dedham, MA). Torres-Puente *et al.* have previously shown that analysis of ~ 10 sequences accurately and reproducibly represents the overall HCV quasispecies diversity within an individual [Torres-Puente et al., 2003], while multiple amplifications of the same template yield similar estimates of quasispecies diversity [Blackard et al., 2004].

Phylogenetic analyses and quantification of quasispecies diversity

All alignments were performed using Clustal X. References used to confirm genotype included D10749 (1a), M62321 (1a), AF511950 (1a), AF177040 (H77; 1a), D90208 (1b), AY460204 (1b), D14853 (1c), AY051292 (1c), AF238481 (2a), AB047639 (2a), AB030907 (2b), AY232746 (2b), D50409 (2c), AF046866 (3a), D17763 (3a), D28917 (3a), D49374 (3b), Y11604 (4a), DQ516084 (4a), FJ025854 (4b), Y13184 (5a), AF064490 (5a), Y12083 (6a), and D84262 (6b). The statistical robustness and reliability of the branching order within each phylogenetic tree were confirmed by bootstrap analysis using 1000 replicates

[Felsenstein, 1985]. Consensus sequences have been submitted to GenBank under accession numbers GU131370, GU131371, GU131375, GU131379, GU131382, GU131384, GU131385, GU131391, and GU580965 – GU580996.

Inpatient genetic distances were calculated by pairwise comparison of nucleotide sequences using the Kimura method of MegAlign (DNASTAR, Inc., Madison, WI). Shannon entropy ($S_n = -\sum(p_i \ln p_i) / \ln N$), where p_i is the frequency of each distinct amino acid sequence, and N is the total number of sequences analyzed, was also calculated. Non-synonymous (dN)-synonymous (dS) values were calculated via the Nei-Gojobori method in MEGA. Individual 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 [Pond and Frost, 2005].

Evaluation of Compartmentalization

Compartmentalization of HCV variants was assessed using Mantel's test as previously described to explore HCV compartmentalization [Afonso et al., 1999; Ducoulombier et al., 2004; Ramirez et al., 2009; Roque-Afonso et al., 2005; Schramm et al., 2008; Sobesky et al., 2007; Zehender et al., 2005]. Briefly, the Kimura two-parameter distance matrix that included all serum and PBMC variants for an individual was compared to a similar matrix in which distances are replaced with 0 if the sequences are from the same compartment and with 1 if the sequences are from distinct compartments. The correlation coefficient was computed for the simple Mantel's test using the Permute! Software version 3.4a9 with 9999 permutations. Compartmentalization was also assessed using the Slatkin-Maddison (S-M) test [Slatkin and Maddison, 1989] for population gene flow as implemented in the HyPhy program [Pond et al., 2005] using 10,000 permutations. P values < 0.05 were considered statistically significant evidence that viral sequences from PBMCs were compartmentalized compared to those from the corresponding serum.

Signature amino acid patterns that distinguish PBMC viral variants from those in the corresponding serum sample were identified using the Viral Epidemiology Signature Pattern Analysis program and a threshold of 70% [Korber and Myers, 1992].

Statistical Analyses

Paired-samples *t* tests were used to compare inpatient variability in serum versus PBMC compartments or non-structural versus structural genes. All P values are two-tailed; only P values < 0.05 were considered statistically significant. All analyses were performed using SAS software (SAS Institute).

Results

Patient characteristics

Thirteen HCV RNA positive, treatment-naïve women from the Providence, Rhode Island site of the HER Study with available serum and PBMC samples were included in the current analysis (Table 1). The mean age at enrollment was 33.6 years. Eight (61.5%) were Caucasian, and injection drug use was the major HIV risk factor for 10 (76.9%). Mean AST and ALT levels were 61.4 U/mL and 43.7 U/mL, respectively. Twelve (92.3%) women were co-infected with HIV; none were on antiretroviral therapy at the time of sample collection. The median CD4 cell count was 380.1 cells/mL, and the median plasma HIV viral load was 3.6 log₁₀ copies/mL (lower limit of detection = 50 copies/mL).

Interpatient variability

Evaluation of E1/HVR1 is frequently utilized to evaluate HCV diversity; however, very limited data on non-structural regions of viral genome are currently available. Thus, it is not known if other, non-structural genomic regions may also influence viral compartmentalization. In the current study, NS5B could not be amplified from the sera of 2 women, while E1/HVR1 could not be amplified from the PBMCs of 4 women. Thus, NS5B PCR products from 11 women were available for comparative analysis in both compartments, while 9 women had E1/HVR1 PCR products from both compartments. As shown in Figure 1, HCV genotypes were similar for both regions and included genotypes 1 (n = 8), 2 (n = 1), and 4 (n = 2). Consensus sequences from different women were distinct from one another, suggesting no epidemiologically-linked infections. Furthermore, no mixed genotype infections were observed, and clonal sequences clustered separately by individual (data not shown). At the nucleotide level, consensus serum-PBMC sequences were identical in 4 women for each region (36.4% for NS5B and 44.4% for E1/HVR1). At the amino acid level, consensus serum-PBMC sequences were identical for 10 (90.9%) and 5 (55.6%) women for NS5B and E1/HVR1, respectively. Among the partial NS5B consensus sequences generated, the functional motifs AD were conserved at the majority of amino acid positions, including the polymerase GDD active site [Qin et al., 2001]. The S282T and C316Y antiviral resistance mutations were not observed in any serum or PBMC samples, although persistence of the C316Y mutation in a treatment naïve woman from the same cohort was previously reported [Dryer et al., 2009].

Intrapatient variability

As consensus sequence analysis may not completely portray the extent of quasispecies diversity within an individual, additional measures of intrapatient variability were also evaluated. As shown in Table 2, median intrapatient genetic distances in the serum were 0.90% (range: 0.30% – 1.20%) for NS5B and 1.52% (range: 0.39% – 5.29%) for E1/HVR1 ($p = 0.0511$). Median genetic distances in the PBMCs were 0.30% (range: 0.00% – 1.20%) for NS5B and 1.59% (range: 0.00% – 4.66%) for E1/HVR1 ($p = 0.0284$). Genetic distances were significantly lower in the PBMC compartment compared to the serum for NS5B ($p = 0.0003$), but not for E1/HVR1. Median intrapatient entropy – which takes into account the relative frequency of distinct variants within the quasispecies [Wolinsky et al., 1996] – was 0.93 (range: 0.53 – 1.00) in the serum for NS5B and 0.88 (range: 0.48 – 1.00) for E1/HVR1. Entropy in PBMCs was 0.53 (range: 0.28 – 1.00) for NS5B and 0.79 (range: 0.14 – 1.00) for E1/HVR1. Entropy was significantly lower in the PBMC compartment compared to the serum for NS5B ($p = 0.0144$), but not for E1/HVR1.

