Clinical Chemistry 51:7
1123–1131 (2005)

Molecular Diagnostics and Genetics

Population Genotyping of Hepatitis C Virus by Matrix-Assisted Laser Desorption/Ionization Time-of-Flight Mass Spectrometry Analysis of Short DNA Fragments

Yoon Jun Kim,[1†] Soo-Ok Kim,[2†] Hyun Jae Chung,[2] Mi Sun Jee,[2] Byeong Gwan Kim,[1] Kang Mo Kim,[1] Jung-Hwan Yoon,[1] Hyo-Suk Lee,[1] Chung Yong Kim,[1] Sukjoon Kim,[2] Wangdon Yoo,[2] and Sun Pyo Hong[2*]

Background: **Identifying hepatitis C virus (HCV) genotypes has become increasingly important for determining clinical course and the outcome of antiviral therapy. Here we describe the development of restriction fragment mass polymorphism (RFMP) analysis, a novel matrix-assisted laser desorption/ionization time-of-flight mass spectrometry (MALDI-TOF MS) assay suitable for high-throughput, sensitive, specific genotyping of multiple HCV species.**

Methods: **The assay is based on PCR amplification and mass measurement of oligonucleotides containing genotype-specific motifs in the 5′ untranslated region, into which a type IIS restriction endonuclease recognition was introduced by PCR amplification. Enzymatic cleavage of the products led to excision of multiple oligonucleotide fragments representing variable regions whose masses were determined by MALDI-TOF MS.**

Results: **The RFMP assay identified viral genotypes present at concentrations as low as 0.5% and reliably determined their relative abundance. When sera from 318 patients were analyzed, the RFMP assay exhibited 100% concordance with results obtained by clonal sequencing and identified mixed-genotype infections in 22% of the samples, in addition to several subtype variants.**

Conclusions: **The RFMP assay has practical advantages over existing methods, including better quantitative detection of mixed populations and detection of genotype variants without need for population-based cloning, enabling reliable viral genotyping in laboratories and efficient study of the relationship between viral genotypes and clinical outcome.**

© 2005 American Association for Clinical Chemistry

Hepatitis C virus (HCV)[3] demonstrates a high degree of sequence variation, which provides the basis for grouping the virus into 6 major genotypes and multiple subtypes *(1, 2)*. These groupings have served as important tools for studying the geographic distribution of HCV genotypes, their routes of transmission, and their association with particular risk groups. Associations between viral genotype, interferon responsiveness, progression of disease, and the likelihood of developing hepatocellular carcinoma have been demonstrated *(3–8)*. The HCV genotype has emerged as an important factor in predicting the response to and determining the duration of antiviral therapy, such as interferon in combination with ribavirin, with genotype 1 infections having less favorable response rates and requiring longer treatment *(9, 10)*. HCV genotyping thus is firmly established as part of the pretreatment evaluation of patients with chronic infections, as proposed by the European Association of Study of Liver consensus conference in 1999 *(11)* and the NIH consensus conference in 2002 *(12)*.

[1] Department of Internal Medicine, Liver Research Institute, Seoul National University College of Medicine, Seoul, Korea.

[2] GeneMatrix, Seoul, Korea.

†These authors contributed equally to this work.

*Address correspondence to this author at: GeneMatrix, 8F, Cancer Research Institute, Seoul National University College of Medicine, 28 Yongon-Dong, Chongno-Gu, Seoul 110-799, Korea. Fax 82-2-764-8076; e-mail sunphong@genematrix.net.

Received December 26, 2004; accepted April 12, 2005.

Previously published online at DOI: 10.1373/clinchem.2004.047506

[3] Nonstandard abbreviations: HCV, hepatitis C virus; MALDI-TOF MS, matrix-assisted laser desorption/ionization time-of-flight mass spectrometry; RFMP, restriction fragment mass polymorphism; UTR, untranslated region; and TEAA, triethylammonium acetate.

Many HCV genotyping methods have been developed, including restriction fragment length polymorphism analysis, genotype-specific PCR, heteroduplex mobility analysis, melting curve analysis with fluorescence resonance energy transfer probes, and line probe assay, a DNA hybridization method; however, nucleotide sequencing of an appropriate subgenomic region remains the most widely accepted method *(13–18)*. Most of these methods, however, are labor-intensive, complex to standardize, and/or ineffective for determining multiple HCV genotypes in an experimental setup. Because of a high rate of mutation, extensive HCV genetic heterogeneity may occur in an individual during the course of infection, and more than 1 genotype may occur in some patients, particularly in individuals who have received multiple transfusions or are intravenous drug users *(19)*. The perceived rate of quasispecies and mixed-genotype infections is highly variable in patients tested by different assays *(20)*, and it is difficult to assess the true prevalence of mixed infection by current methods because they are designed to identify only the predominant HCV genotype in the population. The presence or clinical significance of mixed-genotype infection thus may be underestimated *(19, 21)*.

We have recently developed a matrix-assisted laser desorption/ionization time-of-flight mass spectrometry (MALDI-TOF MS)–based genotyping assay that exploits differences between wild-type and variant sequences in hepatitis B virus *(22)*. In the current study, we used restriction fragment mass polymorphism (RFMP) analysis, an advanced MALDI-TOF MS–based assay, for HCV genotyping capable of differentiating a cluster of multiple variations in the 5′ untranslated region (UTR). We assessed the sensitivity and reliability of the assay by comparing the results with direct or clonal sequencing; we then used the RFMP assay to determine the genotype distribution and prevalence of mixed-genotype infections in 318 chronically infected patients.

Materials and Methods

Samples

Serum samples were obtained from 318 chronically HCV-infected patients treated at the Liver Clinic of Seoul National University Hospital in Korea between March 2002 and July 2003. All of the patients were seropositive for anti-HCV, as tested by enzyme immunoassay (Abbott Diagnostics), and positive for serum HCV RNA, as tested by Cobas Amplicor HCV, Ver. 2.0 (Roche Molecular Systems). Criteria for exclusion included previous immunoglobulin-, interferon-, or other immune- or cytokine-based therapies with possible activity against HCV. Informed consent was obtained from each participant, and the experimental protocol conformed to the ethics guidelines of the 1975 Declaration of Helsinki, as reflected in a priori approval by the Seoul National University Hospital Human Research Committee.

