Hepatitis C virus quasispecies and pseudotype analysis from acute infection to chronicity in HIV-1 co-infected individuals

R. Bridget Ferns[a,d*], Alexander W. Tarr[b], Stephane Hue[a], Richard A. Urbanowicz[b],

C. Patrick McClure[b], Richard Gilson[c], Jonathan K. Ball[b], Eleni Nastouli[d], Jeremy A. Garson[a]

and Deenan Pillay[a,e]

[a] Division of Infection & Immunity, Faculty of Medical Sciences, University College London

[b] School of Life Sciences & NIHR Biomedical Research Unit in Gastrointestinal & Liver

Diseases, Faculty of Medicine and Health Sciences, University of Nottingham

[c] Research Department of Infection and Population Health, University College London

[d] Clinical Microbiology & Virology, UCL Hospital NHS Foundation Trust

[e] Wellcome Trust Africa Centre for Health and Population Sciences, University of KwaZulu-

Natal, South Africa

* Correspondening author: r.ferns@ucl.ac.uk Clinical Microbiology and Virology, UCL

Hospital NHS Foundation Trust, 60 Whitfield Street, London, W1T 4EL, UK

Abstract

HIV-1 infected patients who acquire HCV infection have higher rates of chronicity and liver disease progression than patients with HCV mono-infection. Understanding early events in this pathogenic process is important. We applied single genome sequencing of the E1 to NS3 regions and viral pseudotype neutralization assays to explore the consequences of viral quasispecies evolution from pre-seroconversion to chronicity in four co-infected individuals (mean follow up 566 days). We observed that one to three founder viruses were transmitted. Relatively low viral sequence diversity, possibly related to an impaired immune response, due to HIV infection was observed in three patients. However, the fourth patient, after an early purifying selection displayed increasing E2 sequence evolution, possibly related to being on suppressive antiretroviral therapy. Viral pseudotypes generated from HCV variants showed relative resistance to neutralization by autologous plasma but not to plasma collected from later time points, confirming ongoing virus escape from antibody neutralization.

Introduction

Infection with hepatitis C virus (HCV) becomes chronic in around 60-80% of patients (Santantonio et al., 2008). Only about one fifth of patients are able to spontaneously clear infection within six months. Mechanisms by which HCV is able to evolve and avoid the host defense systems are now beginning to be identified (Jackowiak et al., 2014). Attention has focused on the envelope glycoproteins, E1 and E2, as key mediators of HCV entry. Structural analysis of E2 (Khan et al., 2014, Kong et al., 2013) has revealed antibody and receptor binding sites. Virus entry into hepatocytes occurs as a complex cascade requiring the cellular

factors CD81 (Zhang et al., 2004), scavenger receptor BI (SR-B1, Bartosch et al., 2003c) and two tight junction proteins claudin-1 (Evans et al., 2007) and occludin (Ploss et al., 2009). EGFR and EphA2 also contribute to entry (Lupberger et al., 2011). The two glycoproteins are major targets of neutralizing antibodies (Ball et al., 2014). A broadly-neutralizing antibody response contributes to clearance of acute infection (Dowd et al., 2008, Pestka et al., 2007, Raghuraman et al., 2012) and protection against reinfection (Osburn et al., 2010). Neutralizing activity of serum antibodies is also inversely correlated with viral titer in acute HCV infection (Lavillette et al., 2005). Greater understanding of virus heterogeneity and entry will inform the design of drug entry inhibitors and vaccines for induction of broadly effective immunity.

The ability of the HCV error prone RNA-dependent RNA polymerase to generate genetic diversity is well reported (Abe et al., 1992, Ribeiro et al., 2012, Steinhauer et al., 1992). The generation of many variants (quasispecies) in the presence of host-directed pressure allows the "most fit" species to become dominant (Bull et al., 2011). For identification of genetically divergent quasispecies produced in clinical infections, single genome amplification (SGA, Keele et al., 2008, Simmonds et al., 1990), although labour intensive, is the method of choice and has significant advantages over cloning. DNA polymerases used to amplify mixed populations of variants prior to cloning are able to switch templates during amplification and produce recombinants *in vitro* which do not actually exist in the original sample. By diluting the viral cDNA sample to a level where only a single template/variant molecule is present, the SGA method virtually eliminates this template switching problem. SGA has previously been used to examine HCV founder viruses present during the initial phase of infection (Li et al., 2012, Ribeiro et al., 2012, Stoddard et al., 2015).

Transmission of HCV occurs in high risk groups often already burdened with other infections, notably HIV-1. This has now reached epidemic significance (Arends et al., 2015, Fierer et al., 2008, Gambotti et al., 2005, Gilleece and Sullivan, 2005, Thomson et al., 2009, Turner et al., 2010, Vogel et al., 2005) and the outcome of dual infection may be considerably worse than HCV infection alone (Chen et al., 2014, Thomas, 2002). The factors that influence HCV disease progression in HIV co-infected individuals are still poorly defined.

We hypothesized that most of the sequence evolution would occur in the HCV E2 envelope gene, specifically the hypervariable region 1 (HVR1, Bankwitz et al., 2010, Farci et al., 2000, von Hahn et al., 2007). In addition to E2 we also investigated regions E1, p7, NS2 and NS3 in the plasma from four HCV/HIV co-infected patients at multiple time points from pre-seroconversion to at least eight months, post infection. The capacity of these viruses to induce host antibody neutralization was assessed using the HCV pseudotype particle (HCVpp) expression system (Bartosch et al., 2003a, Bartosch et al., 2003b, Dowd et al., 2009, Tarr et al., 2007, Tarr et al., 2012).

Results

Study design and patient characteristics

Plasma samples from four patients recently infected with HCV were studied; all were HIV-1 infected men who have sex with men (MSM). For each patient a series of plasma samples was collected from pre-seroconversion through to the establishment of chronic infection. Three patients (B, C and M) had longstanding HIV-1 infections while one (D) had a recent HIV-1 seroconversion (40 days before time point 1, Tp1) and may have acquired both viruses as part of the same transmission event. Patient characteristics and laboratory investigations

including HCV seroconversion (anti-HCV), HCV RNA and genotype, ALT, CD4 T-cell counts, HIV-1 viral load and treatment history are summarised in Fig. 1 and Table 1.

HCV infection was acquired by sexual transmission in all patients and there was no history of injecting drug use or other parenteral exposure. None cleared their HCV infection spontaneously during follow up (at least 8 months, Fig.1). For all patients a drop in HCV viraemia at or around anti-HCV seroconversion coincided with a peak ALT level. The latter probably reflected HCV infected cell death at around this time. A second rise in ALT (patient D) between Tp3 and Tp4 (Fig. 1) was investigated by re-genotyping the HCV 5'untranslated region (UTR) region at these time points to exclude superinfection and all UTR sequences were identical. Additionally, comparison of all SGA variants generated from patient D revealed a 92-93 % similarity to the GenBank EU234064.2 sequence confirming the 1a genotyping and providing no evidence of superinfection. However, it is possible that superinfection with the same genotype may have occurred. Chronic HBV infection with low level HBV viraemia was also noted in patient B.

Overall HCV sequence diversity during the course of infection

Variant sequences from the four patients were analysed at several time points from early infection, between 18-42 days post the "estimated time of infection" (ETI- see below for definition) for Tp1, through to chronicity (follow up for 249-930 days; mean 566, Table 1). Only single template variants were included in the final analysis and there was no evidence of recombination (P <0.05) in any of the variant sequences analysed. Due to sequence variability in parts of the genome, several different primers were required to amplify the E1-NS3 regions of HCV from the four patients (Table S1) and thus the precise location and length (bp) of the amplicons differed to some extent (Table S3, column 6). A total of 241

SGA-derived variants were studied with a mean of 16 (range 9-23) sequences per patient per time point (Table S2, column 5). Later time points showed higher numbers of mutations in accordance with the expected increase in genetic complexity over time, especially in variants analysed from patients B and D (Table S2 and highlighter plots, Fig. S1).

