Evolutionary relationship of hepatitis C, pesti-, flavi-, plantviruses, and newly discovered GB hepatitis agents

Ken-ichi Ohba[a], Masashi Mizokami[a,*], Johnson Y.N. Lau[b], Etsuro Orito[a], Kazuho Ikeo[c], Takashi Gojobori[c]

[a]*Second Department of Medicine, Nagoya City University Medical School, Kawasumi, Mizuho, Nagoya 467, Japan*
[b]*Section of Hepatobiliary, Department of Medicine, University of Florida, Gainesville, FL, USA*
[c]*Center for Information Biology, National Institute of Genetics, Mishima 411, Japan*

Received 26 October 1995; revised version received 28 November 1995

Abstract Two flavivirus-like viruses, GB virus-A (GBV-A) and GB virus-B (GBV-B), were recently identified in the GB hepatitis agent, and are distinct from the hepatitis A to E viruses. The putative helicase domain of GBV-A and GBV-B was found to have amino acid sequence homology with hepatitis C virus (HCV), and distantly, is also related to pestiviruses, flaviviruses, and plant viruses. A phylogenetic tree construction showed that GBVs and HCV are closely related, and they are clustered with pestiviruses, flaviviruses and plant viruses in that order.

Key words: GB virus; HCV; Phylogenetic tree; Evolution

1. Introduction

Hepatitis C virus (HCV) has been identified as the major cause of post-transfusion non-A, non-B hepatitis. Still after reliable methods for the detection of hepatitis A to E viruses, it became obvious that 10–20% of non-A, non-B hepatitis was also due to non-C and non-E etiologic agent(s). Recently two flavivirus-like genomes, named GB virus-A (GBV-A) and GB virus-B (GBV-B), were cloned from infectious tamarin serum [1], which originally had been derived from the serum of a surgeon who infected non-A, non-B hepatitis, and which is known to induce and transmit hepatitis in tamarins (previously called GB agents). Further analysis showed that GBV-A and GBV-B are closely related to HCV based on the amino acid sequence homology of the putative RNA helicase (GBV-A 47.2%; GBV-B 57.0%) with a conserved nucleotide triphosphates (NTP) binding motif.

HCV shows an amino acid sequence homology with pestiviruses, flaviviruses, and picornavirus-like and alphavirus-like plant virus supergroups [2]. In particular, the picornavirus-like and alphavirus-like plant viruses are more diverse from HCV and GBV's, it is possible that these plant viruses are the origin of the flavi-like viruses. It is also known that HCV has the internal ribosomal entry site of the translation initiation as like as picornaviruses [3], although the structure of HCV genome resembles that of flaviviruses. An interesting question regards the evolutionary relationship between HCV/GBVs and these viruses. To further elucidate the evolutionary relationship between GBV-A, GBV-B, HCV, and other viruses, we constructed the phylogenetic trees for the putative RNA helicase region.

*Corresponding author. Fax: (81) (52) 852-0849.

2. Materials and methods

We performed homology search for sequences with similarity to HCV sequences in the GenBank and Swiss Protein data bases. A total of 12 viruses were found to share homology with this putative RNA helicase domain, and these viral sequences were extracted from GenBank: two pestiviruses (bovine viral diarrhea virus (BVD) and hog