When measuring selection pressure, dN-dS values > 0 are consistent with positive immune selection pressure, while dN-dS values < 0 indicate purifying selection that preserves functionally constrained genomic regions. Two women – 230 and 289 – had dN-dS_{NS5B} values > 0 in the PBMCs, while no women had dN-dS_{NS5B} values > 0 in sera. The dN-dS_{E1/HVR1} values were > 0 for three women – 106, 116, and 282 – in both the serum and PBMCs. Median intrapatient dN-dS_{NS5B} and dN-dS_{E1/HVR1} values were -0.013 (range: -0.032 – -0.004) and -0.0028 (range: -0.035 – 0.015) in the serum ($p = 0.084$). Median dN-dS_{NS5B} and dN-dS_{E1/HVR1} values were -0.0030 (range: -0.029 – 0.0030) and -0.0056 (range: -0.052 – 0.047) in PBMCs ($p = 0.7771$). Median dN-dS_{NS5B} values were significantly different in the PBMC compartment compared to the serum ($p = 0.0153$), although median dN-dS_{E1/HVR1} values were not significantly different in these compartments ($p = 0.8811$).

Codon selection

Site-specific codon selection was also examined in each patient dataset. Using the FEL method, there were no positively selected amino acids in serum-derived or PBMC-derived NS5B sequences among the women examined (Table 3). There were 4 (1 per woman) and 4 (1 per woman) negatively selected amino acids identified in serum-derived and PBMC-derived NS5B sequences, respectively. By the REL method, there were 7 positively selected amino acids (0 – 6 sites per woman) in serum-derived NS5B sequences but no positively selected amino acids in PBMC-derived NS5B sequences. There were 13 (0 – 8 sites per woman) and 10 (0 – 8 sites per woman) negatively selected amino acids identified in serum-derived and PBMC-derived NS5B sequences, respectively. Two of these amino acids were identified in more than one woman.

By FEL, there were no positively selected amino acids in serum-derived E1/HVR1 sequences and only one positively selected amino acid in PBMC-derived E1/HVR1 sequences among the women examined (Table 4). In contrast, there were 12 (0 – 4 sites per woman) and 8 (0 – 3 sites per woman) negatively selected amino acids identified in serum-derived and PBMC-derived E1/HVR1 sequences, respectively. Two of these amino acids – 370 and 410 of the region analyzed – were identified in more than one woman. Using the REL method, there were 33 positively selected amino acids (0 – 11 sites per woman) in serum-derived E1/HVR1 sequences and 23 positively selected amino acids (0 – 15 sites per woman) in PBMC-derived E1/HVR1 sequences among the women examined. Fifteen of these amino acids were identified in more than one woman. In contrast, there were 7 (0 – 5 sites per woman) and 15 (0 – 9 sites per woman) negatively selected amino acids identified in serum-derived and PBMC-derived E1/HVR1 sequences, respectively. Two of these amino acids – 371 and 410 of the region analyzed – were identified in more than one woman.

Analysis of HCV compartmentalization

HCV compartmentalization in PBMCs has not been explored previously using non-structural viral sequence data. By phylogenetic comparison of viral variants from the serum and PBMC compartments, PBMC-specific variants of NS5B were identified in subject 336 (genotype 2b) who had one PBMC-specific population consisting of 2 variants (Supplementary Figure 1a). The potential compartmentalization of HCV in PBMCs was explored using Mantel's test and the Slatkin-Maddison (S-M) test, which do not rely on interpretation of a phylogenetic tree (Table 5). When comparing matched plasma- and PBMC-derived NS5B variants, Mantel's test was consistent with significant quasispecies compartmentalization for subjects 289 ($p = 0.0001$), 336 ($p = 0.030$), and 389 ($p = 0.0017$), while subject 116 ($p = 0.055$) also showed a trend towards NS5B compartmentalization. The S-M test similarly demonstrated statistically significant evidence of NS5B compartmentalization between plasma and PBMC variants for subjects 116 ($p = 0.0044$), 211 ($p = 0.0406$), 230 ($p = 0.0007$), 336 ($p = 0.0255$), and 389 ($p = 0.0004$), as well as a trend towards compartmentalization for subject 106 ($p = 0.061$).

In the corresponding structural gene sequences, PBMC-specific variants of E1/HVR1 were identified in 3 of 9 (33.3%) women by phylogenetic analysis (Supplementary Figure 1b). By Mantel's test, no evidence of significant compartmentalization of E1/HVR1 was observed for any woman, although S-M test results were statistically significant for E1/HVR1 compartmentalization between plasma and PBMC variants for subjects 269 ($p = 0.0358$) and 389 ($p = 0.0011$). Thus, at least one method demonstrated compartmentalization in NS5B for 6 of 11 (54.5%) subjects and in E1/HVR1 for 2 of 9 (22.2%) subjects.

Signature sequence analysis of PBMC-specific amino acids

To identify specific amino acids associated with compartmentalization in PBMCs, signature sequence analysis was performed using NS5B and E1/HVR1 sequences for matched serum-PBMC samples. Only a single amino acid – identified in subject 389 – distinguished NS5B variants in PBMCs from those in the corresponding serum. For the other 10 women, no other PBMC-specific signature amino acids were identified in the region of NS5B analyzed. For 3 of 9 women (106, 137, and 230), PBMC-specific signature amino acids were detected E1/HVR1 (data not shown). Among these subjects, 4 PBMC-specific signature amino acids were identified.

