Proc. Natl. Acad. Sci. USA
Vol. 91, pp. 11022–11026, November 1994
Microbiology

Hepatitis C virus variants from Vietnam are classifiable into the seventh, eighth, and ninth major genetic groups

(non-A, non-B hepatitis/genetic heterogeneity/virus evolution/taxonomy)

HAJIME TOKITA*, HIROAKI OKAMOTO*, FUMIO TSUDA†, PHAM SONG‡, SUSUMU NAKATA§, TOHRU CHOSA§, HISAO IIZUKA¶, SHUNJI MISHIRO‖, YUZO MIYAKAWA**, AND MAKOTO MAYUMI*††

*Immunology Division, Jichi Medical School, Tochigi-Ken, Japan; †The Viral Hepatitis Research Foundation of Japan, Tokyo, Japan; ‡Institute for Clinical Research in Tropical Medicine, Bach Mai Hospital, Hanoi, Vietnam; §Bureau of International Cooperation, International Medical Center of Japan, Tokyo, Japan; ¶The Japanese Red Cross Saitama Blood Center, Saitama-Ken, Japan; ‖Institute of Immunology, Tokyo, Japan; and **Mita Institute, Tokyo, Japan

Communicated by Robert H. Purcell, July 22, 1994

ABSTRACT **Thirty-four (41%) of 83 hepatitis C virus (HCV) isolates from commercial blood donors in Vietnam were not classifiable into genotype I/1a, II/1b, III/2a, IV/2b, or V/3a; for 15 of them, the sequence was determined for 1.6 kb in the 5′-terminal region and 1.1 kb in the 3′-terminal region. Comparison of the 15 Vietnamese isolates among themselves and with reported full or partial HCV genomic sequences indicated that they were classifiable into four major groups (groups 6–9) divided into six genotypes (6a, 7a, 7b, 8a, 8b, and 9a). Vietnamese HCV isolates of genotypes 7a, 7b, 8a, 8b, and 9a were significantly different from those classified into groups 4, 5, and 6 based on divergence within partial sequences; those of genotype 6a were homologous to a Hong Kong isolate (HK2) of genotype 6a. Phylogenetic trees based on the envelope 1 (E1) gene (576 bp) of 55 isolates and a part of the nonstructural 5 (NS5) region (1093 bp) of 43 isolates revealed at least nine major groups, three of which (groups 7, 8, and 9) were identified only in Vietnamese blood donors. With a prospect that many more HCV isolates with significant sequence divergence will be reported from all over the world, the domain of the HCV genome to be compared and criteria for grouping/typing and genotyping/subtyping will have to be determined, so that they may be correlated with virological, epidemiological, and clinical characteristics.**

Hepatitis C virus (HCV) is a positive-stranded linear RNA virus of ≈9500 nt and is the major causative agent of blood-borne non-A, non-B hepatitis worldwide (for a review, see ref. 1). It has a genomic structure and organization similar to those of pestiviruses and flaviviruses with 5′- and 3′-untranslated regions (UTRs) and a single, long open reading frame coding for core, envelope 1 (E1) glycoprotein, envelope 2/nonstructural 1 (E2/NS1) glycoprotein, and nonstructural 2–5 (NS2–NS5) proteins (1).

At least 16 HCV genomes have been sequenced in their entirety (isolates and accession nos.: HCV-1, M62321; HCV-H, M67463; HC-J1, D10749; HCV-J, D90208; HCV-BK, M58335; HCV-T, M84754; HCV-JK1, X61596; HCV-JT, D01171; HCV-JT′, D01172; HC-J4/83, D01217; HC-J4/91, D10750; HC-C2, D10934; HC-G9, D14853; HC-J6, D00944; HC-J8, D01221; NZL1, D17763). They can be classified into three major groups or types provisionally designated 1, 2, and 3 and divided further into genotypes or subtypes called I/1a, II/1b, 1c, III/2a, IV/2b, and V/3a (2–5). Three additional groups, designated 4–6, are proposed based on divergence within partial sequences (6–8). The range of variability of HCV sequences, however, is not outlined as yet. Presently available HCV sequences were determined on limited isolates from blood donors and patients from restricted areas of the world.

Among 83 HCV isolates from commercial blood donors in Vietnam,‡‡ 34 (41%) were not classifiable into genotype I/1a, II/1b, III/2a, IV/2b, or V/3a by PCR with type-specific primers deduced from the HCV core gene (9). Since the applied method could not identify genotype VI/3b or those of major groups 4, 5, and 6, 32 (94%) of the 34 untypeable Vietnamese HCV isolates were compared with reported isolates of known partial sequences as well as among themselves. The results obtained indicate that there would be at least nine major groups of HCV, three of which are identified in Vietnamese isolates anew and require a proper method and criteria for grouping and genotyping HCV.

