UNIVERSITÀ DEGLI STUDI DI CAGLIARI

Department of Biomedical Sciences and Technologies
Section of General Microbiology and Virology
& Microbial Biotechnologies

Research Doctorate in:

Development and Evaluation of Antiviral Drugs

XX Cycle – A.A. 2004-2007

Coordinator of the Doctorate:
Prof. Paolo La Colla

HCV quasispecies evolution in patients with different outcome of HCV chronic infection

Supervisor:

Dr. Luchino Chessa

Candidate:

Dr. Cinzia Balestrieri

Introduction

Infection with the hepatitis C virus (HCV) is currently a major cause of chronic hepatitis , cirrhosis, end-stage liver disease and hepatocellular carcinoma in most Western Countries where it represents the most frequent reason for liver transplantation. The global prevalence of HCV averages 3%, with an estimated nearly 300 million persons infected worldwide (1,2,3). Although universal screening of blood donors in developed countries and improvements in infection control measure have significantly decrease exposure to the virus, the large reservoir of chronically infected individuals, the high evolutionary potential of the virus, the lack of routine screening of blood products in the developing countries, the use of traditional medicine and tattoing in some culture, the increasing fluxes of immigration from endemic areas to less prevalent regions strongly support the hypothesis that HCV is still spreading throughout the world. One striking characteristic of HCV infection is its high chronicity rate. HCV establishes persistent infection in up to 85% of infected individuals. The clinical features associated with chronic HCV infection are highly heterogeneous and the natural course of the underlying liver disease is also extremely variable and still not fully understood. Some patients remain HCV viremic without biochemical evidence of liver damage; some seem to have a static form of chronic hepatitis characterized by persistently elevated aminotransferase levels without overt symptoms or disease advancement; some progress over a period to histologic fibrosis and cirrhosis; some have long-term stable cirrhosis identified only through liver biopsy; some have progressive cirrhosis that culminates in liver failure; and, finally, some develop hepatocellular carcinoma. It is estimated that 15 to 20 % of chronically infected individuals undergo liver cirrhosis in a decade or so after infection, with hepatocellular carcinoma arising from cirrhosis at an estimated rate of 1 to 4% per year (4).

HCV is an enveloped positive-strand RNA virus of the genus Hepacivirus in the family Flaviviridae. The HCV genome is comprised of a positive-stranded RNA molecule of about 9500 nucleotides (5) containing a single long translational open reading frame (ORF) that encodes a large polypeptide of approximately 3000 amino acids, beginning with the first in-frame methionine codon (6-8) bounded by 5’ and 3’ untranslated regions of approximately 341 nucleotides (fig 1). Analysis of the structure of HCV particles has been

hampered by the low titre of virus in infectious sera and the difficulties of replicating the virus in culture system. It is likely that HCV is an enveloped icosahedral virus. However, detailed structural analysis of HCV virions is still lacking. Particles with a diameter of 45-65 nm have been observed by electron microscopy in human plasma (9) and in chimpanzee and human liver chronically infected with HCV (10,11). HCV-like particles have been detected in experimentally infected (9) or transfected cell lines (12). There is evidence that HCV particles are present in the circulation as immune complexes (13) or in association with serum lipoproteins (14-17). The large ORF extends throughout most of the HCV RNA genomic sequence and encodes a polypeptide of between 3010 and 3033 amino acids, depending on the source of viral isolate. ORF encodes a polyprotein precursor that is processed co- and post-translationally to yield a variety of structural and non- structural (NS) protein (Fig 2). Structural proteins are processed from the N-terminal region of the HCV polyprotein precursor, beginning with an RNA-binding nucleocapsid polypeptide of basic charge (Core; ca. 23 kDa) followed by two glycoproteins-E1 (33-35 kDa) and E2 (68-72 kDa). Additional proteins derived from core, 21 kDa and 17 kDa and a 7 kDa protein (p7) wich lies between E2 and NS2 have also been found. The non-structural (NS) proteins, NS2, NS3, NS4A, NS4B, NS5A, and NS5B, are generated by proteases encoded by the virus. NS2 is a metallo protease that cleaves, in cis, its junction with NS3. NS3 is a serine proteinase that cleaves its junction with NS4A, in cis, the other junctions in trans. The function of NS4B protein is not yet clear. NS5A has been implicated as having a role in inhibiting host cell responses to IFN. NS5B is the RNA-dependent RNA polymerase. The numbering of the NS proteins is derived historically from studies on viruses belonging to the genus flavivirus. HCV and pestivirus do not possess an NS1. The processing of the putative region of the HCV polyprotein is mediated at least in part by the host signal peptidases. E1 and E2 are essential components of the HCV virion envelope and they are necessary for virion entry into host cells by direct receptor binding and possibly for membrane fusion. Internal signal sequences upstream of the E1 and E2 proteins direct the polyprotein precursor to the endoplasmic reticulum (ER), where they are translocated into the lumen, and after signal sequence cleavage (18) they remain anchored inside the lumen (18,19). E1 and E2 are integral membrane proteins and are glycosylated in the ER lumen. E1 and E2 have binding sites for low-density lipoprotein and E1 additionally for high-density lipoprotein. E1 and E2 are cotranslationally separated from each other before they assume their complex conformation as heterodimers in the ER (20). E1 and E2 or their heterodimer are not secret and must be extracted from transfected animal cells (21). It

cannot be excluded that mature HCV particles have a different structure than that deduced from data on vitro expression and from analogies with other genera of the Flaviviridae. The C-terminal region of E1 domain is very hydrophobic and terminates with the hydrophobic E2 signal sequence, which is cleaved by signal peptidase upon translocation into the ER lumen. E1 and E2 proteins can be co-precipitated by antibodies against E1 or E2 when they are expressed together, indicating that these two proteins are associated in HCV-infecetd cells. E1 and E2 were predominantly located in the ER rather than at the cell surface, suggesting that there may be mechanism for retention of these proteins in this compartment and that they are not translocated beyond the Golgi. This fact support the hypothesis that the HCV buds from the endoplasmic reticulum and is released from cells via the endocytosis pathway. E2 protein binds to CD81, a tetraspanin which is present on the surface of B lymphocytes and hepatocytes. Binding of CD 81 by E2 induced the aggregation of lymphocytes and inhibited B-cell proliferation (22) and blocked natural killer cell activation, cytokine production and cytotoxic granular release and proliferation (23,24). This suggests that E2 inhibition of cells of the immune system may be an evasion strategy by HCV to establish chronic infection. But there are indication that CD 81 is not the main/only receptor fro HCV (25,26).

The HCV genome exhibits a considerable degree of sequence variation, and on the basis of these variants, HCV is classified into at least six phylogenetically distinct genotypes which in turn are subdivided into numerous, more than 60, subtypes (27). Viral sequences have been shown to vary up to more than 30% across the entire genome among genotypes, 20% among subtypes, and up to 10% within a subtype(27,28). Moreover, the HCV genome in single hosts is described as a dynamic population of different but closely related genomes designated quasispecies (29). The high variability of HCV is attributed to a limited fidelity of the RNA dependent RNA polymerase encoded by non-structural protein NS5B, and a high rate of viral replication. A model of viral kinetics led to the estimation that more than 10^{12} virions with a virion half-life of 2.7 h are produced each day in an infected person (30). HCV variability plays a crucial role in escaping the host immune surveillance and establishing persistent infection. This quasispecies composition undergoes extensive variations during the course of chronic infection, which may result from the accumulation of random substitutions or from changes in the predominant quasispecies population under the selective pressure of the immune response. Beside the overall variability of the HCV genome, hypervariability exists in the 27 amino acid long hypervariable region-1 (HVR-1) located at the N terminus of glycoprotein E2. HVR-1 displays marked sequence variability

with as many as 1.5 X 10^{-2} excanges/year in HVR-1 and only 0.9 to 5.2 X 10^{-3} in E1 (31). HVR-1 is one of the main targets of immune response against HCV. It is thought to contain a putative immunodominant linearB-cell epitope recognized by neutralizing antibodies (32-35). It has been proposed that escape mutations in HVR-1 evade the limited cross-reactivity of the antibody response and play a key role in the establishment of persistent infection (36,37). However, the biological role of HVR-1 remains unclear.