Discussion

HCV quasispecies diversity is an important predictor of liver disease progression [Arenas et al., 2004; Lyra et al., 2002; Qin et al., 2005; Sullivan et al., 2007] and HCV treatment outcome [Grahovac et al., 2000; Layden-Almer et al., 2005; Pawlotsky, 2000; Sherman et al., 2010]. By definition, viral compartments are characterized by a restriction of viral gene flow between cells or tissues [Nickle et al., 2003] and may occur due to physical isolation of a particular tissue/cell type, selective migration of susceptible cells, different determinants of virulence or viral tropism depending upon the cell type (i.e., distinct levels of co-factors required for replication), and/or localized selection pressures. Importantly, compartmentalization implies that viral adaptation is frequently necessary for efficient infection of and replication within a particular cell/tissue type. Similarly, a non-random distribution of HCV sequences in hepatic and extrahepatic compartments suggests that compartmentalization also occurs during HCV infection [Maggi et al., 1999; Maggi et al., 1997; Roque-Afonso et al., 2005; Roque-Afonso et al., 1999]. For instance, Ducoulombier *et al.* examined compartmentalization in the plasma and multiple peripheral blood mononuclear cell subsets and found that HVR1 sequences were distinct between plasma and at least one cellular compartment in all individuals examined [Ducoulombier et al., 2004]. While multiple studies of chronic HCV infection have demonstrated compartmentalization of the viral quasispecies in extrahepatic cell/tissue types, these studies are limited by the infrequent utilization of appropriate statistical methods, the rare inclusion of HIV/HCV co-infected individuals, and the lack of data on non-structural genomic regions such as NS5B. As a key component of the viral replication machinery that interacts with a variety of cellular co-factors that may not be present – or may differ in absolute level – in distinct cell types, it is reasonable to explore the potential association between non-structural genes and HCV compartmentalization.

As hepatocytes clearly represent the major site of HCV replication, it could be argued that extrahepatic replication of HCV may represent an artifact of selective absorption of virions to the cell surface. Nonetheless, several findings here and elsewhere suggest that this is an unlikely explanation. First, a previous study from the same cohort found that serum- and PBMC-specific HCV RNA levels did not correlate with one another, suggesting that replication may be regulated differently in these compartments [Blackard et al., 2005]. Second, E1/HVR1 compartmentalization was observed in a preliminary analysis of women with detectable negative-strand HCV RNA – that serves as a replication intermediate but is absent from virions – in their PBMCs [Blackard et al., 2007]. Third, in the current study, various measures of E1/HVR1 quasispecies diversity were similar between the serum and PBMC compartments, although all were significantly different between these two compartments for NS5B. While it has been suggested that HCV detection in PBMCs could be due to absorption of circulating viruses onto cell surfaces [Meier et al., 2001; Natarajan et al., 2010], the data presented here strongly suggest that selective absorption of minor viral variants is not a large factor contributing to HCV NS5B compartmentalization since NS5B would not be expected to play a role in initial virus attachment or entry. However, the

possibility that an altered HVR1 sequence permits viral entry into PBMCs but without subsequent viral replication or production of infectious virions cannot be ruled out definitively. Nonetheless, this phenomenon would not explain sequence differences between serum- and PBMC-derived variants in a non-structural genomic region such as NS5B.

Several novel findings can be derived from the current study. First, inpatient genetic diversity was lower for NS5B compared to HVR1. Given its key functional role in the viral life cycle, this is not surprising, particularly as data from this cohort suggest that non-synonymous mutations within NS5B are infrequent but strongly correlated with reduced viral replication *in vivo* [Blackard et al., 2010]. Second, phylogenetic evidence of PBMC compartmentalization was observed for 1 woman in NS5B and for 3 women in E1/HVR1, while statistical measures demonstrated significant compartmentalization of NS5B and of E1/HVR1 for 6 and 5 women, respectively. This disparity between phylogenetic and statistical analyses has been noted previously in the HIV literature in which multiple methods for detecting viral compartmentalization have been utilized extensively. Thus, utilizing several complementary strategies – as performed here – provides the most reliable assessment of viral compartmentalization [Zárate et al., 2007]. Third, while there are considerable data on E1/HVR1 inpatient variability, there are only extremely limited data on inpatient NS5B variability. Thus, for the first time, the data presented in the current study provide the ability to compare NS5B variability in the serum and PBMCs directly. Nonetheless, the codon selection and signature sequence analyses failed to identify a common pattern across all individuals studied, suggesting that viral compartmentalization may be partially host-dependent.

There is accumulating evidence that HCV possesses extensive RNA secondary structure within the NS5B-encoding region that plays an important role in viral replication [Tuplin et al., 2004]. For example, one predicted RNA structure within NS5B has an essential *cis*-acting role in replication of the HCV replicon *in vitro* [You et al., 2004]. Furthermore, long-range RNA-RNA interactions between NS5B and the 5' end of the HCV genome have been identified [Romero-López and Berzal-Herranz, 2011], and minor nucleotide changes within an otherwise considered genomic region that do not alter the amino acid coding sequence may still impact HCV replication. This hypothesis warrants further investigation of additional genomic regions and a larger population size.

Several limitations should also be noted, including the small population size, the possible lack of generalizability across HCV genotypes, no data on the time of HCV infection, and the inability to define more precisely the infected cell population(s) given the limited number of PBMCs available for additional analysis. Moreover, by analyzing only a small portion of the viral genome and by sequencing a limited number of clones, the approach utilized may underestimate the true extent of viral compartmentalization. Nonetheless, previous data suggest that sequencing 10 viral variants is sufficient to evaluate the viral quasispecies [Torres-Puente et al., 2003]. While both positive- and negative-sense HCV RNA can be detected in PBMCs, particularly from HIV/HCV co-infected women in this cohort [Blackard et al., 2005], HCV RNA was not quantified for all study visits from all women, thereby limiting the ability to assess how HCV RNA levels in the PBMC compartment may impact the current findings. Finally, the consequences of the positively or negatively selected amino acids reported in Table 4 on NS5B function are unknown; however, the GDD catalytic site and several functional motifs within the region of NS5B analyzed were well conserved.

In conclusion, evidence of possible compartmentalization in PBMCs is supported further by a previous study demonstrating that short-term propagation of HCV in PBMCs *in vitro* can select for a restricted viral population that is present in PBMCs *in vivo* [Maggi et al., 1999].

