JCM
Journals.ASM.org

Rapid Hepatitis C Virus Divergence among Chronically Infected Individuals

Mayra Cruz-Rivera,[a] Juan Carlos Carpio-Pedroza,[b] Alejandro Escobar-Gutiérrez,[b] Daniela Lozano,[c] Arely Vergara-Castaneda,[d] Pilar Rivera-Osorio,[b] Armando Martinez-Guarneros,[b] Carlos A. Vazquez Chacon,[b] Salvador Fonseca-Coronado,[e] Gilberto Vaughan[b]*

Facultad de Medicina, Universidad Nacional Autónoma de México, Mexico City, Mexico[a]; Instituto de Diagnóstico y Referencia Epidemiológicos, Mexico City, Mexico[b]; Departamento de Biomedicina Molecular, Centro de Investigación y de Estudios Avanzados, Instituto Politécnico Nacional, Mexico City, Mexico[c]; Facultad de Ciencias Químicas, Universidad La Salle, Mexico City, Mexico[d]; Laboratorio de Inmunobiología de Enfermedades Infecciosas, Unidad de Investigación Multidisciplinaria, Facultad de Estudios Superiores Cuautitlán, Universidad Nacional Autónoma de México, Estado de México, Mexico[e]

Here, we analyze the viral divergence among hepatitis C virus (HCV) chronic cases infected with genotype 1. The intrahost viral evolution was assessed by deep sequencing using the 454 Genome Sequencer platform. The results showed a rapid nucleotide sequence divergence. This notorious short-term viral evolution is of the utmost importance for the study of HCV transmission, because direct links between related samples were virtually lost. Thus, rapid divergence of HCV significantly affects genetic relatedness studies and outbreak investigations.

Globally, hepatitis C virus (HCV) infects approximately 130 million people, with 3 million new infections occurring annually (1). Before the arrival of the new direct-acting antivirals (DAAs), the combined therapy based on long-acting pegylated alpha interferon (IFN) and ribavirin (RBV) was successful in only approximately 50% of cases infected with HCV genotype 1 (2). Diverse predictive markers, including virus genotype, viral load, complexity of viral population, and viral genome sequence, have been reported to have some predictive value. Single nucleotide polymorphisms (SNPs) located near the interleukin-28B (IL-28B) gene were found to be associated with the outcome of antiviral treatment, spontaneous viral clearance, and progression to chronicity, which make them useful as predictive factors for personalizing antiviral therapy (3).

HCV is a positive-polarity, single-stranded RNA virus that belongs to the genus *Hepacivirus* in the family *Flaviviridae* (4). The HCV polymerase is an RNA-dependent RNA polymerase lacking proofreading and error correction mechanisms. As a result, HCV replication, like that of many other RNA viruses, is highly error prone (5). The HCV mutation rate has been estimated to be 2.5×10^5 mutations per nucleotide per genome replication (6). This characteristic high mutation rate is consistent with the high degree of intrahost genetic diversity commonly observed among infected individuals (7, 8). The high degree of genetic variability allows rapid evolution that ultimately favors escape from immune and antiviral drug pressures (9, 10). This genetic variation also allows the virus to evolve and optimize its fitness in each host (positive selection) under the functional constraints on replication (11). Furthermore, the high progeny production (approximately 10^{12} virions per day) facilitates rapid viral evolution (12).

Different HCV subgenomic regions exhibit different degrees of variability (13). As a consequence, several regions of the viral genome have been used for molecular studies (7, 14, 15). However, the HCV hypervariable region 1 (HVR1) of the envelope (E) 2 gene has been the region of choice to conduct relatedness studies due to its greater variability (16). Additionally, the HVR1 exhibits more rapid sequence changes over time, a feature that confounds the identification of related HCV cases. Thus, HCV viral divergence further complicates the association of cases to their source of infection (7). Therefore, accurate and reliable methods for the identification of HCV transmission are of critical importance.

Transmission networks are difficult to identify for several reasons, including the lack of laboratory tests capable of distinguishing acute from chronic infections and the typical asymptomatic nature of acute HCV infections (7). Moreover, the characteristically long incubation periods among symptomatic cases make it difficult to link related cases to a common source of infection.

