

# Robust Hepatitis B Virus Genotyping by Mass Spectrometry<sup>∇</sup>

L. Ganova-Raeva,<sup>1\*</sup> S. Ramachandran,<sup>1</sup> C. Honisch,<sup>2</sup> J. C. Forbi,<sup>1</sup> X. Zhai,<sup>3</sup> and Y. Khudyakov<sup>1</sup>

*Division of Viral Hepatitis, Centers for Disease Control and Prevention, Atlanta, Georgia<sup>1</sup>; Sequenom, Inc., San Diego, California<sup>2</sup>; and Jiangsu Centers for Disease Control, Jiangsu, China<sup>3</sup>*

Received 22 April 2010/Returned for modification 13 August 2010/Accepted 25 August 2010

**Genotyping of hepatitis B virus (HBV) is important for tracking HBV infections, prognosticating the development of severe liver disease, and predicting outcomes of therapy. Current genotyping methods can be laborious and costly and rely on subjective data interpretation. To identify less expensive but equally reliable alternatives, we compared “gold standard” sequencing to a novel mass spectrometry approach. Sera from individuals with acute or chronic HBV infection ( $n = 756$ ), representing all genotypes, were used to PCR amplify the HBV S gene. All amplicons were subjected to base-specific cleavage and matrix-assisted laser desorption ionization–time-of-flight mass spectrometry (MALDI-TOF MS). The resulting mass peak patterns were used to identify HBV genotype by automated comparison to peak patterns simulated from reference sets of HBV sequences of known genotypes. The MALDI-TOF MS data and phylogenetic analysis of HBV sequences produced completely concordant results. Several parameters such as genetic relatedness of tested HBV variants to the reference set, chronic infections, and the quality of PCR products can lower the MS score but never affected the accuracy of the genotype call. This new streamlined MS-based method provides for rapid and accurate HBV genotyping, produces automated data reports, and is therefore suitable for routine use in diagnostic settings.**

Hepatitis B virus (HBV) infection is a major cause of morbidity and mortality across the world, 350 million people are currently estimated to be chronically infected with the virus, and 660,000 deaths have been reported from HBV-related liver disease and cancer. HBV is genetically diverse and has been classified into 8 genotypes and 24 subtypes (4, 25). HBV genotyping is an important molecular epidemiological tool for tracking HBV infections. In clinical practice, the HBV genotype has been associated with severity and progression of chronic hepatitis B, the response to anti-HBV therapy, and the risk of liver cancer (13, 19).

Currently, the most accurate approach to HBV genotyping is based on direct dideoxy sequencing of DNA amplified from discrete regions of the HBV genome, followed by phylogenetic analysis of these sequences (10, 16, 29). This approach is considered the “gold standard,” but it is laborious and thus not suitable for use in routine public health or clinical practice. Restriction fragment length polymorphism (RFLP) analysis (11, 27) significantly simplified genotype detection; however, this approach has limited sensitivity. The Trugene method, a standardized and largely automated polymerase gene-sequencing assay, produces only up to 86% of interpretable data (3, 24, 33). Multiplex PCR requires the use of many sets of primers (10, 23), so its efficacy is reduced should the target templates have highly variable sequences. Real-time PCR, although applicable for genotyping, is better suited for viral load quantification (9). The DNA chip (30–32), based on a hybridization technology, efficiently detects HBV genotypes but requires the test sample to carry a viral load greater than 1,000 IU/ml, and

its sensitivity is diminished to ~90% for certain genotypes. It has low throughput and is costly, which limits its wider application (32). The InnoLipa assay, based on reverse hybridization, is one of the more widely used technologies in clinical laboratories (17, 21, 26). It is accurate and capable of discriminating genotypes in mixed-HBV infections; however, it is relatively costly and not open-ended to detect emerging variants.

The MassARRAY system based on nucleic acid analysis by matrix-assisted laser desorption ionization time-of-flight mass spectrometry (MALDI-TOF MS) provides an alternative approach to HBV genotyping. MALDI-TOF MS has the capacity to discriminate single-point mutations. Specific primer extension into the sites of interest has successfully been applied to detect up to 60 single-nucleotide drug resistance mutations in the HBV polymerase gene (16) and to genotype HBV (12, 14). The MassARRAY system of MALDI-TOF MS is capable of detecting wild-type and mutant alleles and can identify mixes if the minority type is present at >10% (2). This technology is less costly and easy to use as it is amenable to automation and has already found application in other disease evaluations, bacterial typing, and cancer marker detection. In this study, we report the development and assessment of a novel approach to HBV genotyping based on the MassARRAY system.

## MATERIALS AND METHODS

**Test specimens.** Serum samples from HBV PCR-positive patients with acute HBV infection collected between 1999 and 2006 in the U.S. ( $n = 705$ ) and chronic hepatitis B collected during 2007 in Central Africa ( $n = 51$ ) and referred to Centers for Disease Control and Prevention in Atlanta, GA, were used. All were also genotyped using DNA dideoxy sequencing followed by phylogenetic analysis (10). The ClustalW algorithm (Lasergene v.8; DNASTAR, Madison, WI) was used for the phylogenetic analysis and to derive the sequence distance data. Normal human sera ( $n = 12$ ) were used as HBV PCR-negative controls.

**PCR, post-PCR processing, and MALDI-TOF MS protocol.** Total nucleic acid was extracted from 100  $\mu$ l of each serum and recovered in 50  $\mu$ l of elution buffer (MagnaPureLC; Roche, Inc., Indianapolis, IN). Ten microliters was used for

\* Corresponding author. Mailing address: Centers for Disease Control and Prevention, 1600 Clifton Rd. NE, MS A-33, Atlanta, GA 30329. Phone: (404) 639-1158. Fax: (404) 639-1563. E-mail: lkg7@cdc.gov.