However, NS5B sequences were not analyzed; thus, there are no data comparing selection of NS5B variants that replicate efficiently *in vitro* with those that are detected *in vivo*. Nonetheless, identification and characterization of the viral and host factors that influence HCV evolution is important for the development of effective strategies to control replication and limit disease progression. A number of *in vitro* studies have identified key positions within NS5B that can dramatically affect HCV replication [Lévêque and Wang, 2002; Lohmann et al., 2000]. Moreover, NS5B variability can impact replicative capacity, as well as the outcome of HCV treatment [Hamano et al., 2005; Itakura et al., 2005; Le Pogam et al., 2008]. Thus, adaptive evolution of non-structural genomic regions such as NS5B may permit viral persistence in a unique cellular microenvironment. Further studies examining the replicative capacity of NS5B variants in extrahepatic cell types compared to hepatocytes are warranted to understand fully the complex interplay between viral diversity and the cell type-specific environment in which the virus replicates.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgments

The authors would like to thank the HER Study staff and participants. The HER Study group consists of Robert S. Klein, MD, Ellie Schoenbaum, MD, Julia Arnsten, MD, MPH, Robert D. Burk, MD, Penelope Demas, PhD, and Andrea Howard, MD, MSc, from Montefiore Medical Center and the Albert Einstein College of Medicine; Paula Schuman, MD, Jack Sobel, MD, Suzanne Ohmit, DrPH, William Brown PhD, Michael Long PhD, Wayne Lancaster PhD, and Jose Vazquez, MD, from the Wayne State University School of Medicine; Anne Rompalo, MD, David Vlahov, PhD and David Celentano, PhD, from the Johns Hopkins University School of Medicine; Charles Carpenter, MD, Kenneth Mayer, MD, Susan Cu-Uvin, MD, Timothy Flanigan, MD, Joseph Hogan, ScD, Valerie Stone, MD, Karen Tashima, MD, and Josiah Rich, MD from the Brown University School of Medicine; Ann Duerr, MD, PhD, Lytt I. Gardner, PhD, Chad Heilig, PhD, Scott D. Holmberg, MD, Denise J. Jamieson, MD, MPH, Janet S. Moore, PhD, Ruby M. Phelps, BS, Dawn K. Smith, MD, MPH and Dora Warren, PhD from the Centers for Disease Control and Prevention; and Katherine Davenny, MPH from the National Institute of Drug Abuse. The findings and conclusions in this report are those of the authors and do not necessarily represent the official position of the Centers for Disease Control and Prevention.

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 Lifespan/Tufts/Brown Center for AIDS Research (NIH P30 AI42853); and the Centers for Disease Control and Prevention (U64/CCU106795 to fund data collection at Brown University). These data were presented at the 5th International Workshop on HIV and Hepatitis Co-Infection held in Lisbon, Portugal in June 2009.

References

- Afonso A, Jian J, Penin F, Tareau C, Samuel D, Petit M, Bismuth H, Dussaix E, Feray C. Nonrandom distribution of hepatitis C virus quasiespecies in plasma and peripheral blood mononuclear cell subsets. *Journal of Virology*. 1999; 73:9213–9221. [PubMed: 10516029]
- Andino R, Böddicker N, Silvera D, Gamarnik A. Intracellular determinants of picornavirus replication. *Trends in Microbiology*. 1999; 7:76–82. [PubMed: 10081085]
- Arenas J, Gallegos-Orozco J, Laskus T, Wilkinson J, Khatib A, Fasola C, Adair D, Radkowski M, Kibler K, Nowicki M, Douglas D, Williams J, Netto G, Mulligan D, Klintmalm G, Rakela J, Vargas H. Hepatitis C virus quasi-species dynamics predict progression of fibrosis after liver transplantation. *Journal of Infectious Diseases*. 2004; 189:2037–2046. [PubMed: 15143471]
- Blackard JT, Hiasa Y, Smeaton L, Jamieson DJ, Rodriguez I, Mayer KH, Chung R. Compartmentalization of hepatitis C virus (HCV) during HCV/HIV coinfection. *Journal of Infectious Diseases*. 2007; 195:1765–1773. [PubMed: 17492592]
- Blackard JT, Kemmer N, Sherman K. Extrahepatic replication of HCV: insights into clinical manifestations and biological consequences. *Hepatology*. 2006; 44:15–22. [PubMed: 16799966]

- Blackard JT, Ma G, Limketkai BN, Welge JA, Dryer PD, Martin CM, Hiasa Y, Taylor LE, Mayer KH, Jamieson DJ, Sherman K. Variability of the polymerase gene (NS5B) in hepatitis C virus-infected women. *Journal of Clinical Microbiology*. 2010; 48:4256–4259. [PubMed: 20810773]
- Blackard JT, Smeaton L, Hiasa Y, Horiike N, Onji M, Jamieson DJ, Rodriguez I, Mayer KH, Chung R. Detection of hepatitis C virus (HCV) in serum and peripheral blood mononuclear cells of HCV-monoinfected and HIV/HCV-coinfected persons. *Journal of Infectious Diseases*. 2005; 192:258–265. [PubMed: 15962220]
- Blackard JT, Yang Y, Bordon P, Sherman KE, Chung R. for the AIDS Clinical Trials Group 383 Study Team. Hepatitis C virus (HCV) diversity in HIV-HCV coinfecting subjects initiating highly active antiretroviral therapy. *Journal of Infectious Diseases*. 2004; 189:1472–1481. [PubMed: 15073685]
- Di Liberto G, Roque-Afonso AM, Kara R, Ducoulombier D, Fallot G, Samuel D, Feray C. Clinical and therapeutic implications of hepatitis C virus compartmentalization. *Gastroenterology*. 2006; 131:76–84. [PubMed: 16831592]
- Dryer PD, Limketkai BN, Martin CM, Ma G, Sherman KE, Taylor LE, Mayer KH, Jamieson DJ, Blackard J. Screening for hepatitis C virus non-nucleotide resistance mutations in treatment naïve women. *Journal of Antimicrobial Chemotherapy*. 2009; 64:945–948. [PubMed: 19767319]
- Ducoulombier D, Roque-Afonso AM, Di Liberto G, Penin F, Kara R, Richard Y, Dussaix E, Feray C. Frequent compartmentalization of hepatitis C virus variants in circulating B cells and monocytes. *Hepatology*. 2004; 39:817–825. [PubMed: 14999702]
- Felsenstein J. Confidence limits on phylogenies: an approach using the bootstrap. *Evolution*. 1985; 39:783–791.
- Forton DM, Karayiannis P, Mahmud N, Taylor-Robinson SD, Thomas H. Identification of unique hepatitis C virus quasispecies in the central nervous system and comparative analysis of internal translational efficiency of brain, liver, and serum variants. *Journal of Virology*. 2004; 78:5170–5183. [PubMed: 15113899]
- Gong G, Lai L, Jiang Y, He Y, Su X. HCV replication in PBMC and its influence on interferon therapy. *World Journal of Gastroenterology*. 2003; 9:291–294. [PubMed: 12532451]
- Grahovac B, Bingulac-Popovic J, Vucelic B, Hrstic I, Ostojic R, Drazic V, Balija M, Grgicevic D. Hypervariable region 1 of hepatitis C virus genome and response to interferon therapy. *Clinical Chemistry and Laboratory Medicine*. 2000; 38:905–910. [PubMed: 11097348]
- Hamano K, Sakamoto N, Enomoto N, Izumi N, Asahina Y, Kurosaki M, Ueda E, Tanabe Y, Maekawa S, Itakura J, Watanabe H, Kakinuma S, Watanabe M. Mutations in the NS5B region of the hepatitis C virus genome correlate with clinical outcomes of interferon-alpha plus ribavirin combination therapy. *Journal of Gastroenterology and Hepatology*. 2005; 20:1401–1409. [PubMed: 16105128]
- Itakura J, Nagayama K, Enomoto N, Hamano K, Sakamoto N, Fanning LJ, Kenny-Walsh E, Shanahan F, Watanabe M. Viral load change and sequential evolution of entire hepatitis C virus genome in Irish recipients of single source-contaminated anti-D immunoglobulin. *Journal of Viral Hepatitis*. 2005; 12:594–603. [PubMed: 16255760]
- Kauder SE, Racaniello V. Poliovirus tropism and attenuation are determined after internal ribosome entry. *Journal of Clinical Investigation*. 2004; 113:1743–1753. [PubMed: 15199409]
- Korber B, Myers G. Signature pattern analysis: a method for assessing viral sequence relatedness. *AIDS Research and Human Retroviruses*. 1992; 8:1549–1560. [PubMed: 1457200]
- Laperche S, Lunel F, Izopet J, Alain S, Dény P, Duverlie G, Gaudy C, Pawlotsky JM, Plantier JC, Pozzetto B, Thibault V, Tosetti F, Lefrère J. Comparison of hepatitis C virus NS5b and 5' noncoding gene sequencing methods in a multicenter study. *Journal of Clinical Microbiology*. 2005; 43:733–739. [PubMed: 15695672]
- Laporte J, Bain C, Maurel P, Inchauspe G, Agut H, Cahour A. Differential distribution and internal translation efficiency of hepatitis C virus quasispecies present in dendritic and liver cells. *Blood*. 2003; 101:52–57. [PubMed: 12393733]
- Laporte J, Malet I, Andrieu T, Thibault V, Toulme J, Wychowski C, Pawlotsky J, Huraux J, Agut H, Cahour A. Comparative analysis of translation efficiencies of hepatitis C virus 5' untranslated

