

Low-Level Persistence of Drug Resistance Mutations in Hepatitis B Virus-Infected Subjects with a Past History of Lamivudine Treatment

Severine Margeridon-Thermet,^a Evguenia S. Svarovskaia,^b Farbod Babrzadeh,^c Ross Martin,^b Tommy F. Liu,^a Mary Pacold,^a Elizabeth C. Reuman,^a Susan P. Holmes,^d Katyna Borroto-Esoda,^b Robert W. Shafer^a

Department of Medicine, Division of Infectious Diseases, Stanford University, Stanford, California, USA^a; Gilead Sciences, Foster City, California, USA^b; Stanford Genome Technology Center, Stanford University, Stanford, California, USA^c; Department of Statistics, Stanford University, Stanford, California, USA^d

We sought to determine the prevalence of hepatitis B virus (HBV) lamivudine (LAM)-resistant minority variants in subjects who once received LAM but had discontinued it prior to virus sampling. We performed direct PCR Sanger sequencing and ultradeep pyrosequencing (UDPS) of HBV reverse transcriptase (RT) of plasma viruses from 45 LAM-naive subjects and 46 LAM-experienced subjects who had discontinued LAM a median of 24 months earlier. UDPS was performed to a depth of ~3,000 reads per nucleotide. Minority variants were defined as differences from the Sanger sequence present in $\geq 0.5\%$ of UDPS reads in a sample. Sanger sequencing identified ≥ 1 LAM resistance mutations (rtL80I/V, rtM204I, and rtA181T) in samples from 5 (11%) of 46 LAM-experienced and none of 45 LAM-naive subjects (0%; $P = 0.06$). UDPS detected ≥ 1 LAM resistance mutations (rtL80I/V, rtV173L, rtL180M, rtA181T, and rtM204I/V) in 10 (22%) of the 46 LAM-experienced subjects, including 5 in whom LAM resistance mutations were not identified by Sanger sequencing. Overall, LAM resistance mutations were more likely to be present in LAM-experienced (10/46, 22%) than LAM-naive subjects (0/45, 0%; $P = 0.001$). The median time since LAM discontinuation was 12.8 months in the 10 subjects with a LAM resistance mutation compared to 30.5 months in the 36 LAM-experienced subjects without a LAM resistance mutation ($P < 0.001$). The likelihood of detecting a LAM resistance mutation was significantly increased using UDPS compared to Sanger sequencing and was inversely associated with the time since LAM discontinuation.

Hepatitis B virus (HBV) drug resistance emerges commonly in subjects receiving monotherapy with the first two licensed anti-HBV nucleoside reverse transcriptase inhibitors (NRTIs): lamivudine (LAM) and adefovir (ADV) (1, 2). Although LAM and ADV monotherapy are no longer recommended for HBV treatment, LAM monotherapy has been used extensively in patients who received it prior to the availability of entecavir (ETV) and tenofovir (TDF), in patients who have lacked access to ETV and TDF, and in patients coinfecting with HIV who received LAM (but not TDF) as part of their antiretroviral treatment (3, 4). HBV genotypic resistance testing is recommended to help guide therapy in patients for whom antiviral therapy has been unsuccessful (2, 3). The most commonly used method of genotypic testing, dideoxynucleotide Sanger sequencing, however, is insensitive for detecting drug-resistant variants present at proportions lower than 20% within a plasma virus population.

Next-generation sequencing is increasingly used to identify low-abundance minority genetic variants within a heterogeneous pool of amplified DNA molecules, such as those within a virus population (5–8). In a previous study, we used 454/Roche ultradeep pyrosequencing (UDPS) (9) to characterize HBV minority NRTI resistance variants in a cross-sectional population of 17 NRTI-naive and 11 NRTI-experienced HBV-infected subjects (5). In that study, we demonstrated that UDPS was sensitive and specific at detecting minority variants present at proportions as low as 1.0% of the plasma viruses within a patient (5).

Here, we apply UDPS to a larger population of subjects who shared a history of receiving LAM and/or interferon (IFN) but who had discontinued these therapies months to years prior to enrolling in a clinical trial of TDF. This study provides insight into the dynamics of LAM resistance mutations following LAM treatment discontinuation.

MATERIALS AND METHODS

Study subjects and samples. Plasma samples were obtained from 101 HIV-1-seronegative HBV-infected subjects before they enrolled in two clinical trials sponsored by Gilead Sciences to assess the efficacy of TDF in HBeAg⁻ (GS-US-174-0102) and HBeAg⁺ (GS-US-174--0103) subjects (10, 11). All subjects were off treatment at the time of sample collection. Samples were selected for UDPS if they had plasma HBV DNA levels of $> 1,725$ IU/ml (ca. $> 10,000$ copies/ml). Among 91 subjects with samples meeting these criteria, 45 were nucleoside RT inhibitor (NRTI) naive, and 46 had received LAM. One subject, who had received emtricitabine, was excluded. The LAM-experienced subjects had received a median of 21 months of therapy but discontinued therapy a median of 24 months prior to plasma sampling. HBV genotypic resistance data at the time of LAM discontinuation were not available. Forty subjects—equally divided between LAM-naive and LAM-experienced individuals—had also been treated with alpha IFN (IFN- α) prior to plasma sampling. Human subjects approval was obtained by Gilead Sciences to collect the samples used for the present study and to perform additional viral sequencing tests. The Stanford University Institutional Review Board provided approval for performing the viral deep sequencing and its associated analyses.

