Biochemistry and Biophysics Reports 32 (2022) 101327

Contents lists available at ScienceDirect

Biochemistry and Biophysics Reports

journal homepage: www.elsevier.com/locate/bbrep

Independent evolution of multi-dominant viral genome species observed in a hepatitis C virus carrier

Tomomi Ando [a], Hideki Aizaki [a], Masaya Sugiyama [b], Tomoko Date [c], Kazuhiko Hayashi [d], Masatoshi Ishigami [e], Yoshiaki Katano [f], Hidemi Goto [d], Masashi Mizokami [c], Masamichi Muramatsu [a], Makoto Kuroda [g], Takaji Wakita [a,*]

[a] Department of Virology II, National Institute of Infectious Diseases, Tokyo, Japan
[b] Department of Viral Pathogenesis and Controls, National Center for Global Health and Medicine, Chiba, Japan
[c] Genome Medical Science Project, National Center for Global Health and Medicine, Chiba, Japan
[d] Department of Gastroenterology and Hepatology, Federation of National Public Service Personnel Mutual Aid Associations, Meijo Hospital, Nagoya, Japan
[e] Department of Gastroenterology and Hepatology, Nagoya University Graduate School of Medicine, Nagoya, Japan
[f] Department of Gastroenterology, Bantane Hospital, Fujita Health University School of Medicine, Nagoya, Japan
[g] Pathogen Genomics Center, National Institute of Infectious Diseases, Tokyo, Japan

ARTICLE INFO

Keywords:
HCV
Quasispecies
Next-generation sequencing
Core 70 polymorphism

ABSTRACT

The viral genome quasispecies composition of hepatitis C virus (HCV) could have important implications to viral pathogenesis and resistance to anti-viral treatment. The purpose of the present study was to profile the HCV RNA quasispecies. We developed a strategy to determine the full-length HCV genome sequences co-existing within a single patient serum by using next-generation sequencing technologies. The isolated viral clones were divided into the groups that can be distinguished by core amino acid 70 substitution. Subsequently, we determined HCV full-length genome sequences of three independent dominant species co-existing in the sequential serum with a 7-year interval. From phylogenetic analysis, these dominant species evolved independently. Our study demonstrated that multiple dominant species co-existed in patient sera and evolved independently.

1. Introduction

Hepatitis C virus (HCV) infection is a global health threat and is a major cause of chronic liver diseases [1]. HCV possesses a positive-strand RNA genome. The HCV genome replicates at an extraordinarily high level by the virus RNA polymerase without proofreading mechanisms and accumulates substitutions termed "quasispecies" during persistent infection [2]. The resultant viral variability in infected individuals contributes to viral pathogenesis, decreased sensitivity to treatment, and the generation of escape mutants from anti-viral agents. It is thus important to discover responsible virus genome variations for such phenotypic changes. However, most HCV studies use a consensus sequence, which is a series of combinations of dominant nucleotide sequences; particular substitutions against the consensus sequence are associated with virus phenotypic changes [3].

To understand the virus quasispecies, we amplified regions of interest and analyzed them by cloning or next-generation sequencing (NGS) technologies [4]. NGS technologies are powerful tools for analysis of virus diversity because they can generate large amounts of sequencing data and detect minor variations [5,6]. It is unclear how virus quasispecies co-exist in the patient. The polymorphism in HCV core amino acid 70 substitution (core 70) is associated with not only susceptibility to IFN therapy but also hepatic steatosis and oxidative stress [7–9], and polymorphisms in the core protein are related to human genetic variations in the IL28B locus [10]. Furthermore, it is likely that polymorphism in core protein is a marker and that the other regions of viral genome are also involved the phenotypic differences. Interestingly, it was reported that some patients have a mixture of core 70; wild (arginine: R) and mutated (glutamine: Q) types [11].

Here, we sequenced the HCV genome in a pair serum sample from a

Abbreviations: core70, the core amino acid 70 substitutions; E, envelope; HVR, hyper variable region; IRRDR, IFN and ribavirin resistance-determining region; ISDR, IFN sensitivity-determining region; NGS, next-generation sequencing; NS, nonstructural; PEG-IFN, pegylated interferon; Q, glutamine; R, arginine; RBV, ribavirin; SVR, sustained virological response.

