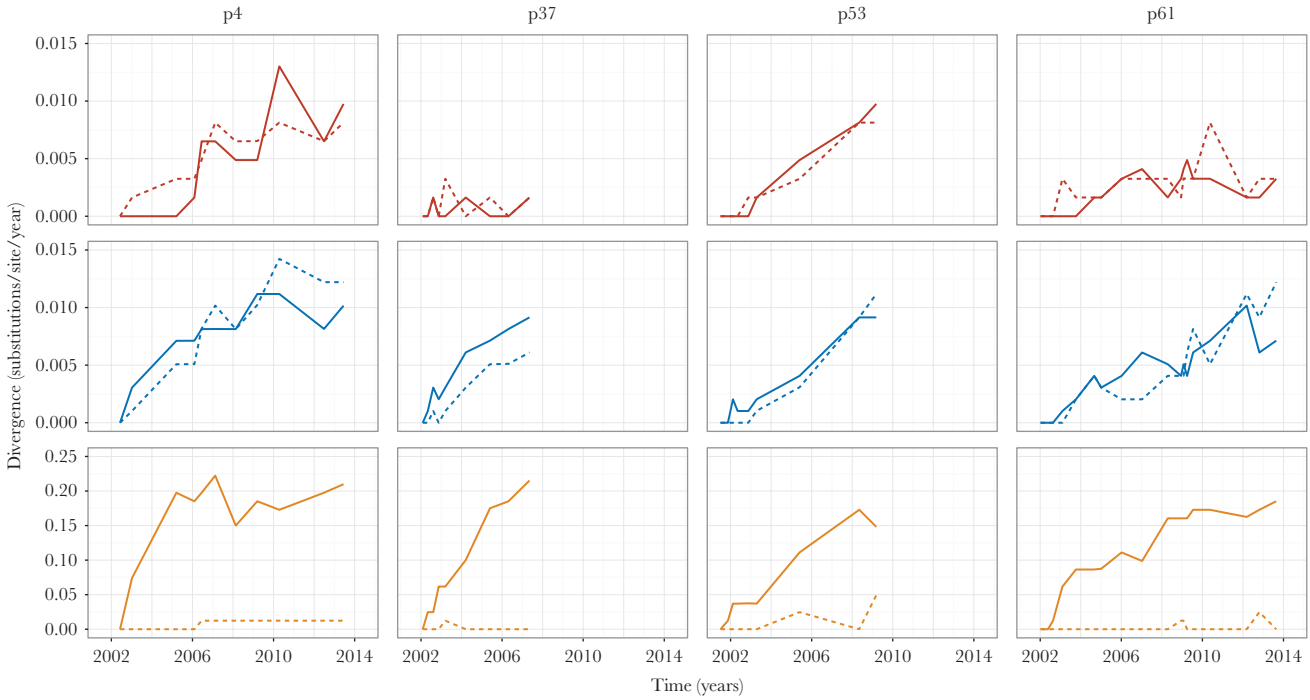

Figure 3. Divergence over time for each gene region (E1, E2, and HVR1 on top, middle, and bottom rows, respectively). Solid lines correspond to nonsynonymous divergence, while dashed lines correspond to synonymous divergence. Note, scales of the y-axes for HVR1 are different from those of E1 and E2.

purifying selection (Figure 3). The apparent lack of synonymous divergence in HVR1 is largely due to the differences in the scales of the y-axes for each subgenomic region. On closer inspection, the synonymous divergence in HVR1 was comparable to E1 and E2, although the pattern through time was more stochastic (Supplementary Figure 1). Lastly, while the gene regions largely diverged at an approximately constant rate over time, in individuals p4 and p61 we found that the nonsynonymous divergence in HVR1 slowed down after 1 to 2 years of infection (Figure 3).

The average within-host recombination rate during HCV infection, across all individuals, was estimated to be 0.28×10^{-7} recombinations/site/day (interquartile range: 0.13–1.05×10^{-7}). This is significantly lower than the inferred within-host substitution rate, which ranged from 2.05 to 8.21×10^{-5} substitutions/site/day. Furthermore, the HCV recombination rate is approximately 2 orders of magnitude lower than that observed for HIV-1 (estimated using the same method) [29]. The low within-host recombination rate indicates that strong linkage effects will influence the viral evolutionary dynamics during infection. Our estimate of recombination rate could be inflated by PCR recombination. If so, the true rate of within-host recombination will be even lower than that estimated here, further supporting our conclusions that strong linkage effects are likely to dominate HCV within-host evolution.

