Evolution of hepatitis C virus variants following blood transfusion from one infected donor to several recipients: a long-term follow-up

Arthur Løve, Vilma Molnegren, Ann-Sofie Månssson, Agnes Smáradóttir, Sigurdur B. Thorsteinsson and Anders Widell

INTRODUCTION

Hepatitis C virus (HCV) is a member of the Hepacivirus genus of the Flaviviridae family. The hepaciviruses are known to cause chronic infections in humans and animals. HCV displays a great genetic variability and HCV strains belong to one of six or more major genotypes, each of which consists of a large number of subtypes (Simmonds et al., 1993; Tokita et al., 1998). Within subtypes there are viral variants, established from the source of infection carrying subsequent mutations following the transmission event (Ross et al., 2000). Finally, as for many RNA viruses causing chronic infections, HCV tends to persist in the host as a dynamic cluster of related viral variants, referred to as ‘quasispecies’ (Forns et al., 1999). The HCV quasispecies phenomenon is most readily observed in hypervariable region 1 (HVR-1) at the N terminus of the larger envelope glycoprotein (Kato et al., 1994; Pawlotsky et al., 1999; Okuda et al., 1999). Quasispecies are present throughout the viral genome, although to a lesser extent, including conserved regions such as the 5′-untranslated region (Soler et al., 2002). Quasispecies are believed to appear by virus replication using a non-proofreading viral RNA polymerase (Holland et al., 1982; Steinhauer & Holland, 1987) and immune selection. Viruses may benefit from such antigenic changes by evading the immune system (Løve et al., 1985).

The transmission of HCV can occur with infectious doses of different sizes, which can confer a low infection rate, e.g. after accidental needle-stick (Arai et al., 1996; Suzuki et al., 1994), sexual (Morisca et al., 2001), perinatal (Rapicetta et al., 2000) or nosocomial exposure (Esteban et al., 1996; Ross et al., 2000; Allander et al., 1995; Widell et al., 1999). Low-level contaminated blood products, such as non-solvent detergent treated Gammagard, were shown to have infected between 5 and 80 % of the recipients, depending upon the batch (Healey et al., 1996; Widell et al., 1997).

Quasispecies evolution of HCV has been reported after administration of anti-D immunoglobulin (McAllister et al., 1998; Casino et al., 1999). These studies demonstrated that
virus variants in the infectious source were relatively homogeneous, and distinct variants in HVR-1 were observed in each anti-D recipient, indicating evolution of the virus in that region. The changes observed were not random but rather strongly constrained, thus conserving the hydrophobicity pattern of HVR-1.

Recipients of blood components such as red cells, platelets and plasma, prior to screening for HCV antibodies among blood donors in the early 1990s (Alter & Houghton, 2000), probably received the largest infectious doses because of the large volume transferred at each transfusion. Thus, transfused patients theoretically should have the greatest potential for broad viral quasispecies. An extensive study by Farci et al. (2000) focusing on early events after transfusion suggests that the magnitude of genetic diversity early after infection is important for the outcome of disease, i.e. the more progressive forms of disease have a larger genetic diversity compared with the resolving HCV cases. In a recent study of HCV from blood donors and their recipients (Lin et al., 2001), the authors observed more extensive recipient sequence divergence with time, reaching the degree seen between unlinked subjects. These analyses included data obtained by direct sequencing of PCR products at one time-point per patient. Similar data have been reported recently by Cantaloube et al. (2003). However, extensive quasispecies data linking a single HCV-infected donor quasispecies to several transfused recipients are not available.

The aim of the present study was to assess the evolution of HCV quasispecies in one blood donor and several recipients infected from his blood components. We had access to frozen sera from 1993 from an HCV-infected blood donor and 13 of his viraemic recipients (Löve et al., 1995). These samples were analysed for viral heterogeneity in 1993 when the donor and the recipients were first found to be infected with HCV. The donor and six viraemic recipients still alive in 1998 were analysed again for quasispecies evolution. The study was done by extensive cloning and sequencing of the HCV HVR-1 region.