Analysis of transmitted/founder (T/F) viruses

Using the Poisson-Fitter model and a mutation rate of 2.5×10^{-5} (Ribeiro et al., 2012) the viral diversity at Tp1 was analysed to examine whether the variants had a star-like (single T/F origin) distribution. One patient (M) revealed a non-star like pattern, patient B a near star-like pattern whereas two (D and C) had star-like phylogeny (Table S3). Unrooted tree analysis (Fig. 2) also supported the random distribution of mutations consistent with a star-like phylogeny for patients B, D and C. For patient M, three groups or clades of virus variants were discernible confirming the non-star like Poisson-Fitter result and suggesting transmission of three T/F viruses. Using the Poisson-Fitter model the days since the most recent common ancestor (MRCA) were estimated (Table S3) and approximately fitted the ETI for patients C, D and M (Table 1). For patient B, instead of 30 days since the estimated time of infection, an MRCA value of 92 days was calculated. However, using a slightly higher mutation rate of 6.2×10^{-5}, but within the reported range (Ribeiro et al., 2012), the two values became compatible (37 days with a range of 33-42).

Pairwise *t*-test analysis of genetic distance (GD) values at Tp1 (Table S2) also confirmed that the population diversity of the virus in patient B was significantly greater (P <0.0001) at the 95% level, than in patients C, D and M. The variants present in patient M were also significantly more diverse than those in patients C and D (P < 0.012, <0.0003, respectively).

Viral diversity and the appearance of a "population bottleneck"

For patient B high variant diversity early on in the infection led, through a purifying selection, to a cluster of mainly homogeneous variants (a "bottleneck") which slowly diversified to give a heterogeneous collection of variants (Tp4) with three branches showing >70% bootstrap support (Fig. 3). For patient D an increase in viral diversity occurred among the Tp4 variants (Fig. 3 and Table S2). In contrast to the phylogenetic trees derived for patients B and D, the trees for patients M and C failed to show evidence of positive selection (Fig. 3). The HCV sequences obtained at different time points from patients C and M were intermingled. It is noteworthy that of the three (T/F) viruses identified in patient M, only one gave rise to the sequences present at Tp2 and Tp3.

Highlighter plots (Fig. S1) and Shannon Entropy data analysis (Fig. S2) illustrate synonymous and non-synonymous mutations, by HCV region, with the highest sequence complexity in the p7 (patient D, Tp4) and the E2 (patient B, Tp4) regions of the virus. All of the variants from patient D had different p7 region nucleotide and amino acid substitutions which failed to become established. The lowest entropy values were obtained from analysis of variants from patient C, especially in the E2 region.

Amino acid changes within immune recognition and CD81 binding sites

To identify locations targeted by the immune response at T helper cell, B cell and CTL epitopes, HCV variant amino acid substitutions where selection occurred were recorded by HCV region (Tables 2 and 3). Single amino acid changes which were not observed at subsequent time points were not recorded as they were considered to have a neutral effect on virus fitness.

Analysis of variants generated from patient B revealed the establishment of six amino acid mutations (V384T, A395T, A396T, I399F, G401S, P405R) within the E2 HVR1 at known T helper cell, B cell and CTL epitopes by the end of follow up. Three other E2 substitutions (outside HVR1) were present in at least 89.5% of variants by Tp4 (K408Q, 94.5%, H444Y 89.5% and I603V, 94.7%). Only one NS2 region mutation, W841R, at a CTL epitope became established. By Tp4 greater than 90% of virus variants from patient B had acquired nine non-synonymous mutations in the E2 region. This patients HIV RNA remained undetectable, on antiretroviral therapy (ART), until Tp4 (2.83 $\log_{10}$ copies/ml).

Substitution at amino acid residue 331 in the E1 region of variants from patient D was at a known T helper cell and CTL receptor recognition site (T331A, 5% at Tp1 increasing to 100% by Tp4). Three E2 substitutions were first seen at Tp4 (930 days post ETI) namely F397S, A475V, V696I. The V696I substitution (70%) which has not been previously reported, was identified by positive selection analysis (positive probability 0.96). A substitution at position 1113, at a CTL epitope in the NS3 region, reverted from a mixed A1113V/T population (Tp1, Tp2, and Tp3) back to A1113 (100%, Tp4). This site was also identified as under positive selection by the HyPhy: FUBAR algorithm (0.98).

For patient C only a few mutations occurred in the NS3 region but none of these variants became established. Thus a relatively stable population of HCV variants were present and the immune selective pressure on the virus was probably low. This patient was followed for only four months after HCV seroconversion and was treated successfully at around Tp4.

No mutations became established in HCV variants from patient M. Only at E837V/K in the NS2 region at a CTL epitope was positive selection seen (0.97 by FUBAR). The E837 polymorphism reverted back to become the dominant variant by last time point. For patient M, variants at later time points were identified as being related to only one of the three T/F viruses (shown at the top of the unrooted tree for patient M, Fig. 2). The other two groups of T/F viruses (bottom left and right Fig. 2) had a Q743P (E2) and/or an E837V (NS2) substitution (Table 2).

Phenotype of the E1 and E2 proteins from majority variants

A single majority variant was chosen as representative of the majority sequences at each time point by distance matrix analysis of the E1/E2 region. This variant was cloned, re-sequenced to confirm no errors had been introduced and expressed on the surface of retroviral HCVpp. Differences were observed in the HCVpp infectivity of the majority clones, with the sequences isolated from patient B resulting in a 10-fold greater infectivity of the Huh7 hepatoma cells than either patient M, patient C or the positive control genotype 1a H77c clone (Yanagi et al., 1997, Fig. 4A).

Neutralization of each majority variant HCVpp sequence was then performed with each Tp plasma from the same patient (Fig.4B). For patient B, plasma (~1:1000 dilution) from Tp3 collected at six months post seroconversion, neutralized (~50% reduction compared to uninhibited control preparation) the B1 HCVpp-Tp1 and B2 HCVpp-Tp2 majority variants. Tp4 plasma (~1:1000), collected at 20 months post seroconversion, neutralized ($\geq$50% reduction compared to uninhibited control preparation) the majority variants from B1 HCV-Tp1, B2 HCVpp-Tp2 and B3 HCV-Tp3. Thus at 1:1000 dilution the antibodies present in plasma Tp3 and plasma Tp4 were able to neutralize earlier HCVpp variants but not the

autologous majority variant. The sequence cloned from Tp4 (B4 HCVpp-Tp4) was as infectious as previous time points in the viral entry assay (Fig. 4A), suggesting that there was no significant fitness cost. Interestingly, the presence of HCV-negative control plasma resulted in significantly enhanced infectivity (above the uninhibited control preparation), most notably for the HCVpp representing Tp4 (Fig. 4B, B4 HCVpp-Tp4). This highlights the potential for selection of glycoproteins that have enhanced ability to infect in the presence of plasma during infection.

   For patients C and M, only one single majority E1/E2 sequence per patient was selected by distance matrix analysis and persisted at each of the time points tested. In patient C, four months after seroconversion only the plasma from Tp4 (~1:50) was able to neutralize this majority sequence, C HCVpp. Similarly, for patient M, none of the earlier time points demonstrated neutralization and only plasma collected 10 months after seroconversion (plasma Tp4; not used for SGA) was able neutralize infectivity of the M HCVpp major variant. In contrast to patients B, C and M, assessment of neutralization in patient D was prevented by the lack of functional glycoprotein clones in the HCVpp assay. Although the variants in patient D were cloned and sequences checked, no expression of the E1 and E2 proteins was observed.

   To test if the HCVpp expressed proteins from different patients were susceptible to neutralization by a well-defined broadly neutralizing antibody, we assessed the ability of mouse monoclonal antibody AP33 (Owsianka et al., 2005) to inhibit entry. Despite the observed differences in their resistance to neutralization by patient plasma, all of the HCVpp tested were neutralized by mAb AP33 (Fig. 4C). However, titration of the monoclonal antibody in these assays revealed subtle differences in the neutralization curves with the virus

from patient B, B4 HCVpp-Tp4 being most resistant and the virus from patient C, C HCVpp, being most sensitive.

