Frequent Hepatitis C Virus Superinfection in Injection Drug Users

Belinda L. Herring, Kimberly Page-Shafer, Leslie H. Tobler, and Eric L. Delwart

Blood Systems Research Institute, and Department of Medicine and Center for AIDS Prevention Studies, University of California, San Francisco

The frequency of hepatitis C virus (HCV) superinfection with a divergent viral strain was determined in a cohort of recently infected young injection drug users (IDUs) with an HCV incidence rate of 25%. HCV was amplified, by use of polymerase chain reaction (PCR), from plasma samples collected from 25 HCV-infected individuals over an average period of 12 months, and their viral sequences were compared. Phylogenetic analysis identified 5 IDUs with superinfection (20%) occurring after seroconversion: 2 IDUs were superinfected with different HCV genotypes, and 3 were superinfected with divergent strains of the same genotype. The superinfecting strains were not detected as minority variants (<0.5%) in the initial plasma HCV quasi species. Extensive measures were taken to exclude PCR contamination and mix-up of samples, and superinfection results were concordant at 2 HCV genetic loci. HCV superinfection in IDUs, both intra- and intergenotype, is therefore a frequent event, with an incidence rate similar to that of de novo infections. These results suggest that no cross-protecting immunity develops during the first year of chronic infection with HCV.

Hepatitis C virus (HCV) is estimated to infect 170 million people worldwide and can cause chronic hepatitis, cirrhosis, liver failure, and hepatocellular carcinoma [1–5]. Spontaneous resolution of HCV viremia occurs in ~15%–45% of infected individuals, and, in both humans and chimpanzees, improved control of viremia is associated with a cytotoxic T lymphocyte (CTL) response directed to numerous epitopes [6–11]. Maintenance of a strong memory CD4+ response is also associated with improved control of HCV viremia [12].

The ability of a previously cleared or ongoing HCV infection to protect against subsequent infection with a divergent strain of HCV (i.e., superinfection) may reflect the ability of adaptive immune responses to provide effective cross-protection. Partial protection against superinfection was reported among HCV-seropositive subjects who had cleared viremia [13].

HCV superinfection has been seen in persons who recently received blood-product transfusions or organ transplants [14–20]. In the chimpanzee model system, previously infected chimpanzees injected with very high viral load inoculums showed transient or chronic viremia with the rechallenge strain [21–25]. In contrast, when chimpanzees that had spontaneously cleared their initial viremia were rechallenged with HCV of different genotypes, they experienced reduced and only transient plasma viremia [26]. In the present study, we measured the frequency of superinfection among highly exposed IDUs, currently the population most at risk for HCV infection.

SUBJECTS, MATERIALS, AND METHODS

Samples. Twenty-five subjects were selected from a cohort of longitudinally followed young (<30 years of age) injection drug users (IDUs) who had recently undergone HCV seroconversion [27] and from whom frozen plasma samples were available. Informed consent was obtained from all subjects. The demographic and behavioral characteristics of this cohort have been de-
scribed elsewhere [27]. Subjects with viremia lasting at least 4 months were selected. HCV RNA was detected by use of the HCV discriminatory version of the Procleix HIV-1/HCV assay (Chiron) [28]. This assay is based on transcription-mediated amplification technology and has a detection rate of 95% for 35 HCV RNA copies/mL [29]. Viral loads were determined by use of the Roche COBAS ampicor HCV monitor 2.0 assay (limit of detection, 600 HCV RNA copies/mL). Seroconversion was determined by use of the Ortho HCV 3.0 ELISA system (Ortho-Clinical Diagnostics). The earliest collected plasma samples used for viral genetic analyses were labeled baseline samples, and the latest collected samples were labeled exit samples. Samples collected between these 2 time points were labeled intermediate samples.

