

Opinion

A New Evolutionary Model for Hepatitis C Virus Chronic Infection

Rebecca R. Gray¹, Marco Salemi², Paul Klenerman³, Oliver G. Pybus^{1*}

1 Department of Zoology, University of Oxford, Oxford, United Kingdom, **2** Department of Pathology, Immunology & Laboratory Medicine, University of Florida, Gainesville, Florida, United States of America, **3** Nuffield Department of Medicine, University of Oxford, Oxford, United Kingdom

Hepatitis C virus (HCV) infects an estimated 3% of humanity [1] and is a leading global cause of liver disease and liver cancer [2]. Intervention is currently limited by the lack of a vaccine and of universally successful drug treatments. Although several next-generation drugs (e.g., direct-acting protease-inhibitors) are already improving outcomes, a number of factors will affect overall treatment success [3]. Among these, viral genetic variation and the emergence of drug resistance are of major importance.

It is now recognized that effective prevention and treatment of HCV infection requires an appreciation of the virus' evolutionary behaviour. HCV evolves very rapidly during infection within a host, resulting in a genetically diverse viral population (often termed a “quasispecies”) whose composition is determined by a combination of evolutionary processes that include mutation, replication rate, natural selection, and random genetic drift. As a consequence, patterns of HCV genetic diversity within infected patients (if correctly analysed) should reveal information about the evolutionary dynamics of infection and could uncover links between viral evolution and the progression of clinical disease.

Studies that have followed this logic are too numerous to review here and vary widely in the number and health status of patients investigated, the timespan over which viruses are sampled, and the quality of viral genetic information obtained (ranging from whole viral genomes to heteroduplex mobility assay data). In most studies, HCV genomic sequences are obtained from peripheral blood and are summarized using two measures: the *diversity* of sequences sampled at any one time and the *divergence* among sequences sampled at different times. No clear-cut patterns in these statistics among patients with HCV have emerged. For example, with respect to patient outcomes, high diversity in the HCV envelope region is associated with progression from acute to chronic infection [4,5] yet also with milder symptoms [6,7] and possibly with poorer outcomes after drug treatment [8].

It is helpful to interpret these results in the context of evolutionary theory. First, it is well-established that much of the evolutionary information inherent in sequence data is lost when they are compressed into summary statistics such as diversity or divergence [9,10]. In retrospect, it is perhaps optimistic to expect one or two numerical values to capture much of the complexity of host–viral interactions during infection. Second, and more importantly, theory tells us that an evolutionary *model* is always required to infer the behaviour of a viral population from sequences sampled from it [11,12]. When previous HCV studies have correlated diversity with patient outcomes, they have, by default, implicitly applied a very simple evolutionary model—one that assumes a single, well-mixed viral population whose members all replicate and evolve identically.

However, phylogenetic analyses of within-patient HCV sequences indicate that this simple model is insufficient to explain HCV evolutionary dynamics during infection. Specifically, phylogenies reconstructed from HCV genomes sampled from multiple time points frequently reveal two or more genetically distinct co-existing viral lineages that are not detected at all sampling times; this is seen in both acute [13] and chronic (e.g., [14–17]) infection. Figure 1a presents an illustrative within-patient phylogeny, in which lineages 3 and 4 co-exist for more than 10 years. Crucially, the composition of viruses sampled from peripheral blood varies greatly; at one time point only lineage 3 is detected, whilst at the next only lineage 4 might be seen; at yet other times both

lineages are observed. The phylogeny thus explains why simple measures of sample diversity oscillate wildly (Figure 1b). By applying more sophisticated methods (see figure legend and [11]), it becomes possible to estimate the genetic diversity of the *whole* viral population, which is comparatively constant through time (Figure 1c). We therefore suggest that the diversity scores (Figure 1b) commonly used in HCV studies do not fully represent the dynamics of infection.