MATERIALS AND METHODS

Sera Containing HCV. Among 83 commercial blood donors in Vietnam positive for serum HCV RNA (4 from Hanoi in the north and 79 from Ho Chi Minh City in the south), 49 carried HCV classifiable into I/1a, II/1b, III/2a, IV/2b, or V/3a by PCR with type-specific primers deduced from the HCV core gene (10, 11), but the remaining 34, including the 4 from Hanoi, could not be classified (9). Of the 34 sera with HCV RNA of unknown genotypes, 32 had high HCV RNA titers by PCR with nested primers deduced from the 5′-UTR (12), and HCV cDNA clones were obtained from them for sequencing.

Determination of Nucleotide Sequences. Nucleic acids were extracted from 0.02–0.1 ml of serum and reverse-transcribed to cDNA with HCV-specific 20-mer primers (122, 299, 337, 316) or a nonspecific 43-mer primer (165) with $(A)_{17}$ as described (13). The extreme 5′-end sequence (nucleotides 1–160) was determined by a single-sided PCR as described (14). Two overlapping 5′-terminal cDNA fragments spanning nucleotides 63–847 (785 bp) and nucleotides 732–1606 (875 bp), as well as two overlapping 3′-terminal fragments spanning nucleotides 8259–8627 (369 bp) and from nucleotide 8531 to the poly(T) tail (889–918 bp), were amplified by PCR with AmpliTaq DNA polymerase (Perkin–Elmer) and the primer pairs 33/122, 50/337, 317/316, and 388/165, respectively (10, 11, 13). Positions were numbered from the putative 5′ end of the HCV genome of genotype II/1b with a known full-length sequence (15). Three cDNA clones were sequenced for each of the five regions, and the consensus sequence was deter-

The publication costs of this article were defrayed in part by page charge payment. This article must therefore be hereby marked "*advertisement*" in accordance with 18 U.S.C. §1734 solely to indicate this fact.

Abbreviations: HCV, hepatitis C virus; UTR, untranslated region; E1, envelope 1; E2, envelope 2; NS, nonstructural.

††To whom reprint requests should be addressed at: Immunology Division, Jichi Medical School, Minamikawachi-Machi, Tochigi-Ken 329-04, Japan.

‡‡The sequences reported in this paper have been deposited in the GenBank data base (accession nos. D14204, D14205, D14207–D14209, D14211, D17470–D17509, D21315–D21328, D30796, and D30797).

mined for at least 2702 nt, covering some 30% of the genomic sequence.

Phylogenetic Trees of HCV. Proposed phylogenetic trees of HCV were constructed by the unweighted pair-group method with the arithmetic mean (16) using a molecular evolutionary analysis system for DNA and amino acid sequences (ODEN; National Institute of Genetics, Mishima, Japan). The results were corrected by the equation of Gojobori *et al.* (17).

RESULTS

Sequence Variations Among Vietnamese HCV Isolates. Of the 83 HCV isolates from Vietnam, genotype I/1a was identified in 24 (29%), II/1b in 19 (23%), and III/2a in 4 (5%); both genotypes I/1a and II/1b were detected in 2 (2%) (9). Genotypes of the remaining 34 isolates were other than I/1a, II/1b, III/2a, IV/2b, and V/3a, which can be distinguished by the applied method (10, 11). Sequences of the NS5b region of 329 bp spanning nucleotides 8279–8607 were determined for 32 of the 34 Vietnamese HCV isolates of untypeable genotypes, and they were compared with those of six HCV genomes of genotype I/1a (HC-J1), II/1b (HC-J4/83), III/2a (HC-J6), IV/2b (HC-J8), V/3a (T-1; ref. 3), and VI/3b (T-9; ref. 3). The 32 Vietnamese isolates showed a similarity of only 61.4–69.3% to any of the six genotypes, indicating that they would belong to genotypes other than I/1a–VI/3b.

Based on two-by-two comparison and phylogenetic analysis of the 32 Vietnamese isolates as well as reasoning discussed below, they were classified into four major groups, 6, 7, 8, and 9, and further into six genotypes provisionally designated 6a, 7a, 7b, 8a, 8b, and 9a (Fig. 1). There were 16 HCV isolates of genotype 6a, 7 of 7a, 1 of 7b, 5 of 8a, 1 of 8b, and 2 of 9a. Some regional differences in the distribution of these genotypes were observed. The 28 isolates from Ho Chi Minh City were of genotype 6a, 7a, or 8a, while the 4 isolates from Hanoi were of 7b, 8b, or 9a.

Determination and Comparison of the 5′- and 3′-Terminal Sequences of Vietnamese HCV Isolates in Major Groups 6–9. For 15 Vietnamese HCV isolates in groups 6–9, the sequence was determined for 2702–2713 bp, which spanned 1580–1584 bp of the 5′-terminal region plus at least 1122 bp of the 3′-terminal region. They included four HCV isolates of genotype 6a (VN506, VN538, VN569, and VN571), four of 7a (VN540, VN787, VN843, and VN998), one of 7b (VN235), three of 8a (VN507, VN530, and VN531), one of 8b (VN405), and two of 9a (VN004 and VN085). HCV isolates of the same genotype showed sequence homology of >92%, and the representative sequences were adopted for genotypes 6a, 7a, 8a, and 9a.

