JOURNAL OF CLINICAL MICROBIOLOGY, Dec. 1997, p. 3156–3162 Vol. 35, No. 12
0095-1137/97/$04.00+0

Determination of Hepatitis C Virus Genotypes in the United States by Cleavase Fragment Length Polymorphism Analysis

DAVID J. MARSHALL,[1] LAURA M. HEISLER,[1] VICTOR LYAMICHEV,[1] CHRISTOPHER MURVINE,[1] D. MICHAEL OLIVE,[1] GARTH D. EHRLICH,[2]† BRUCE P. NERI,[1] AND MONIKA DE ARRUDA[1]*

Third Wave Technologies, Inc., Madison, Wisconsin 53711,[1] and Department of Pathology, University of Pittsburgh School of Medicine, Pittsburgh, Pennsylvania 15261[2]

Received 5 May 1997/Returned for modification 11 July 1997/Accepted 4 September 1997

We describe the application of a new DNA-scanning method, which has been termed Cleavase Fragment Length Polymorphism (CFLP; Third Wave Technologies, Inc., Madison, Wis.), for the determination of the genotype of hepatitis C virus (HCV). CFLP analysis results in the generation of structural fingerprints that allow discrimination of different DNA sequences. We analyzed 251-bp cDNA products generated by reverse transcription-PCR of the well-conserved 5′-noncoding region of HCV. We determined the genotypes of 87 samples by DNA sequencing and found isolates representing 98% of the types typically encountered in the United States, i.e., types 1a, 1b, 2a/c, 2b, 3a, and 4. Blinded CFLP analysis of these samples was 100% concordant with DNA sequencing results, such that closely related genotypes yielded patterns with strong familial resemblance whereas more divergent sequences yielded patterns with pronounced dissimilarities. In each case, the aggregate pattern was indicative of genotypic grouping, while finer changes suggested subgenotypic differences. We also assessed the reproducibility of CFLP analysis in HCV genotyping by analyzing three distinct isolates belonging to a single subtype. These three isolates yielded indistinguishable CFLP patterns, as did replicate analysis of a single isolate. This study demonstrates the suitability of this technology for HCV genotyping and suggests that it may provide a low-cost, high-throughput alternative to DNA sequencing or other, more costly or cumbersome genotyping approaches.

Hepatitis C virus (HCV) was identified less than a decade ago as the primary causative agent of non-A, non-B hepatitis (6). HCV becomes chronically established in 70 to 90% of affected individuals; in 20 to 30% of cases, chronic HCV infection leads to cirrhosis and hepatocellular carcinoma (11, 16). Among the various predictors for development of these sequelae, viral genotype has emerged as a clinically significant variable correlated with liver disease. Viral genotype also appears to be prognostic for hepatocellular carcinoma (15) and for resistance to interferon therapy (3, 12, 26, 28–30). Furthermore, several genotypes have unique geographical origins, making them convenient markers in epidemiological studies (28). Finally, because HCV cannot be grown in culture, it is only by establishing correlations between characteristics like viral genotype and clinical outcome that new antiviral treatments can be evaluated (1, 19, 28).

Because of its sequence conservation, the 5′-noncoding region (5′NCR) is the target of most HCV detection and quantification assays, including the Amplicor HCV detection assay (Roche Diagnostic Systems, Branchburg, N.J.) (27) and the branched-DNA (Chiron, Inc., Emeryville, Calif.) (10) and Amplicor HCV Monitor quantification assays (Roche Diagnostic Systems, Basel, Switzerland) (7), as well as numerous homebrew methods. However, this region also contains genotypically variable sequence positions which permit discrimination of all of the major types and many of the subtypes of HCV.

Techniques for genotyping HCV have relied upon the examination of a limited number of diagnostic sequence polymorphisms in variable regions throughout the 5′NCR. Restriction fragment length polymorphism analysis has been used to discriminate types 1a, 1b, 2a, 3a, 3b, 4, 5, and 6 (8, 9, 14). Allele-specific oligonucleotide analysis such as the line probe assay (LiPA), which was developed by Innogenetics (Ghent, Belgium), uses reverse hybridization to oligonucleotide probes fixed to a solid support (22, 23). The second generation of the LiPA can identify types 1a, 1b, 2a/2c (i.e., these two cannot be distinguished from one another by analysis of the 5′NCR), 2b, 3a, 5a, and 6a, as well as types 4 and 10. All of these techniques have shown good concordance with DNA sequencing for detecting this limited number of types and are considerably simpler to execute than direct sequencing. Nevertheless, these approaches are restricted in terms of the number of types and subtypes that they can discriminate and would be difficult to apply in geographic areas in which diverse genotypes predominate, such as the Middle East or the Far East (4, 12, 24, 25). These limitations, coupled with high material costs, have prompted a number of researchers and reference laboratories to use DNA sequencing for routine analysis of the 5′NCR, despite its relatively high cost and low throughput.

We report here the application of a novel DNA-based technology, Cleavase Fragment Length Polymorphism (CFLP; Third Wave Technologies, Inc., Madison, Wis.) analysis, for the determination of HCV genotype. The CFLP technology relies on the formation of unique secondary structures that result when DNA is allowed to cool following brief heat denaturation (17). These structures serve as substrates for the structure-specific Cleavase I enzyme, which can be used to generate a set of cleavage products from any given DNA fragment (2). Because formation of these secondary structures is exquisitely sensitive to nucleotide sequence, the presence of sequence polymorphisms results in the generation of unique collections of cleavage products, or structural fingerprints, for each sequence analyzed. The similarities and differences of DNA sequences are thus reflected in the characters of the CFLP patterns generated from them. Common sequences re-

* Corresponding author. Phone: (608) 273-8933. Fax: (608) 273-6989. E-mail: Monika_de_Arruda_Indig@twt.com.

† Present address: Center for Genomic Sciences, Allegheny University of the Health Sciences, Pittsburgh, PA.