The behaviour illustrated in Figure 1 requires a more complex model than is commonly assumed. We believe the most important feature missing from current descriptions of HCV dynamics is *population structure*; without this, it is very difficult to explain why viral lineages that remain unobserved for several years do not go extinct before later reappearing (Figure 1). Consequently, we propose a new evolutionary model for HCV, in which the lineages that co-exist during chronic infection represent genetically distinct subpopulations of infected liver cells (Figure 1d; extra-hepatic replication is discussed below). Importantly, this model can reconcile many aspects of HCV within-host genetic data (i.e., the data in Figure 1a–1c).

The observation that only a subset of viral subpopulations is detected in serum at any given time may be explained by one of two modulating mechanisms. First, all viral lineages present in the liver may be shed at a roughly constant rate, but levels of neutralizing antibodies targeting specific epitopes may fluctuate over time, modifying the relative frequency of those lineages

Citation: Gray RR, Salemi M, Klenerman P, Pybus OG (2012) A New Evolutionary Model for Hepatitis C Virus Chronic Infection. PLoS Pathog 8(5): e1002656. doi:10.1371/journal.ppat.1002656

Editor: Glenn F. Rall, The Fox Chase Cancer Center, United States of America

Published: May 3, 2012

Copyright: © 2012 Gray et al. This is an open-access article distributed under the terms of the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.

Funding: This work was funded by The Royal Society and MRC UK. The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

Competing Interests: The authors have declared that no competing interests exist.