**RNA extraction, polymerase chain reaction (PCR) amplification, and sequencing.** HCV RNA was extracted from plasma by use of the QIAamp Viral RNA Mini Kit (Qiagen), in accordance with the manufacturer’s instructions. RNA was reverse transcribed by use of 200 U of Superscript II RNAse H− reverse transcriptase (RT; Invitrogen) and a gene-specific oligonucleotide for either the E1/E2 HVRI region (DT2 5′-GTGCAAGGGTATGCGAGGCTT-3′ [nt 2181–2203]) or the NS5B region (1204 5′-GGAGGGGGGGAATACCTGGTCATAGCCTCCGTGAA-3′ [nt 8616–8650] [for genotype 1] or 1205 5′-TGGAGTTGACGAGCTTGTGCA-3′ [nt 8786–8813] [for genotype 3]). (Nucleotide positions are indi
cated with respect to the HCV H77 genome [GenBank accession number AF009606].) HCV genes were amplified from 5 μL of cDNA by use of nested (n) PCR. Each first-round PCR (50 μL) contained 10 pmol each of a sense and antisense primer, 1.5 mmol/L MgCl2, 200 μmol/L dNTPs, and 1.25 U of Taq polymerase (Promega). Amplification primers for the E1/E2 HVRI and NS5B regions were modified from those described by Wang et al. [30], by the inclusion of degenerate bases: outer primers, BH3 (sense, nt 1291–1311) 5′-GCGGGKGCYAYTGGGGMR-RTC-3′ and BH1 (antisense, nt 2043–2065) 5′-GCGAAACGCK-RTAYGGRTAGTCG-3′; and inner primers, DT3 (sense, nt 1427–1447) 5′-TTCCATGGTGGAGAATCTGGC-3′ and BH2 (antisense, nt 1883–1902) 5′-CATCCAGBTRCCAVCCRACC-3′. The NS5B region was amplified by use of genotype-specific primers and PCR conditions described elsewhere [31]. Amplified products were purified by use of the QIAquick PCR Purification Kit (Qiagen) and directly sequenced by use of the BigDye Dye Terminator Cycle Sequencing Kit 3.0 (Applied Biosystems), on an ABI3700 Sequencer, with sense primers.

**Minority variant detection at baseline.** cDNA was generated by use of genotype 1 specific antisense primer 1204 (for IDus 25, 54, and 59) or random hexamers (for IDus 20 and 32). The number of HCV cDNA molecules generated from each baseline sample was estimated by end-point dilution nPCR, by use of NS5B primers specific for the baseline genotype [31]. cDNA copy number was then estimated by use of the software QUALITY [32]. A minimum of 100 cDNA molecules were tested for the presence of the superinfecting variant in all samples except for the intermediate sample from IDU 32, whose cDNA input copy number was <100 cDNA molecules.

For IDUs 25, 54, and 59, first-round PCR was performed on baseline, intermediate, and exit samples, by use of NS5B primers specific for genotype 1, and second-round PCR was performed by use of superinfecting strain sequence-specific amplification primers (SSAPs), as follows: 25-E-Fwd 5′-CATCTACAAATGTTGTGACCTT-3′, 25-E-Rev 5′-TGCACTCTCTGCTGACCTCG-3′, 54-E-Fwd 5′-ACCCCCAACGCGCTTG-3′, 54-E-Rev 5′-GCTGATCTCCTGGACTCCTGA-3′, 59-E-Fwd 5′-CGTGGGCAATGCTCTCAGCC-3′, and 59-E-Rev 5′-TGCACTCTCTGTGGACCTCG-3′. For IDUs 20 and 32, random primed cDNA from baseline, intermediate, and exit samples was amplified by use of both genotype 3a and 1a NS5B nested primer sets.

**Confirmation of origin of plasma samples.** The origin of each of the samples from IDUs harboring divergent HCV strains at baseline and exit was verified by use of the Miragen Antibody Profile Assay (Miragen; G. I. Industries). This is a serological assay based on the presence, in each individual, of weakly autoreactive antibodies. In brief, a Western blot strip loaded with human cell line antigens was incubated with the subject’s plasma, and bound antibodies were detected colorimetrically. The resulting band pattern is stable over time and capable of discriminating between individuals, since a unique band profile is produced for each individual [33, 34]. Antibody profiles generated from the 5 IDUs suspected of having superinfection were evaluated in a blinded fashion, with respect to the identity of the plasma source, by an independent laboratory (California Department of Public Health Services, Richmond). Western blot strips were then paired on the basis of the similarity of their band profiles.