HVR-1 has usually served as the sequence basis for quasispecies analysis to monitor virus evolution, disease progression, and therapy response. It is important to understand whether there are any genetic differences that might account for the different clinical course of chronic HCV infection. A number of previous study have investigated the HVR-1 quasispecies evolution in different clinical setting and according to therapy outcome. However, the findings of these studies have been somewhat contradictory. In the present study, we evaluated the molecular evolution of HVR-1 quasipsecies in different outcome of HCV infection testing some novel bioinformatics tools.

Patients

We downloaded E1E2 RNA serum sequences from the HCV sequence database of the Los Alamos National Laboratory of 5 patients with different outcome of HCV infection. All the patients were drawing from the cohort studied by Farci et al (38). Briefly, serial serum samples were obtained from 2 patients with a mild and stable disease for more than 20 years (slow progressor), and from 3

patients with severe disease leading to liver-related death within 5 years of the onset of infection (rapid progressor). For all patients in the database were present the first available polymerase chain reaction positive sample (within 2 to 5 weeks of transfusion, time point zero), one sample before antibody seroconversion (time point 1), and one sample after antibody seroconversion (time point 2). For each time point a mean of 9.7 molecular clones were available for the analysis. All 5 patients were infected with HCV genotype 1, 3 patients with subtype 1a (1 slow progressor, and 2 rapid progressor), and the remaining 2 patients were infected with subtype 1b (1 slow and 1 rapid progressor). All the sequences were 558 nucleotides in length, starting from position 1293 and ending to position 1850 (start coordinates based on H77 accession number NC004102).

Computer analysis

Mutation Master: Mutation Master generates profiles of point mutations in a population of sequences and produces a set of visual displays and tables indicating the number, frequency, and character of substitutions. It can be used to analyze hundreds of sequences at a time. Mutation Master sorts and counts the mutations and presents this information in a series of graphs and tables. In the Rank Order and Frequency Plot , the height of the red bar indicates the fraction of the population with the most common substitution. The frequency of the second most common mutation is indicated by the height of a yellow bar. Heights of orange, green, and blue bars indicate the frequency of successively less common mutations. The program also tabulates the percentage of sequences with each substitution (up to a maximum of six). The exact number of different amino acids at each site is given in the “Number of Amino Acids Plot”. This counting feature aids the detection of highly conserved domains and domains with limited diversity. In the BLOSUM Score Plot, Mutation Master identifies each mutation and provides information about its conservative or nonconservative nature. The major substitutions at each position are scored according to the appropriate BLOSUM table. Positive numbers indicate a conservative substitution, and negative numbers a nonconservative substitution. The colour of each box corresponds to that used by the

Rank Order and Frequency Plot to indicate the relative frequency of each substitution. The position of each amino acid is recorded. A second line of numbers indicates the total number of mutant amino acids at each position. BLOSUM tables contain information about the frequency of various mutations in families of related proteins. Commonly observed mutations have BLOSUM scores that are positive numbers or zero; rare mutations have BLOSUM scores that are negative numbers. BLOSUM scores depend upon the overall conservation of the family of proteins under consideration. In the BLOSUM tables, the highest score is typically that of the wild-type amino acid, and the lowest score is that of a very rare substitution that is likely to alter the structure/function of a protein (such as a highly charged basic amino acid in the place of a

bulky hydrophobic amino acid). If a position has more than four mutant amino acids, a blank appears in this line of numbers.

Logos analysis: Sequence logo were invented by Tom Schneider and Mike Stephens to display patterns in sequence conservation, and to assist in discovering and analyzing those patterns. The object of a sequence logo is to visualize the information contained in a set of DNA, RNA, or protein sequences by examining the order and frequency of the chemical subunits which make up the sequences. The name “sequence logo” comes from the fact that a set of sequences is being represented as a single graphic which contains one or more separate elements. The sequence logo functions by graphically representing the conservation of a set of sequences in a clear, concise and mathematically sound manner. Sequence logos are generated by programs which look at the sequences and analyze them using the information theory developed by Claude Shannon. The process of generating a sequence logo is somewhat similar to that of creating a consensus sequence, but unlike a consensus, subtle feature of the data are retained. The logo generation form can process RNA, DNA, or multiple sequence alignments provided in either FASTA or CLUSTAL formats. A logo represents each column of the alignment by a stack of letters, with the height of each letter proportional to the sequence conservation, measured in bits, at that position. The letters of each stack are ordered from most to least frequent, so that one may read the consensus sequence from the tops of the stacks. Shannon entropy is a simple quantitative measure of uncertainty in a data set. One qualitative way to think about it in terms of sequences is that if a sample set is drawn from a large population, the Shannon entropy could be considered as a measure indicative of the ability to guess what amino acids would be in the next sequence took from the population, based on the previous sampling.

Results

The mean number of nucleotides and amino acids mutations did not differ significantly between slow and rapid progressor in the first available PCR-positive sample, in the pre-antibody seroconversion sample and similarly in the sample after the antibody seroconversion. When we focused the analysis on the HVR-1 region a trend toward an increase in the mean number of mutations is noted after antibody seroconversion, both at the nucleotides and amino acids level, despite the difference did not reach the statistical significance ($p=0.09$). To evaluate the intra-sample changes over time of sequence variability, we next calculated the average number of nucleotide and amino acid changes

from the consensus sequence at time point zero with the sequences in each group for each time point. After antibody seroconversion, patients with rapid disease progression showed a statistically significant increase in the mean number of mutations both at nucleotide ($p=0.02$) and amino acid level ($p=0.003$), whereas those with mild and stable disease did not show any significant difference. The mean number of mutations did not differ significantly in the pre-antibody seroconversion sample in both groups. We next looked at differences in Shannon entropy between slow and rapid progressor in HVR-1 region after antibody seroconversion (Fig.3). At 29 nucleotide positions (shown in red in the figure) the two sets differed significantly in the degree of conservation. Interestingly, nine of these sites showed greater conservation in the rapid progressor group. A positive value indicates greater conservation in the rapid progressor sequence set. When the difference in Shannon entropy was assessed in the group of patients with severe and rapidly progressive disease in the first available PCR-positive sample and after antibody seroconversion the differences was statistically significant at 7 nucleotides positions. The position 44 showed a greater conservation after antibody seroconversion (Fig.4). Successively, we investigated the number, frequency, and character of amino acid substitutions in HVR-1 region in both study groups (Fig.5). It allowed to point out some interesting differences that deserve further investigation. Particularly noteworthy the difference at position 12 and 15. In fact at position 15 the consensus amino acid in the set of slow progressor was Glycine , and it is completely conserved throughout each time point. In the set of rapid progressor the consensus amino acid was Glycine again in the first available PCR-positive sample but changed in Serine after antibody seroconversion. Moreover the codon showed an increasing frequency of mutation going to antibody seroconversion. A similar pattern was show at the position 12. That is, in the set of sequences of patients with slow and mild disease the amino acid site is completely stable over time, whereas in the set of sequences from patients with severe and rapidly progressive disease the site undergoes increasing changes. Several amino acid sites were highlighted in the interest score graph in both group as uncommon amino acid pattern. Additionally we assessed the logo analysis (Fig.6). We performed Chi Square test to determine whether there was any position within HVR-1 that showed a composition that differed statistically between the sequences sets. The results for positions 12 and 15 were statistically significant ($p=0.01$ and $p=0.006$, respectively).