<sup>∇</sup> Published ahead of print on 1 September 2010.

nested PCR with HBV-specific primers for the amplification of a 441-bp fragment encompassing the *a*-determinant of the HBV S gene. The thermal conditions for the first-round amplification were initial denaturation at 95°C for 10 min, followed by 35 cycles of 95°C for 30 s, 48°C for 1 min, and 72°C for 1 min, and those for the second-round amplification were initial denaturation at 95°C for 10 min, followed by 30 cycles of 95°C for 30 s, 55°C for 1 min, and 72°C for 1 min. The real-time PCR was conducted by using Fast Start SybrGreen (Roche) on a Mx3005P real-time instrument (Stratagene, Cedar Creek, TX). The first-round PCR primers were 179F (5'-CTA GGA CCC CTC TC GTG TT) and 704R (5'-CAG TCG AAC CAC TGA ACA AAT GGC ACT), and the second-round PCR primers were 217F (5'-CGA TTT AGG TGA CAC TAT AGA AGA GAG GCT GTT GAC AAG AAT CCT CAC AAT ACC) and 658R (5'-CAG TAA TAC GAC TCA CTA TAG GGA GAA GGC GGC TGA GGC CCA CTC CCA TA). The primers listed above contained SP6 (underlined) or T7 (italicized) RNA polymerase promoter and enhancer sequences, as per the MassCLEAVE biochemistry requirements (MassCLEAVE kit—T7/SP6 kit; Sequenom, Inc., San Diego, CA) (6, 28). Application of this protocol typically achieves a sensitivity of 64 genome copies HBV/ml or 11 IU/ml. The viral titer was determined from the real-time cycle threshold ( $C_T$ ) by using a standard curve built by serial dilutions of a positive-control standard of known titer (WHO HBV International Standard [NIBSC code 97/750]). All PCR products were divided into aliquots and stored at -20°C for later analysis. Prior to MALDI-TOF MS analysis, each PCR product was digested with shrimp alkaline phosphatase and split into four reaction wells, each to be subjected to a different *in vitro* transcription-coupled base-specific RNase A cleavage reaction (28). After the cleavage, the products were desalted by adding deionized water and 6 mg clean resin (Sequenom, Inc., San Diego, CA) to each well. Fifteen nanoliters of cleaned product was applied by a Nanodispenser (Sequenom, Inc., San Diego, CA) onto a 384-spot SpectroCHIP (Sequenom, Inc.). Data acquisition was done on a MALDI linear TOF mass spectrometer (MassARRAY Compact Analyzer, Sequenom, Inc., San Diego, CA). All reactions and cleanup steps were automated on a Biomek 3000 robotic station (Beckman-Coulter, Fullerton, CA).

**Reference data set.** Unique S gene sequences ( $n = 153$ ), representing all HBV genotypes, were compiled from available GenBank entries (<http://www.ncbi.nlm.nih.gov/GenBank/index.html>). All of these sequences were annotated to indicate their genotype (6) and used to compile a reference set referred to as the “large reference set.” Their accession numbers are as follows: ay0904543, ay0904573, ay090460, ay090459, ay090461, ay090458, af223963, x69798, ay090455, af223965, af223962, ab036905, ab036913, ab036910, x75663, af241410, af241411, af068756, af223960, ay217376, ay217378, ay217374, x75665, x75656, af411411, d23680, af461357, d50517, m38636/ab050018, af286594, ay066028, af461361, af458664, d16665, d23684, ef690484, ab033553, x04615, af533983, af458665, ab033557, ay057947, ab014378, d23682, ay217373, s75184, m54923, d00331, d23679, d00329, ab010290, ab010291, d23678, ab073854, ab010292, ab073855, af282918, ab073822, ab073826, ab073821, ab073831, ab073824, ab073830, ab073832, ab073836, ab073839, x97850, ab073823, ab073835, ab032431, x75657, ab091255, ab091255, ab091256, x75664, ab033559, ab048701, aj131956, u95551, ab033558, ay233292, ay233294, ay233293/ay233295, ay233291, x65257, ay233296, x85254, x97848, af121240/af1212423, af121239, x02496, af151735, af280817, x65259, x80924, ay057948, aj344116, m32138, af043593, af043594, z35716, x80925, x72702, ay090452, x68292, ab056514, ab064313, ab056515, aj309371, aj309370, aj309369, z35717, x70185, af090839, aj012207, ay128092, af090841, ab014370, af537371, ab064314, z72478, ay233280, af090838, af536524, ay034878, x51970, ay233286, ab076678, ay233288, ay233275, ay233281, ay233284, ay233289, af297623, af090842, m57663, ay233278, ay233276, ay233279, af297621, ay233283, ay233285, ay233277, ay233274, af297625, ay233282, ay233287, af537372, ay233290, ab076679, aj344115. The underlined sequences represent all eight genotypes and some major subtypes; they were used to generate a reference subset referred to as the “minimal reference set” ( $n = 15$ ). Throughout the data analysis, the large reference set was further enriched with new sequences found by the MS pattern analysis software, iSEQ v1 (Sequenom, Inc., San Diego, CA) to generate another reference set, called the “enriched reference set” ( $n = 220$ ). The last reference set, called the “irrelevant reference set” ( $n = 240$ ), was generated with the core gene of HBV (between nucleotide (nt) positions 1852 and 2241 of the genome and of the same length as the S gene amplicon). Those were retrieved from GenBank and represented all genotypes and subtypes.

**Computational analysis.** A preliminary simulation was performed (iSEQ Simulation tool v0.1.2.32; Sequenom, Inc.) to assess *in silico* whether all references in the set can be discriminated by base-specific cleavage and MALDI-TOF MS and to identify the minimal number of reactions required to discriminate all sequences in a given reference set (2). This assessment was used to establish suitability of the reference set for HBV genotyping (28). The tool generates a maximum of four reference patterns per sequence. Two base-specific cleavage

patterns for the cleavage at cytosine (C) and for the cleavage at uracil (T) of the *in vitro*-transcribed RNA plus strand (forward strand; F) were produced. In addition, two corresponding base-specific cleavage patterns, for the cleavage at C and at T for the minus strand (reverse strand; R) were obtained. Distance calculations between peak patterns of different reference sequences allow for peak pattern-based clustering. Values for discriminating power are calculated by summing up intensities of the discriminating features, e.g., peaks present in one reference sequence but absent in the other. Resulting sums are divided into four categories representing the discriminatory power of the features. High (>8 discriminatory features per reaction)-, medium (4 to 8 features)-, and low (2 to 4 features)-power scores were accepted for genotype discrimination, whereas the very-low-power scores (0 to 2 features) were considered insufficient to distinguish one simulated reference pattern from another.

The iSEQ software v 1.0 (Sequenom, Inc., San Diego, CA) was also used for the automatic report generation of the genotype data and extraction of the corresponding nucleotide sequences. The software compares the mass fingerprints acquired by MALDI-TOF MS to mass fingerprints of all reference sequences generated *in silico* and assigns a matching score. The score (0 to 1) is a qualitative measure that reflects how close the detected and unknown mass pattern is in relation to the closest identified match in the reference database; where 0 indicates no similarity and 1 indicates a perfect match. Our results demonstrated that matching score values greater than 0.6 are sufficient for accurate genotype determination.