HCV RNA extraction and cDNA amplification

HCV RNAs were extracted from 140-μL serum samples by the QIAamp® Viral RNA method (Qiagen) according to the manufacturer's instructions. Reverse transcription and PCR were performed in 1-step reactions. Known positive and negative sera and water controls were included in all reactions from the extraction step. The 5′ UTR was amplified with first-round primer pairs described previously *(23)*: KY80 (sense; 5′-GCAGAAAGCGTCTAGCCATGGCGT-3′; nucleotides −274 to −251) and KY78 (antisense; 5′-CTCGCAAGCACCCTATCAGGCAGT-3′; nucleotides −29 to −54). For the second round, we used 2 sets of nested PCR primer. A sense primer, CV160F, specific to bases −199 to −168 (5′-GTGGTCTGCGGAATCCAACGGTGAGTACACCGGAAT-3′) and an antisense primer, CV100R, specific to bases −67 to −92 (5′-TTCGCRACCCAACRCTACTCCAACGGTCCGGCTAG-3′) were used for detection of sequences from −160 to −100 (R = A or G). A sense primer, CV140F, specific to bases −158 to −140 (5′-GACIGGGTCCTGGATGTCTTGGA-3′) and an antisense primer, CV120R, specific to bases −100 to −117 (5′-GCGGGGGCACGGATGCCCAAAT-3′) were used for detection of sequences from −140 to −120. As illustrated in Fig. 1, sequences underlined in the CV160F and CV100R primers were modified to insert *Mme*I and *Mme*I/*Ava*II restriction recognition sites, respectively, in the PCR products, and underlined sequences in the CV140F and CV120R primers were modified to introduce *Fok*I sites in the products.

cDNA amplifications were performed with 20-μL reaction mixtures containing 10 μL of RNA, 20 mM Tris-HCl (pH 9.0), 50 mM KCl, 0.2 mM each deoxynucleotide triphosphate, 1.2 mM $MgSO_4$, 0.4 μM each primer, 0.1 mM dithiothreitol, 40 units of RNase inhibitor (Invitrogen), 200 U of SuperScript™ III Moloney murine leukemia virus reverse transcriptase (Invitrogen), and 2.5 U of Platinum® *Taq* polymerase (Invitrogen). After incubation at 50 °C for 45 min and at 94 °C for 2 min, PCR was performed for 35 cycles; each cycle consisted of denaturation at 94 °C for 15 s, annealing at 55 °C for 15 s, and extension at 72 °C for 20 s. To 17 μL of the second-round PCR mixture, 1 μL of the first-round PCR product was added. The nested PCR mixture was identical to the first-round mixture except for the omission of dithiothreitol, RNase inhibitor, and Moloney murine leukemia virus reverse transcriptase. PCR was conducted with the conditions described above, except that the annealing temperature was 60 °C. Nucleotide sequence positions were numbered according to the prototype HCV type 1 sequence *(24)*.

RFMP assay for HCV genotyping

Restriction enzyme digestion of the nested PCR product obtained with the CV160F and CV100R primer pair was performed by mixing the PCR reaction mixture with 10 μL of reaction buffer containing 50 mmol/L potassium acetate, 20 mmol/L Tris-acetate, 10 mmol/L magnesium