**METHODS**

**Serum samples.** As reported previously (Löve et al., 1995), a look-back study from 1993, done shortly after the introduction of screening of blood donors for HCV antibodies, identified one HCV-infected blood donor (DE). Donor DE proved to be a former intravenous drug user and it was possible to trace 15 recipients (R1–R15) of his blood components from 1988 onwards. All 15 recipients except one (R11), a dialysis patient, had HCV antibodies when traced. The donor and 13 viraemic recipients all shared genotype 1a (Löve et al., 1995). The time between transfusion and the first recipient specimen varied between 0–3 and 5 years. Serum samples from 1993 or extracted RNA was stored at −70 °C. In 1998, new blood samples were obtained from the donor and those recipients who were still alive, and the sera were processed as previously and frozen before being subjected to further analyses.

**HCV diagnosis and RNA detection.** The initial virus diagnosis by detection of HCV antibodies was made with the Ortho second-generation ELISA (Ortho Diagnostic Systems) and confirmed with the second-generation recombinant immunoblot assay (RIBA; Chiron). The HCV diagnosis of the samples taken in 1998 was confirmed with third-generation ELISA and RIBA tests.

Processing and cloning of 1993 and 1998 samples was done completely independently and the guidelines recommended by Kwo & Higuchi (1989) were strictly adhered to to prevent contamination between samples and clones. The molecular studies were based on viral RNA, extracted from 100 µl serum by the method described by Chomczynski & Sacchi (1987). Extracted RNA was divided into three to four aliquots and stored in the presence of RNase inhibitor (Promega) at −70 °C.

For detection of viral RNA, a nested PCR directed to a region in the 5′-untranslated region was used (Widell et al., 1991). The detection limit was about 300 genomic copies ml⁻¹.

**Amplification and direct sequencing of HVR-1.** Viral RNA was reverse transcribed and amplified in a single-tube reaction using primers encompassing the HVR-1 region of genotype 1a (nt 1290–1309 and 1873–1854 for the first PCR, 35 cycles) as described previously (Widell et al., 1997). The enzymes employed were AMV-derived reverse transcriptase (Promega) and AmpliTaq polymerase (Roche). Primers corresponding to nt 1300–1319 and 1870–1848 were used in the second, inner PCR (35 cycles), which also employed AmpliTaq. PCR products were detected by agarose gel electrophoresis and ethidium bromide staining and visualized with ultraviolet light. DNA was purified on 1·0% agarose gels from which the nucleic acid bands were removed and purified by a gel-extraction method (QIAquick; Qiagen) using a microcentrifuge. Subsequently, for direct sequencing of amplifiers, the ABI Prism BigDye Terminator Cycle Sequencing Ready Reaction kit with AmpliTaq DNA polymerase FS (PE Biosystems) was used with either inner PCR primers or special sequencing primers (Widell et al., 1997). Both strand products were then analysed on a gene sequencer (373 Sequenator ABI; PE Biosystems) and edited with software programs (Factura and Sequence Navigator; PE Biosystems). The present study is entirely based upon the 81 nucleotides of HVR-1, the 27 nucleotides upstream and 171 nucleotides downstream were included.)

**Cloning and sequencing of HVR-1.** To study the genomic heterogeneity of HVR-1, PCR products from all viraemic samples were cloned and 8–20 clones picked and subsequently sequenced. For the cloning, a cloning vector, pBAD TOPO TA (Invitrogen), was used in competent cells. Each clone was cultured overnight and the plasmids were purified on a Jet Star plasmid mini-prep kit (Genomed) and cycle-sequenced bi-directionally using vector specific M13-20 forward and M13-reverse primers.

To assess the variability of the PCRs and cloning procedures, donor DE samples from 1993 and 1998 each underwent two independent complete reverse transcription and nested PCRs with subsequent independent cloning procedures, labelled A and B, respectively. A further independent extraction/RT-PCR labelled DE98C (without cloning) was done on the donor 1998 sample.