* Corresponding author. Department of Virology II, National Institute of Infectious Diseases 1-23-1 Toyama, Shinjuku-ku, Tokyo, 162-8640, Japan.
E-mail address: wakita@nih.go.jp (T. Wakita).

https://doi.org/10.1016/j.bbrep.2022.101327
Received 25 June 2022; Received in revised form 7 August 2022; Accepted 8 August 2022

HCV carrier with a mixed type at core70 at 7-year interval to analyze relation between the core70 polymorphism and quasispecies phenotype. Surprisingly, the isolated viral clones were divided into groups that can be distinguished by core70. We then identified three independent dominant species co-existing in the sequential sera with a 7-year interval from one of the patients. Interestingly, the full-length sequences were quite different among the species.

2. Materials and methods

The detailed materials and methods are available in supplementary information.

3. Patients

In this study, viral genome sequences were analyzed in serum samples from a chronic hepatitis patient (Pt1) who was infected with HCV subtype 1b. Two serum samples (sPt1-98 and sPt1-05) were obtained in 1998 and 2005 from Pt1, respectively. This patient had not been treated with IFN and showed mixed types at core 70 Q and R in the previous analysis (data not shown). HCV RNA titers in Pt1-98 and Pt1-05 were determined as 1.2×10^6 and 1.1×10^7 copies/ml, respectively.

3.1. Amplification and sequencing of the core region for population analysis

Complementary DNA (cDNA) was synthesized from total RNA purified from patient serum. A core to E1 region (480–1072) was amplified from each cDNA molecule. The primer sequences are shown in Supplementary Table 1. The amplified products were cloned and 12 clones from each sample were sequenced. The sequences from each clone were aligned, and the positions having variety in more than 10% of the clones were selected in the determined region (500–1040). The positive strand sequences were aligned, and were grouped according to the non-synonymous substitution patterns.

3.2. Short-read DNA sequencing with the illumina Genome Analyzer IIx

The flow of the computational method for determining the consensus sequences using reads from Illumina GAIIx analysis is summarized in Supplementary method and Supplementary Figs. 2 and 3. Double-stranded cDNA was prepared from total RNA to generate the Illumina library. An approximately 300-bp length library was subsequently generated from 50 ng double-stranded cDNA. All sequencing runs were performed with Genome Analyzer IIx (Illumina, San Diego, CA). The human genome–derived sequences were excluded from Illumina reads. Sequence data were analyzed by using MAQ software (ver. 0.7.1) [12] and Edena software (ver. 2.1.1) [13]. Alignment of the Illumina reads was performed by using MAQ [the number of maximum mismatches (-n): 3]. De novo sequence assembly was performed by using Edena software in assembler mode.

3.3. Long-read DNA sequencing with the 454 FLX or 454 jr

PCR primers were designed to exclude the positions that were determined to have variety by the Illumina analysis. The library produced from the PCR amplicon was sequenced in both the forward and reverse directions by using the 454 GS FLX or 454 GS Jr. platform (Roche Diagnostics Japan). After the tag sequences were trimmed, the reads were aligned against the consensus sequence by using BWA-SW (ver. 0.5.9) [14] allowing 5 to 10 nucleotide mismatches within the read sequence. The results of the analysis were displayed by using Integrative Genomics Viewer (IGV; ver. 1.4.2) [15].

3.4. Determination of the multi-dominant viral genome sequence by using species specified PCR

To selectively amplify each species from the cDNA, the primers were designed to include the substitutions, as shown in Supplementary Fig. 4 and Supplementary Table 1. The determined sequences of the PCR products were compared with Illumina reads to distinguish the sequence diversity and PCR error. The 5′ and 3’ terminal end sequences of viral RNA was determined by the method described previously [16].

3.5. Phylogenetic analysis

The phylogenetic tree was constructed by using the neighbor-joining algorithm as described previously [17]. Bootstrap resampling (2000 replicates) was used to measure the statistical support for each node in the phylogenetic trees.

4. Results

4.1. Population analysis in Core70 and the surrounding regions

In this study, we analyzed 2 serum samples (sPt1-98 and sPt1-05) with 7-year interval from the patient Pt1 who have both wild (R) and mutated (Q) types of core70 amino acids. First, we sequenced the core70 polymorphisms and a part of the E1 region of the HCV genome in both samples to analyze the relationship between the substitutions at core 70 and other positions by using conventional PCR, cloning and capillary sequencing. We amplified each HCV genome with primers designed at the conserved regions in the sequences from the database. Twelve clones were isolated from each PCR product and sequenced. The sequence variations of each clone at the positions with variety in more than 10% of the total clone number. Partial viral genome sequences of sPt1-98 and sPt1-05 were analyzed in Fig. 1A and B, respectively. The white boxes represent the nucleotides of major clones with glutamine at core70 (core70 Q, at position 550). The black boxes represent the different nucleotides from the major clones with core70 Q at each position. Every clone was aligned based on the core70 polymorphism. The clones with core70 Q mainly included the white boxes; however, some clones with core70 Q included different numbers of the black boxes. On the other hand, the clones with core70 R mainly included the black boxes and a few white boxes. Furthermore, some of these nucleotide substitutions induced non-synonymous substitutions, as shown in Fig. 1. The clones from sPt1-05 with core70 Q could be separated into two groups, depending on the pattern of the non-synonymous substitutions (Fig. 1B). These observations suggest that multiple dominant viral RNA species are present in each sample and that the species can be distinguished by using core70 Q or R as a marker.