**Phylogenetic analysis.** Sequences generated for the E1/E2 HVRI and NS5B regions were aligned by use of ClustalX, manually edited, and gap-stripped [35]. Phylogenetic analysis was performed by use of PAUP 4.0 [36]. Maximum likelihood (ML) phylogenetic trees were constructed for both regions by use of PAUP 4.0. The model of substitution, base frequencies, gamma distribution, invariant sites, transition/transversion ratios, and rate matrix values were determined by use of Modeltest 3.07 [37]. Bootstrap analysis (100 replicates) was performed on each ML tree.

**Nucleotide sequence accession numbers.** HCV E1/E2 HVRI and NS5B gene sequences were submitted to GenBank (accession numbers AY691950–692022 and AY689289–689329).
RESULTS

IDU population. HCV strains from 25 recently infected, HIV-negative subjects from a cohort of young urban IDUs were genetically characterized at the E1/E2 HVR1 locus by use of plasma samples collected at an average interval of 316 days (range, 134–715 days) (table 1 and Subjects, Materials, and Methods). All longitudinally collected samples from these subjects were found to be HCV RNA positive, at a limit of detection of 35 HCV RNA copies/mL. The baseline samples from 18 of 25 IDUs were found to be HCV seropositive (table 1).

Phylogenetic analysis at the E1/E2 HVR1 region. Viral RNA was extracted from all baseline and exit samples, and the E1/E2 HVR1 region was amplified by use of RT-PCR. Amplification products were directly sequenced and phylogenetically analyzed (figure 1A). HCV strains belonging to genotypes 1 (subtypes 1a and 1b) and 3 were identified (table 1). The baseline and exit sequences from 20 subjects clustered with bootstrap values >88. Five IDUs suspected of having superinfection in whom the baseline and exit sequences did not cluster on the same branch were identified (figure 1A). All 5 IDUs suspected of having superinfection were seropositive at baseline: 2 had cases that involved genotype switches (IDUs 20 and 32), and 3 showed the presence of highly divergent variants of the same genotype in exit samples (IDUs 54, 59, and 25). Three phylogenetic clusters were also identified in this analysis: cluster A, IDUs 2 and 46; cluster B, IDUs 5 and 31; and cluster C, IDUs 13, 23, and 54E (figure 1). IDUs 2 and 46 are self-reported sex partners. These clusters may reflect direct or indirect HCV transmission in epidemiologically linked groups within this geographically restricted San Francisco cohort. To ensure that PCR contamination or mix-up of samples did not account for the HCV sequence changes in the 5 IDUs suspected of having superinfection, new aliquots of plasma were obtained from the cohort’s repository of samples. Viral RNA extraction, RT-nested PCR, sequencing, and phylogenetic analysis for the E1/E2 HVR1 locus were repeated for each additional sample (figure