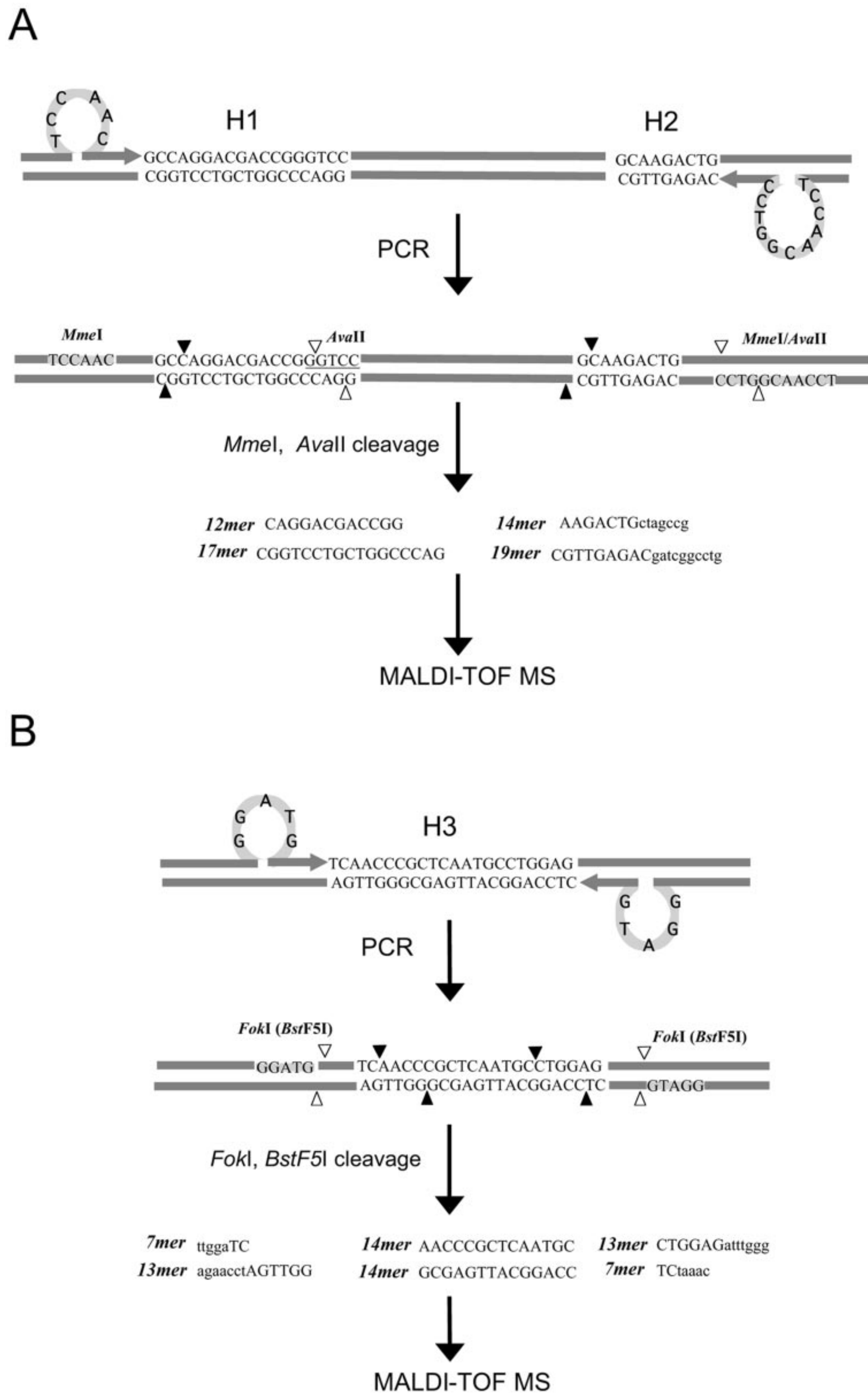

Fig. 1. Schematic diagram of the RFMP genotyping strategy.

From the alignment of the 5′ UTR from published HCV sequences, 3 variable motifs surrounded by conserved regions were chosen: H1, −166 to −150; H2, −101 to −93; and H3, −139 to −118. PCR was performed with primers designed to introduce a type IIS restriction endonuclease recognition sequence (*Mme*I or *Fok*I) ahead of the genotype-specific motifs on amplification. Enzymatic cleavage of the products led to excision of multiple oligonucleotide fragments representing the motifs shown (*uppercase sequences*); the masses of the resulting oligonucleotide fragments were then examined by MALDI-TOF MS. (*A*), genotyping strategy for H1 and H2 motifs: cleavage sites of *Mme*I (*filled arrowheads*) and *Ava*II (*open arrowheads*); artificially introduced recognition sites of *Mme*I and a combined *Mme*I/*Ava*II (*shaded sequences*); and a naturally occurring *Ava*II site (*underlined sequence*). (*B*), genotyping strategy for H3 motif, with cleavage sites of *Fok*I (*filled arrowheads*) and *Bst*F5I (*open arrowheads*), an isoschizomer for *Fok*I, and recognition sites of both restriction endonucleases (*shaded sequences*).

acetate, 1 mmol/L dithiothreitol, and 1 U of *Mme*I (New England Biolabs). The reaction mixture was incubated at 37 °C for 2 h and then incubated at 37 °C for 2 h with *Ava*II (New England Biolabs). Enzyme digestion of the nested PCR products obtained with the CV140F and CV120R primer pair was identical, except for the addition of 50 μmol/L *S*-adenosylmethionine in the mixture, the use of *Fok*I (New England Biolabs), and the subsequent digestion with *Bst*F5I (New England Biolabs). The resulting digest was purified by vacuum filtration through a custom-built 384-well sample preparation plate containing 5 mg of polymeric sorbent per well (Waters) or Oasis® μElution Plates (Waters). Each well was equilibrated with 90 μL of 1 mol/L triethylammonium acetate (TEAA; pH 7.6). Each cleavage reaction mixture was added to 70 μL of 1 mol/L TEAA (pH 7.6) and loaded in a well. After the wells were rinsed 5 times with 85 μL of 0.1 mol/L TEAA (pH 7.0), the plate was reassembled on a vacuum manifold and eluted with 60 μL of 600 mL/L isopropanol in water into a collection plate, which was placed on a heating block at 115 °C for 90 min. The desalted reaction mixtures were resuspended with matrix solution containing 15 g/L 3-hydroxypicolinic acid, 0.023 mol/L ammonium citrate, and 120 mL/L acetonitrile; they were then spotted in 3-μL volumes on a polished MTP AnchorChip™ plate (Bruker Daltonics). Mass spectra were acquired on a Biflex IV linear MALDI-TOF MS (Bruker Daltonics) workstation equipped with a 337 nm nitrogen laser and a nominal ion flight path length of 1.25 m as described previously, with slight modification *(22)*. The samples were analyzed in negative-ion mode with a total acceleration voltage of 20 kV, extraction voltage of 18.25 kV, laser attenuation of 55, and delayed extraction of long time delay. Typically, time-of-flight data from 10 individual laser pulses were recorded and averaged on a transient digitizer with a time base of 2 ns and delay of 24 000 ns, after which the averaged spectra were automatically converted to mass by the accompanying data-processing software (Bruker Daltonics). With such settings, the instrument usually provides mass accuracy of 40–80 ppm (10^{-6}), mass resolution of 1500–2000, and sensitivity of 10–50 fmol in the 2- to 6-kDa mass range for oligonucleotides.

Validation of RFMP analysis for HCV genotyping

We compared the RFMP assay results with the results from both direct sequencing and the clonal sequencing assay. For confirmation of viral quasispecies or mixed-genotype infection, we cloned the PCR products into the PCR-Script Amp cloning vector (Stratagene) for sequence analysis of each clone (at least 10 clones per sample). The commercialized genotype controls, Armored RNA®-HCV genotype 1b and 2a/c (Ambion Diagnostics), were mixed in various ratios and used to evaluate the sensitivity and specificity of the RFMP assay. Sequence analysis was performed by the ABI PRISM 310 Genetic Analyzer (Applied Biosystems). To clarify the relationship among different HCV isolates, we constructed a phylogenetic tree as described elsewhere *(25)*.