DISCUSSION

By examining the evolution of full-length E1E2 sequences from acute to chronic infection, we found that chronic HCV infection was consistent with independently replicating viral subpopulations [10, 12, 32] that were established from a single infecting viral strain, but not all were necessarily detected in the blood at all time points. We also found marked variation in the rates of evolution across the different regions of the envelope, and among individuals, for both synonymous and nonsynonymous mutations, combined with a very low rate of within-host recombination.

Given that longitudinal sampling of both the liver and blood from untreated HCV-infected individuals is unlikely to be feasible, or ethical, it has been difficult to test directly if chronic HCV infection is maintained by a structured population. To address this using an evolutionary approach, we assessed formally whether the observed within-host HCV population dynamics in the 4 individuals support a structured viral population characterized by 2 demes, where only of 1 these subpopulations is observed in the blood at any point in time, or with a single population whose population size varies over time. This analysis was possible due to the comparatively long reads in this dataset, which provided sufficient statistical power to evaluate the 2 alternative hypotheses of HCV population dynamics. For 3 individuals, we found a statistically better fit for a structured population. Although this result overall supports our original hypothesis, it is unclear why we did not find evidence for a structured

population in 1 of the individuals (p4) despite observing cocirculation of multiple viral lineages during chronic infection. The lack of divergence at later time points in this individual might explain this apparent discrepancy. In addition, more complex within-host HCV population dynamics than assumed by the structured coalescent, such as one or more viral subpopulations undergoing changes in population size, could explain the lack of support for a structured viral population. Notably, our finding of significant statistical support for a structured viral population for individual p53, where the within-host phylogeny was characterized by a single viral lineage at any point in time, underlines the importance of formal statistical testing to evaluate alternative hypotheses about the viral population dynamics, and highlights the potential problems of making inferences from visual inspection of within-host phylogenies alone.

Molecular evolutionary analyses of the E1E2 revealed very high rates of nonsynonymous divergence at the HVR1 region, which is consistent with this small genomic region undergoing strong immune-mediated selection [15, 30, 31, 33, 34]. In contrast, the rest of E1E2 is largely characterized by purifying selection. Although similar conclusions have been reported previously [9, 15, 30, 31], the combination of frequent sampling during HCV infection and long-read sequence data has enabled us to robustly compare the rates of molecular evolution among the different gene regions in the HCV envelope both within and between individuals.

In this study we also report, for the first time, an estimate of the within-host recombination rate of HCV during infection that can be directly compared with other evolutionary estimates. In contrast to HIV-1 [29], recombination appears to be a weak evolutionary force during within-host HCV evolution. This helps to explain why considerably fewer circulating recombinant forms are observed at the epidemiological scale for HCV than for HIV-1, even though HCV is more transmissible and mixed infections with distinct genotypes are relatively common. Furthermore, the within-host recombination rate is substantially lower than the overall substitution rate of E1E2, indicating that strong linkage effects are likely to shape the within-host viral evolutionary dynamics; selection is expected to be less effective in nonrecombining populations due to clonal interference [35] and background selection [36–38], thus reducing the rate of fixation of beneficial mutations in the population. A structured population can also limit viral adaptation if migration rates between viral subpopulations are low [39], since beneficial mutations will be restricted to the subpopulations in which they emerged, thereby preventing them from sweeping through the global population. Consequently, these evolutionary dynamics suggests that although drug resistance mutations might emerge during infection, fixation within individuals and transmission between individuals could be restricted.

To fully understand the extent to which within-host HCV populations are structured, and the effect that this, combined with low rates of recombination, has on HCV during infection, will require additional viral sequence data that has been serially sampled from a greater number of individuals. As well as collecting key clinical information, such as HLA background, viral load, and antibody responses, greater priority should be given to long-read, deep-sequenced data that spans the whole virus genome, because this will give the necessary power to determine whether, and how many, distinct viral subpopulations exist. This will be especially important for determining whether viral population structure is associated with strong selection or genetic drift, and will help to elucidate the relative contribution of cell-mediated and humoral immunity during HCV infection.