Based on two-by-two comparisons of the six representative HCV isolates of genotypes 6a (VN506), 7a (VN540), 7b (VN235), 8a (VN507), 8b (VN405), and 9a (VN004), they had 77–85% sequence identity at the nucleotide level. Two isolates of the same major group, VN540 (7a) and VN235 (7b) as well as VN507 (8a) and VN405 (8b), were 83–85% similar. The highest homology (96–100%) was found in the 5′-UTR, and the lowest (48–70%) was in E2/NS1, with only 36–68% of the deduced amino acid sequence being identical. The homology in the 3′-UTR of 29 bp varied from 31% to 83%.

Comparison of 15 Vietnamese HCV Isolates Representing Six Additional Genotypes with Reported Isolates of Defined Genotypes. Nucleotide and deduced amino acid sequences of the 15 Vietnamese isolates of genotypes 6a, 7a, 7b, 8a, 8b, and 9a were compared with each of the known full-length sequences of the six HCV genomes of genotypes I/1a, II/1b, 1c, III/2a, IV/2b, and V/3a. The 15 Vietnamese isolates were similar to them in only 70–75% of the 2702–2706 bp compared. Vietnamese isolates differed from the six HCV genomes particularly in the nucleotide sequence of the E1 gene with a concordance of only 53–66%. Within the 5′-UTR,

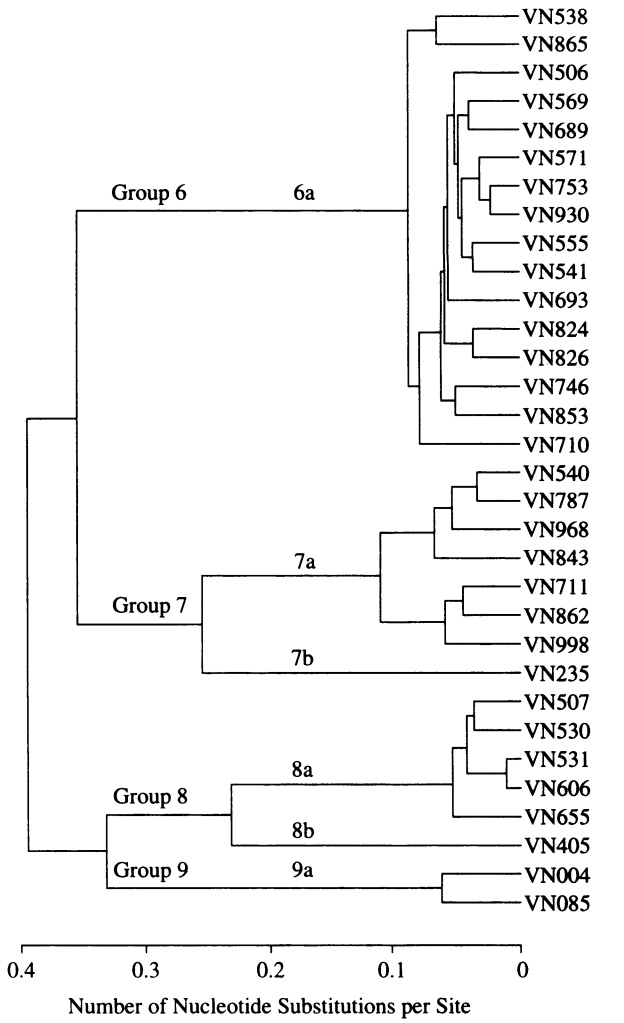

FIG. 1. A phylogenetic tree constructed on 32 Vietnamese HCV isolates of unclassifiable genotypes by comparison of an NS5b sequence (329 bp). The scale is nonlinear, which is also applicable to Figs. 2 and 3.

all the Vietnamese isolates were >96% identical to HCV genomes of group 1 (I/1a, II/1b, and 1c).

Fig. 2 illustrates a proposed phylogenetic tree of HCV based on comparison of the entire nucleotide sequence of the E1 region (576 bp) of 55 isolates. The tree indicates nine major groups of HCV, which are broken down into 23 distinct genotypes separated from each other by an evolutionary distance >0.2. Groups 8 and 9 differed by 0.395, with a distance comparable to that between genotypes III/2a and IV/2b at 0.397. Some intergroup and intergenotypic divergences, therefore, were not clearly separable by comparison of the E1 region.

Another phylogenetic tree was constructed (Fig. 3) based on comparison of a part of the NS5b region (1093 bp, nucleotides 8279–9371), and it classified 43 isolates into the same groups and genotypes as in Fig. 2. In this tree, the smallest intergroup distance was 0.246 between groups 8 and 9, which was significantly larger than the greatest intergenotypic distance of 0.206 between genotypes 7a and 7b.