Discussion

Given the known impact of HIV infection on the natural history of HCV infection, we wished to explore the virus and immune system interactions following acute HCV infection in a group of HIV-1 infected patients. In HCV mono-infected patients, soon after the initial infection, HCV diversity increases especially in the E2 HVR (Brown et al., 2005, Brown et al., 2012, Bull et al., 2011, Casino et al., 1999, D'Arienzo et al., 2013, Farci et al., 2000, Li et al., 2012 Sheridan et al., 2004, Wang et al., 2010). Initial transmission of limited numbers of viruses ("founders") is followed by purifying selection by the immune response and then a slow emergence of variants which become established. In HIV/HCV co-infected patients, Wang and colleagues (2010) reported lower diversity in the HVR1 in a cross sectional study using pyrosequencing.   Reduced HCV genetic complexity was also reported in HVR1 comparing HIV-negative and co-infected patients (Lopez-Labrador et al., 2007, Shuhart et al., 2006). However, the impact of ART in co-infected patients has been reported to lead to increased quasispecies complexity (Rotman and Liang, 2009, Shuhart et al., 2006). Gray et al. (2012) have proposed studying the dynamics of the entire HCV population structure, including liver as well as plasma populations, in order to understand more fully the complexities of HCV evolution.

In our study, soon after initial infection in two of the four patients (D, C) a set of closely related variants were identified. A broader more heterogeneous mix of variants were present in patient B but these three patients all showed a star-like, or near star-like, grouping of variants (suggesting single T/F viruses) using the Poisson-Fitter model (Fig. 2, Table S3).

Patient M however, displayed variants derived from three T/F viruses. The route of HCV infection in all patients was by sexual transmission and these findings of single or few (three) T/F viruses are similar to other reports (Fierer et al., 2014, Li et al., 2012). Using t-test analysis of the GD at Tp1 (baseline) the variant diversity observed in patient B (Table S2), was significantly greater than in the other three patients and subsequently only patient B evolved with a narrowing of diversity or 'bottleneck', (Fig. 3, Table S2) where 62% of the virus variants were identical (Bull et al., 2011, Li et al., 2012, Ribeiro et al., 2012). In agreement with other investigations, our study covering the E1 to NS3 regions revealed an absence of insertions, deletions and recombination hot spots (Fierer et al., 2014, Shi et al., 2012).

Spontaneous HCV clearance has been reported in co-infected patients with high CD4 and ALT values and low viral diversity in the E2 HVR1 region (Thomson et al., 2011). None of the four study patients described here experienced spontaneous HCV clearance even though patients C and M did display low levels of viral diversity throughout follow up. The higher variant diversity which developed by Tp4 in the virus in patient B, may be associated with being on ART and the suppression of HIV replication (Table 1, Table S2, Bernini et al., 2011, Shuhart et al., 2006). This case showed similar diversity to that reported in HCV mono-infected patients (Brown et al., 2005, Brown et al., 2012, Bull et al., 2011, Casino et al., 1999, D'Arienzo et al., 2013, Farci et al., 2000, Li et al., 2012 Sheridan et al., 2004, Shuhart et al., 2006, Wang et al., 2010). ART may have helped to reconstitute the immune response to HCV, even if not sufficient to control HCV replication (Bernini et al., 2011, Herrero-Martinez et al., 2004). CD4 T-cell values throughout follow up were no lower than 190 cells 10^6/L in all patients (Table 1) and were considered not to play a role in the observed differences in diversity progression however the functionality of CD4 cells was not

investigated. Shen and colleagues (2014) also reported that evolution was not influenced by CD4 levels. Lack of positive selection seen in the phylogenetic analysis of viruses from patients C and M may be associated with late seroconversion in (Table 1, 123 and 149 days, post ETI respectively) and/or the high HIV viral load at Tp1 (>5.0 $\log_{10}$ copies/ml). A recent Dutch study (Vanhommerig et al., 2014) reported a mean time to seroconversion of 74 days (IQR 47-125 days) in co-infected patients. Patient D, probably co-infected with HIV and HCV at the same transmission event, also showed a lack of viral sequence evolution. Positive selection was eventually observed two and half years after the ETI, at Tp4. The HIV viral load was > 3.72 $\log_{10}$ copies/ml throughout follow up.   A second elevation in ALT (521 IU/ml) was noted at 814 days, post ETI (Fig. 1). HCV superinfection at this time point was considered but no evidence of a switch in genotype or superinfection with the same genotype detected.

To assess the impact of the host responses on the HCV quasispecies we investigated the selection of amino acids variants that evolved at later time points. Of the four patients studied only patient B showed sufficient immune selection pressure on amino acids in the E2 region to drive change. These changes occurred progressively from Tp3 (270 days, post ETI) to Tp4 (719 days post ETI) and by the end of follow up nine non-synonymous changes in the E2 region (six in the HVR1) were present. These changes were at known B cell, T helper cell, CTL epitopes and at the CD81 binding site. Positively charged amino acids appear to predominate at several E2 locations, H386, R408, H488 and R648 and may modify SR-B1-dependent entry (Koutsoudakis et al., 2012, Sabo et al., 2011). A K408Q substitution evolved in the virus isolated from patient B and perhaps affected SR-B1 mediated entry. Variants with H444Y, at a region involved in CD81 binding, also evolved and may influence HCV replication fitness (Deng et al., 2013a, Keck et al., 2009, Pileri et al., 1998, Tarr et al., 2012).

13

Histidine (H) is unusual at this CD81 location as tyrosine (Y) is more commonly seen and basic residues (K, R and H) may protect against neutralizing antibodies (Boo et al., 2012, Drummer et al., 2006). The positively selected substitution I603V where convergent evolution was detected (Table 2), is present in a ß-sheet of the core E2 domain (Kong et al., 2013) and this site is possibly linked with modulating E1/E2 conformation and virus entry (Douam et al., 2014). Interesting no mutations were detected in the E2 412-424 linear epitope region which has been reported to be a major neutralization target (Owsianka et al., 2005, Tarr et al., 2015). The additional mutations in the viruses from patients C, D and M were dispersed mainly across regions outside of the E2 region. The E1 T331A mutation seen in patient D (referred to as T139A, Douam et al., 2014) may improve virus entry and the NS3 A1113 V/T/A fluctuation was at known CTL epitope. T cell immunity against NS3 was previously shown to be related to recovery from HCV infection (Deng et al., 2013b). In both genotype 1b infections (patients C and M) the variants showed remarkable conservation in the E1 and E2 regions. A limitation of this study was that no HLA typing or investigations of the T cell responses were carried out, which may have provided support for the CTL epitope evolution.

Neutralizing antibodies were observed in all patients from whom functional HCVpp were recovered. Limits to recovery of infectious HCVpp have been previously reported (Dowd et al., 2008, Flint et al., 2004, McKeating et al., 2004). This was the case with virus in patient D. For patient C, 4 months after seroconversion weakly neutralizing antibodies (plasma Tp4 ~1:50) were detected, consistent with the limited selection of variants from the virus population over time and a relatively weak immune response. It was 10 months before the antibodies (plasma Tp4) from patient M were able to neutralize the HCVpp also suggesting a relatively poor humoral immune response. For patient B neutralizing antibodies were present

at six months after seroconversion (plasma Tp3) but later samples (plasma Tp4) were more effective at neutralizing HCV entry. In agreement with our results, von Hahn and colleagues (2007) also reported neutralizing antibodies to lag behind the evolving E2 proteins in a mono-infected patient with chronic HCV. The HVR1 region may also influence the neutralization capacity of the antibodies by obstructing the CD81 binding site and preventing the antibodies binding to the neutralizing epitopes (Bankwitz et al., 2010). Strikingly, control HCV-negative human plasma caused a significant increase in infectivity in this model with the patient B virus at Tp4 (B4 HCVpp-Tp4, Fig. 4B). This is consistent with previous investigations that demonstrated enhancement of HCV entry by human serum (Lavillette et al., 2005, Bartosch et al., 2005). Comparison of the phenotypes of closely-related isolates studied here revealed that the substitution I399F in E2 HVR1 may facilitate this enhancement, contributing to the SR-B1 mediated high-density lipoprotein (HDL) enhancement of infection (Bartosch et al., 2005). This change in phenotype may protect the virus from the autologous neutralizing antibodies and allow it to escape.

Conclusion

   This study, even though small, provides evidence for different patterns or mechanisms of HCV quasispecies evolution after sexual transmission in HIV infected MSM. In all four patients we identified T/F viruses (one to three), but in only one patient (B) was a population bottleneck and subsequent E2 region evolution observed. An improved immune response due to ART with the suppression of HIV replication at the time of HCV infection could have played a role in this instance. For the other three patients a limited increase in virus variant complexity emerged and notably in these cases the baseline HIV viral loads were greater than 4 $\log_{10}$ copies/ml. HCVpp neutralization studies first detected antibody neutralization at four, six and ten months after seroconversion and HCV variants displayed relative resistance to

neutralization by autologous plasma. The greatest neutralization was observed with plasma from later time points. We speculate that after transmission, the evolution of HCV diversity may be influenced by the presence of unsuppressed HIV replication which compromises the host immune response. Further studies on larger patient cohorts would be required to confirm this speculation.