<table>
<thead>
<tr>
<th>Subject</th>
<th>Sex</th>
<th>Subtype(s)</th>
<th>Days between baseline and exit samples (days between baseline and intermediate sample)</th>
<th>Days from preceding HCV RNA–negative sample to baseline sample</th>
<th>Baseline EIA 3.0 result (days until seropositive plasma sample)</th>
</tr>
</thead>
<tbody>
<tr>
<td>2c</td>
<td>M</td>
<td>1a</td>
<td>715</td>
<td>171</td>
<td>+</td>
</tr>
<tr>
<td>3</td>
<td>F</td>
<td>3a</td>
<td>391</td>
<td>166</td>
<td>+</td>
</tr>
<tr>
<td>5c</td>
<td>M</td>
<td>3a</td>
<td>388</td>
<td>0</td>
<td>− (158)</td>
</tr>
<tr>
<td>10</td>
<td>M</td>
<td>1a</td>
<td>186</td>
<td>&gt;441</td>
<td>+</td>
</tr>
<tr>
<td>11</td>
<td>M</td>
<td>1a</td>
<td>519</td>
<td>208</td>
<td>+</td>
</tr>
<tr>
<td>13c</td>
<td>M</td>
<td>1a</td>
<td>199</td>
<td>89</td>
<td>+</td>
</tr>
<tr>
<td>15c</td>
<td>F</td>
<td>1a</td>
<td>448</td>
<td>84</td>
<td>+</td>
</tr>
<tr>
<td>16c</td>
<td>M</td>
<td>1a</td>
<td>160</td>
<td>183</td>
<td>+</td>
</tr>
<tr>
<td>17c</td>
<td>F</td>
<td>1a</td>
<td>404</td>
<td>96</td>
<td>− (117)</td>
</tr>
<tr>
<td>20c</td>
<td>M</td>
<td>3a/1a</td>
<td>208</td>
<td>388</td>
<td>+</td>
</tr>
<tr>
<td>23c</td>
<td>M</td>
<td>1a</td>
<td>434</td>
<td>76</td>
<td>− (345)</td>
</tr>
<tr>
<td>25c</td>
<td>M</td>
<td>1a/1a</td>
<td>167</td>
<td>87</td>
<td>+</td>
</tr>
<tr>
<td>31c</td>
<td>F</td>
<td>3a</td>
<td>247</td>
<td>98</td>
<td>+</td>
</tr>
<tr>
<td>32c</td>
<td>M</td>
<td>1a/1a/3a</td>
<td>236 (62)</td>
<td>144</td>
<td>+</td>
</tr>
<tr>
<td>35</td>
<td>M</td>
<td>3a</td>
<td>398</td>
<td>105</td>
<td>− (73)</td>
</tr>
<tr>
<td>36</td>
<td>F</td>
<td>3a</td>
<td>270</td>
<td>84</td>
<td>+</td>
</tr>
<tr>
<td>40</td>
<td>F</td>
<td>3a</td>
<td>134</td>
<td>259</td>
<td>+</td>
</tr>
<tr>
<td>44</td>
<td>M</td>
<td>3a</td>
<td>254</td>
<td>0</td>
<td>− (151)</td>
</tr>
<tr>
<td>46c</td>
<td>F</td>
<td>1a</td>
<td>172</td>
<td>0</td>
<td>− (90)</td>
</tr>
<tr>
<td>47c</td>
<td>F</td>
<td>1a</td>
<td>325</td>
<td>78</td>
<td>+</td>
</tr>
<tr>
<td>49</td>
<td>F</td>
<td>1a</td>
<td>428</td>
<td>&gt;195</td>
<td>+</td>
</tr>
<tr>
<td>51</td>
<td>F</td>
<td>3a</td>
<td>226</td>
<td>89</td>
<td>+</td>
</tr>
<tr>
<td>53</td>
<td>M</td>
<td>1a</td>
<td>562</td>
<td>0</td>
<td>− (100)</td>
</tr>
<tr>
<td>54c</td>
<td>M</td>
<td>1b/1b/1b/1a</td>
<td>283 (103, 185)</td>
<td>167</td>
<td>+</td>
</tr>
<tr>
<td>59c</td>
<td>F</td>
<td>1a/1a/1a</td>
<td>152 (89)</td>
<td>145</td>
<td>+</td>
</tr>
</tbody>
</table>

NOTE. Superinfected subjects are shown in bold type. +, Positive; −, negative.

* Genotype determined by phylogenetic analysis.
* Individuals whose first collected sample was already HCV RNA positive and EIA negative are labeled 0.
* Samples were analyzed in both E1/E2 HVR1 and NS5B regions.
* Indicates emergence of a highly distinct variant of the same subtype.
Figure 1. Hepatitis C virus (HCV) phylogenetic trees for 25 injection drug users (IDUs), constructed by use of the maximum likelihood method and nucleotide sequences corresponding to the E1/E2 HVR1 (A) and NS5B (B) regions. Baseline time points are labeled “B,” intermediate time points “i” and “ii,” and exit time points “E.” Asterisks denote samples in which a separate aliquot of plasma was sequenced. Bootstrap values >70% are indicated at the nodes of branches. The 5 IDUs showing evidence of HCV superinfection are in bold type.
The newly derived E1/E2 HVR1 sequences from all 5 IDUs suspected of having superinfection were consistent with the original results (figure 1). Intermediate samples were available from 3 of the 5 IDUs suspected of having superinfection and were similarly analyzed at the E1/E2 HVR1 locus. Intermediate sequences clustered with the baseline sequences (figure 1 and table 1).