- regions among intraindividual quasispecies present in chronic infection: opposite behaviors depending on cell type. *Journal of Virology*. 2000; 74:10827–10833. [PubMed: 11044132]
- Layden-Almer JE, Kuiken C, Ribeiro RM, Kunstman KJ, Perelson AS, Layden TJ, Wolinsky SM. Hepatitis C virus genotype 1a NS5A pretreatment sequence variation and viral kinetics in African American and white patients. *Journal of Infectious Diseases*. 2005; 192:1078–1087. [PubMed: 16107963]
- Le Pogam S, Seshadri A, Kosaka A, Chiu S, Kang H, Hu S, Rajyaguru S, Symons J, Cammack N, Nájera I. Existence of hepatitis C virus NS5B variants naturally resistant to non-nucleoside, but not to nucleoside, polymerase inhibitors among untreated patients. *Journal of Antimicrobial Chemotherapy*. 2008; 61:1205–1216. [PubMed: 18343801]
- Lerat H, Shimizu Y, Lemon S. Cell type-specific enhancement of hepatitis C virus internal ribosome entry site-directed translation due to 5' nontranslated region substitutions selected during passage of virus in lymphoblastoid cells. *Journal of Virology*. 2000; 74:7024–7031. [PubMed: 10888641]
- Lévêque VJ, Wang Q. RNA-dependent RNA polymerase encoded by hepatitis C virus: biomedical applications. *Cellular and Molecular Life Sciences*. 2002; 59:909–919. [PubMed: 12169021]
- Lohmann V, Roos A, Körner F, Koch JO, RRB. Biochemical and structural analysis of the NS5B RNA-dependent RNA polymerase of the hepatitis C virus. *Journal of Viral Hepatitis*. 2000; 7:167–174. [PubMed: 10849258]
- Lyra AC, Fan X, Lang DM, Yusim K, Ramrakhiani S, Brunt EM, Korber B, Perelson AS, Bisceglie AD. Evolution of hepatitis C viral quasispecies after liver transplantation. *Gastroenterology*. 2002; 123:1485–1493. [PubMed: 12404223]
- MacParland SA, Pham TN, Guy CS, Michalak T. Hepatitis C virus persisting after clinically apparent sustained virological response to antiviral therapy retains infectivity in vitro. *Hepatology*. 2009; 49:1431–1441. [PubMed: 19177592]
- Maggi F, Fornai C, Morrica A, Vatteroni M, Giorgi M, Marchi S, Ciccorossi P, Bendinelli M, Pistello M. Divergent evolution of hepatitis C virus in liver and peripheral blood mononuclear cells of infected patients. *Journal of Medical Virology*. 1999; 57:57–63. [PubMed: 9890422]
- Maggi F, Fornai C, Vatteroni M, Giorgi M, Morrica A, Pistello M, Cammarota G, Marchi S, Ciccorossi P, Bionda A, Bendinelli M. Differences in hepatitis C virus quasispecies composition between liver, peripheral blood mononuclear cells and plasma. *Journal of General Virology*. 1997; 78:1521–1525. [PubMed: 9225024]
- Meier V, Mihm S, Braun Wietzke P, Ramadori G. HCV-RNA positivity in peripheral blood mononuclear cells of patients with chronic HCV infection: does it really mean viral replication? *World Journal of Gastroenterology*. 2001; 7:228–234. [PubMed: 11819765]
- Natarajan V, Kottitil S, Hazen A, Adelsberger J, Murphy AA, Polis MA, Kovacs J. HCV in peripheral blood mononuclear cells are predominantly carried on the surface of cells in HIV/HCV co-infected individuals. *Journal of Medical Virology*. 2010; 82:2032–2037. [PubMed: 20981790]
- Navas S, Martin J, Quiroga JA, Castillo I, Carreno V. Genetic diversity and tissue compartmentalization of the hepatitis C virus genome in blood mononuclear cells, liver, and serum from chronic hepatitis C patients. *Journal of Virology*. 1998; 72:1640–1646. [PubMed: 9445070]
- Nickle DC, Shriner D, Mittler JE, Frenkel LM, Mullins J. Importance and detection of virus reservoirs and compartments of HIV infection. *Current Opinions in Microbiology*. 2003; 6:410–416.
- Okuda M, Hino K, Korenaga M, Yamaguchi Y, Katoh Y, Okita K. Differences in hypervariable region 1 quasispecies of hepatitis C virus in human serum, peripheral blood mononuclear cells, and liver. *Hepatology*. 1999; 29:217–222. [PubMed: 9862869]
- Oskarsson T, Hreggvidsdóttir HS, Agnarsdóttir G, Matthíasdóttir S, Ogmundsdóttir MH, Jónsson SR, Georgsson G, Ingvarsson S, Andrésson OS, Andréddóttir V. Duplicated sequence motif in the long terminal repeat of maedi-visna virus extends cell tropism and is associated with neurovirulence. *Journal of Virology*. 2007; 81:4052–4057. [PubMed: 17287273]
- Pawlotsky J. Hepatitis C virus resistance to antiviral therapy. *Hepatology*. 2000; 32:889–896. [PubMed: 11050035]
- Payne SL, La Celle K, Pei XF, Qi XM, Shao H, Steagall WK, Perry S, Fuller F. Long terminal repeat sequences of equine infectious anaemia virus are a major determinant of cell tropism. *Journal of General Virology*. 1999; 80:755–759. [PubMed: 10092016]