Material and methods

Ethics statement

This study was conducted on stored plasma samples sent for testing at UCLH as part of clinical care. All patients had been enrolled in the Acute Hepatitis C Cohort at the UCL Centre for Sexual Health and HIV Research Unit and written informed consent was gained from each patient. The samples and clinical data were all anonymized.

Study patients

Four HIV-1 infected patients (B, C, D and M) who were diagnosed with acute/recent HCV infection were included in the study. The HCV diagnosis was based on HCV RNA detection (Daniel et al., 2008) prompted by the investigation of abnormal ALT levels and confirmed by anti-HCV seroconversion (Architect anti-HCV assay B3C370, Abbott diagnostics, Wiesbaden, Germany). Stored plasma samples available in the diagnostic laboratory were tested retrospectively to establish the estimated time of HCV infection (HCV ETI) as part of standard care. HCV ETI was taken as day zero and calculated as the midpoint between the last negative HCV RNA or antibody test date and the first HCV RNA positive test. Time point 1 (Tp1) samples were selected when HCV RNA was first detected and where there was a preceding HCV RNA test negative result within the previous five months. Later time point selection was determined by routine clinic visits and availability of sufficient plasma (>1ml).

HCV RNA extraction and cDNA synthesis

The QIAamp Viral RNA Mini Kit (Qiagen, Manchester, UK) was used to extract RNA from 280 µl of plasma according to the manufacturer's instructions. Care was taken to avoid vigorous vortexing to prevent shearing of long RNA templates. The nucleic acids were eluted into 60 µl of 5mM Tris-HCl, pH8.0 (Tris buffer). For the cDNA step (modified from Fan and colleagues, 2006) 10.6 µl of RNA template was added to 9.4 µl reverse transcription mix consisting of 1x SuperScript® III buffer (Life technologies™, Invitrogen, Paisley, UK), 0.01 M DTT, 1 µM NS3plaOA primer (Table S1), 0.5 mM dNTPs (Invitrogen), 20 units of RNasin ribonuclease inhibitor, 200 units of SuperScript® III RT enzyme and 5 units AMV RT (Promega, Southampton, UK). The reaction was performed by incubation at 50°C for 75 min, followed by heating at 70°C for 15 min.

Single genome amplification

cDNA was serially diluted for use as a template in the first-round of a nested PCR to find the dilution where less than 3/10 of PCR reactions yielded a product. A 10 µl master mix was prepared containing 2 µl of cDNA (or water or Tris buffer), 1x Platinum® taq high-fidelity buffer (Life technologies ™), 2 mM $MgSO_4$, 0.2 mM dNTPs, 0.05 units of Platinum® taq high-fidelity DNA polymerase, 5.9 µl of water and 0.2 µM of the first-round forward and reverse primers (Table S1). The cycling conditions were 94°C for 2 min followed by 40 cycles of 94°C for 20 s, 60°C for 30 s, 68°C for 4 min and a final extension of 68°C for 10 min. This was followed by a second-round PCR where 1 µl of the first round product was transferred to 20 µl reaction mix which contained the second-round primers (genotype dependent, Table S1). The 20 µl reactions were performed with a master mix of 1x Platinum® taq high-fidelity buffer, 2 mM $MgSO_4$, 0.2 mM dNTPs, 0.025 units of Platinum® taq high-

fidelity DNA polymerase, 14.9 µl of water and 0.2 µM of each second-round forward and reverse primers (Table S1). The annealing temperature was changed to 64°C for the second-round PCR. All other cycling conditions remained the same as those described for the first-round PCR. PCR products were analysed using 1.0% agarose E-gels® (Life technologies™). SGA amplicons which were probable single templates were purified (QIAamp purification kit, Qiagen) prior to sequencing. Primers were designed or chosen (Table S1) to amplify the HCV E1 to NS3 regions from each patient. No one set of primers was able to amplify the virus from all four patients, thus products of between 2551-3161 bp were generated. Similarly, when bi-directional sequencing was performed (3130xl Genetic Analyzer machine, Applied Biosystems®) primers were selected (eight used per amplicon) depending on the patient/genotype under investigation (Table S1).

Sequence analysis

Variant sequences were aligned using the Sequencher analysis software® 4.5 (Gene Codes Corporation, Ann Arbor, MI USA http://www.genecodes.com). Any sequences containing mixed bases (apart from at known primer locations) were discarded and recombination checks using the Single Break Point analysis in the HyPhy Package were performed to check for evidence of recombination at the P <0.05 statistical level http://www.datamonkey.org/help/citations.php. Nucleotide highlighter plots analysis (highlighter v2.2.3 http://hcv.lanl.gov/, Keele et al., 2008) was used to visualize the evolution of silent (synonymous, dS) and non-silent mutations (non-synonymous, dN). Consensus Maker v2.0.1 http://www.hiv.lanl.gov/ was used to determine the consensus virus sequence in each patient (Tp1) and this consensus sequence was used in Poisson-Fitter analysis http://www.hiv.lanl.gov/ (Giorgi et al., 2010) and to establish the sequence diversity and phylogeny. A mutation rate of 2.5×10^{-5} mutations per nucleotide per genome replication

(Ribeiro et al., 2012) was used in the sequence analysis for Tp1 samples. Unrooted tree analysis using http://www.hiv.lanl.gov/content/sequence/RAINBOWTREE (Fig. 2) at Tp1 was performed. Phylogenetic trees (Fig. 3) for all variants generated were built using the program RAxML, at the nucleotide level, under the General Reversible Time (GTR) model of nucleotide substitution. A bootstrap analysis with 1000 replicates provided branch support. The genetic distances (GD) were calculated with the HyPhy package, under the GTR model (for nucleotide distances). Positive selection analysis was done using three different algorithms from the package HyPhy: SLAC, FEL and FUBAR. Codon positions were considered positively selected if [dN > dS] at the $P < 0.05$ significance level (SLAC and FEL), or if the posterior probability of [dN > dS] was > 0.95. Shannon entropy values were determined per time point and by HCV region using a sliding window analysis (http://www.hiv.lanl.gov/content/sequence/ENTROPY). HCV B cell, T helper cell and cytotoxic T cell lymphocytes (CTL) epitopes were located using the HCV immunology database   (http://hcv.lanl.gov/content/immuno/immuno-main.html).

Majority sequence selection and HCVpp studies

   Distance matrix analysis was used to identify the majority sequences for each patient at the selected time points. HCVpp expression and neutralization studies were performed as previously described (Tarr et al., 2007). Briefly, the E1/E2 genes were amplified using the primers HCVppF and HCVppR (Table S1) and cloned into vector pcDNA3.1 D/V5-his-TOPO (Life technologies^{TM}). HCVpp were generated in HEK 293T cells by co-transfection of the E1/E2 clones with a Murine Leukemia Virus packaging construct and a luciferase reporter plasmid (Bartosch et al 2003a). Infection assays were performed with Huh7 target cells in the presence of patient heat inactivated plasma and neutralization calculated as a

proportion of luciferase reporter expression compared with the HCV uninhibited control preparation.

Nucleotide sequence accession numbers

Majority sequences used in HCVpp studies were deposited in GenBank under the following accession numbers……..

Appendix A. Supplementary Materials

Attached as PDF

References

Abe, K., Inchauspe, G., Fujisawa, K., 1992. Genomic characterization and mutation rate of hepatitis C virus isolated from a patient who contracted hepatitis during an epidemic of non-A, non-B hepatitis in Japan. J. Gen. Virol. 73, 2725-2729.

Arends., J. E., Lieveld,F. I., Boeijen, L. L., Decanter, C. T. M. M., van Erpecum, K. J., Salmon, D., Hoepelman, A. I. M. , Asselah, T., Ustianowski, A. 2015. Natural history and treatment of HCV/HIV coinfection:Is it time to change paradigms? J. Hepat. 63, 1254-1262.

Ball, J. K., Tarr, A. W., McKeating, J. A., 2014. The past, present and future of neutralizing antibodies for hepatitis C virus. Antiviral Res.105, 100-111.