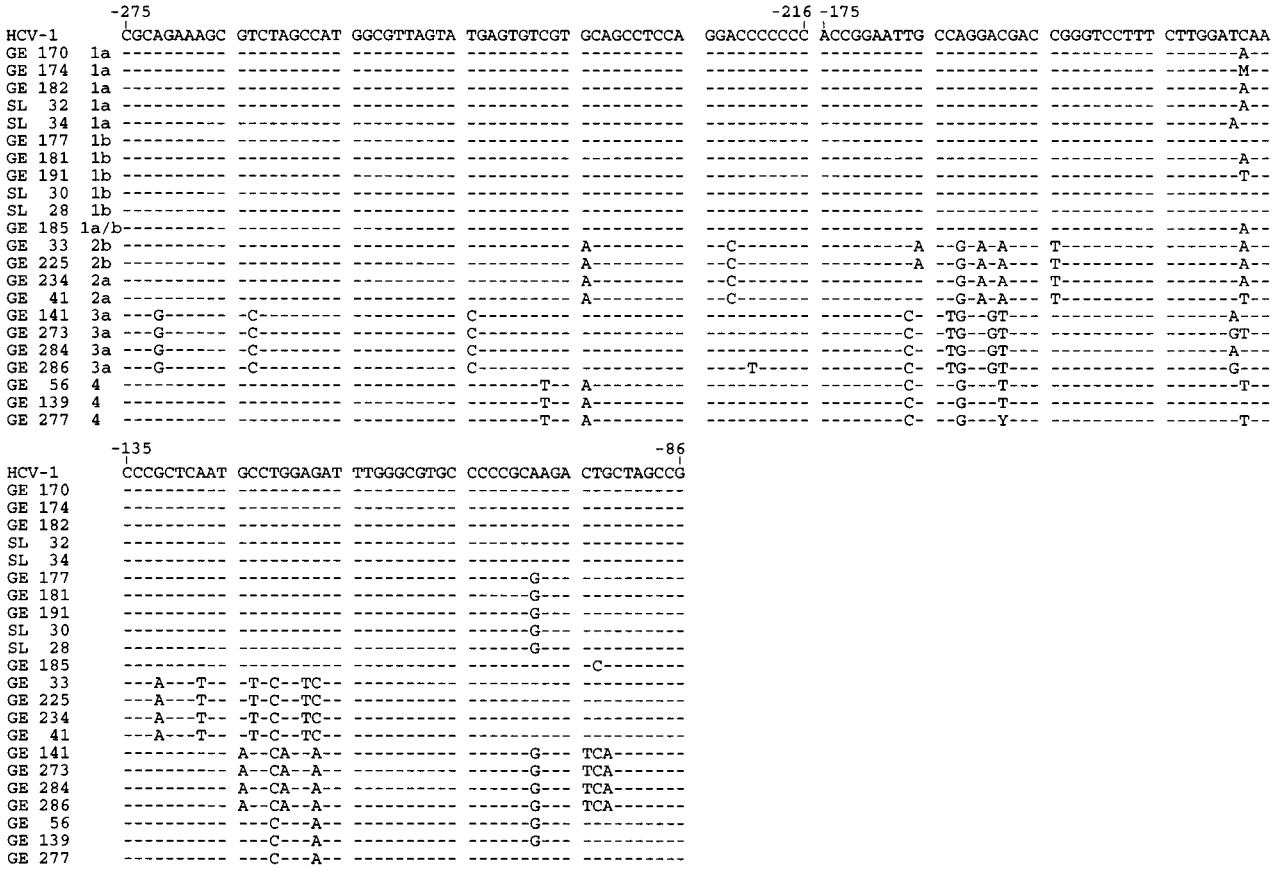

FIG. 1. Sequences of the variable portion of the 5′NCR for HCV isolates of various genotypes. Nucleotide positions reflect their locations relative to the start site of translation as reported elsewhere (20, 22, 23). Regions that showed no variability between isolates are omitted as indicated. The canonical type 1a sequence from HCV isolate 1 (6) is presented at the top. Deviations from this sequence as determined for the isolates presented in Fig. 2 and 3 are indicated; positions of identity with this sequence are indicated by dashes. The isolate number and genotype are indicated to the left of the 5′ end of the reported sequence. M, position containing A and C peaks; Y, position containing T and C peaks.

sult in the generation of common bands. Sequence differences are manifested in the CFLP patterns as one or more of the following: (i) loss or gain of one or more bands, (ii) shifts in position of bands, and (iii) a change in the relative intensities of bands. The structural fingerprints that result from CFLP analysis thus provide a ready means of identifying the extent to which sequences are related to one another.

We have used CFLP analysis for amplicons generated by reverse transcription-PCR (RT-PCR) amplification of the HCV 5′NCR to determine the genotype of the virus contained in HCV-positive plasma specimens. We show here that CFLP analysis exhibited complete correlation to nucleotide sequence data and was able to distinguish viral genotypes differing by as little as a single nucleotide. Blinded analysis of CFLP patterns compared to patterns from samples of known genotype resulted in assignments that were 100% consistent with DNA sequencing at the type and subtype levels. Furthermore, CFLP scanning detected variations within a given genotype, indicating that this method is suitable for identifying the appearance of new or rare HCV variants.

(Third Wave Technologies, Inc., can provide additional technical product availability information for the Cleavase I enzyme and associated reagents.)

MATERIALS AND METHODS

HCV cDNA samples. HCV-positive specimens were obtained by Garth Ehrlich (University of Pittsburgh, Pittsburgh, Pa.) as part of routine care by the Divisions of Transplantation Medicine and Surgery at the University of Pittsburgh Medical Center as described elsewhere (5). Additional HCV-positive plasma samples were obtained courtesy of Pamela Bean, Specialty Labs (Santa Monica, Calif.).