**Phylogenetic analysis.** Sequences obtained either from PCR products or clones were aligned by BioEdit version 4.8.10 (T. A. Hall, University of North Carolina, NC, USA). Pairwise nucleotide p-distances (the proportion of nucleotide sites at which the two sequences are different) between isolates were calculated with the Kimura two-parameter model using the Molecular Evolution Genetic Analysis (MEGA) version 2.1 software (S. Kumar, K. Tamura, I. B. Jakobsen & M. Nei, 2001). MEGA 2.1 was also used to compare...
predicted p-distances regarding synonymous versus non-synonymous changes by the Kumar algorithm and to calculate predicted amino acid differences between clones. In addition, differences between clones in 1993 and 1998 in the same individual were also calculated.

Neighbour-joining trees based on nucleotide p-distances were displayed in circular form when necessary and otherwise in linear form. Robustness of trees was always assessed by bootstrap analysis (1000 resamplings) provided with the MEGA 2.1 software.

**Statistical methods.** Distance and amino acid difference data were transferred via Microsoft Excel 2000 to Statistica 5.1 (StatSoft). Since the number of observed distances varied with the number of clones obtained in each person, we used post hoc comparisons by the Tukey honest significant difference test for unequal N. This test allows for several parallel significance calculations between patients with different numbers of observations. P values of less than 0.05 were considered significant.

**RESULTS**

**Patient exposure time-points and outcome of HCV infection**

Of the 15 traced recipients exposed to donor DE, 13 were, as reported earlier, viraemic at the first available sample (Love et al., 1995). At follow-up 5 years later, it was only possible to analyse samples from six of the individuals by sequencing, since two recipients (R1 and R5) had become PCR negative at that time (Fig. 1) and three others had died. Since blood donations in general are processed into erythrocytes, plasma and thrombocytes (platelets), a number of patients had been exposed to infectious material from the same donation, i.e. R1, R2 and R3; R4 and R5; R7 and R8; and R9 and R10. The type of transfused product is indicated in Fig. 1. Two recipients, R9 and R11, had been exposed twice at 1–2 years apart to blood products from donor DE.

**RNA extraction, reverse transcription, nested PCR of HVR-1, cloning and sequencing**

In both the 1993 and 1998 donor DE samples, the entire procedure from extraction to sequencing of PCR products and multiple clones was repeated twice. Although direct sequencing results (labelled* in Figs 2 and 3) were slightly different, the quasispecies distribution from each extraction was not different. A third extraction that was never cloned yielded a sequence DE98Cdir, which grouped with a number of distant clones, rarely isolated among clones from the previous extractions A and B. In the recipients, the direct sequences (Figs 2 and 3) grouped well with cloned sequences from the same recipient.

---

**Fig. 1.** The horizontal lines show the known course of HCV infection of donor DE and his recipients R1–R15. The initiation of the horizontal recipient lines indicates the time of infection, i.e. the time of transfusion of infected blood from DE. The filled circles designate HCV-viraemic PCR-positive samples and the empty circles non-viraemic PCR-negative samples. The letters in parentheses indicate the types of infecting blood component: R, erythrocytes (packed); P, plasma; T, thrombocytes (platelets); † denotes death of the patient. The following patients had shared components from the same donation: R1, R2 and R3; R4 and R5; R7 and R8; and R9 and R10. Note that R6 and R15 were never found and tested during an HCV-viraemic stage and R1 and R5 were initially HCV viraemic when found but lost their HCV viraemia during the period of observation.
Fig. 2. Phylogenetic relationship between HVR-1 nucleotide sequences in donor DE and 13 recipients in 1993 when the donor DE was first identified as HCV positive and the recipients traced. Data were obtained either by direct sequencing of PCR products (labelled*) or sequencing of multiple clones from the donor and from each recipient. Nucleotide distances were based on the Kimura two-parameter model and a neighbour-joining tree was calculated in which 1000 bootstrap resamplings were done. The tree is depicted as a circular phylogenetic dendrogram with bootstrap frequencies indicated at respective nodes if above 50% (in bold if above 70%). All calculations were done using MEGA 2.1 software. Recipients who were exposed to blood components of the same donation share the same marker colour but differ in the shape of marker symbol. Prototype strains HCV-1 (accession no. M62321) and HCV-H (accession no. M67463) are included as references.
When donor 1993 clones were compared with recipient 1993 clones (1991 for recipient R9), there were donor clones that were closely related to recipient clusters in several instances (Fig. 2), most pronounced in the recipients who had been transfused in recent years. Thus, clones infecting R1 and R2 and those that infected R3 must have been circulating in the donor for at least 1 year between the donor’s donation and the first identification of him as HCV antibody positive when the sample (DE93) was obtained. Likewise, the two independent clones that infected R7 and R8 as well as R9 were recovered in DE93 and hence must have been circulating in the donor’s blood for at least 2 years.