Results

Genotyping strategy and theoretical prediction of database sequences

The RFMP assay is based on mass spectrometric analysis of small DNA fragments containing genotype-specific motifs, as illustrated in Fig. 1. The first step requires PCR amplification using primers flanking the genotype-specific motifs. From the alignment of the 5′ UTR from published sequences, 3 variable regions, H1 to H3, surrounded by conserved regions were chosen as genotype-specific motifs of HCV. Two nested sets of universal primers flanking the variable regions, CV160F/CV100R and CV140F/CV120R, were selected and modified to introduce a type IIS restriction endonuclease recognition sequence ahead of the variable regions on amplification. For interrogating H1 and H2 motifs, CV160F and CV100R primers were designed to introduce the *Mme*I and combined *Mme*I/*Ava*II sites in the amplified product by substituting the restriction recognition sequence TCCAAC for CC (nucleotides −186 to −185) and TCCAACGGTCC for TC (nucleotides −49 to −50), respectively. To interpret the H3 motif, primers CV140F and CV120R were designed to introduce a *Fok*I site (an isoschizomer of *Bst*F5I) in the amplified product by substituting the restriction recognition sequence GGATG for T (nucleotide −147) and G (nucleotide −110), respectively. When the complementary strands are copied, the inserted sequences are incorporated into the product. Type IIS restriction enzymes such as *Mme*I, *Fok*I, and *Bst*F5I cleave DNA outside the recognition sequence: *Mme*I, 20 bases 3′ to the recognition site on one strand and 18 bases from the recognition site on the other strand, leaving a 2-base overhang protruding 3′ end; *Fok*I, 9 bases 3′ to the recognition site on one strand and 13 bases from the recognition site on the other strand, leaving a 4-base overhang protruding 5′ end; *Bst*F5I, 2 bases 3′ to the recognition site on one strand and immediately 3′ to the recognition site on the opposite strand, leaving a 2-base overhang. As outlined in Fig. 1, cleavage of a 142-bp amplified segment with *Mme*I and *Ava*II leads to excision of the 12mer/17mer and 14mer/19mer fragments representing H1 and H3 motifs, respectively, whereas digestion of a 67-bp amplified segment with *Fok*I and *Bst*F5I leads to excision of 2 sets of 7mer/14mer/13mer fragments, reflecting the sequence variation of the H2 motif.

HCV 5′ UTR sequences retrieved from the Entrez Nucleotide database of the National Center for Biotechnology Information were aligned against the HCV prototype 1 sequence (n = 871 as of May 2004; queried by “hcv and 5′ UTR”). From some of the reported sequences covering all 3 variable motifs, we obtained 10 H1, 4 H2, and 28 H3 mass patterns, and a total of 41 combined patterns were assigned to the 6 major HCV genotypes without duplication (see Table 1 of the Data Supplement

that accompanies the online version of this article at http://www.clinchem.org/content/vol51/issue7/).

We found 10 different patterns for samples belonging to type 1, with subtype 1b accounting for the most diverse mass patterns (6 of 10). Genotype 2 could be divided into 8 patterns, with −138A, −132A, and −128T the most distinctive bases among this type. The 2a and 2c subtypes are not completely differentiated, but there is a tendency that TC at nucleotides −117 to −116 is abundant in subtype 2a, whereas CC is abundant in subtype 2c. Most of genotypes 3 and 4 were predicted to show specific patterns in the H1 motif indicative of the existence of −163G and −159T. The patterns for genotype 4 were more complicated and diverse, presenting 12 kinds, among which 4c and 4d could not be differentiated. Genotypes 5 and 6 showed 2 subtype patterns each. From these analyses, 6 major genotypes and 27 subtypes of HCV were predicted to be differentiated into 41 prototype

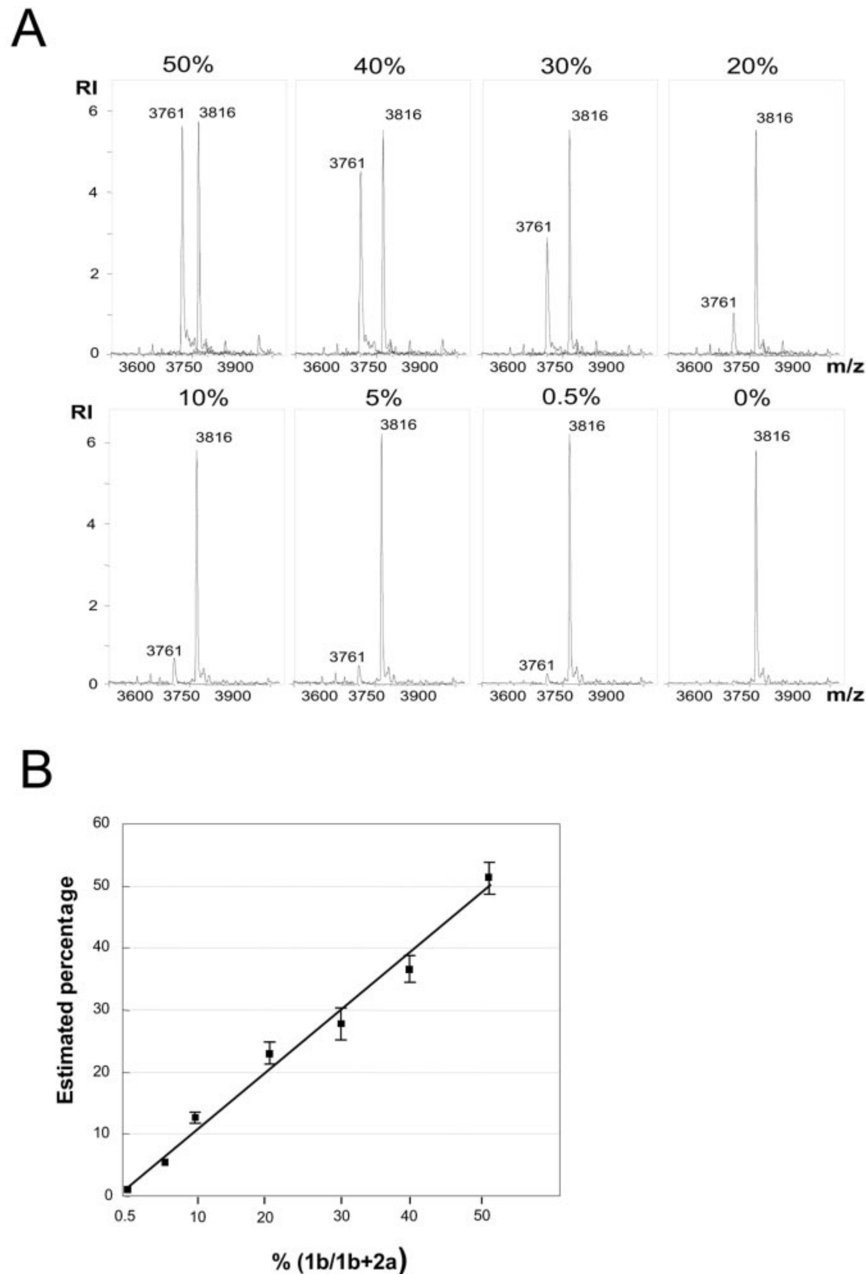

Fig. 2. Evaluation of the sensitivity of the RFMP assay for detection of minor amounts of virus with a defined mixture of genotypes 1b and 2a/c.