## RESULTS

**Simulation experiments.** To examine if the MassARRAY system patterns can discriminate accurately between the different HBV genotypes, we ran simulation experiments with 10 pilot reference sets. Each set was comprised of eight unique sequences, one from each genotype, randomly selected from the large reference set. The divergence between nucleotide sequences in each pilot reference set ranged from 3.2% to 8.9%. These simulations are summarized in Fig. 1 and show that the peak patterns obtained using all four base-specific reactions (TF, TR, CF, and CR) allow for a complete differentiation of all HBV genotypes. Combinations of any two or three mononucleotide cleavage reactions also provided reliable discrimination for all of the samples. If only single reactions were considered, 2.5% of samples were genotyped with a very low power. However, the power value was >0 in all cases.

To examine the utility of the MS patterns to distinguish correctly between subtypes, simulations were also conducted using reference subsets including only unique sequences from a single genotype (for genotype A,  $n = 43$ ; for B,  $n = 23$ ; for C,  $n = 32$ ; for D,  $n = 31$ , for E,  $n = 5$ ; for F,  $n = 13$ ; for G,  $n = 3$ ; and for H,  $n = 3$ ) and covering all possible combinations of any two genotypes. The percentages of divergence between and within all genotypes are shown in Table 1. All low-power discrimination instances were found between sequences of the same genotype with percentages of divergence of the underlying sequences ranging from 0.1% to 2%. Low-power discrimination values were never observed between genotypes. The contribution of the individual cleavage reactions to genotype discrimination did not vary significantly ( $P = 0.4329$ ; analysis of variance [ANOVA]).

The simulation data showed that reliable genotype identification could be achieved with all four reactions. Single successful reactions also produced accurate genotype calls with power >0. This observation indicates redundancy in the MS data and therefore lends confidence to genotype calling based on all four reactions.

**Experimental validation of simulations.** A pilot set of 12 HBV PCR-negative controls and 84 randomly chosen PCR-

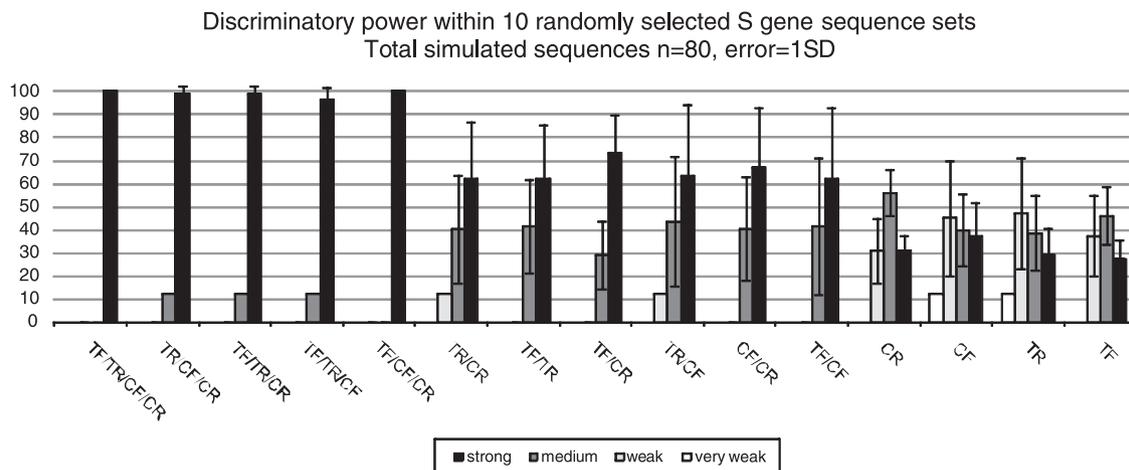


FIG. 1. Simulation results of 10 randomly selected S gene reference sets, each representing all HBV genotypes. The x axis displays the groups of all four reactions (TF, TR, CF, and CR, where “F” is “forward” and “R” is “reverse”) and any combination of three cleavage reactions, two reactions, or one reaction. The y axis displays the average percentage of sequences falling into the individual power of discrimination categories (very weak, no fill; weak, light gray; medium, dark gray; and strong, black). For example, if all four reactions were used, 100% of the data are discriminated with strong power, if only the CF reaction was used, approximately 38% were discriminated with strong power, etc. A threshold of 2 (the equivalent of two full peak changes) is required to distinguish one sequence from another with good confidence.

positive specimens were tested using MS patterns. HBV strains from all positive specimens were genotyped by phylogenetic analysis of their S gene amplicon sequences. Genotyping was conducted using the large reference set of unique S gene sequences. All reported genotypes were correctly identified by using the default cumulative score of all four reactions. All products from negative specimens resulted in scores <0.400 (mean of 0.1413 and standard deviation [SD] of 0.104), based on all four reactions. To confirm that the genotype analysis was sequence specific, we analyzed the spectra from the same chip using the irrelevant reference set consisting of HBV core gene sequences. The scores of all specimens against the irrelevant reference set were <0.4 (mean, 0.145; SD, 0.099) for all specimens. All amplicons tested against the S gene large reference set resulted in cumulative scores of >0.6 (mean, 0.8197; SD, 0.097) using all four reactions. The difference between the mean scores from negative controls and PCR-positive tests was significant ( $P = 0.0000012$ ). Based on these results, the value 0.6 was selected as the cutoff to assign genotype.

The data were also analyzed using the 332 single-nucleotide reactions from positive specimens (84 specimens with each specimen undergoing 4 reactions) separately against the large

reference set. We found that each single reaction also produced the correct genotype call. Nineteen (4.7%) of the single-nucleotide reactions were assigned failing scores by the software. These reactions were not of sufficient quality or were altogether absent. Twelve of the specimens had one failed reaction out of the four, two had two failed reactions, and one had three failed reactions. All of these 15 specimens were analyzed as a separate group (designated “pos\*”). There was no significant difference in their mean cumulative score when compared to the scores for the rest of the positive specimens. All specimens containing failed reactions were genotyped correctly regardless of the number of failed reactions, and their relevant scores are shown in Fig. 2.