RNA extraction and cDNA preparation. The details of the home-brew RT-PCR method used to generate the double-stranded cDNA fragments have been described elsewhere (5). Briefly, RNA was extracted as described previously (5), and the tetrachlorofluorescein (TET)-labeled antisense primer HCV 102 (5′-TACCACAAGCGCTTTCGCGACCCAACACTACTC-3′) was used in a reverse transcriptase reaction to generate a cDNA copy of the antisense strand. Ten microliters of cDNA from each reverse transcriptase reaction was used as a template for subsequent PCRs with an unlabeled sense primer, HCV 101 (5′-CACTCCCCTGTGAGGAACTACTGTCT-3′), and the TET-labeled antisense primer HCV 102. Thermal cycling conditions were 94°C for 10 min for 1 cycle, 94°C for 30 s, 54°C for 1 s, and 72°C for 60 s for 35 cycles, followed by a 7-min extension at 72°C. The resultant 251-bp-long amplicons, which were uniquely labeled on the antisense strand, spanned the 5′NCR from nucleotides (nt) −304 to −54 (Fig. 1). Samples subjected to RT-PCR amplification of an amplicon shifted downstream relative to the 5′ end of the sense strand were generated with the Amplicor RT-PCR detection kit according to the manufacturer's recommendations. This fragment spans nt −274 to −31.

CFLP analysis of HCV cDNA amplicons. DNA fragments were partially purified prior to CFLP analysis according to a variation of procedures as described elsewhere (2). Briefly, RT-PCR products were heated to 70°C for 10 min, cooled to 37°C, and then treated with 1 U of exonuclease I (U.S. Biochemicals-Amersham, Inc., Cleveland, Ohio) per μl of PCR mixture for 30 min at 37°C. The nuclease was inactivated by heating to 70°C for 30 min. The reaction mixtures were then passed over a High Pure PCR Product Purification kit column (Boehringer Mannheim Biochemicals, Indianapolis, Ind.) according to the manufacturer's protocols, except that distilled water was used in place of the standard elution buffer. DNA was quantified by PicoGreen PCR quantification (Molecular Probes, Eugene, Oreg.) according to the manufacturer's recommended procedures. The samples generated by RT-PCR amplification with the Amplicor amplification system were treated similarly, except that these samples were first passed over a High Pure PCR Product Purification column to remove residual

uracil-*N*-glycosylase included in the Amplicor HCV procedure. These column eluates were adjusted to 50 mM KCl, 10 mM Tris-HCl (pH 8.3), and 1.5 mM $MgCl_2$ and then treated with exonuclease I and purified as described above.

CFLP reactions. CFLP analyses were performed as described elsewhere (2). Briefly, optimal reaction time and temperature were determined by examining matrices of reaction times (e.g., 1, 3, and 5 min) and temperatures (45, 50, and 55°C) to yield patterns with even distributions of short and long cleavage products. Approximately 25 to 30 ng (approximately 150 to 180 fmol) of the DNA substrate in a total volume of 10 μl of distilled water was heated to 95°C for 15 s, cooled to 55°C, and mixed with 10 μl of a solution containing 25 U of the Cleavase I enzyme, 2 μl of 10× CFLP buffer (100 mM morpholinepropanesulfonic acid [MOPS; pH 7.5], 0.5% Tween 20, 0.5% Nonidet P-40), and 2 μl of 2 mM $MnCl_2$, and the mixture was then incubated for 90 s at 55°C. The reactions were stopped by the addition of 16 μl of CFLP Fluoro-Stop solution (95% formamide, 10 mM EDTA [pH 8.0], 0.02% methyl violet). The cleavage fragments were resolved by denaturing gel electrophoresis through 10% (19:1 cross-link) polyacrylamide gels (20 by 20 cm by 0.5 mm) containing 7 M urea in 0.5× Tris-borate-EDTA. Fragments containing a 5′ TET label were analyzed by scanning the polyacrylamide gel cassettes on a Hitachi model FMBIO-100 fluorescence image analyzer equipped with a 585-nm filter.

Similar reaction conditions were used for CFLP analysis of the RT-PCR products generated with the Amplicor HCV system, except that digestion with the Cleavase I enzyme was for 4 min at 55°C. The antisense oligonucleotide primer included in the Amplicor system is labeled with a biotin moiety on its 5′ end. Following gel electrophoresis, these fragments were visualized by chemiluminescence detection of the biotin as follows. The DNA fragments were transferred to 0.2-μm-pore-size nylon membranes (Nytran Plus; Schleicher and Schuell, Keene, N.H.) by blotting overnight. The membranes were blocked in 1% Boehringer blocking reagent (Boehringer Mannheim Biochemicals) for 45 min. Streptavidin alkaline phosphatase conjugate (Sigma Chemical Company, St. Louis, Mo.) was added at a 1:15,000 dilution as a secondary detection agent. The membranes were washed three times (10 min each) with 0.1% sodium dodecyl sulfate (SDS) in Genius Buffer I (0.3 M NaCl, 0.2 M Tris-HCl [pH 7.5]), followed by three washes (5 min each) with Genius Buffer III (Boehringer Mannheim Biochemicals). Five milliliters of CDP Star (Tropix, Inc., Bedford, Mass.) was applied to each membrane as a chemiluminescent substrate in the detection system. The membranes were wrapped in cellophane and exposed to X-Omat radiographic film (Kodak, Rochester, N.Y.) for 10 to 30 min and developed manually.

All CFLP analyses were made by an individual blinded to the DNA sequencing results.

DNA sequence analysis. DNA sequencing was performed on an Applied Biosystems 377 DNA sequencer (Applied Biosystems, Foster City, Calif.) with dye-labeled dideoxynucleotide terminators according to the manufacturer's recommendations.