4.2. Determination of the full-length sequences in patient serum

For further detailed analysis of the viral RNA species with different core70 polymorphisms, we attempted to determine the whole viral genome sequences of each dominant viral RNA molecule found in the patient serum. Detailed method to determine the full-length HCV genome sequence in patient serum is described in Supplementary Materials and Methods. First, we listed the whole sequence diversity of the sample. Second, we estimated the combination of the determined substitutions. Subsequently, we determined the lead sequences that included the definite substitution patterns in small regions for distinction of each viral RNA molecule. Finally, we amplified each viral RNA molecule by using the specific primers designed at the lead sequences and determined the sequence of each viral RNA molecule. Our methodology to determine the full-length viral genome sequences is summarized in Fig. 2, and the detailed strategy to establish the Illumina consensus sequence is shown in Supplementary Figs. 2 and 3.

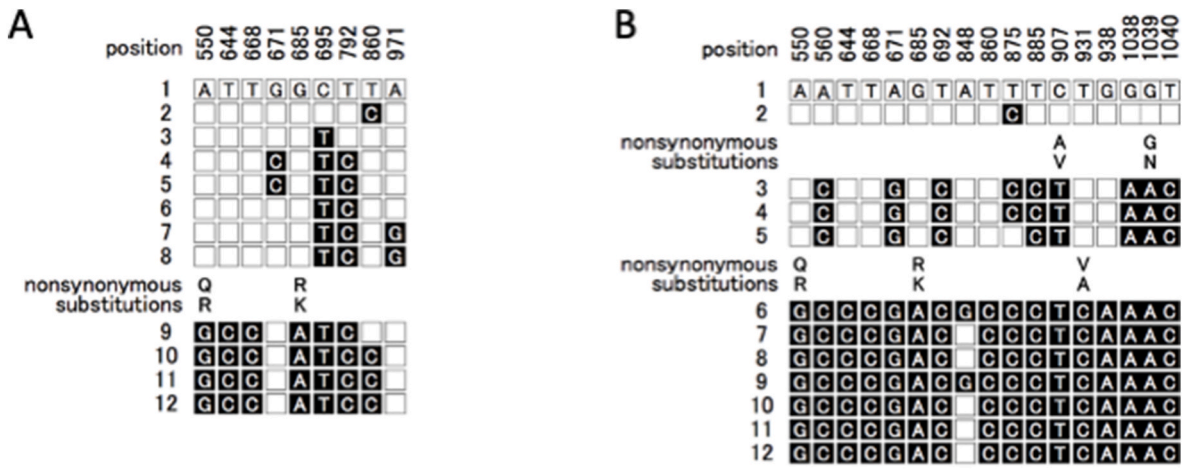

Fig. 1. Population analysis of core70 and the surrounding region. Core70 (nucleotide position: 549–551) and a part of the E1 region (500–1040) were amplified. The PCR products were subjected to TA cloning. Twelve clones were isolated from each PCR product and sequenced. The sequences were grouped based on the core70 sequence. The top numbers show the nucleotide positions having variety in more than 10% of all clones, and the left numbers show the clone numbers. The nucleotides of clone 1 (which has core70 Q) are show in the top line of white boxes. The black boxes represent nucleotides different from clone 1. Non-synonymous substitutions are shown only if the nucleotides were completely different between two groups. (A) sPt1-98, (B) sPt1-05