In addition, molecular approaches required to assess the intrahost viral genetic variation are quite sophisticated, time-consuming, and expensive, further complicating the identification of transmission events (17). The main issues associated with the correct identification of viral variants present in any given clinical specimen are the complexity of such populations and the difficulty in detecting low-frequency variants (16). However, the development of next-generation sequencing (NGS) platforms has significantly improved the analysis of the intrahost viral populations (7, 8), providing a snapshot of the entire virus population, which makes it a viable alternative to other conventional methodologies (18).

The goal of this work was to analyze the viral population divergence among chronic cases of HCV genotype 1 infection in individuals who reported the use of intravenous drugs. All patients were anti-HCV treatment naïve. Subjects were epidemiologically unrelated. The ethical review and informed consent approval was granted by the Ethical Committee of the National Reference Laboratory. Informed consent was obtained from all subjects. The patients' blood samples were collected by venipuncture into PPT

Received 17 November 2012 **Accepted** 26 November 2012
Published ahead of print 5 December 2012
Address correspondence to Gilberto Vaughan, gilvaughan@yahoo.com.
* Present address: Gilberto Vaughan, Division of Viral Hepatitis, Centers for Disease Control and Prevention, Atlanta, Georgia, USA.
M.C.-R. and J.C.C.-P. contributed equally to this work.

doi:10.1128/JCM.03042-12

TABLE 1 Patient characteristics

Patient	Age	Gender	Risk factor	IL-28B genotype	HCV genotype
P-A	28	Male	IDU	TT	1a
P-B	32	Male	IDU	TT	1a
P-C	25	Male	IDU	TT	1a
P-D	31	Male	IDU	TT	1b
C-1	36	Male	IDU	CC	1a
C-2	38	Male	IDU	CT	1a
C-3	63	Male	Transfusion	TT	1a
C-4	41	Female	IDU	TT	1b
C-5	40	Male	IDU	TT	1a
C-6	61	Male	Transfusion	CC	1a
C-7	65	Female	Transfusion	CC	1b
C-8	69	Female	Transfusion	CC	1a
C-9	59	Male	Transfusion	TT	1a

Vacutainer tubes (Becton, Dickinson, Franklin Lakes, NJ) and processed according to the manufacturer's recommendations. IL-28B genotyping was performed as reported previously (19). HCV genotyping was performed as described elsewhere (7).

Initially, HCV chronic cases bearing the non-CC IL-28B genotype were given the option to start conventional dual therapy with IFN and RBV or to postpone treatment until the arrival of the new protease inhibitors. Four individuals reporting the use of intravenous drugs opted to wait for the new therapy regime. These patients (A to D) were sampled twice within 18 months before the start of anti-HCV therapy (18, 14, 12, and 7 months for patients A, B, C, and D, respectively). Three patients were infected with subgenotype 1a (A, B, and C) and one with subgenotype 1b (D). Nine unrelated HCV cases, seven individuals infected with HCV genotype 1a (C1, C2, C3, C5, C6, C8, and C9), and two individuals infected with genotypes 1b (C4 and C7) were also enrolled in the study and were used as controls. The characteristics of all individuals are listed in Table 1.

All serum samples from patients and controls were subjected to ultradeep pyrosequencing to analyze the degree of divergence as previously described (7, 8). A comprehensive nucleotide alignment, containing 358 HCV HVR1 representative sequences from subgenotypes 1a and 1b (140 and 218, respectively), was obtained from the Los Alamos National Laboratory HCV database (http://hcv.lanl.gov/content/sequence/NEWALIGN/align.html) and used as unrelated sequences. The program ModelTest was used to establish the best model of RNA substitution for the data. The general time reversible model was chosen to create maximum-likelihood (ML) trees as implemented in MEGA5 (20). Polymorphism and divergence analyses were conducted using DnaSP software version 5 (21). After denoising the data, phylogenetic analysis (maximum likelihood) was conducted to assess the changes in the architecture of the viral population over time (Fig. 1A). The results showed a high degree of divergence among paired specimens from all four patients (minimal nucleotide distances 22, 17, 32, and 26 for patients A, B, C, and D, respectively). Minimal distances among unrelated cases ranged between 17 and 29 nucleotides. Interhost genetic distances between patients A and B were closer. In con-