Supplementary Data

Supplementary materials are available at *The Journal of Infectious Diseases* online. Consisting of data provided by the authors to benefit the reader, the posted materials are not copyedited and are the sole responsibility of the authors, so questions or comments should be addressed to the corresponding author.

Notes

Financial support. This work was supported by the Wellcome Trust and the Royal Society (grant number 107652/Z/15/Z); and the European Research Council (ERC) under the European Union's Seventh Framework Programme (grant number FP7/2007-2013/ERC grant agreement no. 614725-PATHPHYLODYN).

Potential conflicts of interest. All authors: No reported conflicts of interest. All authors have submitted the ICMJE Form for Disclosure of Potential Conflicts of Interest. Conflicts that the editors consider relevant to the content of the manuscript have been disclosed.

References

1. Gower E, Estes C, Blach S, Razavi-Shearer K, Razavi H. Global epidemiology and genotype distribution of the hepatitis C virus infection. J Hepatol **2014**; 61:S45–57.
2. Pawlotsky JM. New hepatitis C therapies: the toolbox, strategies, and challenges. Gastroenterology **2014**; 146:1176–92.
3. Ramamurthy N, Marchi E, Ansari MA, et al. Impact of IFNL4 genotype on interferon-stimulated gene expression during DAA therapy for hepatitis C. Hepatology [published online ahead of print 13 March 2018]. doi: 10.1002/hep.29877.
4. Smith D, Magri A, Bonsall D, et al. Resistance analysis of genotype 3 hepatitis C virus indicates subtypes inherently resistant to nonstructural protein 5A inhibitors. Hepatology **2018** [published online ahead of print 9 February 2018]. doi: 10.1002/hep.29837.
5. Pybus OG, Rambaut A. Evolutionary analysis of the dynamics of viral infectious disease. Nat Rev Genet **2009**; 10:540–50.
6. Bartlett SR, Wertheim JO, Bull RA, et al. A molecular transmission network of recent hepatitis C infection in people with and without HIV: Implications for targeted treatment strategies. J Viral Hepat **2017**; 24:404–11.