* E-mail: oliver.pybus@zoo.ox.ac.uk

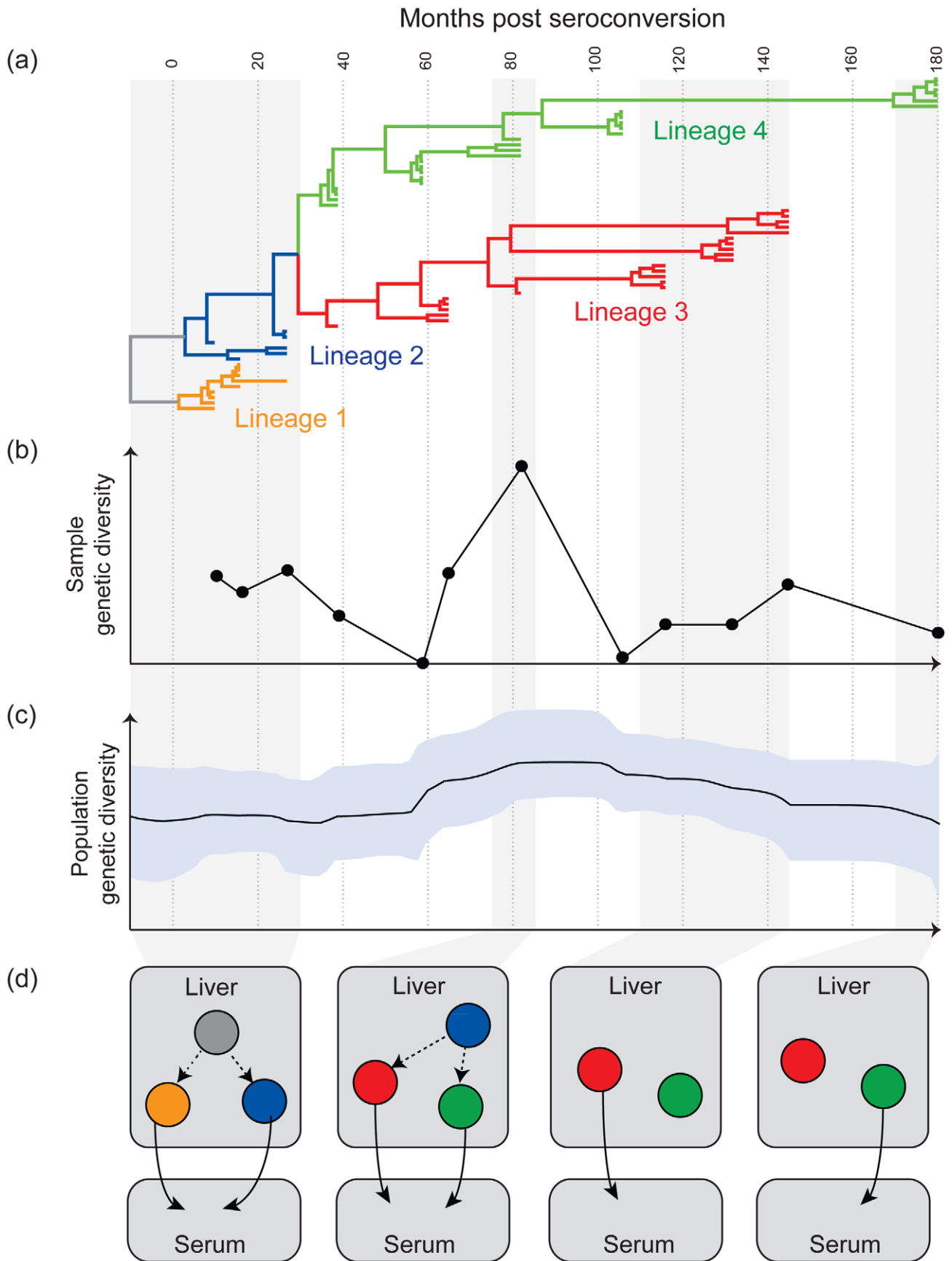


Figure 1. Aspects of the evolutionary dynamics of chronic HCV infection. (a), (b), and (c) show various analyses of HCV gene sequences (E1E2 region) sampled longitudinally from a single infected individual (Pt11 in [7]). Sequences were obtained from serum at twelve occasions over 15 years. All results are placed on the same time scale (top). (a) The phylogeny of the sampled sequences was reconstructed using a molecular clock model, such that branch lengths represent time (calculated using BEAST v1.6.2 [30]). To aid explanation, branches have been grouped into four lineages, indicated by colour. Lineages 3 and 4 co-exist between months 30 and 145. Lineage 4 was present in the patient but undetected at months 65, 116, 132, 145. (b) The average diversity of sequences obtained at each sampling time (mean pairwise genetic distance; calculated using MEGA v4 [31]). The vertical scale represents mean substitutions per site. Sample diversity is notably higher at month 82 because both lineages 3 and 4 are detected. (c) An estimate of the genetic diversity of the whole viral quasispecies through time. This figure was calculated using the Bayesian skyline plot method in BEAST [30,32], which takes into account both sampled and unsampled lineages. The shaded area shows the 95% uncertainty range around the estimate (solid line). (d) A new evolutionary model of HCV infection, as applied to the patient data shown in (a), (b), and (c). Each cartoon represents the state of infection at a different time, as indicated by the shaded areas. Circles represent different sub-populations of HCV-infected cells within the liver (or other sites of extra-hepatic replication), coloured to correspond to the lineages in (a). Solid arrows indicate the fluctuating detection of HCV lineages in serum through time.

doi:10.1371/journal.ppat.1002656.g001

in peripheral blood. Second, the viral subpopulations may replicate or shed virus at different rates. Factors that might generate such variation include the effect of host cell type on viral replication dynamics, interaction with interferons, or viral interference in the cell cycle [18]. The presence of replication rate variation is supported by mathematical models of HCV infection kinetics [19] and may help explain why HCV quasispecies exhibit strong heterogeneity in the rate of molecular evolution [20].

The notion of HCV population structure is consistent with a wide variety of experimental data, including evidence for cell-to-cell viral transmission [21,22], the

observation of hepatic foci of infection [23], and HCV genetic variation within the liver [24]. Although our illustration (Figure 1d) represents viral sub-populations existing in different liver locations, the population structure we propose could equally arise from non-hepatic replication, as detected in monocytes/macrophages [25], lymphocytes [26], and brain tissue [27]. These cells often harbor viruses with distinct genetic signatures [28,29].