Comparison of Vietnamese HCV Isolates in Groups 4–6 with Partially Sequenced Isolates. Many genotypes are proposed based on diversity in partial HCV sequences, and they are provisionally designated 4a–4d, 5a, and 6a (6). Four Vietnamese isolates (VN506, VN571, VN569, and VN538) showed ≥97% homology in the nucleotide sequences of a 5′-terminal region (321 bp) plus the entire E1 gene (576 bp) to

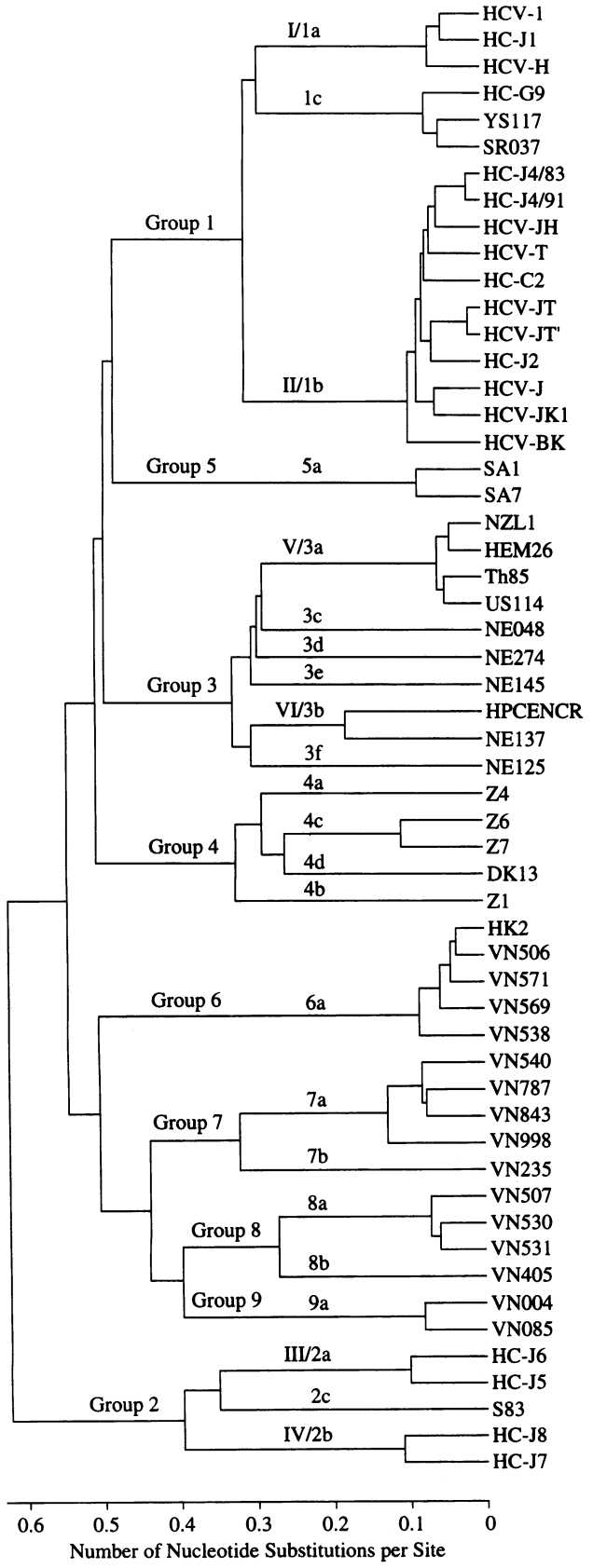

FIG. 2. A phylogenetic tree constructed on 55 HCV isolates by comparison of the nucleotide sequence of the E1 gene (576 bp). Isolates compared include 15 Vietnamese and others with reported sequences (accession nos. given in introduction or refs. 5, 6, 11, and 14). Genotypes 3c–3f have been reported in HCV isolates from Nepal (13).

a Hong Kong isolate of genotype 6a (HK2) reported by Bukh *et al.* (6, 18); therefore, they belonged to 6a. All the other five