Bankwitz, D., Steinmann, E., Bitzegeio, J., Ciesek, S., Friesland, M., Herrmann, E., Zeisel, M. B., Baumert, T. F., Keck, Z. Y., Foung, S. K. H., Pecheur, E. I., Pietschmann, T., 2010. Hepatitis C virus hypervariable region 1 modulates receptor interactions, conceals the CD81 binding site, and protects conserved neutralizing epitopes. J. Virol. 84, 5751–5763.

Bartosch, B., Bukh, J., Meunier, J. C., Granier, C., Engle, R. E., Blackwelder, W. C., Emerson, S. U., Cosset, F. L., Purcell, R. H., 2003a. In vitro assay for neutralizing antibody to hepatitis C virus: Evidence for broadly conserved neutralization epitopes. Proc. Natl. Acad. Sci. USA 100, 14199-14204.

Bartosch, B., Dubuisson, J., Cosset, F. L., 2003b. Infectious hepatitis C virus pseudo-particles containing functional E1-E2 envelope protein complexes. J. Exp. Med.197, 633-642.

Bartosch, B., Vitelli, A., Granier, C., Goujon, C., Dubuisson, J., Pascale, S., Scarselli, E., Cortese, R., Nicosia, A., Cosset, F. L., 2003c. Cell entry of hepatitis C virus requires a set of co-receptors that include the CD81 tetraspanin and the SR-B1 scavenger receptor. J. Biol. Chem. 278, 41624-41630.

Bartosch, B., Verney, G., Dreux, M., Donot, P., Morice, Y., Penin, F., Pawlotsky, J. M., Lavillette. D., Cosset, F.L., 2005. An interplay between hypervariable region 1 of the hepatitis C virus E2 glycoprotein, the scavenger receptor BI, and high-density lipoprotein promotes both enhancement of infection and protection against neutralizing antibodies. J. Virol. 79, 8217–8229.

Bernini, F., Ebranati, E., De Maddalena, C., Shkjezi, R., Milazzo, L., Lo Presti, A., Ciccozzi, M., Galli, M., Zehender, G., 2011. Within-host dynamics of the hepatitis C virus quasispecies population in HIV-1/HCV coinfected patients. PloS One 6, e16551.

Boo, I., teWierik, K., Douam, F., Lavillette, D., Poumbourios, P., Drummer, H. E., 2012. Distinct roles in folding, CD81 receptor binding and viral entry for conserved histidine residues of hepatitis C virus glycoprotein E1 and E2. Biochem. J. 443, 85-94.

Brown, R. J. P., Juttla, V. S., Tarr, A. W., Finnis, R., Irving, W. L., Hemsley, S., Flower, D. R., Borrow, P., Ball, J. K., 2005. Evolutionary dynamics of hepatitis C virus envelope genes during chronic infection. J. Gen. Virol. 86, 1931-1942.

Brown, R. J. P., Hudson, N., Wilson, G., Rehman, S. U., Jabbari, S., Hu, K., Tarr, A. W., Borrow, P., Joyce, M., Lewis, J., Zhu, L. F., Law, M., Kneteman, N., Tyrrell, D. L., McKeating, J. A., Ball, J. K., 2012. Hepatitis C virus envelope glycoprotein fitness defines virus population composition following transmission to a new host. J. Virol. 86, 11956-11966.

Bull, R. A., Luciani, F., McElroy, K., Gaudieri, S., Pham, S. T., Chopra, A., Cameron, B., Maher, L., Dore, G. J., White, P. A., Lloyd, A. R., 2011. Sequential bottlenecks drive viral evolution in early acute hepatitis C virus infection. PloS Pathog. 7, e1002243.

Casino, C., McAllister, J., Davidson, F., Power, J., Lawlor, E., Yap, P. L., Simmonds, P., Smith, D. B., 1999. Variation of hepatitis C virus following serial transmission: multiple mechanisms of diversification of the hypervariable region and evidence for convergent genome evolution. J. Gen. Virol. 80, 717-725.

Chen, J.Y., Feeney, E.R., Chung, R.T., 2014. HCV and HIV co-infection: mechanisms and management. Nat. Rev. Gastroenterol. Hepatol. 11, 362-371.

D'Arienzo, V., Moreau, A., D'Alteroche, L., Gissot, V., Blanchard, E., Gaudy-Graffin, C., Roch, E., Dubois, F., Giraudeau, B., Plantier, J. C., Goudeau, A., Roingeard, P., Brand, D., 2013. Sequence and functional analysis of the envelope glycoproteins of hepatitis C virus variants selectively transmitted to a new host. J. Virol. 87, 13609-13618.

Daniel, H. D., Grant, P. R., Garson, J. A., Tedder, R. S., Chandy, G. M., Abraham, P., 2008. Quantitation of hepatitis C virus using an in-house real-time reverse transcriptase polymerase chain reaction in plasma samples. Diagn. Microbiol. Infect. Dis. 61, 415-420.

Deng, L., Zhong, L. L., Struble, E., Duan, H. Y., Ma, L., Harman, C., Yan, H. L., Virata-Theimer, M. L., Zhao, Z., Feinstone, S., Alter, H., Zhang, P., 2013a. Structural evidence for a bifurcated mode of action in the antibody-mediated neutralization of hepatitis C virus. Proc. Natl. Acad. Sci. USA 110, 7418-7422.

Deng, Y., Guan, J., Wen, B., Zhu, N., Chen, H., Song, J. D., Yang, Y., Wang, Y., Tan, W. J., 2013b. Induction of broadly neutralising HCV antibodies in mice by integration-deficient lentiviral vector-based pseudotyped particles. PloS One 8, e62684.

Douam, F., Dao Thi, V. L., Maurin, G., Fresquet, J., Mompelat, D., Zeisel, M. B., Baumert, T. F., Cosset, F. L., Lavillette, D., 2014. Critical interaction between E1 and E2 glycoproteins determines binding and fusion properties of hepatitis C virus during cell entry. Hepatology 59, 776-788.

Dowd, K. A., Hershow, R. C., Yawetz, S., LaRussa, P., Diaz, C., Landesman, S. H., Paul, M. E., Read, J. S., Lu, M., Thomas, D. L., Netski, D. M., Ray, S. C., 2008. Maternal neutralizing antibody and transmission of hepatitis C virus to infants. J. Infect. Dis. 198, 1651-1655.

Dowd, K. A., Netski, D. M. Wang, X-H., Cox, A. L. Ray S. C. 2009., Selection pressure from neutralizing antibodies drives sequence evolution during acute infection with Hepatitis C Virus. Gastroenterology 136, 2377-2386.

Drummer, H. E., Boo, I., Maerz, A. L., Poumbourios, P., 2006. A conserved Gly(436)-Trp-Leu-Ala-Gly-Leu-Phe-Tyr motif in hepatitis C virus glycoprotein E2 is a determinant of CD81 binding and viral entry. J. Virol. 80, 7844-7853.

Evans, M. J., von Hahn, T., Tscherne, D. M., Syder, A. J., Panis, M., Wolk, B., Hatziioannou, T., McKeating, J. A., Bieniasz, P. D., Rice, C. M., 2007. Claudin-1 is a hepatitis C virus co-receptor required for a late step in entry. Nature 446, 801-805.

Fan, X., Xu, Y., Di Bisceglie, A. M., 2006. Efficient amplification and cloning of near full-length hepatitis C virus genome from clinical samples. Biochem. Biophys. Res. Commun. 346, 1163-1172.

Farci, P., Shimoda, A., Coiana, A., Diaz, G., Peddis, G., Melpolder, J. C., Strazzera, A., Chien, D. Y., Munoz, S. J., Balestrieri, A., Purcell, R. H., Alter, H. J., 2000. The

outcome of acute hepatitis C predicted by the evolution of the viral quasispecies. Science 288, 339-344.

Fierer, D.S., Talal, A. H., Marks, K.M., van Seggelen, W.O., Branch, A.D., Huprikar, S., Wang, S., Shaw, G.S., Li, H., 2014. Assessing HCV acquisition routes in HIV-infected MSM using single genome sequencing-a pilot study. CROI Abstract D674, 434-435.