**Accuracy of HBV amplicon genotyping by MALDI-TOF MS.** A large set of amplicons ( $n = 756$ ), obtained from specimens derived from American and African patients and previously genotyped by standard dideoxy termination sequencing, were genotyped using MS patterns. All PCR-positive specimens had matching scores above the cutoff. The concordance between the genotype calls by MS and phylogenetic analysis of sequences was 100%. Most of the specimens belonged to genotype A ( $n = 538$ ), followed by D ( $n = 119$ ), E ( $n = 59$ ), C

TABLE 1. Inter- and intragenotype divergence among the 153 sequences of the large reference set<sup>a</sup>

Genotype	% nucleotide divergence <sup>a</sup>							
	A	B	C	D	E	F	G	H
A	0.2–5.0	3.7–8.1	3.8–8.8	3.7–7.5	4.0–7.5	5.7–9.4	2.8–5.3	5.4–8.6
B		0.2–3.5	4.2–8.1	4.2–7.2	4.3–6.3	4.9–7.7	5.9–7.2	5.0–6.6
C			0.2–4.3	4.4–7.9	4.7–6.8	4.7–7.2	4.2–7.4	5.0–7.7
D				0.2–3.3	3.2–5.4	4.5–7.4	4.2–5.6	4.9–6.4
E					0.2–0.7	4.9–6.5	4.3–4.9	5.2–5.7
F						0.2–2.8	5.4–6.8	2.3–3.5
G							0.2–0.8	6.1–6.4
H								0.2–0.4

<sup>a</sup> The numbers represent the percentages of nucleotide divergence between and within the genotypes based on a 441-bp fragment of the HBV S gene.

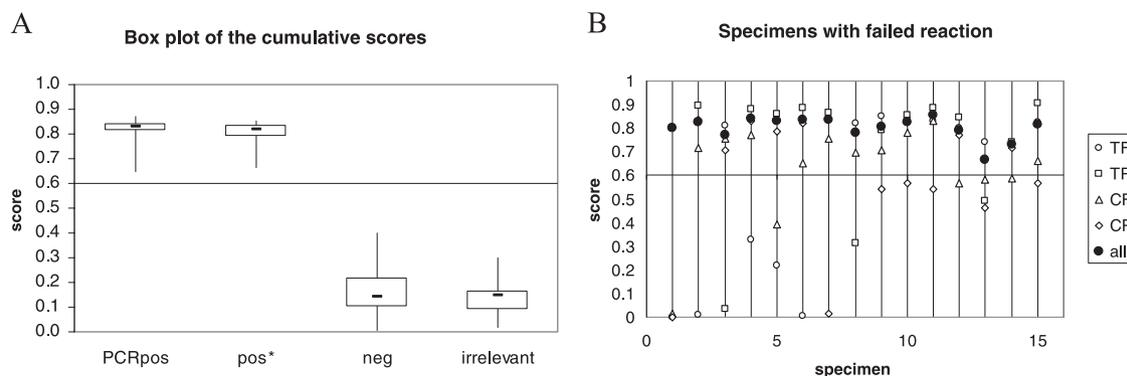


FIG. 2. Pilot experiment data. (A) Box plots of the cumulative scores from all four reactions. The groups are PCR-positive specimens with four successful reactions (PCRpos;  $n = 69$ ), positives with at least one failed reaction (pos\*;  $n = 15$ ), negative controls (neg;  $n = 12$ ), and an irrelevant group consisting of HBV core gene sequences (irrelevant;  $n = 240$ ). The ordinate displays the score values. The box represents the first through the third quartile of the data, the median is denoted with a “–” sign, and the whiskers spread to the minimum and the maximum values. A dotted line is drawn at the genotype cutoff value of 0.6. (B) Score data from the 15 specimens of the pilot experiment that had at least one reaction failure (pos\*). The scores were calculated after analysis of all four mononucleotide reactions, TF, TR, CF, and CR separately (open shapes). The black circles are the cumulative scores acquired using all four reactions and the same reference set. Each vertical gridline contains all five data scores (TF, TR, CF, CR, and cumulative) from one specimen. Dotted line denotes the genotype cutoff value of 0.6. All values below the cutoff represent individual failed nucleotide reactions.

( $n = 21$ ), F ( $n = 8$ ), G ( $n = 5$ ), B ( $n = 3$ ), and H ( $n = 3$ ). The variability in the matching scores, which ranged from 0.601 to 0.978 (mean, 0.820; SD, 0.097) led to further analysis of the factors affecting the score.

**Effect of viral titer.** The sample set had a wide range of viral titers (from  $1E+1$  to  $1E+14$  genome copies/ml) and scores, and this is shown in Fig. 3. The calculated correlation (0.013) between the viral titer and the matching score for all collected data from PCR-positive amplicons ( $n = 756$ ), including tests run in duplicates ( $n = 365$ ; total, 1,121), indicated that the viral titer did not affect the quality of genotype calling.

**Effect of the quality of individual base-specific cleavage reactions.** To evaluate the contribution of the individual reac-

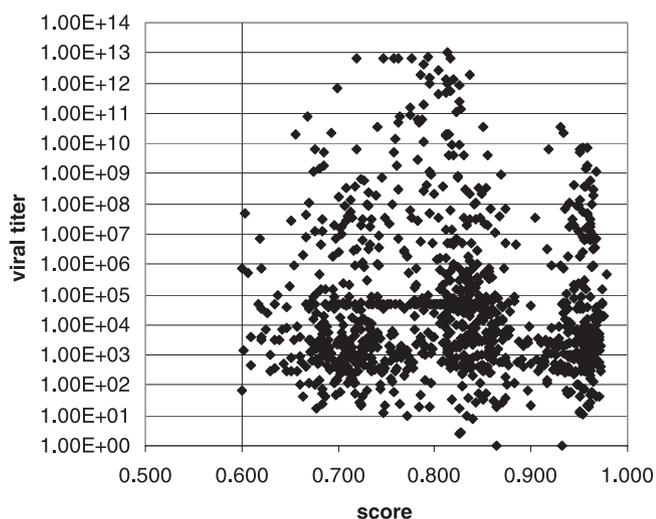


FIG. 3. Matching score values and corresponding viral titers. Shown is a scatter plot of all MS data from all PCR-positive specimens ( $n = 1,121$ ), including duplicates. The scores indicate the matching of a particular specimen to the closest member of the reference set ( $n = 153$ ). The dotted line denotes the genotype cutoff score (0.6).

tions to the accuracy of the genotype identification, the rates of reaction failure and incorrect identification were analyzed. We found that 8.3% of the single reactions failed processing (i.e., had scores below the genotype call cutoff of 0.6). All positive single reactions provided 100% correct genotype calls, when utilized individually for sample identification. At the subtype level, a small call error rate of 4.3% was found to be consistent with the simulated low-power discrimination within the boundaries of any one genotype. All data were separated by groups of tests that had one (17.7%), two (4.6%), or three (2.0%) failed reactions or no failed reactions (75.7%) and reanalyzed. Figure 4A shows box plots comparing the data sets with different reaction success rates. Only the group with three failed reactions revealed a significant decrease in the mean score ( $P = 0.0356$ ). Nevertheless, this decrease in the score did not affect the accuracy of the genotype call.