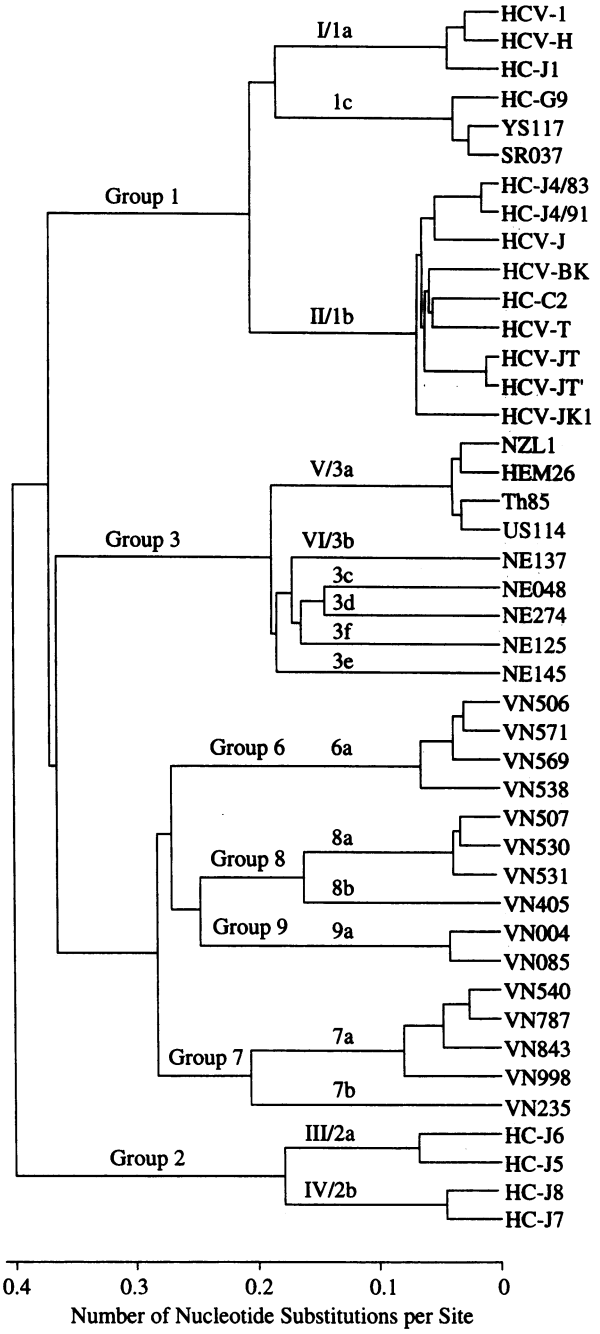

FIG. 3. A phylogenetic tree constructed on 43 HCV isolates by comparison of a part of the NS5b region (1093 bp, nucleotides 8279–9371).

genotypes, 7a, 7b, 8a, 8b, and 9a, represented by Vietnamese isolates VN540, VN235, VN507, VN405, and VN004, respectively, showed a low homology (60–67%) to reported genotypes (4a–4d, 5a, and 6a) in the nucleotide sequence of E1 gene. Sixteen Vietnamese isolates of genotype 6a had 91.8–95.9% similarity to a Hong Kong isolate (HK-2) of type 6 reported by Simmonds *et al.* (7) for 222 bp of the NS5b region. The other 16 Vietnamese isolates of genotypes 7a, 7b, 8a, 8b, and 9a were similar to HCV isolates of types 4–6 in only ≤74.8% within the partial NS5b sequence of 222 bp.

Characteristics of Vietnamese Isolates in Groups 7, 8, and 9. All Vietnamese isolates of genotypes 7a, 7b, 8a, 8b, and 9a, as well as those of genotype 6a, shared the six potential N-glycosylation sites (asparagine positions 196, 209, 234, 250, 305, and 325) and eight cysteine residues (positions 207, 226, 229, 238, 272, 281, 304, and 306) in the E1 protein. In

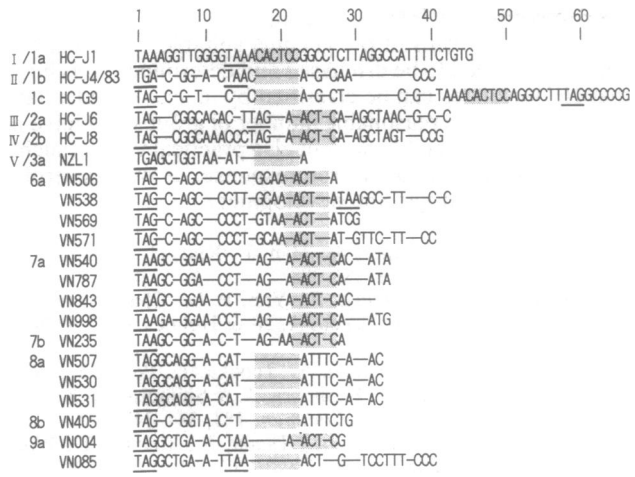

FIG. 4. The 3′-UTR sequences of Vietnamese isolates. Sequences of genotypes 6a, 7a, 7b, 8a, 8b, and 9a are compared with those of the six HCV genomes of distinct genotypes for which the entire sequence is known. The CACTCC motif is shaded, and in-frame stop codons are underlined. Dashes represent the same nucleotides as in HC-J1 of genotype I/1a.

addition, all four Vietnamese isolates of genotype 6a had Cys-344 (6), and the isolate of genotype 7b possessed Cys-373. A central sequence, GHRMAWDMM (aa 315–323), was conserved in all the Vietnamese isolates examined.