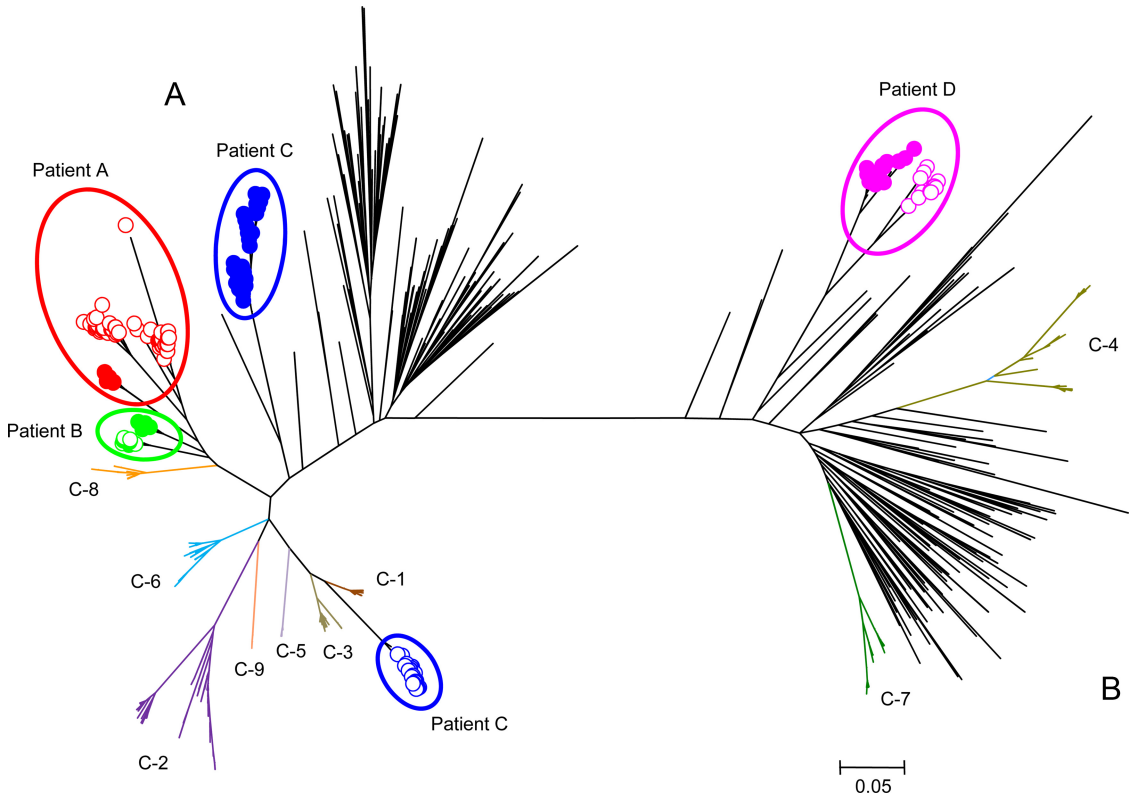

FIG 1 Phylogenetic analysis. Phylogenetic analysis (maximum likelihood) was conducted using the HCV HVR1 sequences from all cases and controls. Sequences from the first (closed circles) and second (open circles) time points are depicted in colors. Sequences from controls are also color coded. Additionally, sequences from the Los Alamos database are shown in black.

trast, patient C displayed the most noticeable divergence among all four of these subjects. The genetic distances between the viral populations from patients and controls belonging to subgenotype 1a did not show statistically significant differences. Finally, the two viral populations sampled from patient D remained within the same sequence space despite the significant sequence divergence observed at the two time points. In summary, all four subjects displayed a remarkable degree of sequence divergence. The phylogenetic and minimal distance analyses showed no significant differences between the paired samples and the nonrelated controls regardless of the infecting viral genotype.

Here, we have shown the rapid, short-term sequence divergence among chronic cases of HCV genotype 1 infection bearing the non-CC IL-28B genotype among those reporting the use of intravenous drugs. Additionally, the results demonstrated that the usage of next-generation platforms provides a comprehensive snapshot of the viral population in clinical samples.