**Effect of amplicon degradation.** Some of the PCR-positive specimens could not be typed with a reproducible score of  $>0.6$  in duplicate experiments when the PCR products were stored longer than 2 months at  $-20^{\circ}\text{C}$  (Fig. 4B). The average scores decreased by 27.2% from 0.938 to 0.683, which represents a significant difference ( $P < 0.000001$ ). We attribute this decline in the score to amplicon degradation during the prolonged storage period before assay. The cases that were not typed by MS after the prolonged storage were found to have viral loads of  $<51,000$  genome copies/ml.

**Effect of the complexity of the reference set.** To test how the content of the reference set affects the score of the genotype calls, the MS output was analyzed with three different reference sets (see Materials and Methods). The minimal reference set ( $n = 15$ ) had a genetic divergence of 0.9% to 7.2%, the large reference set ( $n = 153$ ) had genetic divergence of 0.2% to 11.5%, and the enriched reference set ( $n = 220$ ) had genetic divergence of 0.2% to 11.7%. Figure 5 shows the effect of these reference sets on the scores. No significant difference was observed between scores obtained when using the large and the enriched sets. Their matching scores had means of 0.817 (SD,

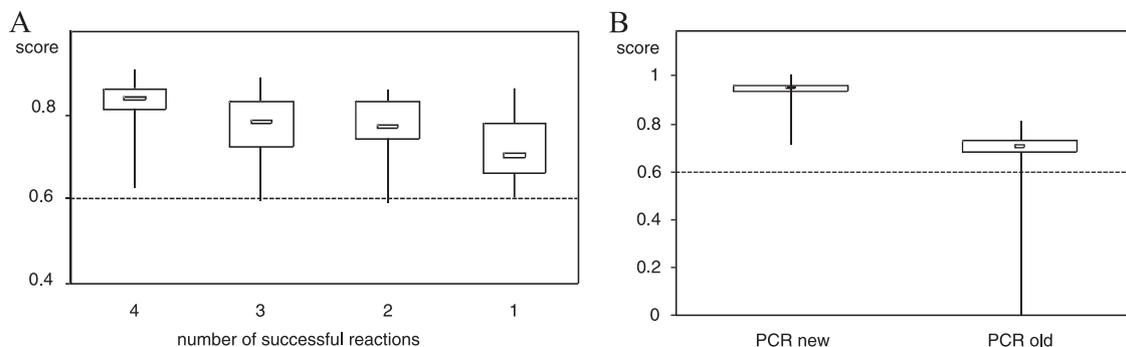


FIG. 4. Box plots showing the effect of PCR quality and the combination of cleavage reactions on score values. The dotted line is placed at the genotype cutoff score (0.6). Score values are on the ordinates. (A) Post-PCR product quality. Scores are grouped as follows by the number of successful reactions per sample: samples with all four successful reactions ( $n = 591$ ), samples with three successful reactions ( $n = 139$ ), samples with two successful reactions ( $n = 36$ ), and samples with one successful reaction ( $n = 16$ ). (B) PCR quality. Shown is a comparison between the scores of fresh PCR products and scores of PCR products stored over 2 months.

0.059) and 0.826 (SD, 0.049), respectively. The matching scores against the minimal reference set had a mean of 0.799 (SD, 0.079). The difference between this mean and the mean of the two larger reference sets is significant ( $P = 0.034$ ). The observed greater spread of the data shows that a lack of close sequences in the minimal reference set is reflected in the score. While the mean score did not undergo significant change between the larger sets, the minimal reference set missed 11 of the PCR-positive specimens, as their scores fell below the genotype cutoff. Eight of the 11 cases, which resulted in the lowered scores, were obtained from amplicons of HBV variants that had identity to the genetically closest sequence from the minimal reference set, ranging from 84.9% to 98.8%.

**Effect of genotype.** To determine whether the genomic context of the genotype itself has any effect on the score, the data for each genotype were analyzed separately. No significant differences in the score means were found for genotypes A, B, C, D, F, and H. Genotypes E and G were, however, associated with lower scores (Fig. 6A). Two confounding factors were noted to affect those scores. First, genotype G sequences from

two specimens contained small insertions differentiating them from the reference sequences, thus resulting in a lower matching score. Because of the small overall number of specimens carrying genotype G HBV, these lower scores skewed the data spread. The second factor was found when examining the genotype E entries. Most of them (86.4%) were collected from chronically infected patients residing in Central Africa, while specimens from all other genotypes were collected from acutely infected U.S. patients. Owing to a longer duration of infection, HBV strains from chronically infected patients have greater opportunity to accumulate mutations resulting in substantial intrahost genetic heterogeneity compared to HBV strains recovered from patients with recently acquired infection. To evaluate this potential factor, we proceeded to analyze the effect of complexity of the intrahost HBV population on the score.

**Effect of intrahost HBV complexity.** The mean score for the acute infection specimens ( $n = 705$ ) was 0.823 (SD, 0.1); however, the chronic infection set ( $n = 51$ ) had a mean score of 0.702 (SD, 0.044). All chronic case specimens were tested in duplicate experiments and yielded completely concordant genotype and score results. Box plots of the data spread are shown in Fig. 6. This statistically significant difference ( $P < 0.00001$ ) prompted further analysis of factors influencing the quality of the score. Using the inherent feature of iSEQ to extract novel sequences, all genotype E sequences were inferred from the MS data and applied to enrich the existing reference set. This addition of sequences that represent a closer match to the MS profiles increased the mean score of the chronic infection set to 0.785 (SD, 0.045); this change was significant ( $P = 0.0383$ ). These results are shown in Fig. 6A and B, and they show that enriching the reference set improved the scoring. However, the enriched reference data set did not provide perfect matches, and the mean remained lower than for HBV genotypes identified in incident cases. Analysis of HBV strains from chronic infection specimens showed the presence of additional genotype G variants in three specimens, genotype D in two specimens, and more than one HBV genotype E strain in the other specimens (data not shown), suggesting that the genetic complexity of samples also affects the score of HBV genotype calls.