(*A*), MALDI-TOF MS spectra resulting from mixed populations of genotypes 1b and 2a/c. Only the mass spectra of the 12mer fragments coming from H1 motifs are shown. Molecular masses of 3761 and 3816 represent 1b- and 2a/c-originated fragments, respectively. Percentages denote [1b/(1b + 2a/c)] × 100. *RI*, relative peak intensity. (*B*), results of test for linearity between the pooling ratios and the corresponding peak-area ratios across artificial pools with different proportions of genotypes ranging from 0.5% to 50%. Five replicates of the artificial pool were genotyped to test the repeatability of the method. *Error bars*, SD. A diagonal line would be expected for complete concordance between measured and real values ($R^2 = 0.992$).

mass patterns as summarized in Table 1 of the online Data Supplement.

DETECTION OF MIXED HCV GENOTYPES

To assess the sensitivity and validity of the RFMP assay for detecting mixed genotypes in a viral population, we performed the assay with samples composed of defined ratios of 5′ UTR sequences of HCV genotype 1b and 2a/c in RNAs. The RNAs were mixed in the following proportions of genotype 1b to the mixture of 2a/c and 1b: 50%, 40%, 30%, 20%, 10%, 5%, 0.5%, and 0%. The RFMP assay detected 0.5% genotype 1b when assessed with 12mer from the H1 motif (Fig. 2A). The DNA sequencing assay could not detect genotype 1b mixed with a 5-fold greater amount of genotype 2a/c, in agreement with previous reports *(26)*.

To test for a relationship between estimated and real proportions, we calculated ratios using the mean peak areas generated from multiple mass spectra and compared them with the defined pooling ratios. We observed a linear correlation of the estimated to the expected proportions between genotypes in the range from 0.5% to 50% (R^2 = 0.992; Fig. 2B). The significant differences obtained in dynamic range did not depend on which spectral band was chosen for quantification among the 10 restriction fragments, except for a pair of 7mers coming from the H3 motif. The 7mers from the H3 motif showed a lower limit of minor genotype detection as low as 0.5% and a dynamic range of 0.5%–50% (R^2 = 0.895). Although the 7mers from the H3 motif seem inferior to other fragments in representing the relative abundance of a mixed-genotype infection, they do not affect the lowest concentrations needed for distinguishing mixed genotypes because the sensitivity was comparable to that for other fragments. The RFMP assay thus could detect mixed genotypes in a viral population and determine their relative abundances.

HCV GENOTYPE DISTRIBUTION IN PATIENTS DETERMINED BY RFMP ANALYSIS

We determined the genotype distribution of the 318 samples from chronically infected patients by RFMP and its representative spectra (Table 1 and Fig. 1 of the online Data Supplement). All of the samples were successfully genotyped by RFMP analysis, and the results agreed with clonal sequence analysis. Consistent with previous reports, genotypes 1b and 2a/c were the most prevalent, constituting 83% of the samples *(27)*. Of interest, 22% of the samples were determined to be infected with mixed genotypes. Similar to single-genotype infections, the 2 predominating genotypes in mixed-genotype infections were 1b and 2a/c, which were found in 77.3% of the samples. Of the mixed-genotype infections, ~30% were determined by direct sequencing to be a single genotype or ambiguous but were further confirmed by clonal sequencing (data not shown). Among samples with mixed-genotype infections, genotype 1b was found in 87.1% (61 of 70) of the individuals, and it was commonly mixed with 2a/c in 42.6% (26 of 61). Of 318 chronic patients with samples analyzed by the RFMP assay, 21 among 38 cases were observed with deviations that did not correspond to the predicted mass pattern from the reference sequences. For example, a motif code pattern (H1/H2/H3) of 2/2/7 was found in 5 cases diagnosed as single genotypes, and in 1 case it was mixed with 2/1/7 (genotype 2a/c). Although the sequences lacked an exact match in Table 1 of the online Data Supplement, they could be inferred by combining the motif sequences corresponding to codes 2, 2, and 7 in motif order, as shown in the first line of genotype 2 in Table 1. The sequences extracted by the above-mentioned approaches fully agreed with the sequencing result. By direct comparison with or construction of a phylogenetic tree linked to the known HCV genotype sequences, the 21 types of mass pattern deviants were assigned to 6 and 15 variants falling into genotypes 1 and 2, respectively (Table 1).

Discussion

Clinical demand for HCV genotyping is increasing because of the significance of HCV genotype as a predictor of treatment response and to enable use of more efficacious therapeutic options. Currently, genotyping is performed largely by sequence analysis and hybridization-based assays *(16–18)*. Sequencing can give information on the majority species present in the viral populations but generally fails to detect species comprising less than 15%–50% of a viral population *(26)*. It is thus necessary to analyze multiple clones to determine the genetic heterogeneity of a population. Assays based on hybridization have been widely used as surrogate genotyping methods, but such assays are labor-intensive, complicate identification of genotype variants or genotype mixtures, and are not suitable for screening large numbers of samples because of the complex protocols involved.