Conclusion

Using several bioinformatics approach, a quasispecies analysis of E1/E2 envelope protein and, in particular, of HVR-1 region is performed in samples from five patients presenting different natural course of HCV chronic infection. We reconfirm a trend toward an increase in the degree of viral diversity after antibody seroconversion in patient that exhibit the more unfavourable outcome. This is particularly relevant in the HVR1 region, where there is a statistically significant association between the mean number of mutation and poor disease outcome. Viral sequence changes likely represent selective transmission or outgrowth of particular clones of quasispecies rather than the novo mutations, with unselective acceptance of synonymous sequence polymorphisms but conservation of the predominant, functional amino acid sequence. The study of Farci et al, source of our patients, already revealed that the mean number of nonsynonymous substitutions per site per week within HVR-1 was significantly higher in progressing hepatitis (38). In expressed genes the rate of nonsynonumpus changes also reflects functional constraints acting in a dynamic interplay between conservative and evolutionary selective pressure. Several studies have analyzed the d_N/d_S ratio from patients with different outcomes and found that patients with persistent disease tend to have a higher d_N/d_S ratio in HVR-1 than patients who resolved infection, suggesting the action of continual immune-driven positive selection. Moreover specific virus-cell receptor interaction could play a crucial role in determining the immune response and the cellular activation. Interestingly, there are some uncommon and specific residues substitutions in HVR-1 region of patients with severe disease that may play an important structural and functional role given that these position are highly conserved among sequences of patients with mild disease. It is reasonable to suggest that the mutation could represent an adaptation of this region in the interaction with other molecules. More intesives studies in a well definite setting of patients are necessary to clear the role of HVR-1 quasispecies evolution during the natural course of HCV infection.

References

1 Alter MJ, Kruszon-Moran D, Nainan OV, McQuillan GM, Gao F, Moyer LA, Kaslow RA, Margolis HS. The prevalence of hepatitis C virus infection in the United States, 1988 through 1994. N Engl J Med. 1999 Aug 19;341(8):556-62.

2 Brown RS Jr, Gaglio PJ. Scope of worldwide hepatitis C problem. Liver Transpl. 2003 Nov;9(11):S10-3.

3 Wasley A, Alter MJ. Epidemiology of hepatitis C: geographic differences and temporal trends. Semin Liver Dis. 2000;20(1):1-16.

4 Lauer GM, Walker BD. Hepatitis C virus infection. N Engl J Med. 2001 Jul 5;345(1):41-52.

5 Choo QL, Kuo G, Weiner AJ, Overby LR, Bradley DW, Houghton M.
Isolation of a cDNA clone derived from a blood-borne non-A, non-B viral hepatitis genome. Science. 1989 Apr 21;244(4902):359-62.

6 Kato N, Hijikata M, Ootsuyama Y, Nakagawa M, Ohkoshi S, Sugimura T, Shimotohno
Molecular cloning of the human hepatitis C virus genome from Japanese patients with non-A, non-B hepatitis. Proc Natl Acad Sci U S A. 1990 Dec;87(24):9524-8.

7 Takamizawa A, Mori C, Fuke I, Manabe S, Murakami S, Fujita J, Onishi E, Andoh T, Yoshida I, Okayama H. Structure and organization of the hepatitis C virus genome isolated from human carriers.J Virol. 1991 Mar;65(3):1105-13.

8 Choo QL, Richman KH, Han JH, Berger K, Lee C, Dong C, Gallegos C, Coit D, Medina-Selby R, Barr PJ, et al. Genetic organization and diversity of the hepatitis C virus. Proc Natl Acad Sci U S A. 1991 Mar 15;88(6):2451-5.

9 Kaito M, Watanabe S, Tsukiyama-Kohara K, Yamaguchi K, Kobayashi Y, Konishi M, Yokoi M, Ishida S, Suzuki S, Kohara M. Hepatitis C virus particle detected by immunoelectron microscopic study. J Gen Virol. 1994 Jul;75 (Pt 7):1755-60.

10 Shimizu YK, Feinstone SM, Kohara M, Purcell RH, Yoshikura H. Hepatitis C virus: detection of intracellular virus particles by electron microscopy. Hepatology. 1996 Feb;23(2):205-9.

11 De Vos R, Verslype C, Depla E, Fevery J, Van Damme B, Desmet V, Roskams T. Ultrastructural visualization of hepatitis C virus components in human and primate liver biopsies. J Hepatol. 2002 Sep;37(3):370-9.

12 Mizuno M, Yamada G, Tanaka T, Shimotohno K, Takatani M, Tsuji T. Virion-like structures in HeLa G cells transfected with the full-length sequence of the hepatitis C virus genome. Gastroenterology. 1995 Dec;109(6):1933-40.

13 Hijikata M, Shimizu YK, Kato H, Iwamoto A, Shih JW, Alter HJ, Purcell RH, Yoshikura H. Equilibrium centrifugation studies of hepatitis C virus: evidence for circulating immune complexes. J Virol. 1993 Apr;67(4):1953-8.

14 Thomssen R, Bonk S, Propfe C, Heermann KH, Köchel HG, Uy A. Association of hepatitis C virus in human sera with beta-lipoprotein. Med Microbiol Immunol. 1992;181(5):293-300.

15 Prince AM, Huima-Byron T, Parker TS, Levine DM. Visualization of hepatitis C virions and putative defective interfering particles isolated from low-density lipoproteins. J Viral Hepat. 1996 Jan;3(1):11-7.

16 Thomssen R, Bonk S. Virolytic action of lipoprotein lipase on hepatitis C virus in human sera. Med Microbiol Immunol. 2002 May;191(1):17-24.

17 Andre P, Komurian-Pradel F, Deforges S, Perret M, Berland JL, Sodoyer M, Pol S, Brechot C, Paranhos-Baccala G, Lotteau V. Characterization of low- and very-low-density hepatitis C virus RNA-containing particles. J Virol. 2002 Jul;76(14):6919-28.

18 Hijikata M, Kato N, Ootsuyama Y, Nakagawa M, Shimotohno K. Gene mapping of the putative structural region of the hepatitis C virus genome by in vitro processing analysis. Proc Natl Acad Sci U S A. 1991 Jul 1;88(13):5547-51.

19 Spaete RR, Alexander D, Rugroden ME, Choo QL, Berger K, Crawford K, Kuo C, Leng S, Lee C, Ralston R, et al. Characterization of the hepatitis C virus E2/NS1 gene product expressed in mammalian cells. Virology. 1992 Jun;188(2):819-30.

20 Dubuisson J, Rice CM. Hepatitis C virus glycoprotein folding: disulfide bond formation and association with calnexin. J Virol. 1996 Feb;70(2):778-86.

21 Ralston R, Thudium K, Berger K, Kuo C, Gervase B, Hall J, Selby M, Kuo G, Houghton M, Choo QL. Characterization of hepatitis C virus envelope glycoprotein complexes expressed by recombinant vaccinia viruses. J Virol. 1993 Nov;67(11):6753-61.

22 Flint M, Maidens C, Loomis-Price LD, Shotton C, Dubuisson J, Monk P, Higginbottom A, Levy S, McKeating JA. Characterization of hepatitis C virus E2 glycoprotein interaction with a putative cellular receptor, CD81. J Virol. 1999 Aug;73(8):6235-44.

23 Tseng CT, Klimpel GR. Binding of the hepatitis C virus envelope protein E2 to CD81 inhibits natural killer cell functions. J Exp Med. 2002 Jan 7;195(1):43-9.