7. Rose R, Lamers SL, Massaccesi G, et al. Complex patterns of hepatitis-C virus longitudinal clustering in a high-risk population. Infect Genet Evol **2018**; 58:77–82.
8. Bezemer D, Cori A, Ratmann O, et al.; ATHENA observational cohort. Dispersion of the HIV-1 epidemic in men who have sex with men in the Netherlands: a combined mathematical model and phylogenetic analysis. PLoS Med **2015**; 12:e1001898; discussion e1001898.
9. Ramachandran S, Campo DS, Dimitrova ZE, Xia GL, Purdy MA, Khudyakov YE. Temporal variations in the hepatitis C virus intrahost population during chronic infection. J Virol **2011**; 85:6369–80.
10. Gray RR, Salemi M, Klenerman P, Pybus OG. A new evolutionary model for hepatitis C virus chronic infection. Plos Pathog **2012**; 8:e1002656.
11. Gismondi MI, Díaz Carrasco JM, Valva P, et al. Dynamic changes in viral population structure and compartmentalization during chronic hepatitis C virus infection in children. Virology **2013**; 447:187–96.
12. Raghwani J, Rose R, Sheridan I, et al. Exceptional heterogeneity in viral evolutionary dynamics characterises chronic hepatitis C virus infection. PLoS Pathog **2016**; 12:e1005894.
13. Ho CK, Raghwani J, Koekkoek S, et al. Characterization of hepatitis C virus (HCV) envelope diversification from acute to chronic infection within a sexually transmitted HCV cluster by using single-molecule, real-time sequencing. J Virol **2017**; 91:pii: e02262-16.
14. Farci P, Quinti I, Farci S, et al. Evolution of hepatitis C viral quasispecies and hepatic injury in perinatally infected children followed prospectively. Proc Natl Acad Sci U S A **2006**; 103:8475–80.
15. Farci P, Shimoda A, Coiana A, et al. The outcome of acute hepatitis C predicted by the evolution of the viral quasispecies. Science **2000**; 288:339–44.
16. Farci P, Wollenberg K, Diaz G, et al. Profibrogenic chemokines and viral evolution predict rapid progression of hepatitis C to cirrhosis. Proc Natl Acad Sci U S A **2012**; 109:14562–7.
17. Alfonso V, Mbayed VA, Sookoian S, Campos RH. Intra-host evolutionary dynamics of hepatitis C virus E2 in treated patients. J Gen Virol **2005**; 86:2781–6.
18. Drummond AJ, Suchard MA, Xie D, Rambaut A. Bayesian phylogenetics with BEAUti and the BEAST 1.7. Mol Biol Evol **2012**; 29:1969–73.
19. Drummond AJ, Ho SYW, Phillips MJ, Rambaut A. Relaxed phylogenetics and dating with confidence. Plos Biol **2006**; 4:e88.
20. Shapiro B, Rambaut A, Drummond AJ. Choosing appropriate substitution models for the phylogenetic analysis of protein-coding sequences. Mol Biol Evol **2006**; 23:7–9.
21. Gill MS, Lemey P, Faria NR, Rambaut A, Shapiro B, Suchard MA. Improving Bayesian population dynamics inference: a coalescent-based model for multiple loci. Mol Biol Evol **2013**; 30:713–24.
22. Yu GC, Smith DK, Zhu HC, Guan Y, Lam TTY. GGTREE: an R package for visualization and annotation of phylogenetic trees with their covariates and other associated data. Methods Ecol Evol **2017**; 8:28–36.
23. Wickham H. ggplot2: elegant graphics for data analysis. New York: Springer-Verlag, **2009**.
24. De Maio N, Wu CH, O'Reilly KM, Wilson D. New routes to phylogeography: a bayesian structured coalescent approximation. Plos Genet **2015**; 11:e1005421.
25. Drummond AJ, Rambaut A, Shapiro B, Pybus OG. Bayesian coalescent inference of past population dynamics from molecular sequences. Mol Biol Evol **2005**; 22:1185–92.
26. Bouckaert R, Heled J, Kuhnert D, et al. BEAST 2: a software platform for bayesian evolutionary analysis. PLoS Comput Biol **2014**; 10:e1003537.
27. Hall MD, Woolhouse MEJ, Rambaut A. The effects of sampling strategy on the quality of reconstruction of viral population dynamics using Bayesian skyline family coalescent methods: a simulation study. Virus Evol **2016**; 2:vew003.
28. Xie W, Lewis PO, Fan Y, Kuo L, Chen MH. Improving marginal likelihood estimation for Bayesian phylogenetic model selection. Syst Biol **2011**; 60:150–60.
29. Neher RA, Leitner T. Recombination rate and selection strength in HIV intra-patient evolution. PLoS Comput Biol **2010**; 6:e1000660.
30. Sheridan I, Pybus OG, Holmes EC, Klenerman P. High-resolution phylogenetic analysis of hepatitis C virus adaptation and its relationship to disease progression. J Virol **2004**; 78:3447–54.
31. Brown RJ, Juttla VS, Tarr AW, et al. Evolutionary dynamics of hepatitis C virus envelope genes during chronic infection. J Gen Virol **2005**; 86:1931–42.
32. Gray RR, Strickland SL, Veras NM, et al. Unexpected maintenance of hepatitis C viral diversity following liver transplantation. J Virol **2012**; 86:8432–9.
33. Gray RR, Parker J, Lemey P, Salemi M, Katzourakis A, Pybus OG. The mode and tempo of hepatitis C virus evolution within and among hosts. BMC Evol Biol **2011**; 11:131.
34. Bailey JR, Laskey S, Wasilewski LN, et al. Constraints on viral evolution during chronic hepatitis C virus infection arising from a common-source exposure. J Virol **2012**; 86:12582–90.
35. Miralles R, Gerrish PJ, Moya A, Elena SF. Clonal interference and the evolution of RNA viruses. Science **1999**; 285:1745–7.
36. Barton NH. Linkage and the limits to natural selection. Genetics **1995**; 140:821–41.
37. Johnson T, Barton NH. The effect of deleterious alleles on adaptation in asexual populations. Genetics **2002**; 162:395–411.
38. Peck JR. A Ruby in the rubbish: beneficial mutations, deleterious mutations and the evolution of sex. Genetics **1994**; 137:597–606.
39. Slatkin M. Fixation probabilities and fixation times in a subdivided population. Evolution **1981**; 35:477–88.