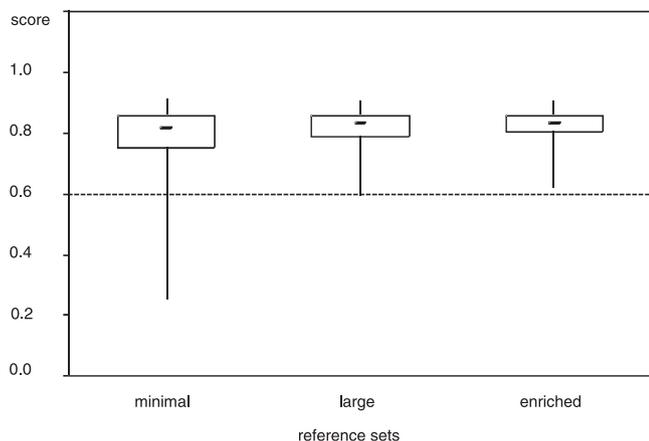


FIG. 5. Effects of the complexity of the reference set on the score. Box plot of the MS data from all cases analyzed ( $n = 756$ ) using a minimal reference set ( $n = 15$ ), the large reference database ( $n = 153$ ), and an enhanced reference set that is enriched with the novel sequences generated automatically during the iSEQ analysis ( $n = 220$ ).

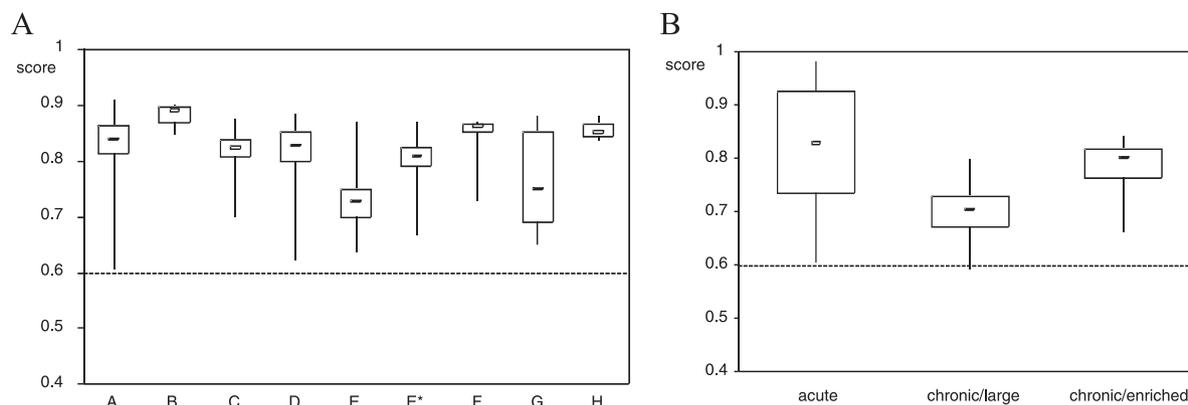


FIG. 6. Effects of the genotype and viral complexity on the score. (A) Box plot of the score data grouped by genotype. Data from PCR-positive specimens ( $n = 756$ ) were analyzed using a large reference set ( $n = 153$ ). The genotypes are indicated on the abscissa. The dotted line is placed at the genotype call cutoff score (0.6). E\* indicates improved scoring of genotype E specimens after analysis with the enriched reference set ( $n = 220$ ). (B) Box plot of the scores of the acute specimens in comparison to the scores from the chronic specimens analyzed against two different reference sets, large ( $n = 153$ ) and enriched ( $n = 220$ ).

## DISCUSSION

The new HBV genotyping approach described here takes advantage of the MassArray technology, which is based on MALDI-TOF MS (28). Using four nucleotide-specific cleavage reactions of RNA molecules transcribed *in vitro* from HBV PCR products, the MassArray system generates MS patterns that are automatically compared to simulated MS patterns of sequences derived from known HBV genotypes. The evaluation determined this assay method to be accurate in identifying all eight HBV genotypes. The data showed complete concordance with results obtained by using phylogenetic analysis of PCR-amplified HBV S gene sequences. Each single-nucleotide reaction provided accurate genotype matching to the reference set, thus, with the added redundancy of three more reactions, contributing to reliability of the assay. Repeated testing of same specimens resulted in accurate genotype detection with a similar score, indicating that the assay is reproducible.

The assay is cost-effective because of its relatively high throughput, and it is approximately half the price of sequencing and is 15 times less expensive than InnoLipa. The assay is rapid: in our laboratory, the technology allows for testing 960 specimens per day without requiring any additional analysis and data interpretation by the operator after the data collection is completed. This technology platform is very versatile; it has been used for bacterial typing and is finding application for cancer single nucleotide polymorphism (SNP) screening. The iSEQ approach has been developed as a resequencing tool capable of detecting single-nucleotide polymorphisms (2). It provides the opportunity to detect the presence of new sequence variants and automatically adds them to the reference database. This important feature makes the assay extremely flexible and immediately suitable for detection of new variants without the need to make any additional changes to the assay itself. The MassCLEAVE approach employed in the HBV genotyping assay is based on interpretation of the patterns of RNA fragments generated from PCR products after transcription and base-specific cleavage with RNase A (28). Thus, any factors that can affect the pattern of the RNA cleavage products may contribute to the decline of the MS signal used for

genotype calling. For example, any RNA molecules generated through transcription of incomplete or degraded PCR products or PCR products of a mixed population of closely related HBV variants could significantly reduce the genotyping score. We examined many of the factors that may influence the outcome of the assay: viral titer, genotype, reference set, PCR and MS product quality, and genetic complexity of the specimen. We found that these factors either do not significantly affect the assay performance (e.g., titer, genotype, or single or double reaction failure) or, if they do, can be reliably controlled (e.g., sample storage, PCR product quality, and content of the reference set).

The major challenge for the assay is ultimately presented by the viral heterogeneity itself. The HBV polymerase lacks proofreading activity and generates multiple variants to combat the immune response of the host and to develop drug resistance and even vaccine escape mutations (5, 15, 20). Consequently, the virus in the host exists as a mixture of sequence variants that are discretely different from each other, with differences becoming more pronounced if viral persistence is maintained or the host comes under the influence of antiviral drugs. These genetic differences may accumulate during long-term infection and contribute to the complexity of the MS patterns detected here. Additionally, HBV variants containing many or rare mutations that are not represented in the reference set may be genotyped with a low MS signal score.