- Pellerin M, Lopez-Aguirre Y, Penin F, Dhumeaux D, Pawlotsky J. Hepatitis C virus quasispecies variability modulates nonstructural protein 5A transcriptional activation, pointing to cellular compartmentalization of virus-host interactions. *Journal of Virology*. 2004; 78:4617–4627. [PubMed: 15078944]
- Pham T, MacParland S, Mulrooney P, Cooksley H, Naoumov N, Michalak T. Hepatitis C virus persistence after spontaneous or treatment-induced resolution of hepatitis C. *Journal of Virology*. 2004; 78:5867–5874. [PubMed: 15140984]
- Pond SL, Frost S. Datamonkey: rapid detection of selective pressure on individual sites of codon alignments. *Bioinformatics*. 2005; 21:2531–2533. [PubMed: 15713735]
- Pond SL, Frost SD, Muse S. HyPhy: hypothesis testing using phylogenetics. *Bioinformatics*. 2005; 21:676–679. [PubMed: 15509596]
- Qin H, Shire NJ, Keenan ED, Rouster SD, Eyster ME, Goedert JJ, Koziel MJ, Sherman KE. HCV quasispecies evolution: association with progression to end-stage liver disease in hemophiliacs infected with HCV or HCV/HIV. *Blood*. 2005; 105:533–541. [PubMed: 15374882]
- Qin W, Yamashita T, Shirota Y, Lin Y, Wei W, Murakami S. Mutational analysis of the structure and functions of hepatitis C virus RNA-dependent RNA polymerase. *Hepatology*. 2001; 33:728–737. [PubMed: 11230755]
- Ramirez S, Perez-Del-Pulgar S, Carrion JA, Costa J, Gonzalez P, Massaguer A, Fondevila C, Garcia-Valdecasas JC, Navasa M, Forns X. Hepatitis C virus compartmentalization and infection recurrence after liver transplantation. *American Journal of Transplantation*. 2009; 9:1591–1601. [PubMed: 19459796]
- Romero-López C, Berzal-Herranz A. The functional RNA domain 5BSL3.2 within the NS5B coding sequence influences hepatitis C virus IRES-mediated translation. *Cellular and Molecular Life Sciences*. 2011:2011. (Epub ahead of print).
- Roque-Afonso A, Ducoulombier D, Di Liberto G, Kara R, Gigou M, Dussaix E, Samuel D, Feray C. Compartmentalization of hepatitis C virus genotypes between plasma and peripheral blood mononuclear cells. *Journal of Virology*. 2005; 79:6349–6357. [PubMed: 15858018]
- Roque-Afonso A, Jiang J, Penin F, Tareau C, Samuel D, Petit M, Bismuth H, Dussaix E, Feray C. Nonrandom distribution of hepatitis C virus quasispecies in plasma and peripheral blood mononuclear cell subsets. *Journal of Virology*. 1999; 73:9213–9221. [PubMed: 10516029]
- Schramm F, Soulier E, Royer C, Weitten T, Fafi-Kremer S, Brignon N, Meyer N, Ellero B, Woehl-Jaegle ML, Meyer C, Wolf P, Doffoël M, Baumert TF, Stoll-Keller F, Schvoerer E. Frequent compartmentalization of hepatitis C virus with leukocyte-related amino acids in the setting of liver transplantation. *Journal of Infectious Diseases*. 2008; 198:1656–1666. [PubMed: 18925843]
- Seeff L. Natural history of viral hepatitis, type C. *Seminars in Gastrointestinal Diseases*. 1995; 6:20–27.
- Sherman KE, Rouster SD, Stanford S, Blackard JT, Shire N, Koziel M, Peters M, Chung RT. for the ACTG 5071 Study Team. Hepatitis C virus (HCV) quasispecies complexity and selection in HCV/HIV coinfecting subjects treated with interferon-based regimens. *Journal of Infectious Diseases*. 2010; 201:712–719. [PubMed: 20105080]
- Shimizu Y, Igarashi H, Kanematu T, Fujiwara K, Wong D, Purcell R, Yoshikura H. Sequence analysis of the hepatitis C virus genome recovered from serum, liver, and peripheral blood mononuclear cells of infected chimpanzees. *Journal of Virology*. 1997; 71:5769–5773. [PubMed: 9223464]
- Simmonds P. Genetic diversity and evolution of hepatitis C virus--15 years on. *Journal of General Virology*. 2004; 85:3173–3188. [PubMed: 15483230]
- Slatkin M, Maddison W. A cladistic measure of gene flow inferred from the phylogenies of alleles. *Genetics*. 1989; 123:603–613. [PubMed: 2599370]
- Smith D, Warne D, Vlahov D, Schuman P, Stein M, Greenberg B, Holmberg S. for the Human Immunodeficiency Virus Epidemiology Research Study Group. 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. *American Journal of Epidemiology*. 1997; 146:459–469. [PubMed: 9290506]