Vietnamese isolates of genotypes 6a, 8a, 8b, and 9a lacked the first amino acid (residue no. 384) of a hypervariable region of 25 aa at the N-terminal end of E2/NS1 region (19, 20). Three amino acids (Gly-389, Phe-403, and Gly-406) were preserved, and there were no cysteine residues in this region as in all reported isolates.

Vietnamese isolates of genotype 6a had two 1-bp deletions in the 5′-UTR (nucleotides 6 and 35) and two insertions (CA after nucleotide 197 and A after nucleotide 206) as a Hong Kong isolate (HK2), and those of genotypes 7a–9a had three 1-bp deletions (nucleotides 6, 15, and 35). A hairpin-like structure as well as direct repeats of CACTCC spanning nucleotides 23–28 and 38–43 were preserved in all of them.

The 3′-UTRs of Vietnamese isolates are shown in Fig. 4 along with those of six genomes of different genotypes with known entire nucleotide sequences. They had 3′-UTRs of variable lengths but shared the sequence of CACTCC with HCV of genotype II/1b (21). They were devoid of the fourth of the four stem–loop structures in the 3′-terminal 300 bp as is the case for HCV isolates of group 3 (4, 13).

DISCUSSION

Since it is an RNA virus, HCV evolves rapidly with a mutation rate estimated at 1.44–1.92 × 10^{-3} base substitutions per site per year (15, 22). Consequently, many variants of HCV arise, and classifying them into distinct genetic groups would have virological, clinical, and epidemiological values. HCV genotypes have different geographic distributions (6–9, 13, 23), and HCV of genotype II/1b is associated with more severe hepatitis as well as lower response to interferons than those of genotypes III/2a and IV/2b (24, 25).

Presently, at least three genotyping systems are proposed, which are based on sequence homology of different domains of the HCV genome. In one system, HCV isolates are classified into genotypes I, II, III, IV, and V based on the comparison of full-length genomic sequences, with intragenotypic divergence of ≤9.2% and intergenotypic divergence of ≥19.9% (4). In another classification scheme based on the phylogenetic analysis of a part of the NS5 region (222 bp) of 76 HCV samples, six major types, 1, 2, 3, 4, 5, and 6, are recognized with a series of subtypes designated a, b, c, etc. (7). Such a two-tiered classification is also advocated by others (8). In a third classification system, at least 12 genotypes are proposed by a phylogenetic analysis of 44 isolates collected worldwide within the sequence of 192 aa encoded by the E1 gene (6).

Since HCV isolates thus far sequenced are from limited areas of the world, many more isolates with substantial sequence variations are expected to occur in areas where HCV variations have not been examined. This proved to be the case in 83 HCV isolates from commercial blood donors in Vietnam, 34 (41%) of which were not typeable into any of the common genotypes (I/1a–V/3a). When 15 of them with high HCV RNA titers were sequenced within the 5′- and 3′-terminal regions covering some 30% of the genome, six additional genotypes were recognized, five of which have not been previously reported. Phylogenetic trees based on comparison of the entire E1 sequence of 576 bp and an NS5b sequence of 1093 bp indicated that most of them would belong to additional major groups, provisionally designated 7, 8, and 9. Altogether, there would be at least nine major groups of HCV, which break down into 23 genotypes.

Local differences in the distribution of the additional genotypes were noted in Vietnam. Three new genotypes were discovered in the four Hanoi samples, which were not found in any of 28 Ho Chi Minh City samples. This would indicate that many more genotypes would prevail in Hanoi, leaving the possibility for coinfection by two new genotypes to occur in some Hanoi samples.

Significant sequence divergence from the other groups notwithstanding, all Vietnamese HCV isolates of groups 7–9 had many characteristics in common with reported isolates within the E1 protein as well as in the 5′-UTR and 3′-UTR. Vietnamese HCV isolates of major groups 7, 8, and 9 highlighted issues that deserve consideration in classifying HCV isolates so far described and the many more likely to be reported from all over the world in the future.

First, the domain(s) of the HCV genome to be compared for classification needs to be determined. Apparently, domains of highly conserved sequences, typified by the 5′-UTR, or those of highly variable sequences such as the 3′-UTR and E2/NS1 region, are not suitable for comparison. As has been proposed by Bukh *et al.* (6) and reinforced in the present study, the E1 region may be a reasonable choice as the basis of classification. This domain has conserved structures, such as six possible glycosylation sites and eight cysteine residues, as well as a central sequence (GHRMAWDMM), which were identified in all the 11 Vietnamese isolates of groups 7, 8, and 9. The classification based on the E1 sequence would have an advantage of possible correlation with serotypes.