HCV evolution is a critical factor that in many ways dictates the outcome of the disease. The mechanisms of persistence, a key feature of the biology of the virus, are not fully understood. The high mutation rate, characteristic of HCV infection, typically leads to a continuous diversification of the viral population (22). As a consequence, HCV circulates *in vivo* as a complex mixture of viral variants closely related to each other with a dynamic distribution subjected to a continuous process of genetic variation, competition, and selection. This genomic heterogeneity allows rapid adaptation to changes in the microenvironment when the viral population is subjected to selective forces (host immune response, antiviral therapy, etc.). Thus, the enormous variability of HCV represents a great challenge for disease control.

The evolution of HVR1 among these individuals highlights some of the most important limitations encountered when dealing with outbreak investigations. The results clearly showed that the high degree of divergence makes the identification of related cases extremely challenging. The remarkable changes in the composition of the viral populations at different times observed among these patients could be the result of potential reinfection or superinfection. Indeed, both phenomena have been known to occur in individuals reporting the use of drugs (23). Thus, one could expect that the dramatic changes in the architecture of the viral population may be the result of viral replacement by related strains rather than by molecular evolution. The close relatedness among strains can be explained because networks of intravenous drug use (IDU) are "tight-knit" groups that can maintain certain viral lineages that might have an advantage (fitness) in this particular setting.

The use of minimal genetic distances has been shown to be an appropriate approach to address relatedness among HCV cases (7). However, the high degree of divergence observed among these cases in such a short period of time prevents the correct assignment of cases to clusters of transmission because the minimal distances overlap those obtained from unrelated cases. Alternatively, and as shown in Fig. 1, a phylogeny relationship can be used to establish transmission links. Additionally, the use of other informative regions (i.e., the NS5A gene) or larger segments of the genome can help overcome the limitations of current approaches used in molecular epidemiological studies. Nevertheless, overcoming HCV molecular evolution is not an easy task. As seen with patient D, in some cases phylogeny is not sufficient to correctly link cases. As a consequence, the time of identification plays an important role in the accurate identification of cases and their source of infection. Therefore, the opportune identification of cases in a timely manner will still play a pivotal role in the recognition of transmission networks.

All four patients displayed the TT IL-28B genotype. Others have suggested the participation of the IL-28B genotype in the evolution of HCV (24).Whether the host IL-28B genotype is responsible for the rapid sequence divergence observed among these patients requires further research. One of the limitations of our study was the lack of CC IL-28B genotype HCV-infected subjects. However, the ethical restrictions surrounding this issue prevent the inclusion of individuals with that particular IL-28B genotype, as these patients are likely to respond to the antiviral treatment. Therefore, it is highly improbable to carry out prospective studies looking at the molecular evolution of HCV among CC patients. The second critical shortcoming of this study was the small number of individuals. It is important to mention that with the arrival of the DAAs, the success rates have been improved significantly, and, therefore, keeping patients from being treated is not justifiable. In this particular study, the individuals were originally enrolled before the licensing of the new antiviral drugs. Thus, the possibilities of enrolling a large number of individuals suitable for long prospective follow-up studies are rather limited.

Recently, a study has shown the advantages of using next-generation sequencing platforms over conventional methods for the analysis of the HCV intrahost variability (8). Ultradeep sequencing is a powerful technology that allows for the analysis of the viral population in great detail, providing a comprehensive picture of the architectural arrangement of the viral population. Several reports have described the usage of this methodology in the study of HCV intrahost viral evolution (25–29). Thus, usage of more advanced technology, such as amplicon deep sequencing, is more appropriate for a thorough evaluation of the viral population in clinical samples (7, 30).

In conclusion, HCV can exhibit a high degree of sequence divergence. These findings have important implications for outbreak investigations and genetic relatedness studies, as links to related cases can be lost over a short period of time. Additionally, the use of NGS technologies will most likely help to better understand the molecular basis of HCV evolution.

ACKNOWLEDGMENTS

This work was partially supported by Proyecto PAPIIT TA200112, Dirección General de Asuntos del Personal Academico, Universidad Nacional Autónoma de México.