**Phylogenetic analysis at the NS5B region.** To confirm these initial results, further analyses were performed at an additional viral locus. Samples from the 5 IDUs suspected of having superinfection, all 7 members of the 3 phylogenetic clusters, and 6 IDU control subjects were genetically analyzed at the NS5B locus. NS5B phylogenetic analysis corroborated (1) the 5 cases of superinfection observed at the E1/E2 HVR1 locus, (2) the phylogenetic linkage of the 3 San Francisco HCV clusters, and (3) the nonsuperinfected status of the other 6 IDU control subjects (figure 1B).

**Confirmation of origin of plasma samples.** Aliquots of the baseline and exit samples from the 5 IDUs suspected of having superinfection were also analyzed in a blinded panel, for possible mislabeling or mix-up of samples (see Subjects, Materials, and Methods). The baseline and exit samples from the same IDUs were all unambiguously paired on the basis of the similarity of their autoreactive antibody profiles (data not shown). This result eliminated the possibility that samples from different subjects were mistakenly labeled as originating from the same subject.

**Viral loads.** Plasma viral loads fluctuated widely, with 17 IDUs showing a decreasing and 8 IDUs showing an increasing viral load between the baseline and exit samples (figure 2). The highest viral loads at baseline were measured in 4 of the 7 samples collected before seroconversion (table 1 and figure 2). After superinfection, viral loads increased in 1 and decreased in the other 4 IDUs. No correlation was observed between superinfection and viral load.

**Testing for HCV coinfection.** The presence of the superinfecting strain in the baseline and intermediate samples would reflect coinfection rather than superinfection. Same-genotype superinfections (IDUs 25, 54, and 59) were tested by use of strain-specific primers, and different-genotype superinfections (IDUs 20 and 32) were tested by use of genotype-specific primers (see Subjects, Materials, and Methods).

Each primer set was initially shown to be capable of detecting the superinfecting exit variant in reconstituted PCR mixtures when present at a frequency of \( \geq 0.5\% \) (figure 3). The number of cDNA molecules in the baseline sample from each of the 5 IDUs suspected of being superinfected was then determined by end-point dilution and nPCR (see Subjects, Materials, and Methods), to ensure sufficient sampling of the quasi species. At least 100 cDNA molecules from the baseline time point were therefore used as input for PCR, with strain- or genotype-specific primers specific for the exit variant. No evidence for the presence of the superinfecting exit variants was found in baseline or intermediate samples (table 1 and figure 3).

**Genetic distances.** To further verify the 2 cases of same-subtype (1a) superinfections (i.e., those in IDUs 25 and 59),
Figure 3. Testing for coinfection in baseline (B) plasma samples. Mixtures of first-round polymerase chain reaction products from the B and exit (E) samples were amplified with E strain sequence-specific amplification primers (A) and E strain genotype-specific primers (B), to determine the limit of detection for the superinfecting strain in the B sample. Positive controls of B samples amplified with B genotype-specific primers are labeled “C.” IDU, injection drug user.

Figure 4. Within- and between-subject pairwise genetic distances, at the E1/E2 HVR1 (A) and NS5B (B) regions. Intrapatient (black bars) and interpatient (gray bars) distances were calculated by use of maximum likelihood settings identical to those used for phylogenetic analysis. Between-subject distances for subjects whose strains phylogenetically clustered are excluded. Intrapatient genetic distance for injection drug users 25 and 59 are indicated.

we compared these IDUs’ within-subject HCV genetic distances, at both loci sequenced, with those observed in non-superinfected subjects (figure 4). The average within-subject genetic distances in the E1/E2 HVR1 and NS5B regions for IDUs whose intrasubject sequences clustered together were 1.75% (range, 0%–5.82%) and 0.17% (range, 0%–0.06850%), respectively (figure 4). The average within-subject genetic distances for IDUs 25 and 59 were 13.9% and 18.4%, respectively, for the E1/E2 HVR1 locus, and 4.1% and 6.1%, respectively, for the NS5B locus. The genetic distances for these 2 IDUs were within the range observed between phylogenetically unlinked IDUs (6.6%–24.8% in the E1/E2 HVR1 region and 2.3%–7.8% in the NS5B region). The high levels of sequence divergence seen in IDUs 25 and 59 were therefore consistent with superinfection with distinct viral strains of the same subtype, rather than HCV evolution of the same strain.