Nucleotide sequence accession numbers. The 5′NCR sequences presented in Fig. 1 will appear in the GenBank/EMBL/DDBJ database under the following accession numbers: isolate GE 33, AF021883; GE 41, AF021884; GE 225, AF021885; GE 234, AF021886; GE 170, AF021887; GE 174, AF021888; GE 182, AF021889; GE 177, AF021890; GE 181, AF021891; GE 191, AF021892; GE 185, AF021893; GE 141, AF021894; GE 273, AF021895; GE 284, AF021896; GE 286, AF021897; GE 56, AF021898; GE 139, AF021899; GE 277, AF021900; SL 28, AF021901; SL 30, AF021902; SL 32, AF021903; and SL 34, AF021904.

RESULTS

Sequence analysis of the 5′NCR of HCV cDNA fragments. A total of 198 HCV-positive serum specimens were available for this study. Of these, 63 were not suitable for DNA sequencing or CFLP analysis due to inadequate RT-PCR amplification. All of the remaining 135 samples were subjected to CFLP analysis without reference to the sequences; of these, 87 were confirmed by DNA sequencing. The results of the analyses performed on these 87 samples are summarized in Table 1, and the sequences of representative isolates are presented in Fig. 1. The majority of these isolates belonged to type 1, with type 1a being the most commonly encountered type. Type 1b, as defined by a polymorphism relative to type 1a at position −99 (A to G), was the second most commonly encountered. Among these, we encountered six variants of type 1a that contained a polymorphism at position −235, and three variants of 1b with this polymorphism. While these patterns showed slight differences from those obtained from samples with the canonical sequences for these types, they were clearly recognizable and were correctly called in a blind analysis. We also encountered three type 1 samples that contained a T-to-C polymorphism at position −94 that has been described for both type 1a and 1b isolates (1 a/b [23]). Less frequently encountered isolates included types 2, 3, and 4, with different subtypes being represented as indicated in Table 1. As with the type 1 samples, variants of some of these types containing sequence polymorphisms relative to the canonical sequences for these types were observed (one for type 3 and two for type 4). While these changes did not affect the ability to determine genotype in a blinded analysis, some of these polymorphisms were reflected in minor variations in the resultant CFLP patterns, such as those among the type 3 isolates in Fig. 2A, lanes 12 to 15.

TABLE 1. Genotype classifications of the samples characterized in the present study as determined by DNA sequencing[a]

Type and subtype	No. in present study	Percentage of total	Avg percentage in U.S. population
1	66	76	80
1a	35	40	58
1b	28	32	22
1a/b	3	3	10–20
2	10	11	14
2a/c	4	4	2
2b	6	7	12
3	5	6	4
4	6	7	1

[a] The percentages of the total cohort of samples represented by each genotype are indicated, as are the U.S. averages for each genotype as described elsewhere (28).

CFLP-based genotyping of RT-PCR products. We used CFLP analysis to determine the genotype of HCV cDNA fragments generated by RT-PCR. The structural fingerprint generated by CFLP analysis of a given molecule results from cleavage of localized structures (2). Two DNA molecules containing both unique and common nucleic acid sequences will retain the same structural fingerprint for the shared sequences. Thus, closely related HCV genotypes would be expected to have a familial relationship with slight pattern differences, reflective of sequence variations, embedded in a common overall structural fingerprint. The overall pattern would be indicative of genotypic grouping (i.e., type and subtype), while finer pattern changes may reflect isolate or subgenotypic variations. Comparison of the patterns in Fig. 2A, lanes 1 to 3, to those in lanes 4 to 7 reveals that while the patterns derived from subtypes 1a and 1b are closely related (compare bands A to G in all 7 lanes), these two subtypes are readily distinguishable from one another by the presence of pattern elements characteristic of each subtype. In particular, we observe differences in the relative intensities of bands H to J in the patterns from the subtype 1a isolates (lanes 1 to 3) compared to those generated from the 1b isolates (lanes 4 to 6). Figure 2A, lane 7, contains a structural fingerprint that is clearly recognizable as belonging to type 1 but that does not precisely match those generated from the samples in lanes 1 to 6. Sequence analysis of this isolate reveals that it contains an additional polymorphism at position −94 found in 10 to 20% of type 1 viruses (type 1a/b) (23), which results in the appearance of bands K and L. While the pattern change in lanes 4 to 6 versus lanes 1 to 3 is highly reproducible, it is subtle. Figure 2B, which contains CFLP patterns spanning a slightly different region, more clearly demonstrates the differences between these two subtypes. Specifically, 1b samples are readily differentiated from 1a by the presence of several bands in region I and pronounced shifts in intensity of several bands in region II.