24 Crotta S, Stilla A, Wack A, D'Andrea A, Nuti S, D'Oro U, Mosca M, Filliponi F, Brunetto RM, Bonino F, Abrignani S, Valiante NM. Inhibition of natural killer cells through engagement of CD81 by the major hepatitis C virus envelope protein. J Exp Med. 2002 Jan 7;195(1):35-41.

25 Petracca R, Falugi F, Galli G, Norais N, Rosa D, Campagnoli S, Burgio V, Di Stasio E, Giardina B, Houghton M, Abrignani S, Grandi G. Structure-function analysis of hepatitis C virus envelope-CD81 binding. J Virol. 2000 May;74(10):4824-30.

26 Masciopinto F, Freer G, Burgio VL, Levy S, Galli-Stampino L, Bendinelli M, Houghton M, Abrignani S, Uematsu Y. Expression of human CD81 in transgenic mice does not confer susceptibility to hepatitis C virus infection. Virology. 2002 Dec 20;304(2):187-96.

27 Simmonds P, Alberti A, Alter HJ, Bonino F, Bradley DW, Brechot C, Brouwer JT, Chan SW, Chayama K, Chen DS, et al. A proposed system for the nomenclature of hepatitis C viral genotypes.Hepatology. 1994 May;19(5):1321-4.

28 Simmonds P, Smith DB, McOmish F, Yap PL, Kolberg J, Urdea MS, Holmes EC. Identification of genotypes of hepatitis C virus by sequence comparisons in the core, E1 and NS-5 regions. J Gen Virol. 1994 May;75 (Pt 5):1053-61.

29 Martell M, Esteban JI, Quer J, Genescà J, Weiner A, Esteban R, Guardia J, Gómez J. Hepatitis C virus (HCV) circulates as a population of different but closely related genomes: quasispecies nature of HCV genome distribution. J Virol. 1992 May;66(5):3225-9.

30 Neumann AU, Lam NP, Dahari H, Gretch DR, Wiley TE, Layden TJ, Perelson AS.Hepatitis C viral dynamics in vivo and the antiviral efficacy of interferon-alpha therapy. Science. 1998 Oct 2;282(5386):103-7.

31 Hohne M, Schreier E, Roggendorf M. Sequence variability in the env-coding region of hepatitis C virus isolated from patients infected during a single source outbreak. Arch Virol. 1994;137(1-2):25-34.

32 Farci P, Alter HJ, Wong DC, Miller RH, Govindarajan S, Engle R, Shapiro M, Purcell RH. Prevention of hepatitis C virus infection in chimpanzees after antibody-mediated in vitro neutralization. Proc Natl Acad Sci U S A. 1994 Aug 2;91(16):7792-6.

33 Farci P, Shimoda A, Wong D, Cabezon T, De Gioannis D, Strazzera A, Shimizu Y, Shapiro M, Alter HJ, Purcell RH. Prevention of hepatitis C virus infection in

chimpanzees by hyperimmune serum against the hypervariable region 1 of the envelope 2 protein. Proc Natl Acad Sci U S A. 1996 Dec 24;93(26):15394-9.

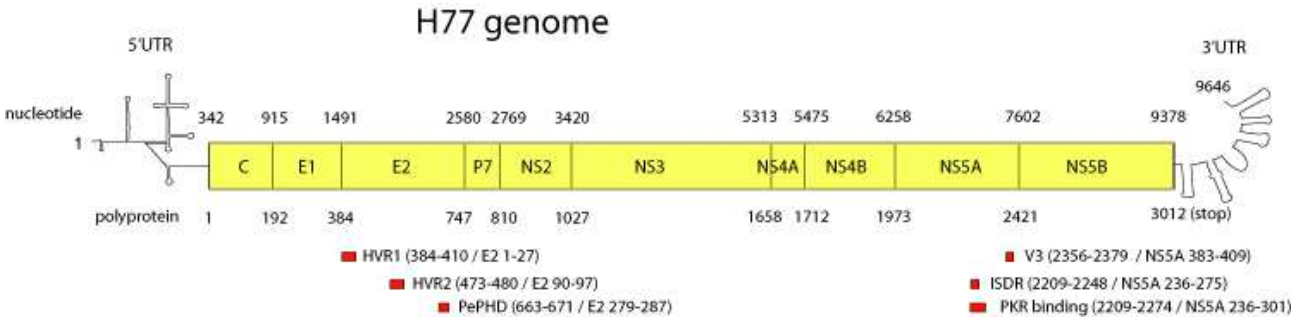
34 Shimizu YK, Hijikata M, Iwamoto A, Alter HJ, Purcell RH, Yoshikura H.Neutralizing antibodies against hepatitis C virus and the emergence of neutralization escape mutant viruses. J Virol. 1994 Mar;68(3):1494-500.

35 Shimizu YK, Igarashi H, Kiyohara T, Cabezon T, Farci P, Purcell RH, Yoshikura H. A hyperimmune serum against a synthetic peptide corresponding to the hypervariable region 1 of hepatitis C virus can prevent viral infection in cell cultures. Virology. 1996 Sep 15;223(2):409-12.

36 Booth JC, Kumar U, Webster D, Monjardino J, Thomas HC. Comparison of the rate of sequence variation in the hypervariable region of E2/NS1 region of hepatitis C virus in normal and hypogammaglobulinemic patients. Hepatology. 1998 Jan;27(1):223-7.

37 Roccasecca R, Ansuini H, Vitelli A, Meola A, Scarselli E, Acali S, Pezzanera M, Ercole BB, McKeating J, Yagnik A, Lahm A, Tramontano A, Cortese R, Nicosia A. Binding of the hepatitis C virus E2 glycoprotein to CD81 is strain specific and is modulated by a complex interplay between hypervariable regions 1 and 2. J Virol. 2003 Feb;77(3):1856-67.

38 Farci P, Shimoda A, Coiana A, Diaz G, Peddis G, Melpolder JC, Strazzera A, Chien DY, Munoz SJ, Balestrieri A, Purcell RH, Alter HJ. The outcome of acute hepatitis C predicted by the evolution of the viral quasispecies. Science. 2000 Apr 14;288(5464):339-44.



Start coordinates, based on H77 (accession NC_004102), from Fields' Virology. Structures of 3'UTR (based on Kolykhalov 1996) and 5'UTR (Honda 1996) adapted from a figure provided by Stuart Ray. Other features based on Major & Feinstone, 1997 and You et al, 2004.