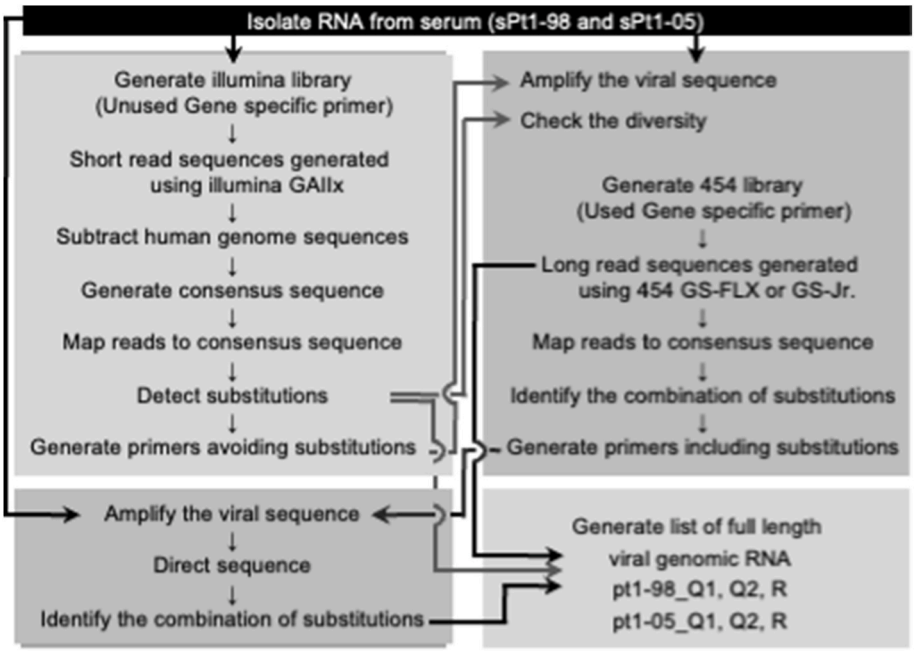

Fig. 2. Flowchart of the method for determining full-length sequences in patient serum, Pt1-98 and Pt1-05. Ful-length HCV seuences were determined by the method as described in this flowchart. Illumina GAIIx is suitable for determining consensus sequence and generating substitutions list because generating library for Illumina analysis does not need sequence specific amplification. 454 GS FLX is suitable for identifying the combination of substitutions in small regions because the read length is relatively long (−500 mer), though generating library for 454 analysis needs sequence specific amplification. Using primers avoiding substitutions, we could amplify the viral sequence maintaining the diversity detected by Illumina analysis for generating 454 libraly. After identifying small regions in which reads can be separated to some groups depend on the substitutions pattern using 454 analysis, we designed species specific primers including substitutions. Using these primers, we could amplify each viral sequence specifically. The PCR products were sequenced using capillary sequencing. The determined sequences were verified using every data from Illumina and 454 analysis, and then we determined the full-length viral genomic RNA sequences, pt1-98_Q1, Q2, R and pt1-05_Q1, Q2, R.

4.3. Sequence comparison of dominant species in a single patient

By using the method described above and supplementary information, we identified three independent dominant species co-existing in sera collected from patient Pt1 at two time points (1998: pt1-98, and 2005: pt1-05). The positions of substitutions against the consensus sequence determined by Illumina analysis are indicated by vertical lines in Fig. 3 (pt1-98_MUT and pt1-05_MUT, upper: nucleotide substitutions (nt), lower: amino acid substitutions (aa)), and the numbers of mutations were 231 nt positions and 48 aa positions in pt1-98_MUT, and 320 nt positions and 73 aa positions in pt1-05_MUT. At both time points, the sequences with core70 Q or R were clearly distinguished. Almost all substitutions against the consensus sequence were independently distributed among the species, and the numbers and positions of the substitutions were indicated in Fig. 3 and Supplementary Table 2. Interestingly, diversity in the IRRDR region was only found in pt1-98_R and pt1-05_R but not in pt1-98_Q and pt1-05_Q, and there was no diversity in the ISDR region. El-Shamy et al. reported that more than six substitutions against the standard sequence (HCV-J) in the IRRDR region is related to a favorable response for IFN treatment [18]. To compare with HCV-J, five substitutions in the core70 Q species and eight substitutions in the core70 R species were found in the IRRDR. This result indicates that the polymorphisms in the different genome regions related to the favorable response to IFN were gathered in one species. We cloned each full-length species detected in the pt1-98 serum into plasmid. For confirmation of the selectivity of the primer sets used in Supplementary Fig. 4, we amplified the HCV genome on each plasmid by using each primer set as shown in Supplementary Fig. 6.