Our data show that HBV heterogeneity can be accommodated by the MALDI-TOF assay without affecting its accuracy. For example, the lower MS scores detected for HBV genotypes E and G compared to the other six genotypes were found to be related to genetic differences between the tested and reference sequences for genotype G and to intrasample genetic complexity for genotype E. The lower scores obtained for HBV genotypes insufficiently represented in the reference set can be improved with future submissions of novel sequences to the set. However, the intrasample heterogeneity may be accommodated differently. With all other aforementioned factors being effectively controlled, scores between 0.6 and 0.8 suggest increased complexity of the sample (i.e., the presence of viral

variants or of different coinfecting genotypes or the presence of sequence that is not close to any of the reference sequences used). Samples from acute cases are generally expected to be associated with lower intrahost HBV heterogeneity than chronic cases because of the transmission bottlenecks and limited intrahost evolution. Accordingly, lower scores for acute cases were found to be related to mismatches between tested and reference sequences. When the chronic cases were additionally analyzed by the endpoint limiting-dilution PCR and sequencing approach (22), it was shown that indeed all of these cases contained many genetically distant intrahost HBV variants of genotype E as well as five specimens that were found to contain additional HBV variants of different genotypes (data not shown). In terms of data interpretation, the matching of these mixed populations to a single consensus sequence from the reference set inevitably results in lowered scores, because the consensus match cannot contain all of the single-nucleotide variability existing in a given specimen. However, the important observation was that genotype calls were always correct, even if only single reaction was used. The MS assay can undoubtedly benefit from more extended testing of chronic cases representing all genotypes, so that accurate assessment could be made on the effect of genetic heterogeneity on the assay accuracy. The assay could also benefit from further validation with more specimens from patients infected with different HBV genotypes.

The HBV genotype assay developed in this study does not discriminate among HBV variants that may be found in a given specimen and therefore can genotype only the dominant HBV variants. Mixed-genotype infections may result in a lower score of genotype calling. However, the data obtained here showed that the genotyping score was always above the cutoff for all specimens, including those that were found to contain additional HBV genotypes at <10% of the population. The reported incidence of infection with more than one HBV genotype varies from 0.8% to 17.5%, and the variation depends on the method used for genotype identification (7, 25). Frequent detection of genotype G in the presence of genotype A or, occasionally, genotype D (8, 18) raises the possibility that the MS assay may produce the genotype G calls with lower scores. The problem of mixed-genotype infections is intrinsic to all HBV genotyping techniques, some of which tend to underreport (1) and others of which tend to overestimate the incidence of mixed-genotype infections (17).

The novel MALDI-TOF MS assay described here allows for rapid and accurate HBV genotyping. This medium-throughput assay is highly reproducible because of the built-in redundancy that ensures accurate genotype calling. It is amenable to the detection of any novel HBV variant. The assay produces automatic data reports and does not require any additional data interpretation. The combination of all of these characteristics lends the assay to more widespread, routine application in public health and diagnostic laboratories.