Figure 1 Annoteted genomic map of HCV
From **http://hcv.lanl.gov**

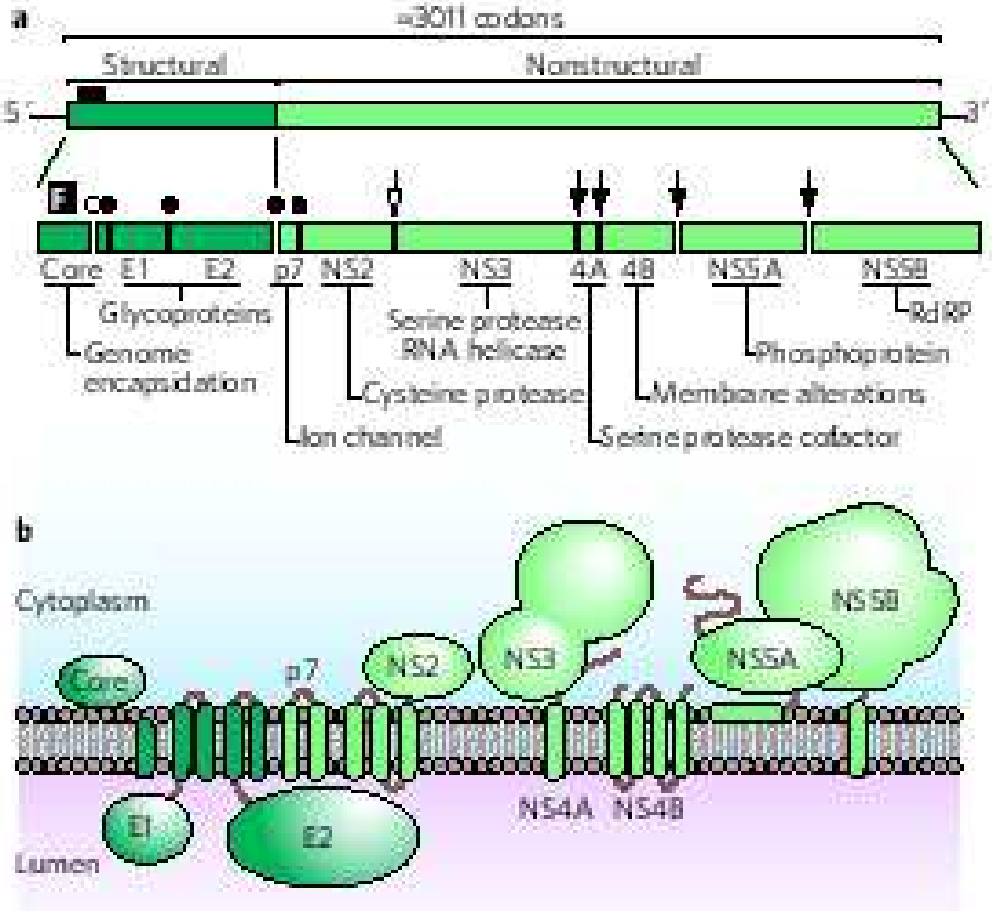


Figure 2 HCV genes and gene products
From: Lindenbach BD, Rice CM. Unravelling hepatitis C virus replication from

Entropy difference in the HVR-1 nt sequences between Slow and Rapid Progressor after antibody seroconversion

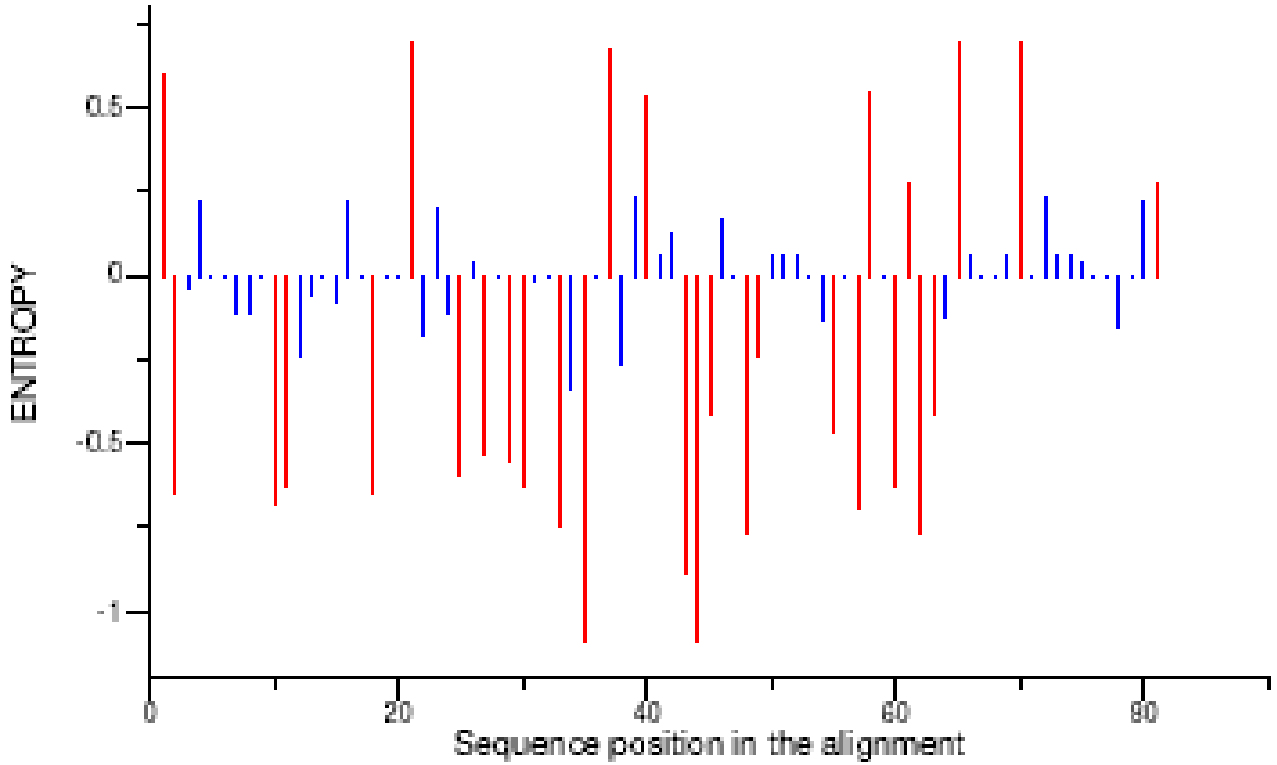
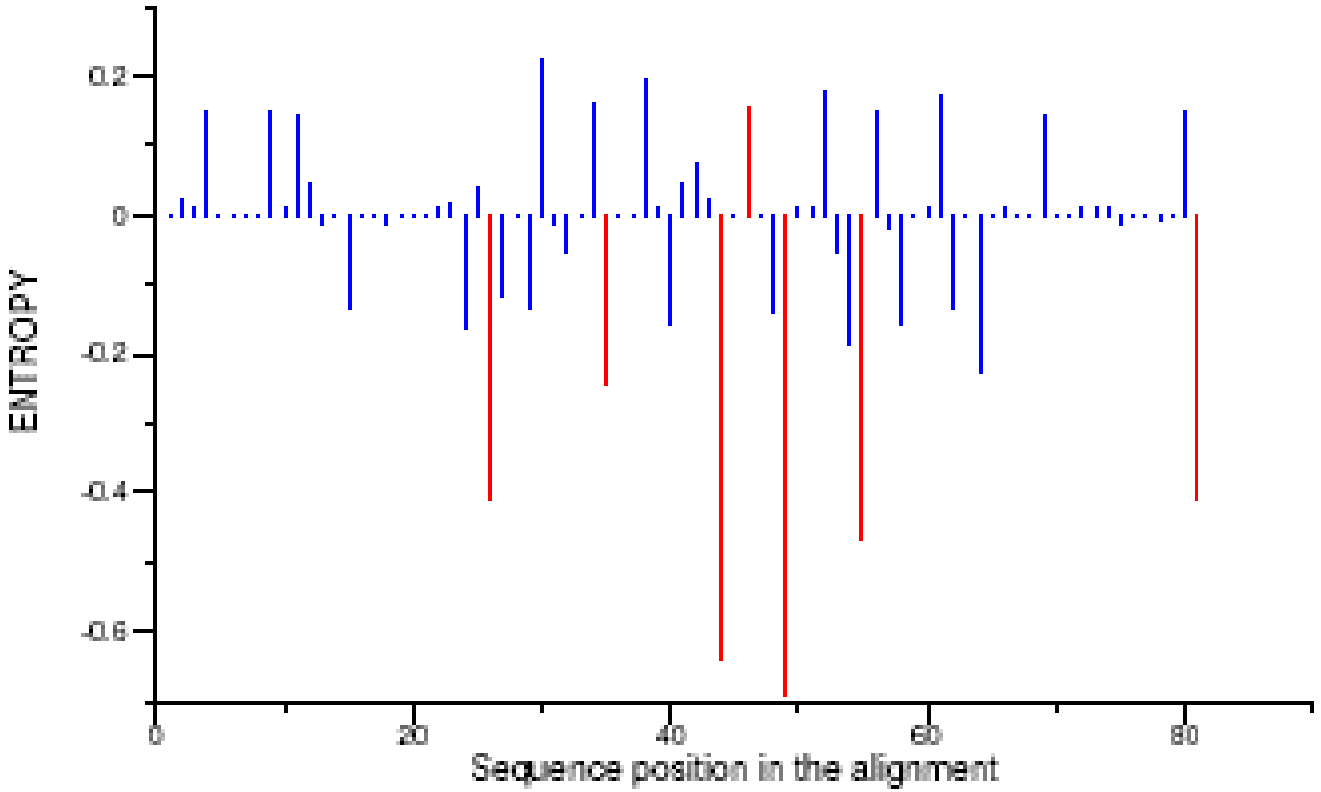