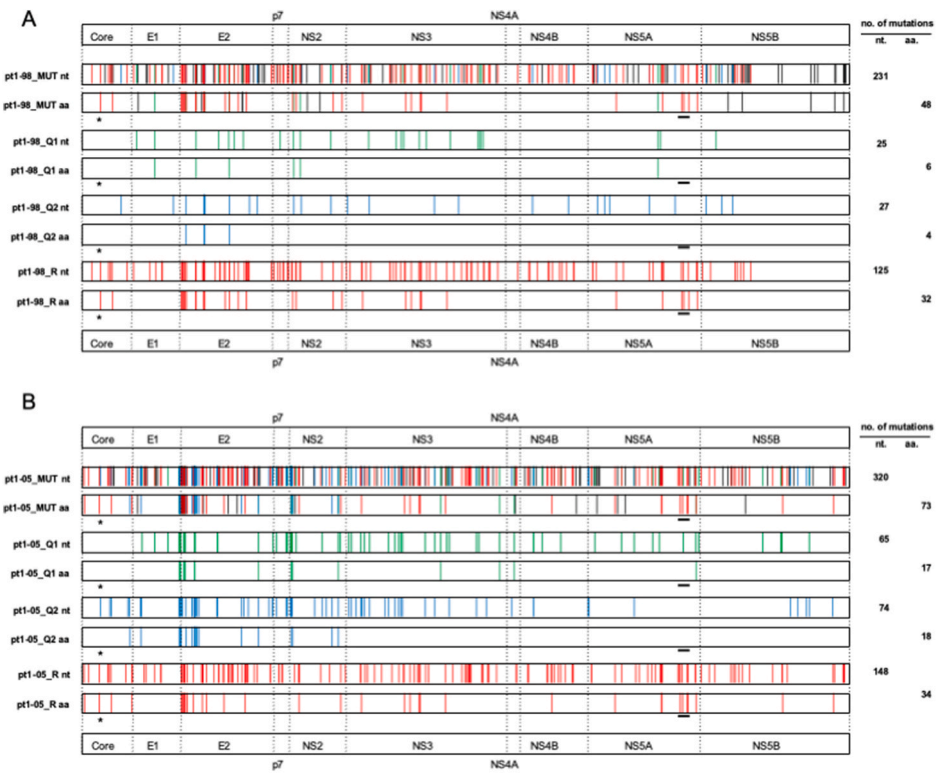

Fig. 3. The position of substitutions in each dominant species against the consensus sequence. From each serum sPt1-98 (A) and sPt1-05 (B), we determined the positions of substitutions (pt1-98_MUT and pt1-05_MUT) against the consensus sequence determined by Illumina analysis and three independent dominant species (pt1-98_Q1, Q2, R and pt1-05_Q1, Q2, R) by the method in Fig. 2. The upper and lower horizontal boxes show the nucleotide (nt) and amino acid (aa) sequence of open reading frame in HCV genome. The vertical lines show the positions of each substitution against the Illumina consensus sequence. The red vertical lines indicate the positions of mutations in the species with core70 R, and the green and blue lines indicate the species with core70 Q (Q1 and Q2, respectively). Black vertical lines in pt1-98_MUT and pt1-05_MUT show the mutations not found in the species with core70 R, Q1 and Q2. Dotted lines are the borders for dividing each virus protein. pt1-98_MUT and pt1-05_MUT show every position of sequence variation determined by Illumina analysis. *, position of core70; heavy black line, position of IRRDR regions. (For interpretation of the references to colour in this figure legend, the reader is referred to the Web version of this article.)

4.4. Phylogenetic analysis with Co-existent virus genome sequences in one patient

We discovered co-existence of independent dominant species in the patient serum samples. For analysis of the relationship among the virus species, we constructed phylogenetic trees by using full-length, core, and NS5B sequences of each species. The sequences derived from the database that have more than 94% homology with pt1 were also included in the tree. If the patient had two independent infections, the sequences might be quite different and form different clusters. However, as shown in Fig. 4, all the species isolated from Pt1 formed one cluster, although the closed strains found in the database were included. The virus species with core70 Q and another species having core70 R formed different clusters sequentially. However, it is not clear the evolutional relations between Q1 and Q2 of pt1-98 and pt1-05. These results suggests that this patient was not infected twice and the species evolved independently in this patient.

5. Discussion

The HCV genome accumulates substitutions and forms quasispecies during persistent infection [2]. In the present study, we found multiple dominant species co-existing in each serum sample, and the sequences were quite unique based on the core70 amino acid sequence in the initial population analysis (Fig. 1). We thus developed a method to determine the dominant full-length HCV sequences co-existing in serum sample (Fig. 2). By using this method, we identified three independent dominant species co-existing in sequential serum samples with a 7-year interval from a single patient, Pt1 (Fig. 3). Interestingly, the polymorphisms related to the favorable response to IFN therapy (core70 R and diversity in the IRRDR) were found in one species, and not found in two other species (Fig. 3). The full-length sequences were clearly distinguished among the species. From the phylogenetic analysis, the one species with core70 R and the two other species with core70 Q seemed to evolve independently during the 7-year period (Fig. 4).