Although fewer clones (about 10 per recipient) were analysed in the recipients, it was evident that the clones of each recipient displayed a very limited quasispecies distribution, despite the fact that 0.3–5 years had elapsed from transfusion to the diagnosis of HCV infection. The most diverse were R12 and R14. The narrow pattern among recipients with nucleotide p-distances of 0.044 (R1), 0.006 (R2), 0.010 (R3), 0.005 (R4), 0.004 (R5), 0.013 (R7), 0.013 (R8), 0.003 (R9), 0.013 (R10), 0.012 (R11), 0.062 (R12), 0.012

---

**Fig. 3.** Phylogenetic relationship after 5 years of follow-up (1998) between donor DE and six recipient sequences, analysed by direct sequencing and sequencing of multiple clones. Nucleotide p-distances were calculated by the Kimura two-parameter model. The neighbour-joining tree after 1000 bootstrap resamplings is depicted as a circular phylogenetic dendrogram. Bootstrap values are indicated at respective nodes if above 50% (in bold if above 70%). Donor DE and recipients R7, R8, R10, R11, R13 and R14 share the same colour code as in Fig. 2. *, Sequences obtained by direct sequencing. Prototype strains HCV-1 and HCV-H are included as references.
(R13) and 0.037 (R14) was in clear contrast to donor DE whose mean nucleotide p-distances among clones was 0.130. All differences between the donor and each recipient were highly significant by post hoc testing \( (P<0.00003) \). In addition, R12 significantly differed \( (P<0.005) \) from all other recipients except R1 and R14. Recipient R1 was a special case since one clone (R0193clo10) was different from all his other clones and closely related, but not identical, to a sequence found in the donor 1998 sample (data not shown). Overall, quasispecies diversity in 1993 was not different in the two individuals (R1 and R5) whose HCV infection resolved spontaneously compared with those who remained chronically HCV infected.

In 1993, samples from recipients exposed to the components from the same donation differed in sequence clustering, such as R1 and R2 in contrast to R3; R4 in contrast to R5; R7 in contrast to R8; and R9 in contrast to R10. This patient grouping was in general supported by high bootstrap values. The opposite outcome was also observed where recipients were exposed to different blood donations obtained several months apart, for example R8 and R9 shared HVR-1 sequences and R10 and R11 had closely related HVR-1 sequences, despite years elapsing since transfusion and first diagnosis of HCV infection.

**Donor DE 1998 sequences compared with the sequences of six recipients at follow-up in 1998**

In concordance with the donor 1993 sequence diversity, the quasispecies in the donor DE 1998 sample was wider (mean nucleotide p-distance 0.172; \( P<0.0005 \)) and contained several groups supported by high bootstrap values (Fig. 3). For the six recipients being followed up and included in the same dendrogram, quasispecies diversity was limited in R7, R8, R10 and R13 (0.010–0.016) and somewhat broader in R11 (0.040; \( P<0.05 \)) and R14 (0.086; \( P<0.001 \)). R14 still showed a bimodal pattern of distribution, supported by high bootstrap values for each lineage. Despite original exposure to donor DE with broad quasispecies, no recipient changed to the clades of any other recipient.

To facilitate longitudinal comparisons over 5 years within the same person, individual neighbour-joining trees based on nucleotide p-distances between clones were calculated for DE, R7, R8, R10, R11, R13 and R14 (Fig. 4). For DE, several 1998 variants were related to 1993 precursors (upper half of tree), but several lineages were lost and a new cluster had appeared. In the recipients, the nucleotide evolution was much more restricted, in particular for R7, R8, R10 and R11. Recipients R13 and R14 changed more, leaving the R13 1998 quasispecies as relatively restricted while the bimodal pattern in R14 had evolved.