The under-appreciated complexity of chronic HCV evolution has several practical consequences. First, simple statistics of viral genetic variation (sample diversity, divergence) will be of limited utility; we instead recommend that analyses begin

with a phylogenetic approach. Second, our model suggests that viruses sampled from peripheral blood give an incomplete picture of HCV infection dynamics. Where possible, future studies should incorporate additional sites, including liver biopsies, explant livers, and other cell types. Third, population structure can maintain high viral diversity within a patient even when serum viraemia and diversity are low, perhaps contributing to the evolution of drug resistance and to treatment failure. The emergence of new anti-HCV drugs therefore makes the understanding of HCV evolutionary dynamics a research priority.

References

- WHO (2010) HCV: surveillance and control. Available: <http://www.who.int/csr/disease/hepatitis/whocdscslyo2003/en/index4.html>. Accessed 6 April 2012.
- CDC (2010) Hepatocellular carcinoma - United States, 2001–2006. *MMWR Morb Mortal Wkly Rep* 59: 517–552.
- Kuntzen T, Timm J, Berical A, Lennon N, Berlin AM, et al. (2008) Naturally occurring dominant resistance mutations to hepatitis C virus protease and polymerase inhibitors in treatment-naïve patients. *Hepatology* 48: 1769–1778.
- Farci P, Shimoda A, Coiana A, Diaz G, Peddis G, et al. (2000) The outcome of acute hepatitis C predicted by the evolution of the viral quasispecies. *Science* 288: 339–344.
- Thomson EC, Fleming VM, Main J, Klenerman P, Weber J, et al. (2011) Predicting spontaneous clearance of acute hepatitis C virus in a large cohort of HIV-1-infected men. *Gut* 60: 837–845.
- Sullivan D, Bruden D, Deubner H, McArdle S, Chung M, et al. (2007) Hepatitis C virus dynamics during natural infection are associated with long-term histological outcome of chronic hepatitis C disease. *J Infect Dis* 196: 239–248.
- Farci P, Quinti I, Farci S, Alter HJ, Strazzera R, et al. (2006) Evolution of hepatitis C viral quasispecies and hepatic injury in perinatally infected children followed prospectively. *Proc Natl Acad Sci U S A* 103: 8475–8480.
- Morishima C, Polyak SJ, Ray R, Doherty MC, Di Bisceglie AM, et al. (2006) Hepatitis C virus-specific immune responses and quasi-species variability at baseline are associated with nonresponse to antiviral therapy during advanced hepatitis C. *J Infect Dis* 193: 931–940.
- Felsenstein J (1992) Estimating effective population size from samples of sequences: inefficiency of pairwise and segregating sites as compared to phylogenetic estimates. *Genet Res* 59: 139–147.
- Steel MA, Hendy MD, Penny D (1988) Loss of information in genetic distances. *Nature* 336: 118.
- Lemey P, Rambaut A, Pybus OG (2006) HIV evolutionary dynamics within and among hosts. *AIDS Rev* 8: 125–140.
- Wilson DJ, Falush D, McVean G (2005) Germs, genomes and genealogies. *Trends Ecol Evol* 20: 39–45.
- Smith JA, Aberle JH, Fleming VM, Ferenci P, Thomson EC, et al. (2010) Dynamic coinfection with multiple viral subtypes in acute hepatitis C. *J Infect Dis* 202: 1770–1779.
- Alfonso V, Mbayed VA, Sookoian S, Campos RH (2005) Intra-host evolutionary dynamics of hepatitis C virus E2 in treated patients. *J Gen Virol* 86: 2781–2786.