Fierer, D. S., Uriel, A. J., Carriero, D. C., Klepper, A., Dieterich, D. T., Mullen, M. P., Thung, S. N., Fiel, M. I., Branch, A. D., 2008. Liver fibrosis during an outbreak of acute hepatitis C virus infection in HIV-infected men: A prospective cohort study. J. Infect. Dis. 198, 683-686.

Flint, M., Logvinoff, C., Rice, C. M., McKeating, J. A., 2004. Characterization of infectious retroviral pseudotype particles bearing hepatitis C virus glycoproteins. J. Virol. 78, 6875-6882.

Gambotti, L., Batisse, D., Colin-de-Verdiere, N., Delaroque-Astagneau, E., Desenclos, J. C., Dominguez, S., Dupont, C., Duval, X., Gervais, A., Ghosn, J., Larsen, C., Pol, S., Serpaggi, J., Simon, A., Valantin, M. A., Velter, A., 2005. Acute hepatitis C infection in HIV positive men who have sex with men in Paris, France, 2001-2004. Euro. Surveill. 10, 115-117.

Gilleece, Y., Sullivan, A., 2005. Management of sexually transmitted infections in HIV positive individuals. Cur. Opin. Infect. Dis. 18, 43-47.

Giorgi, E. E., Funkhouser, B., Athreya, G., Perelson, A. S., Korber, B. T., Bhattacharya, T., 2010. Estimating time since infection in early homogeneous HIV-1 samples using a Poisson model. BMC Bioinformatics 11, 532.

Gray, R. R., Salemi, M., Klenerman, P., Pybus, O. G., 2012. A new evolutionary model for Hepatitis C virus chronic infection. PloS Pathog. 8, e1002656.

Herrero-Martínez, E., Sabin, C. A., Lee, C. A., Jones, I. M., Pillay, D., Emery, V. C., 2004. The effect of highly active antiretroviral therapy for HIV on the anti-HCV specific humoral immune response. J. Med. Virol. 72, 187-193.

Jackowiak, P., Kuls, K., Budzko, L., Mania, A., Figlerowicz, M., Figlerowicz, M., 2014. Phylogeny and molecular evolution of the hepatitis C virus. Infect. Gen. Evol. 21, 67-82.

Keck, Z. Y., Li, S. H., Xia, J., von Hahn, T., Balfe, P., McKeating, J. A., Witteveldt, J., Patel, A. H., Alter, H., Rice, C. M., Foung, S. K. H., 2009. Mutations in hepatitis C virus E2 located outside the CD81 binding sites lead to escape from broadly neutralizing antibodies but compromise virus infectivity. J. Virol. 83, 6149-6160.

Keele, B. F., Giorgi, E. E., Salazar-Gonzalez, J. F., Decker, J. M., Pham, K. T., Salazar, M. G., Sun, C. X., Grayson, T., Wang, S. Y., Li, H., Wei, X. P., Jiang, C. L., Kirchherr, J. L., Gao, F., Anderson, J. A., Ping, L. H., Swanstrom, R., Tomaras, G. D., Blattner, W. A., Goepfert, P. A., Kilby, J. M., Saag, M. S., Delwart, E. L., Busch, M. P., Cohen, M. S., Montefiori, D. C., Haynes, B. F., Gaschen, B., Athreya, G. S., Lee, H. Y., Wood, N., Seoighe, C., Perelson, A. S., Bhattacharya, T., Korber, B. T., Hahn, B. H., Shaw, G. M., 2008. Identification and characterisation of transmitted and early founder virus envelopes in primary HIV-1 infection. Proc. Natl. Acad. Sci. USA 105, 7552-7557.

Khan, A. G., Whidby, J., Miller, M. T., Scarborough, H., Zatorski, A. V., Cygan, A., Price, A. A., Yost, S. A., Bohannon, C. D., Jacob, J., Grakoui, A., Marcotrigiano, J., 2014. Structure of the core ectodomain of the hepatitis C virus envelope glycoprotein 2. Nature 509, 381-384.

Kong, L., Giang, E., Nieusma, T., Kadam, R. U., Cogburn, K. E., Hua, Y. Z., Dai, X. P., Stanfield, R. L., Burton, D. R., Ward, A. B., Wilson, I. A., Law, M., 2013. Hepatitis C virus E2 envelope glycoprotein core structure. Science 342, 1090-1094.

Koutsoudakis, G., Dragun, J., Perez-del-Pulgar, S., Coto-Llerena, M., Mensa, L., Crespo, G., Gonzalez, P., Navasa, M., Forns, X., 2012. Interplay between basic residues of hepatitis C virus glycoprotein E2 with viral receptors, neutralizing antibodies and lipoproteins. PloS One 7, e52651.

Lavillette, D., Tarr, A. W., Voisset, C., Donot, P., Bartosch, B., Bain, C., Patel, A. H., Dubuisson, J., Ball, J. K., Cosset, F. L., 2005. Characterization of host-range and cell entry properties of the major genotypes and subtypes of hepatitis C virus. Hepatology 41, 265-274.

Li, H., Stoddard, M. B., Wang, S. Y., Blair, L. M., Giorgi, E. E., Parrish, E. H., Learn, G. H., Hraber, P., Goepfert, P. A., Saag, M. S., Denny, T. N., Haynes, B. F., Hahn, B. H., Ribeiro, R. M., Perelson, A. S., Korber, B. T., Bhattacharya, T., Shaw, G. M., 2012. Elucidation of hepatitis C virus transmission and early diversification by single genome sequencing. PloS Pathog. 8, e1002881.

Lopez-Labrador, F. X., Dove, L., Hui, C. K., Phung, Y., Kim, M., Berenguere, M., Wright, T. L., 2007. Trends for genetic variation of hepatitis C virus quasispecies in human immunodeficiency virus-1 coinfected patients. Virus Research 130, 285-291.

Lupberger, J., Zeisel, M. B., Xiao, F., Thumann, C., Fofana, I., Zona, L., Davis, C., Mee, C. J., Turek, M., Gorke, S., Royer, C., Fischer, B., Zahid, M. N., Lavillette, D., Fresquet, J., Cosset, F. L., Rothenberg, S. M., Pietschmann, T., Patel, A. H., Pessaux, P., Doffoel, M., Raffelsberger, W., Poch, O., McKeating, J. A., Brino, L., Baumert, T. F., 2011. EGFR

and EphA2 are host factors for hepatitis C virus entry and possible targets for antiviral therapy. Nature Medicine 17, 589-595.

McKeating, J. A., Zhang, L. Q., Logvinoff, C., Flint, M., Zhang, J., Yu, J., Butera, D., Ho, D. D., Dustin, L. B., Rice, C. M., Balfe, P., 2004. Diverse hepatitis C virus glycoproteins mediate viral infection in a CD81-dependent manner. J. Virol. 78, 8496-8505.

Osburn, W. O., Fisher, B. E., Dowd, K. A., Urban, G., Liu, L., Ray, S. C., Thomas, D. L., Cox, A. L. 2010. Spontaneous control of primary hepatitis C virus infection and immunity against persistent reinfection. Gastroenterology 138, 315-324.

Owsianka, A., Tarr, A. W., Juttla, V. S., Lavillette, D., Bartosch, B., Cosset, F. L., Ball, J. K., Patel, A. H., 2005. Monoclonal antibody AP33 defines a broadly neutralizing epitope on the hepatitis C virus E2 envelope glycoprotein. J.Virol. 79, 11095-11104.

Pestka, J, M, L., Zeisel, M. B., Bläser, E., Schürmann, P., Bartosch, B., Cosset, F. L., Patel, A. H., Meisel, H., Baumert, J., Viazov, S., Rispeter, K., Blum, H. E., Roggendorf, M., Baumert, T. F., 2007. Rapid induction of virus-neutralizing antibodies and viral clearance in a single-source outbreak of hepatitis C. Proc. Natl. Acad. Sci. USA 104, 6025-6030.

Pileri, P., Uematsu, Y., Campagnoli, S., Galli, G., Falugi, F., Petracca, R., Weiner, A. J., Houghton, M., Rosa, D., Grandi, G., Abrignani, S., 1998. Binding of hepatitis C virus to CD81. Science 282, 938-941.

Ploss, A., Evans, M. J., Gaysinskaya, V. A., Panis, M., You, H. N., de Jong, Y. P., Rice, C. M., 2009. Human occludin is a hepatitis C virus entry factor required for infection of mouse cells. Nature 457, 882-886.