REFERENCES

1. **Alter MJ.** 2007. Epidemiology of hepatitis C virus infection. World J. Gastroenterol. **13**:2436–2441.
2. **Teoh NC, Farrell GC, Chan HL.** 2010. Individualisation of antiviral therapy for chronic hepatitis C. J. Gastroenterol. Hepatol. **25**:1206–1216.
3. **Ge D, Fellay J, Thompson AJ, Simon JS, Shianna KV, Urban TJ, Heinzen EL, Qiu P, Bertelsen AH, Muir AJ, Sulkowski M, McHutchison JG, Goldstein DB.** 2009. Genetic variation in IL28B predicts hepatitis C treatment-induced viral clearance. Nature **461**:399–401.
4. **Lemon SM, Walker CM, Alter MJ, Yi M.** 2007. Hepatitis C virus, p 1253–1304. *In* Knipe DM, Howley PM (ed), Fields virology, vol 1. Lippincott Williams & Wilkings, Philadelphia, PA.
5. **Moradpour D, Penin F, Rice CM.** 2007. Replication of hepatitis C virus. Nat. Rev. Microbiol. **5**:453–463.
6. **Ribeiro RM, Li H, Wang S, Stoddard MB, Learn GH, Korber BT, Bhattacharya T, Guedj J, Parrish EH, Hahn BH, Shaw GM, Perelson AS.** 2012. Quantifying the diversification of hepatitis C virus (HCV) dur-