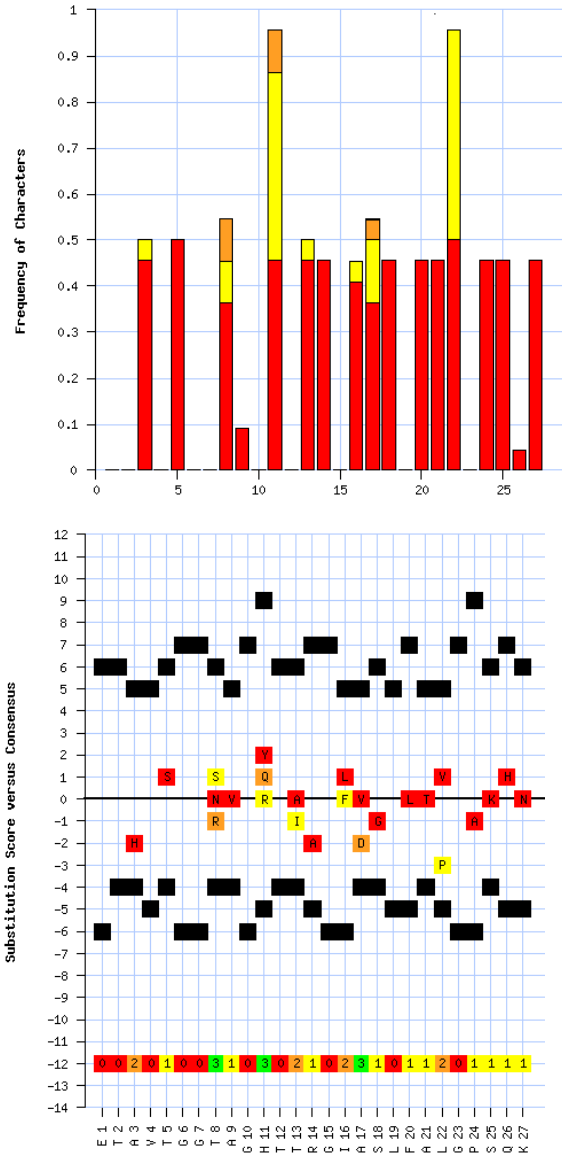
Fig. 3: A positive value indicates greater conservation in the Rapid progressor sequence set. The Slow progressor set contained 18 clones from 2 patients. The Rapid progressor set contained 28 clones from 3 patients

Entropy difference in the HVR-1 nt sequences between Rapid Progressor at different time point



he Rapid progressor set contained 30 clones at time point zero and 28 clones at time point 2

Slow progressor time point zero



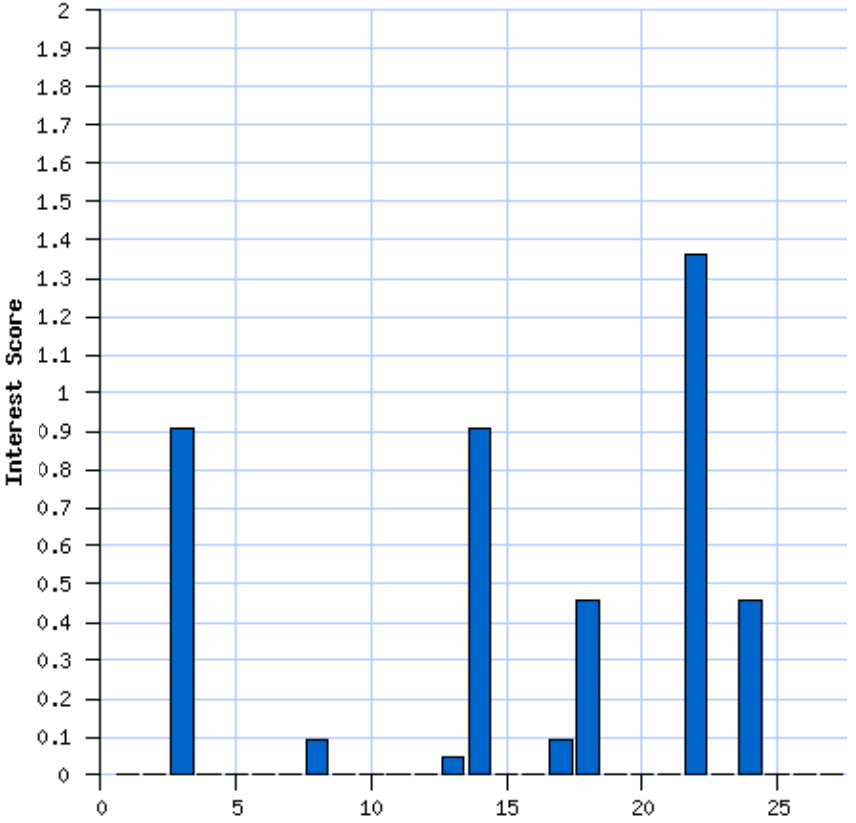
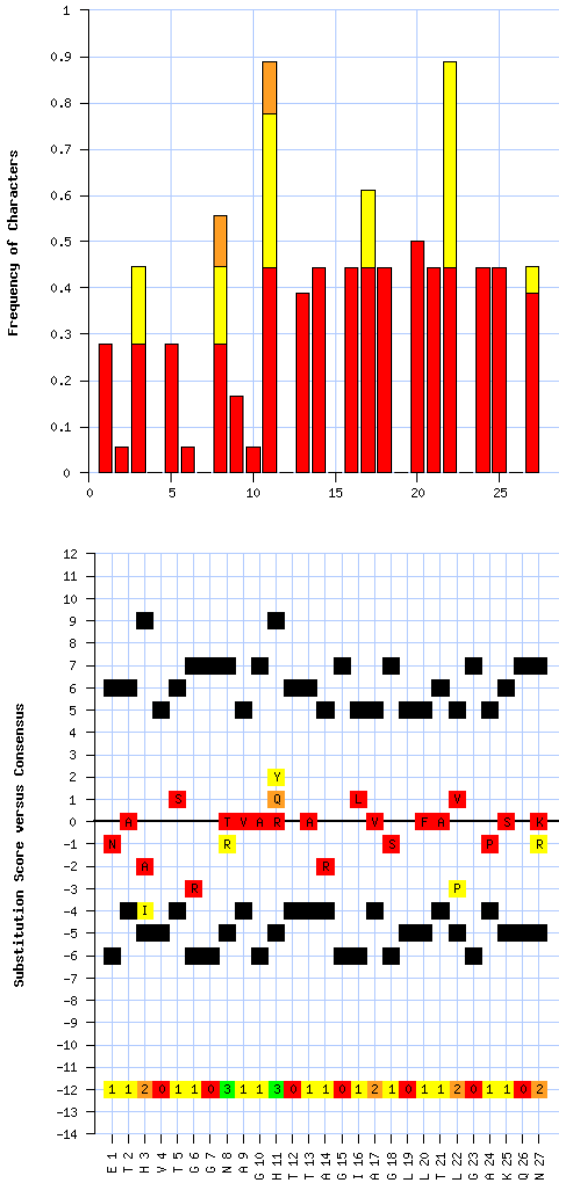