**Evolution of predicted amino acids**

Further patterns emerged during follow-up when synonymous and non-synonymous substitutions were analysed from all clones and compared, both as p-distances and as the mean number of predicted amino acid differences. Distance and difference data from each individual were calculated at the two time-points (1993 and 1998) (Table 1). In addition, the evolution within each individual was calculated as distances and differences between the 1993 and 1998 clones, the latter procedure giving indications of amino acid changes between the two time-points in each study subject. As shown, the two subjects with the greatest nucleotide heterogeneity (DE and R14) also had a strong predominance of non-synonymous compared with synonymous substitutions in HVR-1. In agreement, DE showed the highest mean amino acid difference from 1993 to 1998, with a mean value of 9.2 amino acid differences during the this period, compared with a mean value of 6.1 differences in R14. Recipient R13, who had relatively homogeneous quasispecies in both 1993 and again in 1998, did not display a high non-synonymous substitution rate at either of the time-points, but nevertheless had a mean of 5.4 different amino acids over the 5-year period. In contrast, the substitutions occurring in the four recipients R7, R8, R10 and R11 were predominantly synonymous and led to neither amino acid heterogeneity at either time-point nor changes in amino acids during follow-up from 1993 to 1998. The differences described were supported by high significance levels, as shown in Table 1 (right column).

**DISCUSSION**

This study was based predominantly on samples from 1993 and represents a transmission situation that was common in the past where single active HCV-infected blood donors often infected many recipients. HCV transmission ceased with HCV antibody testing of blood donors (Alter & Houghton, 2000). Our molecular follow-up represents the natural course of infection before antiviral treatment was generally implemented.

Several main points of this study should be emphasized. First, and the key finding of our study, is that a substantial fraction (here four of six) of HCV-infected recipients, despite original exposure to a broad quasispecies, exhibited a narrow quasispecies pattern over at least a 5-year period of observation. In one case (R11) the recipient never developed HCV antibodies, whereas all the others did. This may suggest that the humoral immunity affects quasispecies evolution minimally. The lack of viral drift in several cases indicates that, despite an error prone viral RNA polymerase, viral variability and quasispecies development may be limited in a substantial number of patients and in this respect HCV infection may resemble GBV-C infection. Indeed several studies (Ross et al., 2000; Allain et al., 2000) mention that a few patients show this conserved pattern, which is in contrast to the favoured view of a generally rapidly mutating HVR-1 region in HCV infection (Kato et al., 1994). The few mutations that occurred during follow-up in recipients R7, R8, R10 and R11 were mainly synonymous. In contrast, donor DE and recipient R14...
Fig. 4. Longitudinal nucleotide changes in donor DR and the six recipients who were followed from 1993 to 1998 depicted as linear neighbour-joining trees, one for each individual. Quasispecies clones in the 1993 samples are shown as filled black circles and those from 1998 as open circles. Each tree includes prototype strains HCV-1 and HCV-H which are displayed as filled boxes at the bottom of the tree. All trees are presented at the same scale.
Table 1. Amino acid substitutions in the HVR-1 within donor DE and his six recipients at two time-points (1993 and in 1998) and evolution in each individual between 1993 and 1998