ing primary infection: estimates of the in vivo mutation rate. PLoS Pathog. **8**:e1002881. doi:10.1371/journal.ppat.1002881.
7. **Escobar-Gutierrez A, Vazquez-Pichardo M, Cruz-Rivera M, Rivera-Osorio P, Carpio-Pedroza JC, Ruiz-Pacheco JA, Ruiz-Tovar K, Vaughan G.** 2012. Identification of hepatitis C virus transmission using a next-generation sequencing approach. J. Clin. Microbiol. **50**:1461–1463.
8. **Fonseca-Coronado S, Escobar-Gutierrez A, Ruiz-Tovar K, Cruz-Rivera MY, Rivera-Osorio P, Vazquez-Pichardo M, Carpio-Pedroza JC, Ruiz-Pacheco JA, Cazares F, Vaughan G.** 2012. Specific detection of naturally occurring hepatitis C virus mutants with resistance to telaprevir and boceprevir (protease inhibitors) among treatment-naive infected individuals. J. Clin. Microbiol. **50**:281–287.
9. **Forns X, Purcell RH, Bukh J.** 1999. Quasispecies in viral persistence and pathogenesis of hepatitis C virus. Trends Microbiol. **7**:402–410.
10. **Rong L, Dahari H, Ribeiro RM, Perelson AS.** 2010. Rapid emergence of protease inhibitor resistance in hepatitis C virus. Sci. Transl. Med. **2**:30ra32.
11. **Ray SC, Fanning L, Wang XH, Netski DM, Kenny-Walsh E, Thomas DL.** 2005. Divergent and convergent evolution after a common-source outbreak of hepatitis C virus. J. Exp. Med. **201**:1753–1759.
12. **Martell M, Esteban JI, Quer J, Genesca J, Weiner A, Esteban R, Guardia J, Gomez J.** 1992. Hepatitis C virus (HCV) circulates as a population of different but closely related genomes: quasispecies nature of HCV genome distribution. J. Virol. **66**:3225–3229.
13. **Vizmanos JL, Gonzalez-Navarro CJ, Novo FJ, Civeira MP, Prieto J, Gullon A, Garcia-Delgado M.** 1998. Degree and distribution of variability in the 5′ untranslated, E1, E2/NS1 and NS5 regions of the hepatitis C virus (HCV). J. Viral Hepat. **5**:227–240.
14. **Hong SK, Cho SI, Ra EK, Kim EC, Park JS, Park SS, Seong MW.** 2012. Evaluation of two hepatitis C virus genotyping assays based on the 5′-untranslated region: the limitations of 5′ UTR based assays and the need for a supplementary sequencing-based approach. J. Clin. Microbiol. **50**:3741–3743.
15. **Ramachandran S, Campo DS, Dimitrova ZE, Xia GL, Purdy MA, Khudyakov YE.** 2011. Temporal variations in the hepatitis C virus intrahost population during chronic infection. J. Virol. **85**:6369–6380.
16. **Ramachandran S, Xia GL, Ganova-Raeva LM, Nainan OV, Khudyakov Y.** 2008. End-point limiting-dilution real-time PCR assay for evaluation of hepatitis C virus quasispecies in serum: performance under optimal and suboptimal conditions. J. Virol. Methods **151**:217–224.
17. **Escobar-Gutierrez A, Cruz-Rivera M, Ruiz-Tovar K, Vaughan G.** 2012. Genetic relatedness among Japanese HAV isolates, 2010. J. Clin. Virol. **54**:287–288.
18. **Beerenwinkel N, Zagordi O.** 2011. Ultradeep sequencing for the analysis of viral populations. Curr. Opin. Virol. **1**:413–418.
19. **Fonseca-Coronado S, Vaughan G, Cruz-Rivera MY, Carpio-Pedroza JC, Ruiz-Tovar K, Ruiz-Pacheco JA, Escobar-Gutierrez A.** 2011. Interleukin-28B genotyping by melt-mismatch amplification mutation assay PCR analysis using single nucleotide polymorphisms rs12979860 and rs8099917, a useful tool for prediction of therapy response in hepatitis C patients. J. Clin. Microbiol. **49**:2706–2710.
20. **Tamura K, Peterson D, Peterson N, Stecher G, Nei M, Kumar S.** 2011. MEGA5: molecular evolutionary genetics analysis using maximum likelihood, evolutionary distance, and maximum parsimony methods. Mol. Biol. Evol. **28**:2731–2739.
21. **Librado P, Rozas J.** 2009. DnaSP v5: a software for comprehensive analysis of DNA polymorphism data. Bioinformatics **25**:1451–1452.
22. **Farci P.** 2011. New insights into the HCV quasispecies and compartmentalization. Sem. Liver Dis. **31**:356–374.
23. **Grebely J, Pham ST, Matthews GV, Petoumenos K, Bull RA, Yeung B, Rawlinson W, Kaldor J, Lloyd A, Hellard M, Dore GJ, White PA, Grp AS.** 2012. Hepatitis C virus reinfection and superinfection among treated and untreated participants with recent infection. Hepatology **55**:1058–1069.
24. **Yuan HJ, Adams-Huet B, Petersen T, Attar N, Lee WM, Jain MK.** 2012. A single nucleotide polymorphism in IL28B affects viral evolution of hepatitis C quasispecies after pegylated interferon and ribavirin therapy. J. Med. Virol. **84**:1913–1919.
25. **Bull RA, Luciani F, McElroy K, Gaudieri S, Pham ST, Chopra A, Cameron B, Maher L, Dore GJ, White PA, Lloyd AR.** 2011. Sequential bottlenecks drive viral evolution in early acute hepatitis C virus infection. PLoS Pathog. **7**:e1002243. doi:10.1371/journal.ppat.1002243.
26. **Hiraga N, Imamura M, Abe H, Hayes CN, Kono T, Onishi M, Tsuge M, Takahashi S, Ochi H, Iwao E, Kamiya N, Yamada I, Tateno C, Yoshizato K, Matsui H, Kanai A, Inaba T, Tanaka S, Chayama K.** 2011. Rapid emergence of telaprevir resistant hepatitis C virus strain from wild-type clone in vivo. Hepatology **54**:781–788.
27. **Merani S, Petrovic D, James I, Chopra A, Cooper D, Freitas E, Rauch A, di Iulio J, John M, Lucas M, Fitzmaurice K, McKiernan S, Norris S, Kelleher D, Klenerman P, Gaudieri S.** 2011. Effect of immune pressure on hepatitis C virus evolution: insights from a single-source outbreak. Hepatology **53**:396–405.
28. **Verbinnen T, Van Marck H, Vandenbroucke I, Vijgen L, Claes M, Lin TI, Simmen K, Neyts J, Fanning G, Lenz O.** 2010. Tracking the evolution of multiple *in vitro* hepatitis C virus replicon variants under protease inhibitor selection pressure by 454 deep sequencing. J. Virol. **84**:11124–11133.
29. **Wang GP, Sherrill-Mix SA, Chang KM, Quince C, Bushman FD.** 2010. Hepatitis C virus transmission bottlenecks analyzed by deep sequencing. J. Virol. **84**:6218–6228.
30. **Lauck M, Alvarado-Mora MV, Becker EA, Bhattacharya D, Striker R, Hughes AL, Carrilho FJ, O'Connor DH, Pinho JRR.** 2012. Analysis of hepatitis C virus intrahost diversity across the coding region by ultradeep pyrosequencing. J. Virol. **86**:3952–3960.