MS generates precise information on the molecular mass of the analyte, both DNA strands can be analyzed in parallel, and the procedure can be fully automated. MS directly assesses the nature of the PCR products, either as a whole or as fragmented oligonucleotides, whereas other technologies only indirectly measure PCR products, either through hybridization or by primer-extended minisequencing reactions, which use PCR products as templates. Genotyping by creating or abolishing recognition sites for restriction enzymes, similar to conventional restriction fragment length polymorphism analysis, has been used in combination with MALDI MS detection *(28)*. Either naturally occurring restriction sites were used or base changes were incorporated in one of the PCR primers to create a recognition site with one of the alleles of the polymorphic site. Only a small number of polymorphisms, however, will alter known restriction sites, and the design of amplification primers to create restriction sites in connection with 1 allele is not straightforward in

Table 1. Genotype distribution, as determined by RFMP analysis, in single- and mixed-genotype infections in 318 patients with chronic HCV infections.

	Motif			Incidence, n (%)	
Genotype	H1	H3[a]	H2	Single-genotype infection	Mixed-genotype infection
1a	GCCAGGACGACCGGGTC	TCAA CCCGCTCAATGCCTGGAG	GCAAGACTGCTAGCCG	10 (4)	5 (3.5)
1a/b	-----------------	---- ------------------	-------C--------		2 (1.4)
1b	-----------------	---- ------------------	--G-------------	102 (41.1)	54 (38.3)
	-------T---------	---- ------------------	--G-------------	4 (1.6)	2 (1.4)
	-----------------	-T-- ------------------	--G-------------	13 (5.2)	5 (3.5)
1c	-----------------	-T-- ------------------	----------------	1 (0.4)	
1	-----------------	---- -------T----------	--G-------------	1 (0.4)	2 (1.4)
	-----------------	---- -------G----------	--G-------------	1 (0.4)	
	-----------------	---- -------------C----	--G-------------	1 (0.4)	
	-----------------	-T-- ------------------	-------C--------	1 (0.4)	
	-----------------	-A-- ------------------	--G-------------	1 (0.4)	
	-----------------	-TC- ------------------	----------------	1 (0.4)	
2a/c	---G---A---T-----	-A-- ---A---T-----C--TC	----------------	26 (10.5)	19 (13.5)
	---G---A---T-----	-A-- ---A---T-----C--CC	----------------	61 (24.6)	29 (20.6)
2b	A--G-A-A---T-----	-A-- ---A---T---T-C--TC	----------------	2 (0.8)	
2d	-------A---T-----	-A-- ---A---T--------TC	----------------	1 (0.4)	5 (3.5)
2e	---G---A---T-----	-A-- ---A---T-----C--TC	-------C--------	1 (0.4)	7 (5.0)
2	---G---A---T-----	-A-- ---A---T-----C--CC	-------C--------	5 (2.0)	1 (0.7)
	---G---A---T-----	-A-- ---A---TG----C--CC	-------C--------	1 (0.4)	
	---G---A---T-----	A--- ---A---TG----C--CC	----------------	1 (0.4)	4 (2.8)
	---G---A---T-----	G---A---A---T-----C--CC	----------------	1 (0.4)	
	---G---A---T-----	A--- ---A---T-----C--CC	----------------	1 (0.4)	1 (0.7)
	-------A---T-----	-A-- ---A---T-----C--CC	-------C--------	1 (0.4)	1 (0.7)
	---G---A---T-----	-A-- ---A---T-----C--TC	--G-------------	1 (0.4)	
	--TG---G---T-----	-A-- ---A---C-----C--CC	----------------	1 (0.4)	
	-------A---T-----	-A-- ---A---T--------TC	-------C--------	1 (0.4)	1 (0.7)
	-------A---T-----	-A-- ---A---T--------CC	-------C--------	1 (0.4)	2 (1.4)
	---G---A---T-----	CA-- ---A---C-----C--CC	----------------	1 (0.4)	
	---G---A---T-----	-A-- ---A---T--------CC	----------------	1 (0.4)	1 (0.7)
	---G---A---T-----	-A-C ---A---C-----C--CC	----------------	1 (0.4)	
	---G---A---T-----	-G-- ---A---T-----C--CC	----------------	1 (0.4)	
	A--G-A-A---T-----	-AC- ---A---T---T-C--TC	-------C--------	1 (0.4)	
3e	--TG--GT---------	A--- ----------A--CA--A	--G---TCA-------	2 (0.8)	
Total				248 (100)	141 (100)
Total no. of individuals				248	70

[a] The 1-base insertion between −136 and −135 in the H3 motif is blanked for easy comparison.

most cases, reducing the usefulness of this approach to very special circumstances.

In this study, we established an RFMP assay for HCV genotyping that exploits differences among HCV genotypes in the molecular masses of oligonucleotides comprising the viral 5′ UTR. The assay is based on amplification and mass detection by MALDI-TOF MS of oligonucleotides excised from type IIS enzyme digestions (Fig. 1). The use of a type IIS restriction enzyme makes this assay independent of the chance occurrence of restriction sites because these enzymes have cleavage sites distal to their recognition sites. Recognition sites are incorporated into the amplification primers, and short fragments that contain the polymorphisms can be generated for mass spectrometric analysis. In addition to its speed and high-throughput capacity, the RFMP assay is very sensitive. It distinguishes between mixed-genotype infections when the minor species constitutes as little as 0.5% of the population (Fig. 2), revealing clear correlations between peak ratios and relative genotype concentrations in mixed populations. For samples from 318 chronically infected patients, the RFMP assay showed excellent concordance with clonal sequencing and detected diverse polymorphisms whose sequences deviated from reported HCV genotypes (Table 1). The significant incidence of mixed-genotype infections was detectable because of the assay's higher sensitivity, leading us to conclude that HCV

mixed-genotype infections are more common than previously estimated.

Taken together with previous findings that mixed infections seem to be associated with unfavorable disease outcomes and that detection of minor genotypes before antiviral therapy may be important for choosing an adequate regimen *(21, 29–31)*, our results indicate that further testing of mixed-genotype infections is needed to expand current treatment algorithms that rely on specific, single genotypes. This research should be facilitated by the high sensitivity of the RFMP assay for detecting mixed infections and the capacity of the RFMP assay to determine relative abundances among viral species. In terms of cost-effectiveness, we estimated the direct cost per test (reagents and labor) of the RFMP assay to be ~$30.00 (US), including viral nucleic acid extraction, PCR, restriction digestion, restriction fragment purification, and matrix, which is much less expensive than the line probe assay at ~$130.00 (US) per test. Assuming that the capital equipment costs for the Biflex IV in our laboratory, estimated at $50 000.00 including annual amortization and maintenance, are similarly applicable, clinical laboratories that expect to test >500 specimens referred for HCV genotyping per year could consider the RFMP assay as an alternative to the line probe assay. In laboratories already using MALDI-TOF MS for clinical genotyping or for clinical proteomics, these costs could be avoided or minimized because the same instrumentation is used for the RFMP assay.