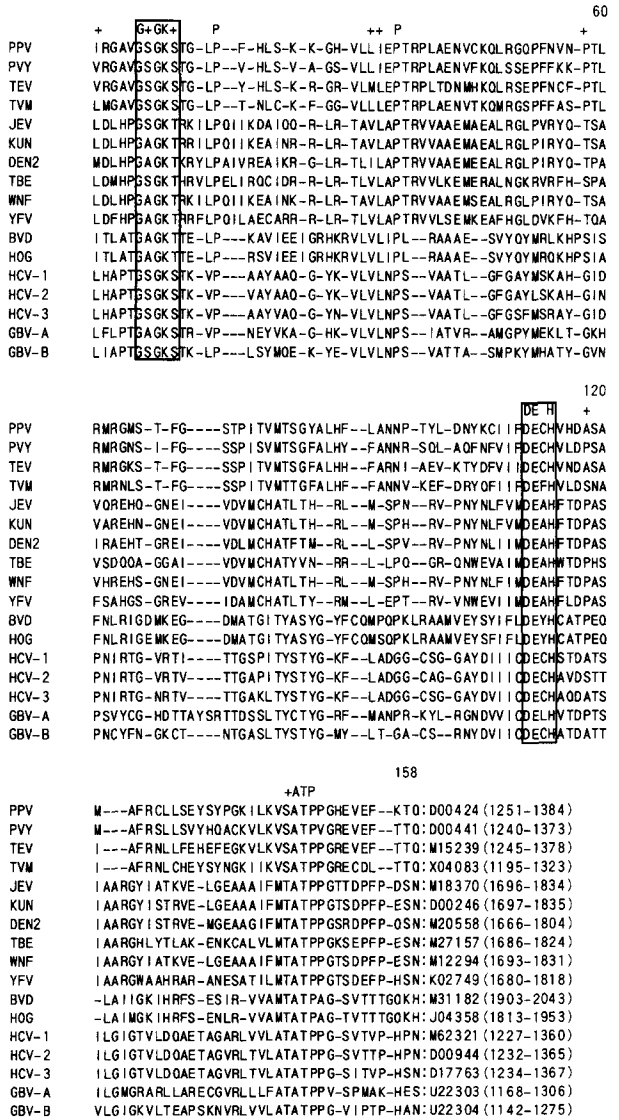

Fig. 1. Amino acid alignment of putative RNA helicase with NTP-binding motif. Accession numbers of these viruses on GenBank and Amino acid positions of these viral polyprotein.

SSDI 0014-5793(95)01441-1

cholera virus (HOG)), six flaviviruses (tick-borne encephalitis virus (TBE), yellow fever virus (YFV), dengue virus type 2 (DEN2), Japanese encephalitis virus (JEV), West Nile fever virus (WNF), and Kunjin virus (KUN)), four plant viruses (tobacco vein mottling virus (TVM), tobacco etch virus (TEV), potato virus Y (PVY), and plum pox virus (PPV)). Amino acid sequences with the putative RNA helicase domain of these viruses were maximally aligned together with GBV-A and GBV-B as well as HCV type 1 (HCV-1, isolate HCV1), HCV type 2 (HCV-2, isolate HC-J6), and HCV type 3 (HCV-3, isolate NZL1).

Using the program ODEN ver. 1.1.1 [4], the number of amino acid substitutions per site between all possible pairs of these viruses was estimated using Kimura's formula $d = -\ln(1 - p - 0.2p^2)$, where p is the fraction of different amino acids [5]. Using these values, phylogenetic trees were constructed by the unweighted pair-group method with arithmetic mean (UPGMA) [6] and the neighbor-joining (NJ) method [7]. We also used the PHYLIP ver. 3.5 [8] to elucidate evolutionary distance, the genetic distances between pairs of amino acid sequences were estimated using the PROTDIDT program with Dayhoff PAM matrix, and the phylogenetic tree was constructed using the programs NEIGHBOR and DRAWTREE.

3. Results and discussion

The partial amino acid alignment of the putative RNA helicase region is given in Fig. 1. It shows two conserved NTP-binding motifs, G+GK+ and DE H (denoted with boxes). The homologies for the amino acid sequence between HCV and