Several studies are published to determine whole HCV genome sequence using novel NGS methods including recently developed third generation sequencing technologies which generate long reads. For example, nanopore method offer the first opportunity to sequence whole viral genomes as single reads, thereby potentially enabling detailed and reliable characterization of viral variants. However, two studies with nanopore method have done with HCV cDNA samples amplified by long PCR method which may induce sequence deviation during the amplification step [19,20]. Another study using nanopore technology determined the whole HCV genome without PCR amplification [21]. However, nanopore reads have a high error rate compared to paired-end short reads generated on the Illumina platform which limits the reliability and usefulness of its long reads. Further study will be necessary to discover the method how these third generation sequencing technology can apply for precise HCV genome characterization.

The HCV genome accumulates substantial variations during its high level replication and establish its quasispecies nature during persistent infection [2]. It is thus difficult to identify the functionally viable consensus HCV viral genome sequence. For the virological analysis, however, it is important to discover the functional virus genome sequence. Previously, we have established JFH-1 strain which can replicate in the cultured cell line and produce infectious virus [22]. On the other hand, it has been difficult to isolate replication competent viral genome from chronic hepatitis patients. Since JFH-1 strain was isolated from a patient with fulminant hepatitis the virus was almost monoclonal in the patient [23,24]. On the other hand, we discovered the Illumine consensus full-length HCV genome and determined independent dominant species co-existing in the patient serum. For the virological studies, it is thus important to test the dominant species sequences found in this study for the replication if they are functional in cell culture.

We discovered co-existence of independent dominant species in one patient serum sample. These species could be classified with two prediction factors for IFN therapy in the core and NS5A polymorphisms (Fig. 3). It is interesting that some viral species with different phenotypes appear to co-exist in one patient. There are three possible explanations for this observation: two species of viruses co-infected at the initial infection; one species of virus infected at the initial infection and

Figure 4.

Fig. 4. Phylogenetic trees of subtype 1b strains estimated from (A) core, (B) NS5B, and (C) full-genome gene sequence. Bootstrap values based on 2000 replicates each are shown for branches with more than 50% bootstrap support. The scale bars are in units of nucleotide substitutions per site. The trees were rooted with subtype 1a strains (H77; not shown). The red lines indicate the sequences identified from serum samples sPt1-98 and sPt05. (For interpretation of the references to colour in this figure legend, the reader is referred to the Web version of this article.)

then evolved into two distinguished species of viruses; or, the patient had two separate infections with different species of viruses. From the phylogenetic analysis in Fig. 4, all the species from Pt1 formed one cluster and the results suggests that this patient was not infected twice and the species evolved independently in this patient. The patient P1 had not been received IFN therapy, however, the infected virus population was consisted with different dominant species with IFN treatment resistant and favorable phenotype such as core 70 Q and R. It is interesting if this co-existence and co-evolution of different species may contributes for persistent viral infection or pathogenesis of chronic hepatitis. However, it is still unclear when this patient was infected with these two kinds of virus species, or with one virus that evolved into two species.

The core70 polymorphism influences not only IFN sensitivity but also hepatic steatosis and oxidative stress [7–9]. Although the current standard therapy has been DAA therapy, IFN and RBV were used to treat HCV carriers. The presence of core70 Q in the patient before starting IFN and RBV combination therapy was associated with an increased risk of treatment failure [25]. Thus, the co-existence of virus species with different core70 sequences in one patient may contribute to treatment susceptibility, HCV replication, and pathogenesis. At least in the case of Pt1, entire sequences of species with core 70 Q or R were quite unique (Fig. 3), and each species seemed to evolve individually (Fig. 4). It is possible that the contributions of the mixed core70 polymorphism are not only by the core70 amino acid mutation but also from the full-length species with different characteristics.

The ratio of core70 polymorphism was changed dramatically in the previous report [25].[25] Therefore, it is considered that some dominant viral species that might be related within a patient play a prominent role in the pathogenesis or persistent infection of HCV. Further analysis with sequences of each species co-existing in patient serum samples may provide us additional information regarding the viral phenotypes. In the present study, we analyzed a single patient with variations at core 70. However, we also analyzed some patient sera with or without variations at core70. Their virus sequences also exhibited multiple dominant viral species. Interestingly, in virus species from without variation at core 70, there was no specified amino acid to separate the species (data not shown). Further analysis will be necessary to address this point.