#### REFERENCES

- Ding, X., H. Gu, Z. H. Zhong, X. Zilong, H. T. Tran, Y. Iwaki, T. C. Li, T. Sata, and K. Abe. 2003. Molecular epidemiology of hepatitis viruses and genotypic distribution of hepatitis B and C viruses in Harbin, China. *Jpn. J. Infect. Dis.* **56**:19–22.
- Ehrlich, M., S. Boecker, and D. van den Boom. 2005. Multiplexed discovery of sequence polymorphisms using base-specific cleavage and MALDI-TOF MS. *Nucleic Acids Res.* **33**:e38.
- Gintowt, A. A., J. J. Germer, P. S. Mitchell, and J. D. Yao. 2005. Evaluation of the MagNA Pure LC used with the TRUGENE HBV genotyping kit. *J. Clin. Virol.* **34**:155–157.
- Gish, R. G., and S. Locarnini. 2007. Genotyping and genomic sequencing in clinical practice. *Clin. Liver Dis.* **11**:761–795, viii.
- Gunther, S., F. von Breuning, T. Santantonio, M. C. Jung, G. B. Gaeta, L. Fischer, M. Sterneck, and H. Will. 1999. Absence of mutations in the YMDD motif/B region of the hepatitis B virus polymerase in famciclovir therapy failure. *J. Hepatol.* **30**:749–754.
- Honisch, C., Y. Chen, C. Mortimer, C. Arnold, O. Schmidt, D. van den Boom, C. R. Cantor, H. N. Shah, and S. E. Gharbia. 2007. Automated comparative sequence analysis by base-specific cleavage and mass spectrometry for nucleic acid-based microbial typing. *Proc. Natl. Acad. Sci. U. S. A.* **104**:10649–10654.
- Kao, J. H., P. J. Chen, M. Y. Lai, and D. S. Chen. 2001. Acute exacerbations of chronic hepatitis B are rarely associated with superinfection of hepatitis B virus. *Hepatology* **34**:817–823.
- Kato, H., E. Orito, R. G. Gish, F. Sugauchi, S. Suzuki, R. Ueda, Y. Miyakawa, and M. Mizokami. 2002. Characteristics of hepatitis B virus isolates of genotype G and their phylogenetic differences from the other six genotypes (A through F). *J. Virol.* **76**:6131–6137.
- Laperche, S., V. Thibault, F. Bouchardeau, S. Alain, S. Castelain, M. Gassin, M. Gueudin, P. Halfon, S. Larrat, F. Lunel, M. Martinot-Peignoux, B. Mercier, J. M. Pawlotsky, B. Pozzetto, A. M. Roque-Afonso, F. Roudot-Thoraval, K. Saune, and J. J. Lefrere. 2006. Expertise of laboratories in viral load quantification, genotyping, and precore mutant determination for hepatitis B virus in a multicenter study. *J. Clin. Microbiol.* **44**:3600–3607.
- Lim, C. K., J. T. Tan, A. Ravichandran, Y. C. Chan, and S. H. Ton. 2007. Comparison of PCR-based genotyping methods for hepatitis B virus. *Malays. J. Pathol.* **29**:79–90.
- Lindh, M., A. S. Andersson, and A. Gusdal. 1997. Genotypes, nt 1858 variants, and geographic origin of hepatitis B virus—large-scale analysis using a new genotyping method. *J. Infect. Dis.* **175**:1285–1293.
- Luan, J., J. Yuan, X. Li, S. Jin, L. Yu, M. Liao, H. Zhang, C. Xu, Q. He, B. Wen, X. Zhong, X. Chen, H. L. Chan, J. J. Sung, B. Zhou, and C. Ding. 2009. Multiplex detection of 60 hepatitis B virus variants by MALDI-TOF mass spectrometry. *Clin. Chem.* **55**:1503–1509.
- Mahtab, M. A., S. Rahman, M. Khan, and F. Karim. 2008. Hepatitis B virus genotypes: an overview. *Hepatobiliary Pancreat. Dis. Int.* **7**:457–464.
- Malakhova, M. V., E. N. Ilina, V. M. Govorun, S. A. Shutko, K. R. Dudina, O. O. Znoyko, E. A. Klimova, and N. D. Iushchuk. 2009. Hepatitis B virus genotypic typing using mass-spectrometry. *Bull. Exp. Biol. Med.* **147**:220–225.
- Margeridon-Thermet, S., N. Shulman, A. Ahmed, R. Shahriar, T. Liu, C. Wang, S. Holmes, F. Babrzadeh, B. Gharizadeh, B. Hanczaruk, B. Simen, M. L. Egholm, and R. Shafer. 2009. Ultra-deep pyrosequencing of hepatitis B virus quasispecies from nucleoside and nucleotide reverse-transcriptase inhibitor (NRTI)-treated patients and NRTI-naive patients. *J. Infect. Dis.* **199**:1275–1285.
- Niesters, H. G., S. Pas, and R. A. de Man. 2005. Detection of hepatitis B virus genotypes and mutants: current status. *J. Clin. Virol.* **34**(Suppl. 1):S4–S8.
- Osiowy, C., and E. Giles. 2003. Evaluation of the INNO-LiPA HBV genotyping assay for determination of hepatitis B virus genotype. *J. Clin. Microbiol.* **41**:5473–5477.
- Osiowy, C., D. Gordon, J. Borlang, E. Giles, and J. P. Villeneuve. 2008. Hepatitis B virus genotype G epidemiology and co-infection with genotype A in Canada. *J. Gen. Virol.* **89**:3009–3015.
- Palumbo, E. 2007. Hepatitis B genotypes and response to antiviral therapy: a review. *Am. J. Ther.* **14**:306–309.
- Pujol, F. H. 2005. Genotypic variability of hepatitis viruses associated with chronic infection and the development of hepatocellular carcinoma. *J. Clin. Gastroenterol.* **39**:611–618.
- Qutub, M. O., J. J. Germer, S. P. Rebers, J. N. Mandrekar, M. G. Beld, and J. D. Yao. 2006. Simplified PCR protocols for INNO-LiPA HBV genotyping and INNO-LiPA HBV PreCore assays. *J. Clin. Virol.* **37**:218–221.
- Ramachandran, S., G. L. Xia, L. M. Ganova-Raeva, O. V. Nainan, and Y. Khudyakov. 2008. End-point limiting-dilution real-time PCR assay for evaluation of hepatitis C virus quasispecies in serum: performance under optimal and suboptimal conditions. *J. Virol. Methods* **151**:217–224.
- Repp, R., S. Rhiel, K. H. Heermann, S. Schaefer, C. Keller, P. Ndumbe, F. Lampert, and W. H. Gerlich. 1993. Genotyping by multiplex polymerase chain reaction for detection of endemic hepatitis B virus transmission. *J. Clin. Microbiol.* **31**:1095–1102.
- Roque-Afonso, A. M., M. P. Ferey, V. Mackiewicz, L. Fki, and E. Dussaix. 2003. Monitoring the emergence of hepatitis B virus polymerase gene variants during lamivudine therapy in human immunodeficiency virus coinfecting patients: performance of CLIP sequencing and line probe assay. *Antivir. Ther.* **8**:627–634.
- Schaefer, S. 2007. Hepatitis B virus taxonomy and hepatitis B virus genotypes. *World J. Gastroenterol.* **13**:14–21.
- Sertoz, R. Y., S. Erensoy, S. Pas, U. S. Akarca, F. Ozgenc, T. Yamazhan, T. Ozacar, and H. G. Niesters. 2005. Comparison of sequence analysis and

- INNO-LiPA HBV DR line probe assay in patients with chronic hepatitis B. *J. Chemother.* **17**:514–520.
27. **Sertoz, R. Y., S. Erensoy, S. Pas, T. Ozacar, and H. Niesters.** 2008. Restriction fragment length polymorphism analysis and direct sequencing for determination of HBV genotypes in a Turkish population. *New Microbiol.* **31**:189–194.
  28. **Stanssens, P., M. Zabeau, G. Meersseman, G. Remes, Y. Gansemans, N. Storm, R. Hartmer, C. Honisch, C. P. Rodi, S. Bocker, and D. van den Boom.** 2004. High-throughput MALDI-TOF discovery of genomic sequence polymorphisms. *Genome Res.* **14**:126–133.
  29. **Takahashi, K., Y. Akahane, K. Hino, Y. Ohta, and S. Mishiro.** 1998. Hepatitis B virus genomic sequence in the circulation of hepatocellular carcinoma patients: comparative analysis of 40 full-length isolates. *Arch. Virol.* **143**:2313–2326.
  30. **Tang, X. R., J. S. Zhang, H. Zhao, Y. H. Gong, Y. Z. Wang, and J. L. Zhao.** 2007. Detection of hepatitis B virus genotypes using oligonucleotide chip among hepatitis B virus carriers in Eastern China. *World J. Gastroenterol.* **13**:1975–1979.
  31. **Tran, N., R. Berne, R. Chann, M. Gauthier, D. Martin, M. A. Armand, A. Ollivet, C. G. Teo, S. Ijaz, D. Flichman, M. Brunetto, K. P. Bielawski, C. Pichoud, F. Zoulim, and G. Vernet.** 2006. European multicenter evaluation of high-density DNA probe arrays for detection of hepatitis B virus resistance mutations and identification of genotypes. *J. Clin. Microbiol.* **44**:2792–2800.
  32. **Vernet, G.** 2002. DNA-chip technology and infectious diseases. *Virus Res.* **82**:65–71.
  33. **Woo, H. Y., H. Park, B. I. Kim, W. K. Jeon, and Y. J. Kim.** 2008. Evaluation of dual priming oligonucleotide-based multiplex PCR for detection of HBV YMDD mutants. *Arch. Virol.* **153**:2019–2025.