Raghuraman, S., Park, H., Osburn, W. O., Winkelstein, E., Edlin, B. R., Rehermann, B.,
2012. Spontaneous clearance of chronic hepatitis C virus infection is associated with
appearance of neutralizing antibodies and reversal of T-cell exhaustion. J. Infect. Dis.
205, 763-771.

Ribeiro, R. M., Li, H., Wang, S. Y., Stoddard, M. B., Learn, G. H., Korber, B. T.,
Bhattacharya, T., Guedj, J., Parrish, E. H., Hahn, B. H., Shaw, G. M., Perelson, A. S.,
2012. Quantifying the diversification of hepatitis C virus (HCV) during primary
infection: estimates of the in vivo mutation rate. PloS Pathog. 8, e1002881.

Rotman, Y., Liang, T. J., 2009. Coinfection with hepatitis C Virus and human
immunodeficiency virus: virological, immunological, and clinical outcomes. J. Virol. 83,
7366-7374.

Sabo, M. C., Luca, V. C., Prentoe, J., Hopcraft, S. E., Blight, K. J., Yi, M., Lemon, S. M.,
Ball, J. K., Bukh, J., Evans, M. J., Fremont, D. H., Diamond, M. S., 2011. Neutralizing
monoclonal antibodies against hepatitis C virus E2 protein bind discontinuous epitopes
and inhibit infection at a postattachment step. J. Virol. 85, 7005-7019.

Santantonio, T., Wiegand, J., Gerlach, J. T., 2008. Acute hepatitis C: current status and
remaining challenges. J. Hepatol. 49, 625-633.

Shen, C. L., Gupta, P., Xu, X. C., Sanyal, A., Rinaldo, C., Seaberg, E., Margolick, J. B.,
Martinez-Maza, O., Chen, Y., 2014. Transmission and evolution of hepatitis C virus in
HCV seroconverters in HIV infected subjects. Virology 449, 339-349.

Sheridan, I., Pybus, O. G., Holmes, E. C., Klenerman, P., 2004. High-resolution phylogenetic
analysis of hepatitis C virus adaptation and its relationship to disease progression. J.
Virol. 78, 3447-3454.

Shi, W. F., Freitas, I. T., Zhu, C. D., Zheng, W., Hall, W. W., Higgins, D. G., 2012. Recombination in Hepatitis C Virus: Identification of Four Novel Naturally Occurring Inter-Subtype Recombinants. PloS One 7, e41997.

Shuhart, M. C., Sullivan, D. G., Bekele, K., Harrington, R. D. Kitahata, M. M., Mathisen, T. L., Thomassen, L. V., Emerson, S. S. Gretch, D. G., 2006. HIV infection and antiretroviral therapy:effect on hepatitis C virus quasispecies variability. J. Infect. Dis. 193, 1211-1218.

Simmonds, P., Balfe, P., Ludlam, C. A., Bishop, J. O., Brown, A. J. L., 1990. Analysis of sequence diversity in hypervariable regions of the external glycoprotein of Human-Immunodeficiency-Virus Type-1. J. Virol. 64, 5840-5850.

Steinhauer, D. A., Domingo, E., Holland, J. J., 1992. Lack of evidence for proofreading mechanisms associated with an RNA virus polymerase. Gene 122, 281-288.

Stoddard, M. B., Li, H., Wang, S., Saeed, M., Andrus, L., Ding, W., Jiang, X., Learn, G. H., von, S. M., Wen, J., Goepfert, P. A., Hahn, B. H., Ploss, A., Rice, C. M., Shaw, G. M., 2015. Identification, molecular cloning, and analysis of full-length hepatitis C virus transmitted/founder genotypes 1, 3, and 4. MBio. 6:e02518.

Tarr, A. W., Owsianka, A. M. Jayaraj, D., Brown, R. J. P., Hickling, T. P., Irving, W. L., Patel, A, H., Ball, J. K., 2007. Determination of the human antibody response to the epitope defined by the hepatitis C virus neutralizing monoclonal antibody AP33. J. Gen. Virol. 88, 2991-3001.

Tarr, A. W., Urbanowicz, R. A., Jayaraj, D., Brown, R. J. P., McKeating, J. A., Irving, W. L., Ball, J. K., 2012. Naturally occurring antibodies that recognize linear epitopes in the amino terminus of the hepatitis C virus E2 protein confer noninterfering, additive neutralization. J. Virol. 86, 2739-2749.

Tarr, A. W., Khera, T., Hueging, K., Sheldon, J., Steinmann E., Pietschmann T., Brown, R. J. P., 2015. Genetic diversity underlying the envelope glycoproteins of Hepatitis C Virus: Structural and functional consequences and the implications for vaccine design. Viruses 7, 3995-4046.

Thomas, D. L., 2002. Hepatitis C and human immunodeficiency virus infection. Hepatology 36, S201-S209.

Thomson, E. C., Fleming, V. M., Main, J., Klenerman, P., Weber, J., Eliahoo, J., Smith, J., McClure, M. O., Karayiannis, P., 2011. Predicting spontaneous clearance of acute hepatitis C virus in a large cohort of HIV-1-infected men. Gut 60, 837-845.

Thomson E.C., Nastouli, E, Main J, Karayiannis, P, Eliahoo, J., Muir, D., McClure, M.O. 2009. Delayed anti-HCV antibody response in HIV-positive men acutely infected with HCV. AIDS. 23, 89-93.

Turner, J., Bansi, L., Gilson, R., Gazzard, B., Walsh, J., Pillay, D., Orkin, C., Phillips, A., Easterbrook, P., Johnson, M., Porter, K., Schwenk, A., Hill, T., Leen, C., Anderson, J., Fisher, M., Sabin, C., 2010. The prevalence of hepatitis C virus (HCV) infection in HIV-positive individuals in the UK - trends in HCV testing and the impact of HCV on HIV treatment outcomes. J. Viral Hepatitis 17, 569-577.

Vanhommerig J. W., Thomas, X. V., van der Meer, J. T., Geskus, R. B., Bruisten, S. M., Molenkamp, R., Prins, M., Schinkel, J.; MOSAIC (MSM Observational Study for Acute

Infection with hepatitis C) Study Group. 2014. Hepatitis C virus (HCV) antibody dynamics following acute HCV infection and reinfection among HIV-infected men who have sex with men. Clin. Infect. Dis. 59, 1678-1685.

von Hahn, T., Yoon, J. C., Alter, H., Rice, C. M., Rehermann, B., Balfe, P., McKeating, J. A., 2007. Hepatitis C virus continuously escapes from neutralizing antibody and T-cell responses during chronic infection in vivo. Gastroenterology 132, 667-678.

Wang, G. P., Sherrill-Mix, S. A., Chang, K. M., Quince, C., Bushman, F. D., 2010. Hepatitis C virus transmission bottlenecks analyzed by deep sequencing. J. Virol. 84, 6218-6228.

Yanagi, M., Purcell, R. H., Emerson, S. U., Bukh, J., 1997. Transcripts from a single full-length cDNA clone of hepatitis C virus are infectious when directly transfected into the liver of a chimpanzee. Proc. Natl. Acad. Sci. USA 94, 8738-8743.

Zhang, J., Randall, G., Higginbottom, A., Monk, P., Rice, C. M., McKeating, J. A., 2004. CD81 is required for hepatitis C virus glycoprotein-mediated viral infection. J. Virol. 78, 1448-1455.

Table 1

Patient characteristics and laboratory data.

Patient ID	Age	HCV Genotype	Time point (Tps)	Sampling days at Tps since HCV ETI	Estimated sero-conversion day since HCV ETI	HIV RNA at Tps (log$_{10}$ copies/ml)	CD4 T-cell count 10^6/L	ART[a]
B[b]	37	1a	1	30	86	Ndet[c]	290	Y
			2	136		Ndet	440	Y
			3	270		Ndet	410	Y
			4	719		2.83	390	N
D[d]	31	1a	1	29	29	4.17	400	N
			2	113		3.72	290	N
			3	181		4.23	580	N
			4	930		5.34	490	N
C[e]	24	1b	1	18	123	5.23	280	N
			2	66		4.70	350	N
			3	151		Ndet	390	Y
			4	249		Ndet	530	Y
M	48	1b	1	42	149	5.59	190	N
			2	149		Ndet	390	Y
			3	366		Ndet	300	Y

[a] Y indicates treatment with antiretroviral treatment (ART) and N indicates no treatment.