We identified 10 type 2 isolates in this sample cohort. In-

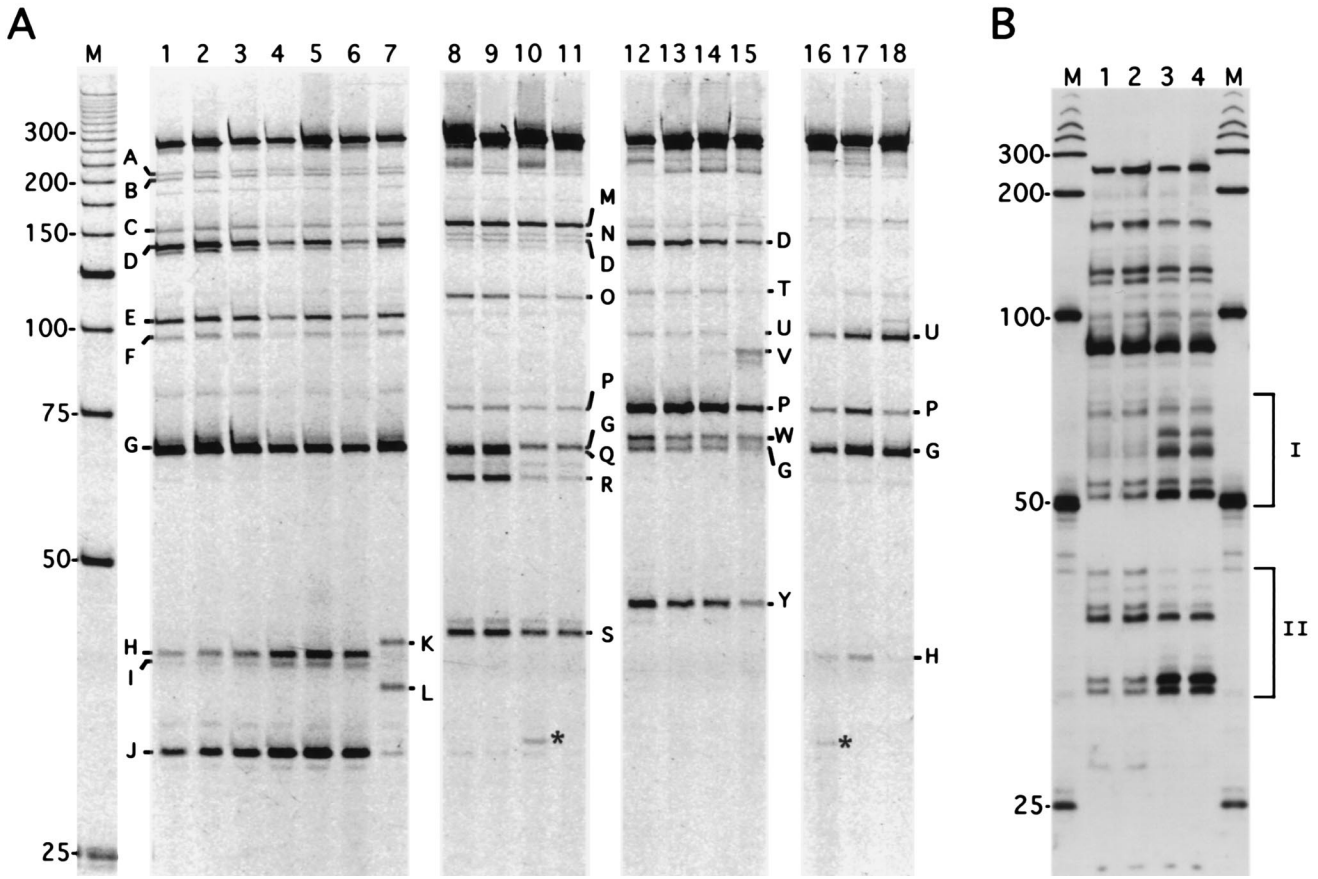

FIG. 2. (A) CFLP analysis of RT-PCR amplicons of 5′NCR from select HCV isolates. CFLP patterns were generated from RT-PCR amplicons generated from the 5′NCR of HCV isolates of various genotypes, as described in Materials and Methods. Lanes: M, molecular weight markers as indicated; 1 to 18, CFLP patterns from sample GE 170 (type 1a), GE 174 (type 1a), GE 182 (type 1a), GE 177 (type 1b), GE 181 (type 1b), GE 191 (type 1b), GE 185 (type 1a/b), GE 33 (type 2b), GE 225 (type 2b), GE 41 (type 2a/c), GE 234 (type 2a/c), GE 141 (type 3a), GE 273 (type 3a), GE 284 (type 3a), GE 286 (type 3a), GE 56 (type 4), GE 139 (type 4), and GE 277 (type 4), respectively. This image was obtained by scanning a polyacrylamide gel cassette on a Hitachi FMBIO 100 fluorescence image analyzer equipped with a 585-nm filter according to the manufacturer's recommended procedures. The image was labeled with Canvas 5.0 (Deneba Systems, Miami, Fla.) and printed on a Kodak 8650 PS dye sublimation printer. (B) CFLP analysis of RT-PCR products generated by the Amplicor HCV RT-PCR system. CFLP patterns were generated from RT-PCR amplicons generated from the 5′NCR of HCV isolates of genotypes 1a and 1b, as described in Materials and Methods. Lanes: M, molecular weight markers as indicated; 1 to 4, samples SL 34 and SL 32 (type 1a) and SL 30 and SL 28 (type 1b), respectively. The sequences of the RT-PCR products generated from these samples are presented in Fig. 1. The image was taken from an autoradiogram which was scanned with a Hewlett Packard ScanJet 4C, labeled with Canvas 5.0 (Deneba Systems), and printed on a Kodak 8650 PS dye sublimation printer.

spection of the CFLP patterns generated from four of these samples (Fig. 2A, lanes 8 to 11) indicates the presence of shared elements common to both type 1 and type 2 samples (viz., bands A to C, G, and J), which is reflective of the high degree of sequence similarity characteristic of the 5′NCR. Nonetheless, differences in numerous bands (D, M to R) make discrimination of these two types straightforward. Furthermore, sequence analysis revealed that the samples in lanes 8 and 9 belonged to type 2b and that those in lanes 10 and 11 belonged to type 2a/c. Comparison of the CFLP patterns generated from these samples indicates differences in intensity in bands Q and R. This correlation suggests that, upon analysis of an expanded sample set, CFLP structural fingerprinting may prove to be sufficiently sensitive to discriminate between type 2 subtypes. The sporadic appearance of pronounced bands migrating at more than 200 nt, such as in lanes 8, 10, and 12, is due to slight renaturation of double-stranded DNA during gel electrophoresis and is disregarded. Similarly, the bands in lanes 10 and 16 marked by asterisks are due to impurities in the cDNA sample itself and are seen in undigested controls (data not shown).

We encountered five type 3 and six type 4 samples, which is a finding consistent with the infrequent occurrence of these types in U.S. populations (29). Analysis of the sequences obtained from these two types reveals pronounced similarities; yet in each case, we identified characteristic elements distinguishing each type. Specifically, comparison of the CFLP patterns of representative type 3 and 4 samples (Fig. 2A, lanes 12 to 15 and 16 to 18, respectively) indicates the presence of pattern elements common to both types (bands G, P, and U), as well as one shared with types 1 and 2 (band G). Nonetheless, differences in numerous bands (D, W, G, and Y) make the discrimination of types 3 and 4 straightforward. The pattern in lane 15 contains differences relative to those in lanes 12 to 14, notably the loss of band U and the gain of band V, that may be attributable to a sequence difference at position −221. The presence of a band H in the type 4 samples in lanes 16 and 17 that is more prominent than that in lane 18 may reflect the sequence of these samples at position −99; the samples in lanes 16 and 17 contain the A-to-G polymorphism seen in patterns obtained from 1b isolates, whereas the sample in lane 18 does not.