Accurate determination of ion masses is achievable because of high mass resolution and low ppm precision in the peak centroid positions for the raw flight time data *(32, 33)*, which are necessary features for resolving adjacent peaks, particularly with increasing oligonucleotide fragment size. Although the absolute error, on average, was less than +0.5 Da and the width of the mass peak at the half maximum was as low as 1–4 Da for 2- to 6-kDa mass ranges in our experimental setup, concerns could arise relating to distinctions between multiple fragments of extremely similar masses, such as codes 1 and 4 of motif H2, and consequently between genotypes 1a (codes 1/1/1) and 1b (1/4/1) if present in the same specimen. The resolution achieved by the RFMP assays reported here should be achievable with MALDI-TOF MS instruments with specifications and settings comparable to or beyond those used here. A limitation of the RFMP assay, however, is that it determines only molecular mass and, therefore, will fail to detect alterations in sequences that do not lead to a change in molecular mass. To provide certainty in such cases, anticipated from the reported sequences as exemplified in Table 1 of the online Data Supplement for codes 3 and 6 of motif H1 and code 1 of motif H2, we designed the assay so that mass patterns obtained from other motifs would enable exact genotype differentiation. The presence of a novel genotype variant with sequence alterations that do not cause a deviation from the molecular mass pattern predicted from up-to-date HCV databases is possible, but we believe it would be extremely rare, as corroborated by the fact that we found no such case in 318 patients by either RFMP or sequencing assays. Alternative mass spectrometric techniques to consider involve collision-induced dissociation of polyprotonated oligonucleotides produced by electrospray ionization introduced to allow de novo sequencing of DNA oligomers up to 19 bases in length *(34)*. Even longer fragments could be handled if the sequences differ at not more than 2 nucleotide positions from a reference sequence with mass-to-charge values predicted by established fragmentation pathways *(35)*. These techniques could be considered to investigate potential sequence variations leading to no change in molecular mass, although whether they are amenable to high-throughput genotyping of multiple variations for clinical applications is not yet established.

In conclusion, the RFMP genotyping method uses the mass difference of oligonucleotides, requires the simple steps of single PCR amplification and restriction enzyme digestion, and is amenable to high-throughput analyses. The RFMP method has practical advantages over existing methods, including more sensitive and specific detection of mixed populations, as well as primary screening of novel viral genotype variants leading to deviation from the predicted reference mass patterns without need for cloning or multiple PCR steps. This screening method allows reliable viral genotyping in laboratories and efficient analysis of large samples and viral quasispecies to address the dynamics of HCV infections and the relationship between HCV genotypes and clinical outcome. The RFMP genotyping assay is easily adaptable for the detection of other genetic variations, including polymorphisms, deletions, and additions, as well as motifs with multiple variations, because the range and sizes of the resulting restriction fragments analyzed in MALDI-TOF MS can be chosen according to the investigator's needs. We believe, therefore, that the RMFP genotyping assay will be quite useful for developing molecular diagnostics needed to differentiate clusters of multiple variations, such as virus genotyping, and for conducting pharmacogenomic studies of multiple single-nucleotide polymorphisms.

We gratefully acknowledge the valuable assistance of Dr. William Folk at the University of Missouri in editing the manuscript. This work was supported by grants from the Ministry of Commerce, Industry and Energy, Republic of Korea (Project No.10008427) and from Seoul National University Hospital Research Fund (04-2003-053-0).

References

1. Simmonds P. Variability of hepatitis C virus. Hepatology 1995;21: 570–83.

2. Simmonds P, Alberti A, Alter HJ, Bonino F, Bradley DW, Brechot C, et al. A proposed system for the nomenclature of hepatitis C viral genotypes. Hepatology 1994;19:1321–4.

3. Gayowski T, Singh N, Marino IR, Vargas H, Wagener M, Wannstedt C, et al. Hepatitis C virus genotypes in liver transplant recipients: impact on posttransplant recurrence, infections, response to interferon-α therapy and outcome. Transplantation 1997;64: 422–6.

4. Hopf U, Berg T, Konig V, Kuther S, Heuft HG, Lobeck H. Treatment of chronic hepatitis C with interferon α: long-term follow-up and prognostic relevance of HCV genotypes. J Hepatol 1996;24:67–73.

5. Martinot-Peignoux M, Marcellin P, Pouteau M, Castelnau C, Boyer N, Poliquin M, et al. Pretreatment serum hepatitis C virus RNA levels and hepatitis C virus genotype are the main and independent prognostic factors of sustained response to interferon α therapy in chronic hepatitis C. Hepatology 1995;22:1050–6.

6. Mizokami M, Orito E, Gibo Y, Suzuki K, Ohba K, Ohno T, et al. Genotype, serum level of hepatitis C virus RNA and liver histology as predictors of response to interferon-α 2a therapy in Japanese patients with chronic hepatitis C. Liver 1996;16:23–7.

7. Tsubota A, Chayama K, Ikeda K, Yasuji A, Koida I, Saitoh S, et al. Factors predictive of response to interferon-α therapy in hepatitis C virus infection. Hepatology 1994;19:1088–94.

8. Zein NN, Persing DH. Hepatitis C genotypes: current trends and future implications. Mayo Clin Proc 1996;71:458–62.

9. Poynard T, McHutchison J, Goodman Z, Ling MH, Albrecht J. Is an "a la carte" combination interferon alfa-2b plus ribavirin regimen possible for the first line treatment in patients with chronic hepatitis C? The ALGOVIRC Project Group. Hepatology 2000;31: 211–8.