Fig.5a

Slow progressor after antibody seroconversion



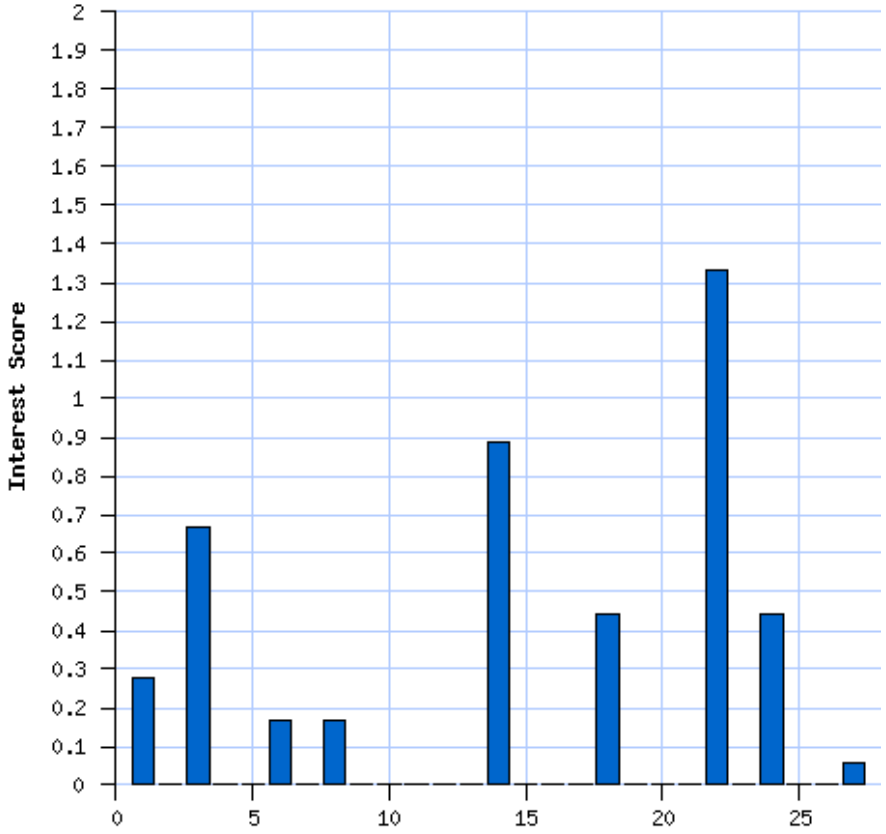
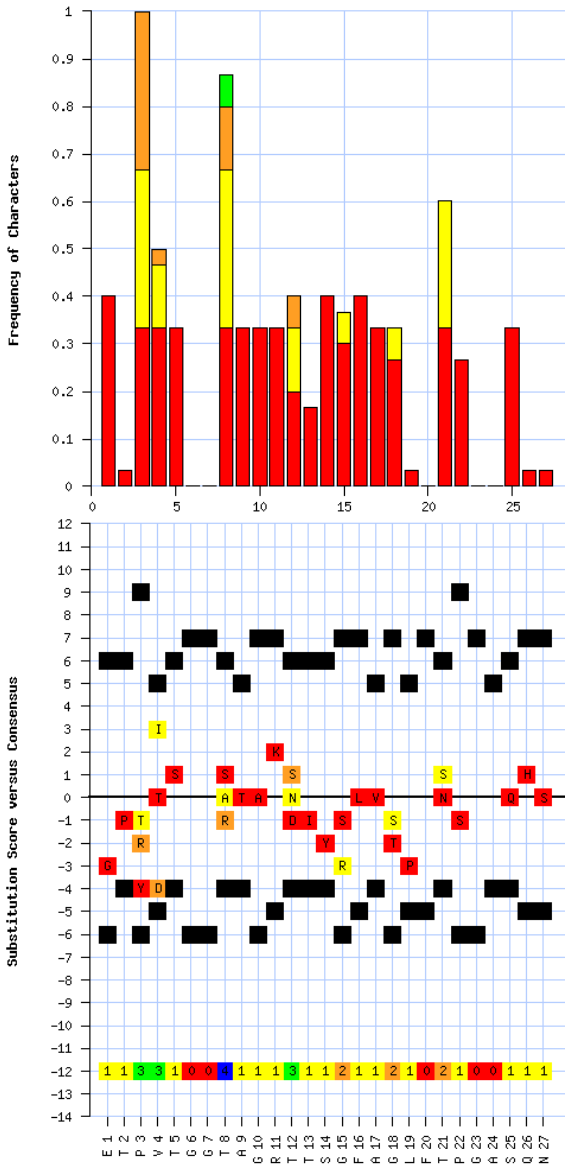


Fig.5b

Rapid progressor time point zero



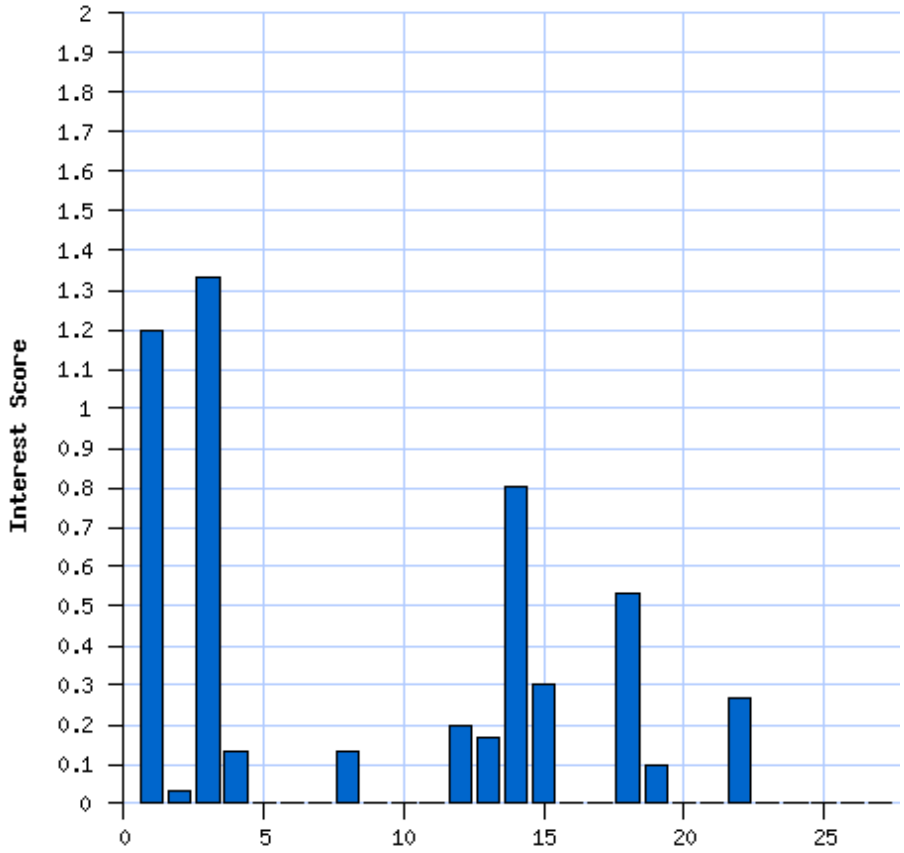
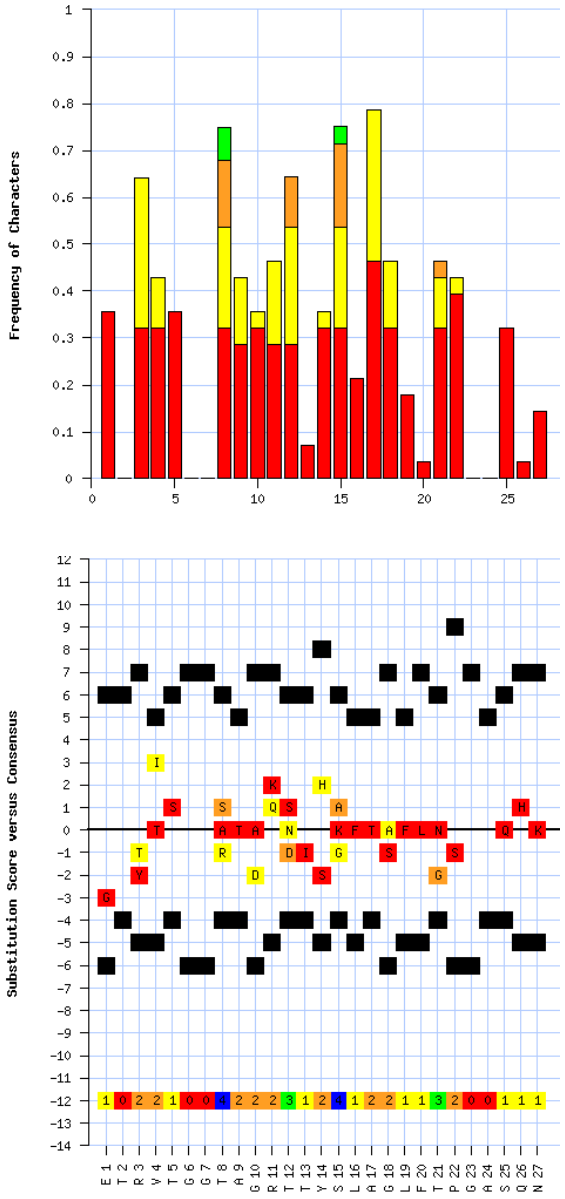