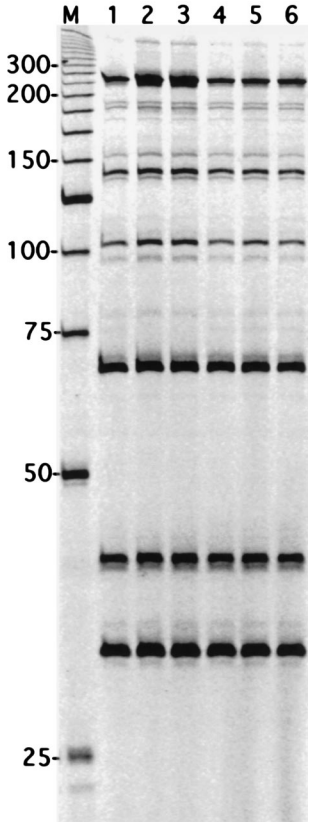

FIG. 3. Reproducibility of CFLP analysis. CFLP patterns were generated as described in Materials and Methods. Lanes: M, molecular weight markers as indicated; 1 to 3, products of CFLP reactions done on three different isolates of genotype 1b (samples GE 177, GE 181, and GE 191, respectively); 4 to 6, replicate CFLP reactions done on a single isolate (sample GE 177). This image was obtained by scanning a polyacrylamide gel cassette on a Hitachi FMBIO 100 fluorescence image analyzer equipped with a 585-nm filter according to the manufacturer's recommended procedures. The image was labeled with Canvas 5.0 (Deneba Systems) and printed on a Kodak 8650 PS dye sublimation printer.

Reproducibility of HCV CFLP patterns. Subtle variations among different isolates of a given HCV genotype could potentially lead to ambiguities in genotype determination. Therefore, we examined the ability of CFLP analysis to produce a reproducible structural fingerprint from three patients known to be infected with HCV type 1b. As shown in Fig. 3, lanes 1 to 3, the CFLP patterns were indistinguishable for the three isolates examined. The reproducibility of CFLP structural fingerprinting was examined further by generating multiple patterns from a single viral isolate (lanes 4 to 6). These data indicate that CFLP analysis is highly reproducible and that the genotypic assignments are not affected by strain variations within a given genotype. Finally, comparison of the patterns in Fig. 3 with those in Fig. 2A (lanes 1 to 6 versus lanes 4 to 6 in Fig. 2A) further underscores the reproducibility of this method when samples belonging to the same genotype are analyzed on different polyacrylamide gels.

DISCUSSION

The highly conserved 5′NCR of HCV can be routinely analyzed to determine genotype by a number of analytical methods. Techniques such as RFLP and LiPA are effective in identifying the types that occur most frequently in North America, Europe, and Japan (9, 14, 22, 23). However, DNA sequencing is currently the only means of discriminating types endemic to other geographical regions (3, 20). Recent reports indicate that diverse genotypes are beginning to emerge in the United States and Europe with greater frequency as a result of considerable immigration and travel, necessitating a more broadly based means of genotyping samples isolated in those locales (28). Indeed, in the present study, 13% of the isolates belonged to neither type 1 nor 2. Furthermore, HCV is a growing problem in much of the developing world, where diverse types predominate (12, 24–26, 28). Because routine direct sequencing is unfeasible in the typical clinical laboratory, the ability to obtain genotype information in this setting has been limited.

In this report, we describe the application of CFLP analysis, a new DNA-scanning method, to genotyping HCV in human plasma samples. CFLP digestion creates a structural fingerprint of a given DNA molecule which is the result of the cleavage of localized structures and has been applied to the analysis of numerous loci (2, 18). Therefore, two DNA molecules containing unique as well as common nucleic acid sequences will retain the same basic structural fingerprint for the shared sequences, while displaying distinct differences. Thus, when the relatively conserved 5′NCR is used for comparative purposes, HCV genotypes would be expected to have a familial relationship with pattern differences embedded in a common overall structural fingerprint. The aggregate pattern is, therefore, indicative of genotypic grouping, while finer changes suggest subgenotypic and strain differences.

In this report, HCV-positive plasma samples were categorized in terms of genotype and subgenotype by DNA sequencing of RT-PCR products generated from the 5′NCR. Variations in the CFLP patterns relative to those generated from samples comprised of the canonical type 1a sequence were in agreement with sequence differences between types and subtypes. For example, all type 1 samples resulted in closely related CFLP patterns yet contained differences reflecting the sequence elements typical of subtype 1a versus 1b, i.e., the −99 A-to-G polymorphism characteristic of type 1b. Furthermore, CFLP structural fingerprinting was able to differentiate type 1 variants with an alternative polymorphism (T-to-C transition at nt −94; type 1a/b) (23).

It is noteworthy that the difference between these two type 1 subtypes is more readily appreciated when a slightly different DNA fragment is analyzed. This fragment encompasses additional sequences 3′ of the polymorphism at position −99 (Fig. 2B). It is likely that the additional length downstream of the variable region permits these fragments to assume alternative structures that more definitively reflect the sequence differences among these subtypes. For this reason, the use of this amplicon appears to be preferable when subtype information is needed or when more facile analysis of patterns is required, such as in a high-throughput clinical laboratory setting. Sreevatsan et al. describe the use of this amplicon for CFLP analysis of a large number of clinical samples of diverse genotypes (21).