<table>
<thead>
<tr>
<th>Type of comparison*</th>
<th>DE</th>
<th>R7</th>
<th>R8</th>
<th>R10</th>
<th>R11</th>
<th>R13</th>
<th>R14</th>
<th>Statistical comparisons†</th>
</tr>
</thead>
<tbody>
<tr>
<td>1993</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Synonymous substitutions</td>
<td>0.090</td>
<td>0.023</td>
<td>0.037</td>
<td>0.019</td>
<td>0.027</td>
<td>0.043</td>
<td>0.006</td>
<td>DE vs all six recipients, ( P \leq 0.03 )</td>
</tr>
<tr>
<td>Non-synonymous substitutions</td>
<td>0.186</td>
<td>0.010</td>
<td>0.004</td>
<td>0.014</td>
<td>0.007</td>
<td>0.000</td>
<td>0.058</td>
<td>DE vs all six recipients, ( P \leq 0.00003 ) and R14 vs five recipients, ( P &gt; 0.24 )</td>
</tr>
<tr>
<td>Amino acid differences</td>
<td>6.8</td>
<td>0.47</td>
<td>0.18</td>
<td>0.55</td>
<td>0.33</td>
<td>0.00</td>
<td>2.7</td>
<td>DE vs all six recipients, ( P \leq 0.00003 )</td>
</tr>
<tr>
<td>1998</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Synonymous substitutions</td>
<td>0.131</td>
<td>0.035</td>
<td>0.054</td>
<td>0.008</td>
<td>0.017</td>
<td>0.011</td>
<td>0.121</td>
<td>DE vs all recipients except R11, ( P \leq 0.00003 ), with DE vs R11, ( P = 0.43 ); R11 vs R7, R8, R10 and R13, ( P \leq 0.002 )</td>
</tr>
<tr>
<td>Non-synonymous substitutions</td>
<td>0.258</td>
<td>0.003</td>
<td>0.000</td>
<td>0.005</td>
<td>0.017</td>
<td>0.011</td>
<td>0.121</td>
<td>DE vs all six recipients, ( P \leq 0.00003 ) and R14 vs all other recipients, ( P \leq 0.003 )</td>
</tr>
<tr>
<td>Amino acid differences</td>
<td>8.7</td>
<td>0.15</td>
<td>0.00</td>
<td>0.27</td>
<td>0.82</td>
<td>0.62</td>
<td>4.7</td>
<td>DE vs all six recipients, ( P \leq 0.00003 ); R14 vs all other recipients, ( P \leq 0.00004 )</td>
</tr>
<tr>
<td>1993 to 1998</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Synonymous substitutions</td>
<td>0.134</td>
<td>0.073</td>
<td>0.097</td>
<td>0.094</td>
<td>0.105</td>
<td>0.048</td>
<td>0.046</td>
<td>DE vs all six recipients, ( P \leq 0.00003 ) and R7 or R13 vs five recipients, ( P \leq 0.037 ); R8 vs R10 vs R11, ( P &gt; 0.7 ); and R13 vs R14, ( P = 1.0 )</td>
</tr>
<tr>
<td>Non-synonymous substitutions</td>
<td>0.255</td>
<td>0.007</td>
<td>0.022</td>
<td>0.040</td>
<td>0.015</td>
<td>0.116</td>
<td>0.146</td>
<td>DE vs all six recipients, ( P \leq 0.00003 ); R13 and R14 vs all other recipients, ( P \leq 0.00003 ); but R13 vs R14, ( P = 0.3 ); and R7 vs R8 vs R10 vs R11, ( P &gt; 0.075 )</td>
</tr>
<tr>
<td>Amino acid differences</td>
<td>9.2</td>
<td>0.35</td>
<td>1.1</td>
<td>1.2</td>
<td>0.72</td>
<td>5.4</td>
<td>6.1</td>
<td>DE vs all six recipients, ( P \leq 0.00003 ); R13 and R14 vs R7, R8, R10 and R11, within R7, R8, R10, R11 ( P &gt; 0.11 ), R13 vs R14, ( P = 0.44 )</td>
</tr>
</tbody>
</table>

*Comparisons were made on sequences from clones both by p-distances for synonymous and non-synonymous substitutions (Kumar) and for amino acid differences by MEGA 2.1.
†The statistical comparisons were made by Tukey honest significant difference test for unequal N.
steady state can be maintained for long periods. This is partially in contrast with the common view of a rapidly evolving HVR-1 in the HCV and should be evaluated further. The question of the role of viral variants in the evolution of the HCV still remains to be answered, as well as their significance in the pathogenesis of disease.

ACKNOWLEDGEMENTS

This study was supported by grants from the Medical Faculties of Lund University (ALF), Malmö University Hospital Cancer Foundation and Alfred Osterlund Foundations as well as the Science Foundation of the National University Hospital, Iceland. The authors want to thank Jan-Åke Nilsson, Malmö, Sweden, for help with the statistical calculations.

REFERENCES