- Sobesky R, Feray C, Rimlinger F, Derian N, Dos Santos A, Roque-Afonso AM, Samuel D, Bréchet C, Thiers V. Distinct hepatitis C virus core and F protein quasispecies in tumoral and nontumoral hepatocytes isolated via microdissection. *Hepatology*. 2007; 46:1704–1712. [PubMed: 17935229]
- Stieler K, Schulz C, Lavanya M, Aepfelbacher M, Stocking C, Fischer N. Host range and cellular tropism of the human exogenous gammaretrovirus XMRV. *Virology*. 2010; 399:23–30. [PubMed: 20110097]
- Stover C, Smith D, Schmid D, Pellett P, Stewart J, Klein R, Mayer K, Vlahov D, Shchuman P, Cannon M. for the HIV Epidemiology Research Study Group. Prevalence of and risk factors for viral infections among human immunodeficiency virus (HIV)-infected and high-risk HIV-uninfected women. *Journal of Infectious Diseases*. 2003; 187:1388–1396. [PubMed: 12717619]
- Sullivan DG, Bruden D, Deubner H, McArdle S, Chung M, Christensen C, Hennessy T, Homan C, Williams J, McMahon BJ, Gretch D. Hepatitis C virus dynamics during natural infection are associated with long-term histological outcome of chronic hepatitis C disease. *Journal of Infectious Diseases*. 2007; 196:239–248. [PubMed: 17570111]
- Thomas D, Rich J, Schuman P, Smith D, Astemborski J, Nolt K, Klein R. for the HIV Epidemiology Research (HER) Study Group. Multicenter evaluation of hepatitis C RNA levels among female injection drug users. *Journal of Infectious Diseases*. 2001; 183:973–976. [PubMed: 11237816]
- Torres-Puente M, Bracho M, Jimenez N, Garcia-Robles I, Moya A, Gonzalez-Candelas F. Sampling and repeatability in the evaluation of hepatitis C virus genetic variability. *Journal of General Virology*. 2003; 84:2343–2350. [PubMed: 12917454]
- Tuplin A, Evans DJ, Simmonds P. Detailed mapping of RNA secondary structures in core and NS5B-encoding region sequences of hepatitis C virus by RNase cleavage and novel bioinformatic prediction methods. *Journal of General Virology*. 2004; 85:3037–3047. [PubMed: 15448367]
- Vera-Otarola J, Barría MI, León U, Marsac D, Carvallo P, Soza A, López-Lastra M. Hepatitis C virus quasispecies in plasma and peripheral blood mononuclear cells of treatment naïve chronically infected patients. *Journal of Viral Hepatitis*. 2009; 16:633–643. [PubMed: 19254350]
- Wolinsky S, Korber B, Neumann A, Daniels M, Kuntsman K, Whetsell A, Furtado M, Cao Y, Ho D, Safrit J. Adaptive evolution of human immunodeficiency virus-type 1 during the natural course of infection. *Science*. 1996; 272:537–542. [PubMed: 8614801]
- You S, Stump DD, Branch AD, Rice C. A cis-acting replication element in the sequence encoding the NS5B RNA-dependent RNA polymerase is required for hepatitis C virus RNA replication. *Journal of Virology*. 2004; 78:1352–1366. [PubMed: 14722290]
- Zárate S, Pond SL, Shapshak P, Frost S. Comparative study of methods for detecting sequence compartmentalization in human immunodeficiency virus type 1. *Journal of Virology*. 2007; 81:6643–6651. [PubMed: 17428864]
- Zehender G, De Maddalena C, Bernini F, Ebranati E, Monti G, Pioltelli P, Galli M. Compartmentalization of hepatitis C virus quasispecies in blood mononuclear cells of patients with mixed cryoglobulinemic syndrome. *Journal of Virology*. 2005; 79:9145–9156. [PubMed: 15994809]

Figure 1a – NS5B

Table 1

Demographic and clinical characteristics of HCV-positive women. ALT – alanine aminotransferase; AST – aspartate aminotransferase; SD – standard deviation

Variable	Value (percentage) [*]
Mean age at enrollment (years) ± SD	33.6 ± 5.7 [21.8, 43.2]
Race	
White	8 (61.5%)
African American	3 (23.1%)
Hispanic	2 (15.4%)
Risk factor	
Injection drug use	10 (76.9%)
Heterosexual contact	3 (23.1%)
Mean AST (IU/mL) ± SD [#]	61.4 ± 34.8 [29, 142]
Mean ALT (IU/mL) ± SD [#]	43.7 ± 17.7 [17, 78]
HIV status	12 (92.3%)
HIV positive/on antiretroviral therapy	0 (0.0%)
Median HIV viral load (log ₁₀ copies/mL) ± SD [#]	3.6 ± 0.9 [2.4, 5.2]
Median CD4 cell count (cells/mL) ± SD [#]	380.1 ± 229.5 [156.4, 881.2]

^{*} Data are numbers (percentage) unless otherwise noted; ranges are given in brackets [minimum, maximum]

[#] N = 12 women with available data

Table 2

Comparison of median inpatient genetic distance, entropy, and dN-dS in NS5B (left) and E1/HVR1 (right) sequences from serum (grey columns) and PBMC (white columns) samples.

ID	Genetic distance	Entropy	dN-dS
106	1.20%	0.9167	-0.0220
116	1.20%	1.0000	-0.0180
149	0.60%	0.8019	-0.0040
151	0.90%	1.0000	-0.0170
211	0.90%	1.0000	-0.0320
230	0.45%	1.0000	-0.0110
269	1.20%	0.8796	-0.0100
282	0.30%	1.0000	-0.0040
289	0.60%	0.8750	-0.0120
336	1.20%	0.5331	-0.0260
389	0.90%	0.9299	-0.0130
Median	0.90%	0.9299	-0.0130
106	0.30%	0.5331	-0.0080
116	0.90%	0.9070	-0.0270
149	0.30%	0.4729	-0.0030
151	0.00%	0.3693	-0.0030
211	0.45%	0.7632	-0.0040
230	0.00%	0.3112	0.0030
269	0.90%	1.0000	0.0000
282	0.00%	0.2775	0.0000
289	0.00%	0.2775	0.0020
336	1.20%	0.8796	-0.0290
389	0.60%	0.7592	-0.0020
Median	0.30%	0.5331	-0.0030
ID	Genetic distance	Entropy	dN-dS
106	5.29%	0.8895	0.0151