CFLP analysis of cDNAs from viruses belonging to types 2, 3, and 4 were clearly distinguished from each other and from the type 1 samples. Within type 2, we were able to distinguish subtype a/c from b with the small number of type 2 samples in the available sample cohort. Five samples classified as type 3a and six samples classified as type 4 by DNA sequencing gave closely related patterns but contained slight differences. However, the sequence elements characteristic of different type 4 subtypes have been insufficiently described to permit assignment of subtype based on the polymorphisms detected in this study. The results from this initial study strongly suggest that CFLP analysis can be used to differentiate type 3 and 4 sam-

ples. However, the definitive demonstration of this capability will require the analysis of additional samples belonging to these types which are rarely encountered in the United States (21). Such studies are presently under way.

The reproducibility of CFLP analysis and the consistency of individual genotype patterns are important considerations in determining the practical application of CFLP technology to HCV genotype analysis. While the 5′NCR is the most-conserved region of the HCV genome, there is sufficient variability within sequences analyzed for any given genotype to complicate classification of a particular sample. However, examination of multiple samples shows that the structural fingerprints obtained for a given genotype are consistent. Minor sequence variations did not obscure genotypic classification by CFLP analysis. Three distinct samples belonging to HCV type 1b gave indistinguishable CFLP patterns, reinforcing this observation. Furthermore, when a single type 1b sample was analyzed multiple times, the resultant structural fingerprints were indistinguishable. The ability to generate identical patterns reliably from multiple samples belonging to the same type as well as from multiple analyses of a single sample confirms the reproducibility of CFLP structural fingerprinting as a genotyping method. That the presence of a sequence polymorphism among these three type 1b samples did not alter the CFLP pattern suggests that this change at position −138 (Fig. 1) does not alter the structure of the cDNA in this region. This observation is consistent with the hypothesis that it is the sequence changes that result in structure changes in the 5′NCR that are characteristic of different genotypes (1).

While we have not examined all of the major genotypes described for HCV, those examined in the study represent 98% of the isolates encountered in the United States (9, 13, 28, 29). The CFLP method clearly categorized HCV into the major types and subtypes as well as identified single nucleotide variations within a particular genotype. The observation that CFLP structural fingerprinting can distinguish sequence changes as subtle as a single nucleotide suggests that the CFLP method should be able to categorize samples belonging to the other genotypes. This versatility should serve to make the CFLP technology a useful tool both for identifying emerging genotypes and for monitoring the natural history of HCV infection within a patient.

The CFLP assay system described herein for application to RT-PCR products can be carried out in 2 h or less at a total material cost per sample substantially lower than that for other genotyping methods, including DNA sequencing (21). CFLP analysis is a rapid, cost-effective method for HCV genotype determination and should contribute to more-effective and timely patient management.

ACKNOWLEDGMENTS

This work was supported by grants from the National Institutes of Standards and Technology Advanced Technology Program (94-05-0012).

We acknowledge Mary Oldenburg and Mary Ann Brow for technical assistance and comments on the manuscript and Pamela Bean of Specialty Laboratories for HCV-positive plasma specimens.