Fig.5c

Rapid progressor after antibody seroconversion



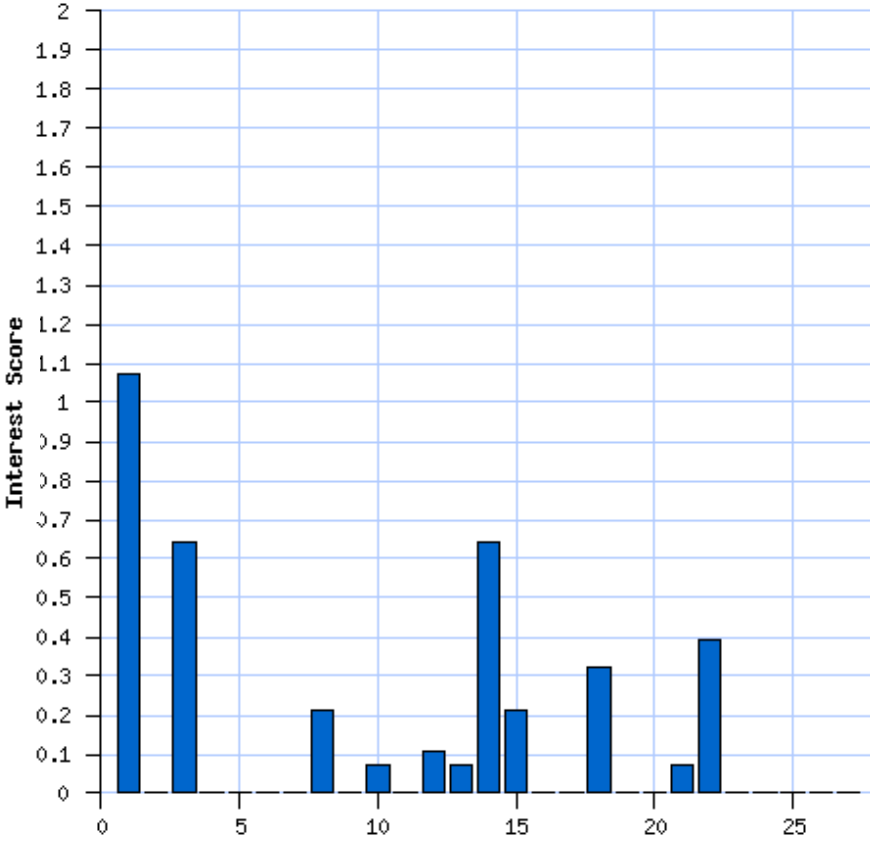


Fig.5d

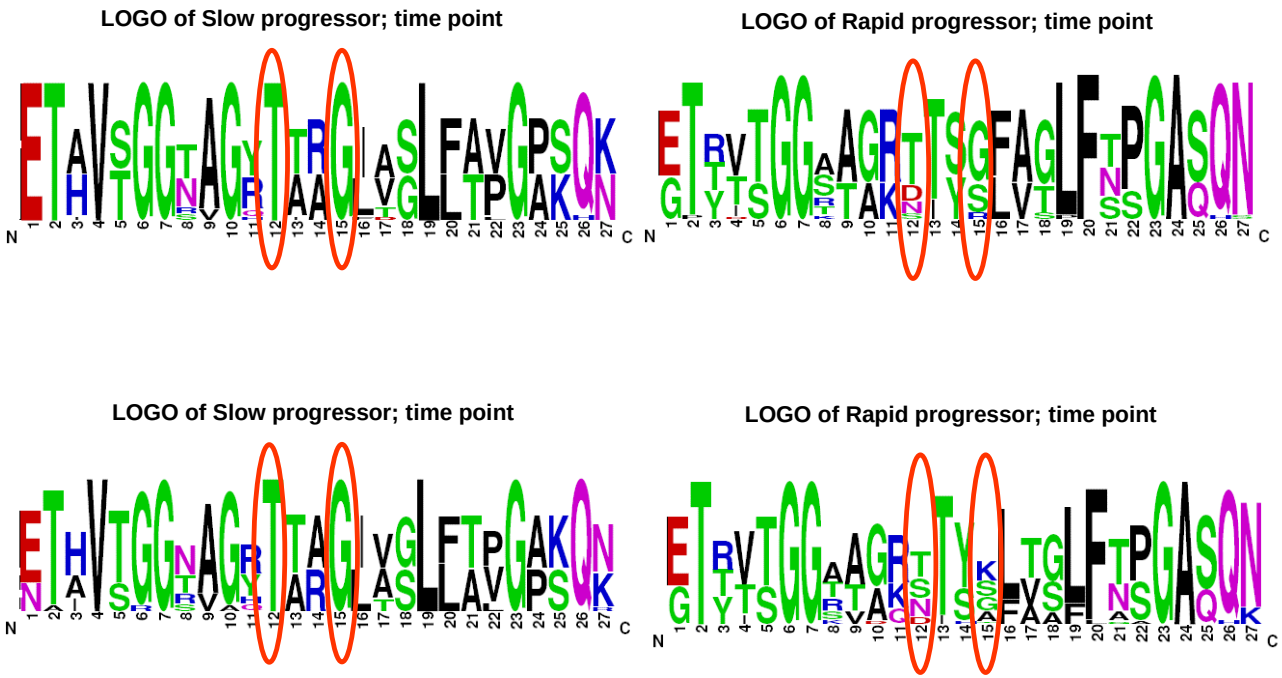


Fig.6 : LOGO of the HVR-1 protein. The taller the letter, the higher the frequency of the amino acid in that position