[b] HBV co-infection (3.0 log$_{10}$ IU/ml) at time point 1.

[c] Ndet indicates not detected and this was where virus levels were below the level of detection (1.7 HIV log$_{10}$ copies/ml).

[d] Recent anti-HIV seroconversion.

[e] 24 weeks of HCV treatment started on day 249 since HCV ETI with pegylated interferon α and ribavirin.

Table 2
Evolution of amino acid substitutions in HCV variants analysed from patients B, D, C and M. Single non-synonymous mutations where amino acids failed to establish at subsequent time points are not shown.

Patient ID and Tp	Days since HCV ETI	E1 192-383	E2 384-746			NS2 810-1026			NS3 1027-1223		
			Mutation	Variants with Mutation	% of variants with mutation	Mutation	Variants with Mutation	% of variants with mutation	Mutation	Variants with mutation	% of variants with mutation
B Tp1	30	Tp 1 dominant sequence identified	V384T I603V E717D	1/23 4/23 4/23	4.3 17.4 17.4						
B Tp2	136	No change	V384T G401D	21/21 1/21	100 4.8	W841R	21/21	100			
B Tp3	270	No change	V384T G401S P405Q K408Q K408M H444Y A563V	29/29 1/29 1/29 13/29 7/29 1/29 3/29	100 3.4 3.4 44.8 24.1 2.4 10.3	W841R L864 frame shift	29/29 1/29	100 3.4	R1187G	1/29	3.4
B Tp4	719	No change	V384T A395T A396T I399F G401S P405R K408Q H444Y N446K I603V	19/19 19/19 19/19 19/19 19/19 19/19 18/19 17/19 2/19 18/19	100 100 100 100 100 100 94.7 89.5 10.5 94.7	W841R L864 frame shift V895I	19/19 1/19 9/19	100 5.3 47.3	R1187G	1/19	5.3

Underlined mutation at I603V undergoes convergent evolution with P >0.001 using SLAC and FEL algorithms [dN>dS] where P <0.05 values are significant.

Patient ID and Tp	Days since HCV ETI	E1 192-383			E2 384-746			NS2 810-1026			NS3 1027-1223		
		Mutation	Variants with mutation	% of variants with mutation	Mutation	Variants with mutation	% of variants with mutation	Mutation	Variants with mutation	% of variants with mutation	Mutation	Variants with mutation	% of variants with mutation
D Tp1	29	T331A	1/20	5							A1113V A1113T	1/20 1/20	5 5
D Tp2	113	T331A	10/18	55.5				T812A	1/18	5.5	A1113V A1113T	8/18 1/18	44.4 5.5
D Tp3	181	T331A	10/11	90.9							A1113V A1113T	9/11 1/11	81.8 9.1
D Tp4	930	T331A	10/10	100	F397S A475V V696I	10/10 7/10 7/10	100 70 70	T812A A929V	4/10 7/10	40 70	ALL NOW A1113	10/10	100

Underlined mutations at V696I and A1113V/T undergo convergent evolution with a posterior probability 0.96 and 0.98 respectively using FUBAR algorithm of [dN>dS] >0.95.

Patient ID and Tp	Days since HCV ETI	NS3 1027-1223		
		Mutation	Variants with mutation	% of variants with mutation
C Tp1	18			
C Tp2	66			
C Tp3	151	A1113S	4/10	40
C Tp4	249	A1113S	1/10	10

Patient ID and Tp	Days since HCV ETI	E2 384-746			NS2 810-1026		
		Mutation	Variants with mutation	% of variants with mutation	Mutation	Variants with Mutation	% of variants with mutation
M Tp1	42	Q743P	5/20	25	<u>E837V</u>	10/20	50
M Tp2	149				<u>E837K</u> G1017D	1/11 1/11	9.1 9.1
M Tp3	366				<u>E837K</u> I1010V G1017D	1/9 4/9 3/9	11.1 44.4 33.3

Underlined mutation at <u>E837V/K</u> undergoes convergent evolution
-posterior probability 0.97 using FUBAR algorithm of [dN>dS] >0.95.

Table 3
HCV epitope locations and CD81 binding sites where amino acids evolved.

Patient ID	Possible B cell epitope location (neutralizing)	Possible CTL epitope location	Possible T helper cell location	CD81[a] binding sites
B	384 E2 HVR1	395 E2	395 E2	444 E2
	395 E2 HVR1	396 E2	396 E2	446 E2
	396[b] E2 HVR1	399 E2	399 E2	
	399[b] E2 HVR1	401 E2	401 E2	
	401[b] E2 HVR1	405 E2	405 E2	
	405[b] E2 HVR1	408 E2	408 E2	
	408 E2	717 E2	1187 NS3	
	444 E2	841 NS2		
	446 E2	1187 NS3		
D	397[b] E2 HVR1	331 E1	331 E1	
	475 E2 HVR2	475 E2	397 E2	
		1113 NS3		
C		1113 NS3		
M		837 NS2	743 E2	
		1010 NS2	1010 NS2	
		1017 NS2	1017 NS2	

[a] Reference, Pileri et al., 1998.
[b] SR-B1 binding site location.
HCV B cell, T helper cell and cytotoxic T cell (CTL) epitopes were located using the HCV immunology database http://hcv.lanl.gov/content/immuno/immuno-main.html.

1

Figure captions

Fig. 1. Laboratory parameters at the time points (Tps) studied. The solid lines show HCV
RNA values in IU/ml and the Tps at which SGA was performed are labelled. ALT values
(IU/L) are indicated by dashed lines. The normal range of ALT is 10-50 IU/L. The arrows
define the time of anti-HCV seroconversion and the open triangle (patient C, Tp4) indicates
start of successful treatment. HCV ETI to Tp1 for patient B was 30 days, patient D 29 days,
patient C 18 days and patient M 42 days.

Fig. 2. Unrooted phylogenetic tree analysis of SGA-derived sequences at Tp1 using
http://www.hiv.lanl.gov/content/sequence/RAINBOWTREE. Sequences from patients B, C
and D clustered around a central node indicating that the productive HCV infection was due
to a single virus whereas M displayed three separate groups/lineages suggesting infection
with three different founder viruses.

Fig. 3. Phylogenetic trees of patient derived SGA E1-NS3 sequences. Variants at the
different time points are colour coded (Tp1- yellow, Tp2- green, Tp3- red and Tp4 -blue).
The trees where rooted against a sequence from another patient and the branches represent
nucleotide substitutions per site. A bootstrap analysis, with 1000 replicates, was done to
assess branch support. Branches with a support > 70% are labelled with a star (∗).

Fig. 4. HCVpp infectivity results where all reactions were performed in triplicate. (A)
HCVpp infectivity levels displayed by the expression of the majority E1/E2 sequences from
patients C, (C HCVpp), M, (M HCVpp) and B1 (B1 HCVpp-Tp1), B2 (B2 HCVpp-Tp2), B3
(B3 HCVpp-Tp3), B4 (B4 HCVpp-Tp4) in relation to controls H77c and VSV-G (measured
in relative light units, RLU). (B) Each graph represents the infectivity of the HCV

pseudoparticles (HCVpp) derived from patient B, C and M in the presence of patient plasma (● plasma Tp1, ■ plasma Tp2, ▲ plasma Tp3, ▼ plasma Tp4) compared to the signal achieved with an uninhibited control preparation of each HCVpp. Each HCVpp represents the majority E1/E2 amino acid sequence sampled at each time point (B1 HCVpp-Tp1, B2 HCVpp-Tp2, B3 HCVpp-Tp3, B4 HCVpp-Tp4, C HCVpp, M HCVpp). An HCV-negative control plasma was also included (○). Greatest neutralization was seen with the later time point plasma samples. (C) Neutralization of the HCVpp expressed E1/E2 proteins, selected from three patients (from left to right, B1, 2, 3, 4 HCVpp; C HCVpp; M HCVpp) by the broadly reactive AP33 monoclonal antibody. Data are expressed as a proportion of the signal observed with the uninhibited HCVpp. Serial dilution of the AP33 antibody revealed varying degrees of neutralization; B4 HCVpp derived from patient B at Tp4 being the most resistant to neutralization and that from patient C HCVpp the least resistant.