REFERENCES

1. **Altamirano, M., A. Delaney, A. Wonge, J. Marostenmaki, and D. Pi.** 1995. Identification of hepatitis C virus genotypes among hospitalized patients in British Columbia, Canada. J. Infect. Dis. **171:**1034–1038.
2. **Brow, M. A., M. Oldenberg, V. Lyamichev, L. Heisler, N. Lyamicheva, J. Hall, N. Eagan, D. M. Olive, L. Smith, L. Fors, and J. Dahlberg.** 1996. Differentiation of bacterial 16S rRNA genes and intergenic regions and *Mycobacterium tuberculosis katG* genes by structure-specific endonuclease cleavage. J. Clin. Microbiol. **34:**3129–3137.
3. **Bukh, J., R. Miller, and R. Purcell.** 1995. Genetic heterogeneity of hepatitis C virus: quasispecies and genotypes. Semin. Liver Dis. **15:**41–63.
4. **Cha, T., E. Beall, B. Irvine, J. Kolberg, D. Chien, G. Kuo, and M. Urdea.** 1992. At least five related, but distinct, hepatitis C viral genotypes exist. Proc. Natl. Acad. Sci. USA **89:**7144–7148.
5. **Chen, Y., D. L. Cooper, and G. D. Ehrlich.** 1996. Comparative analysis of three nucleic acid-based detection systems for hepatitis C virus RNA in plasma from liver transplant recipients. Mol. Cell. Probes **10:**331–336.
6. **Choo, Q.-L., G. Kuo, A. J. Weiner, L. R. Overby, D. W. Bradley, and M. Houghton.** 1989. Isolation of a cDNA clone derived from a blood-borne non-A, non-B viral hepatitis genome. Science **244:**359–362.
7. **Colucci, G., and K. Gutenunst.** 1997. Development of a quantitative PCR assay for monitoring HCV viraemia levels in patients with chronic hepatitis C. J. Viral Hepatol. **4**(Suppl. 1)**:**75–78.
8. **Constantine, N., M. Abdel-Hamid, and D. Oldach.** 1995. Rapid genotyping of hepatitis C virus. N. Engl. J. Med. **333:**800.
9. **Davidson, F., P. Simmonds, J. Ferguson, L. Jarvis, B. Dow, E. Follett, C. Seed, T. Krusius, C. Lin, G. Medgyesi, H. Kiyokawa, G. Olim, G. Duraisamy, et al.** 1995. Survey of major genotypes and subtypes of hepatitis C virus using RFLP of sequences amplified from the 5′ non-coding region. J. Gen. Virol. **76:**1197–1204.
10. **Detmer, J., R. Lagier, J. Flynn, C. Zayati, J. Kolberg, M. Collins, M. Urdea, and R. Sanchez-Pescador.** 1996. Accurate quantification of hepatitis C virus (HCV) RNA from all HCV genotypes by using branched-DNA technology. J. Clin. Microbiol. **34:**901–907.
11. **Donohue, J. G., A. Munoz, P. M. Ness, D. E. Brown, Jr., D. H. Yawn, H. A. McAllister, Jr., B. A. Reitz, and K. E. Nelson.** 1992. The declining risk of posttransfusion hepatitis C virus infection. N. Engl. J. Med. **327:**369–373.
12. **Dusheiko, G., H. Schmilovitz-Weiss, D. Brown, F. McOmish, P.-L. Yap, S. Sherlock, N. McIntyre, and P. Simmonds.** 1994. Hepatitis C virus genotypes: an investigation of type-specific differences in geographic origin and disease. Hepatology **19:**13–18.
13. **Lau, J., G. Davis, L. Prescott, G. Maertens, K. Lindsay, K.-P. Qian, M. Mizokami, P. Simmonds, and the H. I. T. Group.** 1996. Distribution of hepatitis C virus genotypes determined by line probe assay in patients with chronic hepatitis C seen at tertiary referral centers in the United States. Ann. Intern. Med. **124:**868–876.
14. **McOmish, F., S.-W. Chan, B. Dow, J. Gillon, W. Frame, R. Crawford, P.-L. Yap, E. Follett, and P. Simmonds.** 1993. Detection of three types of hepatitis C virus in blood donors: investigation of type-specific differences in serological reactivity and rate of alanine aminotransferase abnormalities. Transfusion **33:**7–13.
15. **Nousbaum, J.-B., S. Pol, B. Nalpas, P. I. Landais, P. Berthelot, C. Brechot, and C. S. Group.** 1995. Hepatitis C virus type 1b (II) infection in France and Italy. Ann. Intern. Med. **122:**161–168.
16. **Olynyk, J. K., and B. R. Bacon.** 1995. Hepatitis C: recent advances in understanding and management. Postgrad. Med. **98:**79–92.
17. **Orita, M., Y. Suzuki, T. Sekiya, and K. Hayashi.** 1989. A rapid and sensitive detection of point mutations and genetic polymorphisms using polymerase chain reaction. Genomics **5:**874–879.
18. **Rossetti, S., S. Englisch, E. Bresin, P. F. Pignatti, and A. E. Turco.** 1997. Detection of mutations in human genes by a new rapid method: cleavage fragment length polymorphism analysis (CFLPA). Mol. Cell. Probes **11:**155–160.
19. **Sharara, A., C. Hunt, and J. Hamilton.** 1996. Hepatitis C. Ann. Intern. Med. **125:**658–668.
20. **Simmonds, P., F. McOmish, P. L. Yap, S.-W. Chan, C. Lin, G. Dusheiko, A. Saeed, and E. Holmes.** 1993. Sequence variability in the 5′ non-coding region of hepatitis C virus: identification of a new virus type and restrictions on sequence diversity. J. Gen. Virol. **74:**661–668.
21. **Sreevatsan, S., J. B. Bookout, F. M. Ringpis, M. R. Pottathil, D. J. Marshall, M. de Arruda, C. Murvine, L. Fors, R. M. Pottathil, M. Eisenberg, T. Alcorn, M. Lai-Goldman, and R. R. Barathur.** Algorithmic approach to high-throughput molecular screening for alpha interferon-resistant genotypes in hepatitis C patients. Submitted for publication.
22. **Stuyver, L., R. Rossau, A. Wyseur, M. Duhamel, B. Vanderborght, H. Heuverswyn, and G. Maertens.** 1993. Typing of hepatitis C virus isolates and characterization of new subtypes using a line probe assay. J. Gen. Virol. **74:**1093–1102.
23. **Stuyver, L., A. Wyseur, W. van Arnheim, F. Hernandez, and G. Maertens.** 1996. Second-generation line probe assay for hepatitis C virus genotyping. J. Clin. Microbiol. **34:**2259–2266.
24. **Tokita, H., H. Okamoto, H. Iizuka, J. Kishimoto, F. Tsuda, L. Lesmana, Y. Miyakawa, and M. Mayumi.** 1996. Hepatitis C virus variants from Jakarta, Indonesia classifiable into novel genotypes in the second (2e and 2f), tenth (10a) and eleventh (11a) genetic groups. J. Gen. Virol. **77:**293–301.
25. **Tokita, H., H. Okamoto, F. Tsuda, P. Song, S. Nakata, T. Chosa, H. Iizuka, S. Mishiro, Y. Miyakwa, and M. Mayumi.** 1994. Hepatitis C virus variants from Vietnam are classifiable into the seventh, eighth, and ninth major genetic groups. Proc. Natl. Acad. Sci. USA **91:**11022–11026.
26. **Tsubota, A., K. Chayama, K. Ikeda, A. Yasuji, I. Koida, S. Saitoh, M. Hashimoto, S. Iwasaki, M. Kobayashi, and K. Hiromitsu.** 1994. Factors

predictive of response to interferon-alpha therapy in hepatitis C virus infection. Hepatology **19:**1088–1094.
27. **Young, K. K. Y., R. M. Resnick, and T. W. Myers.** 1993. Detection of hepatitis C virus RNA by a combined reverse transcription-polymerase chain reaction assay. J. Clin. Microbiol. **31:**882–886.
28. **Zein, N. N., and D. H. Persing.** 1996. Hepatitis C genotypes: current trends and future implications. Mayo Clin. Proc. **71:**458–462.
29. **Zein, N. N., J. Rakela, E. L. Krawitt, K. R. Reddy, T. Tominaga, D. H. Persing, and the C. S. Group.** 1996. Hepatitis C virus genotypes in the United States: epidemiology, pathogenicity, and response to interferon therapy. Ann. Intern. Med. **125:**634–639.
30. **Zeuzem, S., A. Franke, J.-H. Lee, G. Herrmann, B. Ruster, and W. K. Roth.** 1996. Phylogenetic analysis of hepatitis C virus isolates and their correlation to viremia, liver function tests, and histology. Hepatology **24:**1003–1009.