Second, the extent of divergence within groups and genotypes will have to be defined. Fig. 5 illustrates ranges of sequence divergence among HCV isolates within the entire genome and various domains used as the basis for classification. The two-tiered nature of sequence differences proposed by Simmonds *et al.* (7), based on a 222-bp portion (nucleotides 8313–8534) of the NS5b region, is no longer justified, because group and genotype divergences overlap when HCV isolates of new groups and genotypes are taken into account. Such an overlap is also observed by comparing the nucleotide sequence of the E1 gene or the amino acid sequence of E1 protein used by Bukh *et al.* (6). The overlap can be avoided when another partial sequence of 329 bp (nucleotides 8279–8607) or, better still, that of 1093 bp (nucleotides 8279–9371) in the NS5b region is compared. The borderline between group and genotypic differences is set at 19.7% in the comparison of an NS5b sequence (1093 bp), which will be subject to changes when many more sequence data accumulate. Thus, for a reasonable classification of

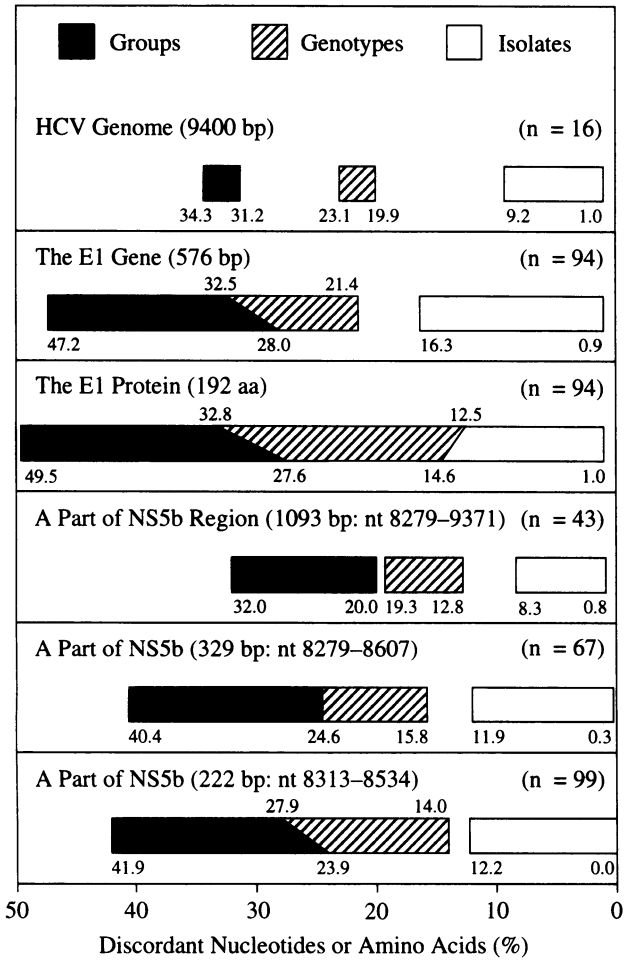

FIG. 5. Ranges of sequence divergence among groups, genotypes, and isolates of the same genotype when they are compared within the entire HCV genome and various domains. Two-by-two comparisons were performed on HCV isolates for which comparable sequences are available (the 32 Vietnamese isolates in this study are included).

HCV isolates, HCV domains would have to be selected and standardized to avoid further confusion.

Third, a prototype HCV genome representing each newly identified genotype should be sequenced in its entirety as the reference, which will define similarity and difference among HCV isolates of distinct genotypes with precision. This has been accomplished only for HCV genomes in groups 1–3 and of genotypes I/1a, II/1b, 1c, III/2a, IV/2b, and V/3a. A wide range of sequence divergence among HCV isolates of the same genotype would require selecting a representative genome before determining the entire sequence. Ideally, several HCV isolates of the target genotype would be sequenced partly within domains like the E1 gene and NS5b region, to select an isolate in the center of distribution curve. Such tactics could minimize costs and labor of sequencing the standard genome to be used as the reference for each HCV genotype. The serum or plasma from which the HCV prototypes are propagated and sequenced must be in quantities large enough for distribution upon request. Even better, they would be passaged through chimpanzees for cloning, until HCV is propagated in culture.

The three groups of HCV identified by sequencing Vietnamese isolates may herald many more groups of HCV from other unexamined areas. The standard nomenclature for the classification of HCV must be agreed upon immediately to keep attracting the interest of clinicians and epidemiologists. For a start, domains of the genome suitable as the basis for comparison need to be discussed and determined. It is hoped that HCV isolates can be classified by an easily understandable system into a reasonable number of groups that have not only virological significance but also practical values, such as disease severity and interferon response, as well as serological diagnosis and vaccine development.