In conclusion, we found that the viral clones were divided into two or three groups clearly distinguished by core70 sequences in the analyzed serial patient serum samples. To generalize our finding to the full-length HCV genome, we constructed a method to determine the full-length viral genome RNA species from patient sera by a combination of next-generation sequencing and capillary sequencing. By using this method, we identified multiple dominant species co-existing in the patient serum samples that evolved independently. Further studies are required to understand the contributions or roles of these co-existent viruses within patient serum with different characteristics in the virus lifecycle or pathogenesis.

Accession numbers

DDBJ/EMBL/GenBank accession numbers: pt1-98_R: AB828699, pt1-98_Q1: AB828700, pt1-98_Q2: AB828701, pt1-05_R: AB828702, pt1-05_Q1: AB828703, pt1-05_Q2: AB828704.

Writing assistance

None.

Author contributions

Study concept and design: TA and TW; Acquisition, analysis and interpretation of data: TA, MS, TD, MK and TW; Drafting of the manuscript: TA and TW; Study supervision: HK, IM, KY, GH, HA, MM and MK.

Grant support

This study was partially supported by The Japan Society for the Promotion of Science KAKENHI (No. JP207K08369) and The Japan Agency for Medical Research and Development, AMED (Nos. JP21fk02100539j0103, 21fk0310103j0305 and 21fk0210065j1002). HG received a research grant from AstraZeneca, Astellas Pharma, Ajinomoto Pharmaceutical Co, Bristol-Myers Squibb, Chugai Pharmaceutical Co, Daiichi Sankyo, Dainippon Sumitomo Pharma, Eisai, Mitsubishi Tanabe Pharma, MSD, Otsuka Pharmaceutical Co and Takeda Pharmaceutical Co. KH received a research grant from AstraZeneca, Eisai, and MSD.

Declaration of competing interest

The authors declare the following financial interests/personal relationships which may be considered as potential competing interests:

Hidemi Goto reports a relationship with AstraZeneca, Astellas Pharma, Ajinomoto Pharmaceutical Co, Bristol-Myers Squibb, Chugai Pharmaceutical Co, Daiichi Sankyo, Dainippon Sumitomo Pharma, Eisai, Mitsubishi Tanabe Pharma, MSD, Otsuka Pharmaceutical Co and Takeda Pharmaceutical Co that includes: funding grants. Kazuhiko Hayashi reports a relationship with AstraZeneca, Eisai, and MSD that includes: funding grants.

Data availability

Data will be made available on request.

Appendix A. Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.bbrep.2022.101327.