10. McHutchison JG, Gordon SC, Schiff ER, Shiffman ML, Lee WM, Rustgi VK, et al. Interferon alfa-2b alone or in combination with ribavirin as initial treatment for chronic hepatitis C. Hepatitis Interventional Therapy Group. N Engl J Med 1998;339:1485–92.

11. European Association of Study of Liver. International consensus conference on hepatitis C. Paris, 26–28 February 1999, Consensus statement. J Hepatol 1999;30:956–61.

12. NIH Consensus Statement on management of hepatitis C: 2002. NIH Consens State Sci Statements 2002;19:1–46.

13. Nakao T, Enomoto N, Takada N, Takada A, Date T. Typing of hepatitis C virus genomes by restriction fragment length polymorphism. J Gen Virol 1991;72:2105–12.

14. Ohno O, Mizokami M, Wu RR, Saleh MG, Ohba K, Orito E, et al. New hepatitis C virus (HCV) genotyping system that allows for identification of HCV genotypes 1a, 1b, 2a, 2b, 3a, 3b, 4, 5a, and 6a. J Clin Microbiol 1997;35:201–7.

15. White PA, Zhai X, Carter I, Zhao Y, Rawlinson WD. Simplified hepatitis C virus genotyping by heteroduplex mobility analysis. J Clin Microbiol 2000;38:477–82.

16. Schroter M, Zollner B, Schafer P, Landt O, Lass U, Laufs R, et al. Genotyping of hepatitis C virus types 1, 2, 3, and 4 by a one-step LightCycler method using three different pairs of hybridization probes. J Clin Microbiol 2002;40:2046–50.

17. Stuyver L, Wyseur A, van Arnhem W, Hernandez F, Maertens G. Second-generation line probe assay for hepatitis C virus genotyping. J Clin Microbiol 1996;34:2259–66.

18. Bullock GC, Bruns DE, Haverstick DM. Hepatitis C genotype determination by melting curve analysis with a single set of fluorescence resonance energy transfer probes. Clin Chem 2002; 48:2147–54.

19. Lau JY, Mizokami M, Kolberg JA, Davis GL, Prescott LE, Ohno T, et al. Application of six hepatitis C virus genotyping systems to sera from chronic hepatitis C patients in the United States. J Infect Dis 1995;171:281–9.

20. Forns X, Maluenda MD, Lopez-Labrador FX, Ampurdanes S, Olmedo E, Costa J, et al. Comparative study of three methods for genotyping hepatitis C virus strains in samples from Spanish patients. J Clin Microbiol 1996;34:2516–21.

21. Eyster ME, Sherman KE, Goedert JJ, Katsoulidou A, Hatzakis A. Prevalence and changes in hepatitis C virus genotypes among multitransfused persons with hemophilia. The Multicenter Hemophilia Cohort Study. J Infect Dis 1999;179:1062–9.

22. Hong SP, Kim NK, Hwang SG, Chung HJ, Kim S, Han JH, et al. Detection of hepatitis B virus YMDD variants using mass spectrometric analysis of oligonucleotide fragments. J Hepatol 2004;40: 837–44.

23. Young KK, Resnick RM, Myers TW. Detection of hepatitis C virus RNA by a combined reverse transcription-polymerase chain reaction assay. J Clin Microbiol 1993;31:882–6.

24. Choo QL, Richman KH, Han JH, Berger K, Lee C, Dong C, et al. Genetic organization and diversity of the hepatitis C virus. Proc Natl Acad Sci U S A 1991;88:2451–5.

25. Mizokami M, Gojobori T, Lau JY. Molecular evolutionary virology: its application to hepatitis C virus. Gastroenterology 1994;107: 1181–2.

26. Hu YW, Balaskas E, Kessler G, Issid C, Scully LJ, Murphy DG, et al. Primer specific and mispair extension analysis (PSMEA) as a simple approach to fast genotyping. Nucleic Acids Res 1998;26: 5013–5.

27. Kim YS, Ahn YO, Lee HS. Risk factors for hepatitis C virus infection among Koreans according to the hepatitis C virus genotype. J Korean Med Sci 2002;17:187–92.

28. Liu YH, Bai J, Zhu Y, Liang X, Siemieniak D, Venta PJ, et al. Rapid screening of genetic polymorphisms using buccal cell DNA with detection by matrix-assisted laser desorption/ionization mass spectrometry. Rapid Commun Mass Spectrom 1995;9:735–43.

29. Schroter M, Feucht HH, Zollner B, Schafer P, Laufs R. Multiple infections with different HCV genotypes: prevalence and clinical impact. J Clin Virol 2003;27:200–4.

30. Hu YW, Balaskas E, Furione M, Yen PH, Kessler G, Scalia V, et al. Comparison and application of a novel genotyping method, semiautomated primer-specific and mispair extension analysis, and four other genotyping assays for detection of hepatitis C virus mixed-genotype infections. J Clin Microbiol 2000;38:2807–13.

31. Tuveri R, Rothschild C, Pol S, Reijasse D, Persico T, Gazengel C, et al. Hepatitis C virus genotypes in French haemophiliacs: kinetics and reappraisal of mixed infections. J Med Virol 1997; 51:36–41.

32. Loboda AV, Krutchinsky AN, Bromirski M, Ens W, Standing KG. A tandem quadrupole/time-of-flight mass spectrometer with a matrix-assisted laser desorption/ionization source: design and performance. Rapid Commun Mass Spectrom 2000;14:1047–57.

33. Moskovets E, Karger BL. Mass calibration of a matrix-assisted laser desorption/ionization time-of-flight mass spectrometer including the rise time of the delayed extraction pulse. Rapid Commun Mass Spectrom 2003;17:229–37.

34. Ni J, Mathews MA, McCloskey JA. Collision-induced dissociation of polyprotonated oligonucleotides produced by electrospray ionization. Rapid Commun Mass Spectrom 1997;11:535–40.

35. Oberacher H, Mayr BM, Huber CG. Automated de novo sequencing of nucleic acids by liquid chromatography-tandem mass spectrometry. J Am Soc Mass Spectrom 2004;15:32–42.