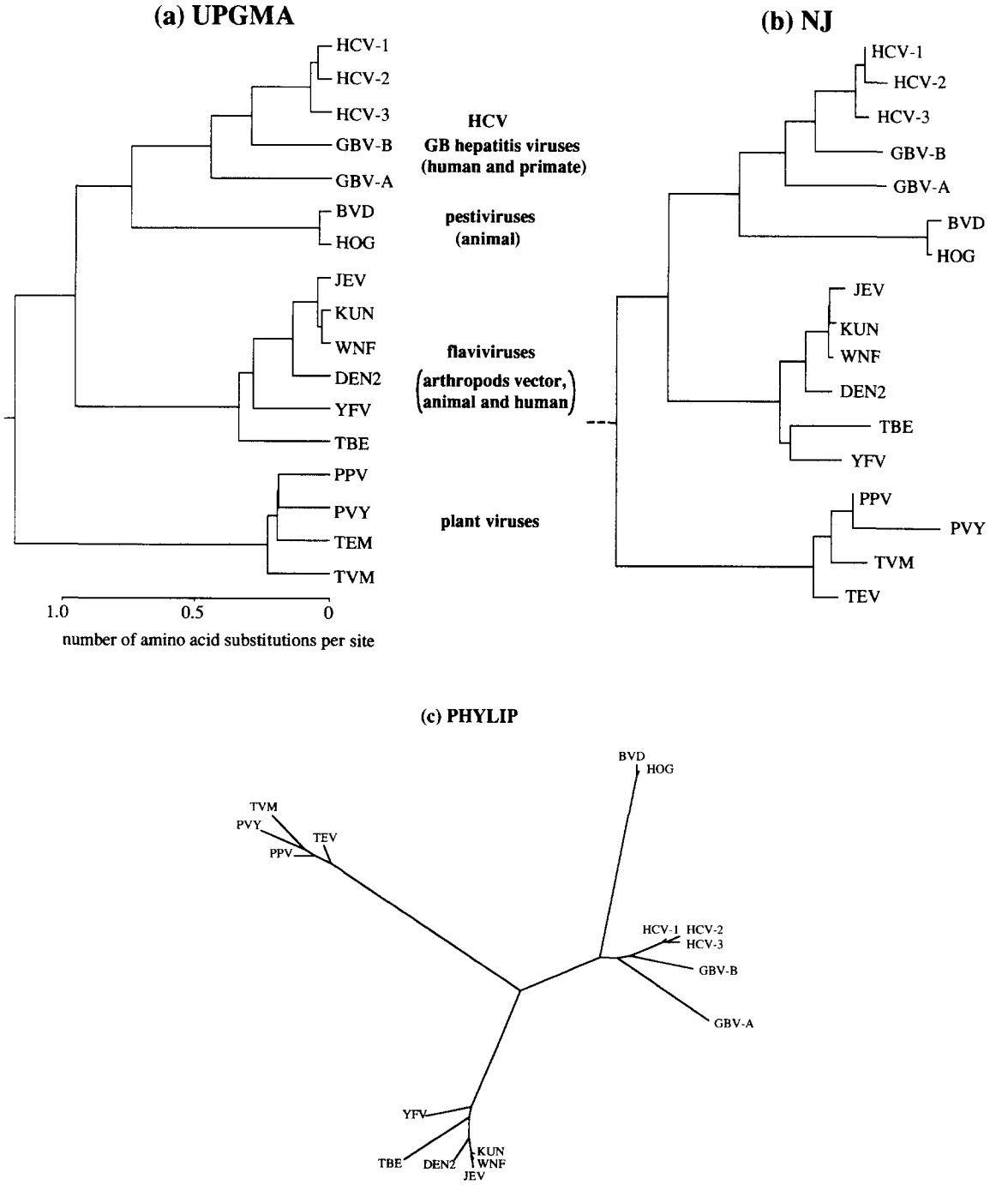

Fig. 2. Phylogenetic trees for the putative RNA helicase region by the UPGMA (a), and NJ (b), and PHYLIP (c) methods.

these viruses ranged from 20–30%. It has been thought that viral proteins containing NTP-binding motif may be subunits of RNA helicases involved in the unwinding of double-stranded replication forms during viral RNA replication, and in recombination between RNA genomes [9–11].