References

[1] World Health Organization (WHO), Interim Guidance for Country Validation of Viral Hepatitis Elimination, World Health Organization, Geneva, Switzerland, 2021.
[2] J. Gomez, M. Martell, J. Quer, et al., Hepatitis C viral quasispecies, J Viral Hepat 6 (1999) 3–16.
[3] X. Forns, R.H. Purcell, J. Bukh, Quasispecies in viral persistence and pathogenesis of hepatitis C virus, Trends Microbiol 7 (1999) 402–410.
[4] N. Hiraga, M. Imamura, H. Abe, et al., Rapid emergence of telaprevir resistant hepatitis C virus strain from wildtype clone in vivo, Hepatology 54 (2011) 781–788.
[5] S. Chevaliez, C. Rodriguez, J.M. Pawlotsky, New virologic tools for management of chronic hepatitis B and C, Gastroenterology 142 (2012) 1303–1313 e1301.
[6] V. Suppiah, M. Moldovan, G. Ahlenstiel, et al., IL28B is associated with response to chronic hepatitis C interferon-α and ribavirin therapy, Nature Genetics 41 (2009) 1100–1104.
[7] Y. Tachi, Y. Katano, T. Honda, et al., Impact of amino acid substitutions in the hepatitis C virus genotype 1b core region on liver steatosis and hepatic oxidative stress in patients with chronic hepatitis C, Liver Int 30 (2010) 554–559.
[8] Y. Sumida, K. Kanemasa, T. Hara, et al., Impact of amino acid substitutions in hepatitis C virus genotype 1b core region on liver steatosis and glucose tolerance in non-cirrhotic patients without overt diabetes, J Gastroenterol Hepatol 26 (2011) 836–842.
[9] S. Nakamoto, F. Imazeki, K. Fukai, et al., Association between mutations in the core region of hepatitis C virus genotype 1 and hepatocellular carcinoma development, J Hepatol 52 (2010) 72–78.
[10] K. Chayama, C.N. Hayes, Hepatitis C virus: how genetic variability affects pathobiology of disease, J Gastroenterol Hepatol 26 (Suppl 1) (2011) 83–95.
[11] M. Sato, S. Maekawa, N. Komatsu, A. Tatsumi, M. Miura, M. Muraoka, Y. Suzuki, F. Amemiya, [11, S. Takano, M. Fukasawa, Y. Nakayama, T. Yamaguchi, T. Uetake, T. Inoue, T. Sato, M. Sakamoto, A. Yamashita, K. Moriishi, N. Enomoto, Deep sequencing and phylogenetic analysis of variants resistant to interferon-based protease inhibitor therapy in chronic hepatitis induced by genotype 1b hepatitis C virus, J Virol 89 (11) (2015 Jun) 6105–6116.
[12] H. Li, J. Ruan, R. Durbin, Mapping short DNA sequencing reads and calling variants using mapping quality scores, Genome Res 18 (2008) 1851–1858.
[13] D. Hernandez, P. Francois, L. Farinelli, et al., De novo bacterial genome sequencing: millions of very short reads assembled on a desktop computer, Genome Res 18 (2008) 802–809.
[14] H. Li, R. Durbin, Fast and accurate short read alignment with Burrows-Wheeler transform, Bioinformatics 25 (2009) 1754–1760.
[15] J.T. Robinson, H. Thorvaldsdottir, W. Winckler, et al., Integrative genomics viewer, Nat Biotechnol 29 (2011) 24–26.
[16] T. Masaki, R. Suzuki, M. Saeed, et al., Production of infectious hepatitis C virus by using RNA polymerase I-mediated transcription, J Virol 84 (2010) 5824–5835.
[17] T. Nakano, L. Lu, Y. He, et al., Population genetic history of hepatitis C virus 1b infection in China, J Gen Virol 87 (2006) 73–82.
[18] A. El-Shamy, I. Shoji, S.R. Kim, et al., Sequence heterogeneity in NS5A of hepatitis C virus genotypes 2a and 2b and clinical outcome of pegylated-interferon/ribavirin therapy, PLoS One 7 (2012), e30513.
[19] R.A. Bull, A.A. Eltahla, C. Rodrigo, S.M. Koekkoek, M. Walker, M.R. Pirozyan, B. Betz-Stablein, A. Toepfer, M. Laird, S. Oh, C. Heiner, L. Maher, J. Schinkel, A. R. Lloyd, F. Luciani, A method for near full-length amplification and sequencing for six hepatitis C virus genotypes, BMC Genomics 17 (2016) 247.
[20] N. Riaz, P. Leung, K. Barton, M.A. Smith, S. Carswell, R. Bull, A.R. Lloyd, C. Rodrigo, Adaptation of Oxford Nanopore technology for hepatitis C whole genome sequencing and identification of within-host viral variants, BMC Genomics 22 (1) (2021) 148.
[21] Y. Uchida, J.I. Kouyama, K. Naiki, H. Uemura, S. Tsuji, K. Sugawara, M. Nakao, D. Motoya, N. Nakayama, Y. Imai, T. Tomiya, S. Mochida, A case of genotype-3b hepatitis C virus in which the whole genome was successfully analyzed using third-generation nanopore sequencing, Hepatol Res 49 (9) (2019) 1083–1087.
[22] T. Wakita, T. Pietschmann, T. Kato, et al., Production of infectious hepatitis C virus in tissue culture from a cloned viral genome, Nat Med 11 (2005) 791–796.
[23] T. Kato, A. Furusaka, M. Miyamoto, T. Date, K. Yasui, J. Hiramoto, K. Nagayama, T. Tanaka, T. Wakita, Sequence analysis of hepatitis C virus isolated from a fulminant hepatitis patient, J Med Virol 64 (3) (2001) 334–339.
[24] T. Kato, T. Date, M. Miyamoto, et al., Efficient replication of the genotype 2a hepatitis C virus subgenomic replicon, Gastroenterology 125 (2003) 1808–1817.
[25] F. Kurbanov, Y. Tanaka, K. Matsuura, et al., Positive selection of core 70Q variant genotype 1b hepatitis C virus strains induced by pegylated interferon and ribavirin, The Journal of Infectious Diseases 201 (2010) 1663–1671.