- Li H, McMahon BJ, McArdle S, Bruden D, Sullivan DG, et al. (2008) Hepatitis C virus envelope glycoprotein co-evolutionary dynamics during chronic hepatitis C. *Virology* 375: 580–591.
- Wang X, Netski DM, Astemborski J, Mehta SH, Torbenson MS, Thomas DL, et al. (2007) Progression of fibrosis during chronic hepatitis C is associated with rapid virus evolution. *J Virol* 81: 6513–6522.
- Ramachandran S, Campo DS, Dimitrova ZE, Xia GL, Purdy MA, et al. (2011) Temporal variations in the hepatitis C virus intrahost population during chronic infection. *J Virol* 85: 6369–6380.
- Kannan RP, Hensley LL, Evers LE, Lemon SM, McGovern DR, et al. (2011) Hepatitis C virus infection causes cell cycle arrest at the level of initiation of mitosis. *J Virol* 85: 7989–8001.
- Neumann AU, Lam NP, Dahari H, Gretch DR, Wiley TE, et al. (1998) Hepatitis C viral dynamics in vivo and the antiviral efficacy of interferon-alpha therapy. *Science* 282: 103–107.
- Gray RR, Parker J, Lemey P, Salemi M, Katzourakis A, et al. (2011) The mode and tempo of hepatitis C virus evolution within and among hosts. *BMC Evol Biol* 11: 131.
- Liang Y, Shilgard T, Xiao SY, Snyder N, Lau D, et al. (2009) Visualizing hepatitis C virus infections in human liver by two-photon microscopy. *Gastroenterology* 137: 1448–1458.
- Brimacombe CL, Grove J, Meredith LW, Hu K, Syder AJ, et al. (2011) Neutralizing antibody-resistant hepatitis C virus cell-to-cell transmission. *J Virol* 85: 596–605.
- Stiffler JD, Nguyen M, Sohn JA, Liu C, Kaplan D, et al. (2009) Focal distribution of hepatitis C virus RNA in infected livers. *PLoS ONE* 4: e6661.
- Sobesky R, Feray C, Rimlinger F, Derian N, Dos Santos A, et al. (2007) Distinct hepatitis C virus core and E protein quasispecies in tumoral and nontumoral hepatocytes isolated via microdissection. *Hepatology* 46: 1704–1712.
- Laskus T, Radkowski M, Piasek A, Nowicki M, Horban A, et al. (2000) Hepatitis C virus in lymphoid cells of patients coinfecting with human immunodeficiency virus type 1: evidence of active replication in monocytes/macrophages and lymphocytes. *J Infect Dis* 181: 442–448.
- Laskus T, Operskalski EA, Radkowski M, Wilkinson J, Mack WJ, et al. (2007) Negative-strand hepatitis C virus (HCV) RNA in peripheral blood mononuclear cells from anti-HCV-positive/HIV-infected women. *J Infect Dis* 195: 124–133.
- Radkowski M, Wilkinson J, Nowicki M, Adair D, Vargas H, et al. (2002) Search for hepatitis C virus negative-strand RNA sequences and analysis of viral sequences in the central nervous system: evidence of replication. *J Virol* 76: 600–608.
- Ducoulombier D, Roque-Alfonso AM, Di Liberto G, Penin F, Kara R, et al. (2004) Frequent compartmentalization of hepatitis C

- virus variants in circulating B cells and monocytes. *Hepatology* 39: 817–825.
29. Roque-Afonso AM, Ducoulombier D, Di Liberto G, Kara R, Gigou M, et al. (2005) Compartmentalization of hepatitis C virus genotypes between plasma and peripheral blood mononuclear cells. *J Virol* 79: 6349–6357.
 30. Drummond AJ, Rambaut A (2007) BEAST: Bayesian evolutionary analysis by sampling trees. *BMC Evol Biol* 7: 214.
 31. Tamura K, Dudley J, Nei M, Kumar S (2007) MEGA4: Molecular Evolutionary Genetics Analysis (MEGA) software version 4.0. *Mol Biol Evol* 24: 1596–1599.
 32. Drummond AJ, Rambaut A, Shapiro B, Pybus OG (2005) Bayesian coalescent inference of past population dynamics from molecular sequences. *Mol Biol Evol* 22: 1185–1192.