Fig. 2a,b,c showed the phylogenetic trees for the putative RNA helicase domain of these viruses generated from UPGMA, NJ, and PHYLIP, respectively. The evolutionary relationship between different viruses in these trees were essentially the same, attesting to the reliability of our estimation. As illustrated in these figures, HCVs and newly discovered GBV-B and GBV-A, derived from human and primate, were clustered first. Second, these viruses form a group with pestiviruses which infect animals. The evolutionary position of the newly discovered GBV's, closely related to HCV, is a mid-point between animal pestivirus and HCV, which is the main causative human pathogen of post-transfusion non-A, non-B hepatitis. Third, these viruses clustered with flaviviruses which are transmitted by blood-sucking arthropods. Finally, these viruses were joined by plant viruses. It is interesting to note that the host branching order of these viruses is plant, insect, animal, primate, human. It is tempting to postulate that these viruses evolved from an ancestor virus in this evolutionary direction. Alternatively, it is also possible that new viruses (e.g. HCV and GBVs) evolved through molecular recombination events between plant, insect, and animal viruses. Moreover, it is surprising that the genetic diversity of HCV genotypes is greater than that of BVD and HOG which infect different hosts. It may be suggested that newly discovered GBV-B and GBV-A also have a great genetic variability. The seroprevalence in 1300 samples from West Africa was 7.7% ~ 13.9% and a survey of blood donors in the USA indicated positive anti-GBV reaction in 2% donors [12]. The seroepidemiology and genomic sequences of GBVs will be clarified the clinical and public health significance using the same strategy of HCV. Recently Muerhoff et al. [13] reported detailed analyses of the genomic organization both GBVs and phylogenetic trees for helicase and RNA dependent RNA polymerase regions by the PHYLIP method. Our short analysis supports their analysis. It is interesting to think the origin of HCV and GB hepatitis viruses through the relationship of these viruses from an evolutionary viewpoint. Recently Simons et al. [14] reported novel virus-like sequences, GB virus C (GBV-C) associated with human hepatitis, which were different from GBV-A and GBV-B. Unfortunately, the published amino acids sequences were the downstream region, where we constructed the alignment and could not make the confident alignment of amino acids sequences were the downstream region, where we constructed the alignment and could not make the confident alignment of amino acids sequences including GBV-C. Moreover hepatitis G virus (HGV), a new hepatitis virus, of the full-length sequence was cloned and sequenced. The global identity between HGV and GBV-C was 95% at amino acids level suggesting that GBV-C was an isolate of HGV (Kim, J.P. personal communication, Genelabs Technologies, Inc., CA, USA). Kim et al. [15] reported HGV RNA was found in 1.7% donors with normal aranine transaminase (ALT) and in 2.3% donors with elevated ALT. However, the role of the GBV-C and HGV in the etiology and pathogenesis of chronic liver disease remains to be determined.

References

[1] Simons, J.N., Pilot-Matias, T.J., Leary, T.P., Dawson, G.J., Desai, S.M., Schlauder, G.G., Muerhoff, A.S., Erker, J.C., Buijk, S.L., Charlmers, M.L., Van Sant, C.L. and Mushahwar, I.K. (1995) Proc. Natl. Acad. Sci. USA 92, 3401–3405.
[2] Miller, R.H. and Purcell, R.H. (1990) Proc. Natl. Acad. Sci. USA 87, 2057–2061.
[3] Tsukiyama-Kohara, K., Iizuka, N., Kohara, M. and Nomoto, A. (1992) J. Virol. 66, 1476–1483.
[4] Ina, Y. (1991) ODEN ver. 1.1.1, Molecular evolutionary analysis system for DNA and amino acid sequences, National Institute of Genetics, Mishima, Japan.
[5] Kimura, M. (1983) The neutral theory of molecular evolution, Cambridge University Press, Cambridge.
[6] Nei, M. (1975) Molecular population genetics and evolution, North-Holland, Amsterdam.
[7] Saitou, N. and Nei, M. (1987) Mol. Biol. Evol. 4, 406–425.
[8] Felsenstein, J. (1993) PHYLIP inference package, ver. 3.5. Department of genetics, University of Washington, Seattle.
[9] Hodgman, T.C. (1988) Nature 333, 22–23.
[10] Gorbalenya, A.E. and Koonin, E.V. (1989) Nucleic Acids Res. 17, 8413–8440.
[11] Wassarman, D.A. and Steitz, J.A. (1991) Nature 349, 463–464.
[12] Zuckerman, A.J.Z. (1995) Lancet 345, 1453–1454.
[13] Muerhoff, A.S., Leary, T.P., Simons, J.N., Pilot-Matias, T.J., Dawson, G.J., Erker, J.C., Chalmers, M.L., Schlauder, G.G., Desai, S.M. and Mushauhwar, I.K. (1995) J. Virol. 69, 5621–5630.
[14] Simons, J.N., Leary, T.P., Dawson, G.J., Pilot-Matias, T.J., Muerhoff, A.S., Schlauder, G.G., Desai, S.M. and Mushahwar, I.K. (1995) Nature Med. 1, 564–567.
[15] Kim, J.P., Linnen, J., Wages, J., Shih, J., Nakatsuji, Y., Fry, K., Zhang-Keck, Z.-Y., Young, L., Gallagher, P., Krawczynski, K., Ismay, S., Hyams, C., Lifson, J., Alter, M., Margollis, H., Foung, S., Bradley, D. and Alter, H. (1995) Hepatology 22 (suppl.), 218A.