DISCUSSION

Partial protection from recurring HCV viremia and reduced viral loads have been reported in seropositive subjects and chimpanzees who cleared viremia [13, 26]. In the present study, we sought to determine whether such protection from superinfection could be detected in viremic subjects with a high incidence rate of HCV infection. We found that the frequency of infection with highly divergent strains of HCV was similar to the incidence rate of new infection, indicating that little or no protection was afforded by seroconversion and persistent viremia.

The 5 cases of superinfection identified here were not the result of coinfection with multiple HCV strains, since the superinfecting strains were not detected in the baseline samples (limit of detection, ≥0.5%). The absence of the superinfecting strains from the intermediate samples allowed a more precise timing of superinfection. IDU 32 was superinfected at least 62 days, IDU 59 at least 89 days, and IDU 54 at least 185 days after the initial seropositive sample was obtained (figure 1 and table 1). Therefore, although superinfections occurred in HCV-seropositive subjects, it is still possible that the relatively immature immune responses in these individuals may have been responsible for the absence of protection against superinfection detected here [38].

Mixed-genotype infections have been reported by use of non-sequencing methods [39–41]; however, the methods used may have resulted in an overestimation of the frequency of mixed-
genotype infections, since other studies found cases of mixed-genotype infections to be rare [42–45]. A case report of cross-genotype superinfection after clearance of initial infection/viremia in an IDU was described; however, the data were not supported by detailed phylogenetic analysis or stringent controls for PCR contamination and mix-up of samples [46].

The high frequency of HCV superinfections that we detected among young IDUs indicates the ease with which a new viral strain can surmount immune responses directed at the resident strain and rapidly become the dominant quasi species. HCV superinfection was observed between viral strains belonging to the same, as well as different, genotypes. This result may reflect the difficulty for adaptive immune responses to recognize divergent strains, even those belonging to the same HCV genotype or subtype. The partial protection from reinfection reported in subjects who had apparently cleared their initial viremia may reflect a stronger immune response in this subset of individuals [13]. Chimpanzees with lower viremia have shown stronger CTL responses targeting more epitopes [10].

Although examples of HIV superinfection have recently been reported [47–50], longitudinal sequence analyses have reported a low frequency of superinfection with strains of the same HIV subtypes [51, 52]. The higher probability of HCV infection after parenteral exposure, compared with that for HIV [53, 54] as well as the greater level of genetic variability of HCV, may, in part, account for the large difference between frequencies of HIV and HCV superinfection.

The present report has described the frequent occurrence of HCV superinfection in a cohort of highly exposed subjects. Superinfection in the 5 IDUs presented here was confirmed by sequencing at 2 HCV genetic loci, combined with detailed phylogenetic analysis, careful quality controls for PCR contamination and mix-up of samples, and exclusion of initial coinfection. The high frequency (20%) of superinfection in the cohort in the present study is similar to the underlying HCV incidence rate and, therefore, indicates that little or no protection is provided by seroconversion and ongoing viremia. The result of this natural-history experiment in HCV challenge indicates that successful vaccination against this highly diverse virus may prove difficult, except for the minority of individuals likely to otherwise spontaneously clear HCV [13, 26]. Although the clinical consequences of HCV superinfection remain unknown and no apparent effect on plasma viral load was detected here, such consequences might include a decreased probability of spontaneous HCV clearance and a greater resistance to therapy. At the very least, our data suggest that already-infected subjects and those whose viremia was cleared by antiviral therapy should protect themselves against subsequent HCV exposure.

Acknowledgments

We thank Lubov Pitina for processing of plasma samples and Michael Busch (Blood Systems Research Institute) and Stewart Cooper (University of California, San Francisco) for helpful comments.

References


54. Logvinoff C, Major ME, Oldach D, et al. Neutralizing antibody re-