1. Houghton, M., Weiner, A., Han, J., Kuo, G. & Choo, Q. L. (1991) *Hepatology* **14,** 381–388.
2. Simmonds, P., McOmish, F., Yap, P. L., Chan, S. W., Lin, C. K., Dusheiko, G., Saeed, A. A. & Holmes, E. C. (1993) *J. Gen. Virol.* **74,** 661–668.
3. Mori, S., Kato, N., Yagyu, A., Tanaka, T., Ikeda, Y., Petchclai, B., Chiewsilp, P., Kurimura, T. & Shimotohno, K. (1992) *Biochem. Biophys. Res. Commun.* **183,** 334–342.
4. Sakamoto, M., Akahane, A., Tsuda, F., Tanaka, T., Woodfield, D. G. & Okamoto, H. (1994) *J. Gen. Virol.* **75,** 1761–1768.
5. Okamoto, H., Kojima, M., Sakamoto, M., Iizuka, H., Hadiwandowo, S., Suwignyo, S., Miyakawa, Y. & Mayumi, M. (1994) *J. Gen. Virol.* **75,** 629–635.
6. Bukh, J., Purcell, R. H. & Miller, R. H. (1993) *Proc. Natl. Acad. Sci. USA* **90,** 8234–8238.
7. Simmonds, P., Holmes, E. C., Cha, T. A., Chan, S. W., McOmish, F., Irvine, B., Beall, E., Yap, P. L., Kolberg, J. & Ureda, M. S. (1993) *J. Gen. Virol.* **74,** 2391–2399.
8. Stuyver, L., Rossau, R., Wyseur, A., Duhamel, M., Vanderborght, B., Van, H. H. & Maertens, G. (1993) *J. Gen. Virol.* **74,** 1093–1102.
9. Song, P., Duc, D. D., Hien, B., Nakata, S., Chosa, T., Watanabe, J., Tsuda, F., Murata, K. & Okamoto, H. (1994) *Clin. Diag. Lab. Immunol.* **1,** 413–418.
10. Okamoto, H., Sugiyama, Y., Okada, S., Kurai, K., Akahane, Y., Sugai, Y., Tanaka, T., Sato, K., Tsuda, F., Miyakawa, Y. & Mayumi, M. (1992) *J. Gen. Virol.* **73,** 673–679.
11. Okamoto, H., Tokita, H., Sakamoto, M., Horikita, M., Kojima, M., Iizuka, H. & Mishiro, S. (1993) *J. Gen. Virol.* **74,** 2385–2390.
12. Nagayama, R., Tsuda, F., Okamoto, H., Wang, Y., Mitsui, T., Tanaka, T., Miyakawa, Y. & Mayumi, M. (1993) *J. Clin. Invest.* **92,** 1529–1533.
13. Tokita, H., Shrestha, S. M., Okamoto, H., Sakamoto, M., Horikita, M., Iizuka, H., Shrestha, S., Miyakawa, Y. & Mayumi, M. (1994) *J. Gen. Virol.* **75,** 931–936.
14. Okamoto, H., Kurai, K., Okada, S., Yamamoto, K., Iizuka, H., Tanaka, T., Fukuda, S., Tsuda, F. & Mishiro, S. (1992) *Virology* **188,** 331–341.
15. Okamoto, H., Kojima, M., Okada, S., Yoshizawa, H., Iizuka, H., Tanaka, T., Muchmore, E. E., Peterson, D. A., Ito, Y. & Mishiro, S. (1992) *Virology* **190,** 894–899.
16. Nei, M. & Gojobori, T. (1986) *Mol. Biol. Evol.* **3,** 418–426.
17. Gojobori, T., Ishii, K. & Nei, M. (1982) *J. Mol. Evol.* **18,** 414–423.
18. Bukh, J., Purcell, R. H. & Miller, R. H. (1992) *Proc. Natl. Acad. Sci. USA* **89,** 4942–4946.
19. Hijikata, M., Kato, N., Ootsuyama, Y., Nakagawa, M., Ohkoshi, S. & Shimotohno, K. (1991) *Biochem. Biophys. Res. Commun.* **175,** 220–228.
20. Weiner, A. J., Brauer, M. J., Rosenblatt, J., Richman, K. H., Tung, J., Crawford, K., Bonino, F., Saracco, G., Choo, Q.-L., Houghton, M. & Han, J. H. (1991) *Virology* **180,** 842–848.
21. Takamizawa, A., Mori, C., Fuke, I., Manabe, S., Murakami, S., Fujita, J., Onishi, E., Andoh, T., Yoshida, I. & Okayama, H. (1991) *J. Virol.* **65,** 1105–1113.
22. Ogata, N., Alter, H. J., Miller, R. H. & Purcell, R. H. (1991) *Proc. Natl. Acad. Sci. USA* **88,** 3392–3396.
23. Cha, T. A., Beall, E., Irvine, B., Kolberg, J., Chien, D., Kuo, G. & Urdea, M. S. (1992) *Proc. Natl. Acad. Sci. USA* **89,** 7144–7148.
24. Takada, N., Takase, S., Enomoto, N., Takada, A. & Date, T. (1992) *J. Hepatol.* **14,** 35–40.
25. Yoshioka, K., Kakumu, S., Wakita, T., Ishikawa, T., Itoh, Y., Takayanagi, M., Higashi, Y., Shibata, M. & Morishima, T. (1992) *Hepatology* **16,** 293–299.