ID	Genetic distance	Entropy	dN-dS
116	0.80%	0.6949	0.0095
137	1.52%	0.8750	-0.0114
149	2.78%	0.9535	-0.0016
211	0.39%	0.4836	-0.0044
230	1.20%	1.0000	-0.0028
269	1.97%	0.5930	-0.0354
282	0.78%	0.7592	0.0004
389	1.63%	1.0000	-0.0190
Median	1.52%	0.8750	-0.0028
106	4.02%	0.8750	0.0474
116	1.59%	0.7699	0.0089
137	4.66%	1.0000	-0.0177
149	2.38%	0.8750	-0.0003
211	0.78%	0.7898	-0.0120
230	1.97%	1.0000	-0.0518
269	0.00%	0.2503	-0.0056
282	0.00%	0.1412	0.0011
389	0.00%	0.3368	-0.0058
Median	1.59%	0.7898	-0.0056

Table 3

NS5B codons under positive or negative selection pressure in serum and/or PBMCs. Shown are codons with a significance of $p \leq 0.05$ (FEL method) or posterior probability ≥ 0.99 (REL method). Codons under selection pressure in more than one woman are underlined.

Patient	Compartment	Positive selection (FEL)	Negative selection (FEL)	Positive selection (REL)	Negative selection (REL)
106	Serum	None	None	None	None
116*	Serum	None	2711	None	None
149	Serum	None	None	None	None
151	Serum	None	2702	2652, 2664, 2689, 2708, 2733, 2751	None
211	Serum	None	2701	None	2659, 2683, 2690, 2694, 2701, 2709, <u>2716</u> , 2721, 2726, 2735, 2738, 2750, <u>2755</u>
230*	Serum	None	2655	None	None
269	Serum	None	None	2747	None
282	Serum	None	None	None	None
289*	Serum	None	None	None	None
336	Serum	None	None	None	None
389	Serum	None	None	None	None
106	PBMC	None	None	None	None
116*	PBMC	None	2718	None	None
149	PBMC	None	2673	None	2673, 2757
151	PBMC	None	None	None	None
211	PBMC	None	None	None	None
230*	PBMC	None	None	None	None
269	PBMC	None	None	None	None
282	PBMC	None	2759	None	None
289*	PBMC	None	None	None	None
336	PBMC	None	2719	None	2680, 2687, 2703, 2707, <u>2716</u> , 2719, 2723, <u>2755</u>
389	PBMC	None	None	None	None

* denotes African American women.

Table 4

E1/HVR1 codons under positive or negative selection pressure in serum and/or PBMCs. HVR1 corresponds to amino acids 384 – 415. Shown are codons with a significance of $p < 0.05$ (FEL method) or posterior probability > 0.99 (REL method). Codons under selection pressure in more than one woman are underlined, while codons under selection pressure in both serum and PBMCs from the same woman are denoted by asterisks.

Patient	Compartment	Positive selection (FEL)	Negative selection (FEL)	Positive selection (REL)	Negative selection (REL)
106	Serum	None	348	<u>392</u> , 402, 405	None
116*	Serum	None	None	<u>386</u> , <u>404</u> *	None
137	Serum	None	None	<u>399</u> *, <u>403</u> , 407	None
149	Serum	None	335, 375	<u>330</u> , <u>334</u> , 358, <u>359</u> , 378, <u>392</u> , <u>397</u> , <u>399</u> , <u>403</u> , <u>404</u> , <u>410</u>	375*
211	Serum	None	369	<u>348</u> *, <u>357</u> , <u>390</u> , <u>403</u> , <u>410</u>	None
230*	Serum	None	<u>371</u> *, <u>410</u> *	None	<u>351</u> *, <u>371</u> *, <u>400</u> *, <u>405</u> *, <u>410</u> *
269	Serum	None	358, <u>370</u> , <u>403</u> , <u>410</u>	None	None
282	Serum	None	None	<u>384</u>	None
389	Serum	None	<u>380</u> , <u>410</u>	<u>330</u> , <u>340</u> , <u>357</u> , <u>359</u> , <u>384</u> , 391, 396, <u>408</u>	380
106	PBMC	None	<u>370</u>	<u>334</u> , 335, 338, 341, 349, 350, 354, 374, 379, <u>385</u> , <u>390</u> , <u>397</u> , <u>399</u> , <u>400</u> , <u>408</u>	370
116*	PBMC	404	None	<u>404</u> *	<u>371</u>
137	PBMC	None	331, 355	<u>356</u> , <u>385</u> , <u>386</u> , <u>399</u> *	None
149	PBMC	None	350	None	350, 375*, 391, 396
211	PBMC	None	None	<u>348</u> *, 387, 394	None
230*	PBMC	None	<u>371</u> *, <u>410</u> *, <u>411</u>	None	<u>351</u> *, <u>371</u> *, 378, 394, <u>400</u> *, <u>405</u> *, <u>408</u> , <u>410</u> *, <u>411</u>
269	PBMC	None	388	None	None
282	PBMC	#	#	#	#
389	PBMC	None	None	None	None

* denotes African American women.

Table 5

P values from Mantel's test and the Slatkin-Maddison test for NS5B and E1/HVR1 compartmentalization between the serum and PBMCs. NS, non-significant ($p > 0.10$). # denotes phylogenetic evidence of E1/HVR1 compartmentalization, while * denotes phylogenetic evidence of NS5B compartmentalization.

Patient ID	NS5B Mantel's test	S-M test [inferred number of migration events]	NS5B Mantel's test	E1/HVR1 Mantel's test	S-M test [inferred number of migration events]	E1/HVR1
106 [#]	NS		0.061 [4]	NS		NS [6]
116 [#]	0.0547		0.0044 [4]	NS		NS [7]
137 [#]	Not done		Not done	NS		NS [8]
149	NS		NS [7]	NS		NS [7]
151	NS		NS [7]	Not done		Not done
211	NS		0.0406 [4]	NS		NS [7]
230	NS		0.0007 [2]	NS		NS [5]
269	NS		NS [5]	NS		0.0358 [5]
282	NS		NS [6]	NS		NS [7]
289	0.0001		NS [6]	Not done		Not done
336 *	0.0300		0.0255 [4]	Not done		Not done
389	0.0017		0.0004 [2]	NS		0.0011 [3]

NS, non-significant ($p > 0.10$)

* denotes phylogenetic evidence of NS5B compartmentalization

[#] denotes phylogenetic evidence of E1/HVR1 compartmentalization