Sanger sequencing and mutation definitions. HBV DNA was extracted from 200 μ l of serum using a QIAamp DNA blood minikit (Qiagen, Germantown, MD). Full-length HBV RT coding area was am-

Received 2 August 2012 Returned for modification 31 August 2012

Accepted 25 October 2012

Published ahead of print 31 October 2012

Address correspondence to Severine Margeridon-Thermet, smargeridon@hotmail.com, or Robert W. Shafer, rshafer@stanford.edu.

Supplemental material for this article may be found at <http://dx.doi.org/10.1128/AAC.01601-12>.

Copyright © 2013, American Society for Microbiology. All Rights Reserved.

doi:10.1128/AAC.01601-12

plified and bidirectionally sequenced as previously described (12). The HBV genotypes of Sanger population sequences were determined using the National Center for Biotechnology Information viral genotyping resource (13). Mutations were defined as differences from the genotype-specific consensus amino acid at each position. A list of the genotype-specific consensus amino acid sequences can be found in appendix 1 at <http://hivdb.stanford.edu/HBV/releaseNotes>. HBV NRTI resistance mutations were defined as rtL80I/V, rtI169T, rtV173L, rtL180M, rtA181T/V, rtT184S/A/I/L/F/G, rtS202G/I, rtM204I/V/S, rtN236T, and rtM250V (1, 2). rtM204I/V/S and rtI181T/V were defined as primary LAM resistance mutations because each mutation alone reduces LAM susceptibility (2). rtL180M, rtL80I/V, and rtV173L were defined as secondary compensatory LAM resistance mutations. Minority variant mutations were defined as mutations detected by UDPS that were not detected by direct PCR Sanger sequencing.

UDPS. HBV DNA was extracted from plasma by using a QIAamp Ultrasens virus minikit (Qiagen). Limiting-dilution PCR of the full-length HBV RT, performed to estimate the number of amplifiable viral DNA templates per sample, demonstrated that samples with $>1,725$ IU/ml consistently had more than 100 amplifiable DNA templates in 200 μ l to 1 ml of plasma.

Four pairs of primers tailed with UDPS adaptors and a subject-specific barcode were used to amplify four overlapping fragments of ~ 400 bp encompassing the complete HBV RT coding area as described previously (5). Thirty cycles of PCR were performed using Expand High Fidelity^{plus} PCR system reagents (*Taq* plus proofreading enzyme *Pwo*; Roche Applied Sciences, Indianapolis, IN). PCR products were purified using Agencourt AMPure XP system (Beckman Coulter Genomes, Danvers, MA) and quantified using Quant-iT Picogreen dsDNA reagent (Invitrogen, Carlsbad, CA). DNA libraries were obtained by pooling at equimolar concentrations 16 amplicons obtained from four HBV samples. The DNA libraries were pyrosequenced using the 454/Roche GS FLX platform producing reads of 250 bp on average. The 91 plasma samples and 9 HBV control plasmids were sequenced in nine GS-FLX sequencing runs.

Standard Flowgram Format files were processed to generate paired files containing FASTA sequence reads and Phred-equivalent quality scores for each sequence library. To reduce sequence artifact, we excluded reads shorter than 200 nucleotides and removed reads containing one or more bases with a quality score of <10 ($>10\%$ probability of sequence error) or a mean quality score of <25 ($>0.3\%$ probability of sequence error per base per sequence read) (14). Sequence reads were then demultiplexed using the 5' primer and barcode sequences, resulting in the assignment of each read to a subject sample and primer-pair.

Each UDPS read was aligned to a consensus genotype sequence using MosaikAligner (<http://bioinformatics.bc.edu>). Minor variant mutations were defined as mutations that were not present in the Sanger sequence. Because extensive APOBEC-mediated G-to-A editing leads to mutational impairment of viruses (15, 16), distinguishing hypermutated sequence reads from nonhypermutated reads is required for accurate analysis of virus quasispecies. Therefore, following alignment we excluded sequence reads considered to be hypermutated by one of the following three criteria (17): (i) $\geq 10\%$ of nucleotide positions with a G in the population sequence but an A in the UDPS read, (ii) ≥ 2 unusual HBV RT variants caused by G-to-A differences (an unusual variant was defined as a variant present in fewer than three sequences from $\sim 4,000$ different individuals with HBV RT sequences in GenBank) (18), or (iii) a stop codon or one of the following active site mutations rtD83N, rtD205N, and rtD206N.

Statistical analyses performed to identify authentic minority variants. On each of the nine picotiter plates, we also sequenced PCR-amplified DNA from a plasmid HBV clone. The median coverage for these nine control experiments was 3,219 reads (range, 1,443 to 4,341). The median proportion of errors per plate, defined as differences between the plasmid sequence and UDPS read, was 0.082% (range, 0.020 to 0.125%).

To distinguish authentic minority variants from technical artifacts resulting from the cumulative effect of PCR and pyrosequencing errors,

we estimated the technical error rate and identified a threshold above which mutations detected by UDPS were unlikely to have resulted from technical artifact. Mutations detected by UDPS in a plasma sample were reported only if the mutation's proportion in the plasma sample was $\geq 0.5\%$ and more than five times the median frequency with which that mutation occurred in the nine control plasmid sequences. For example, 94 mutations at 88 positions had a median plasmid error rate $\geq 0.10\%$, including 8 mutations at 8 amino acid positions with a median plasmid error rate of $\geq 0.20\%$; (range, 0.20% to 0.38%). Each of the NRTI resistance mutations had median plasmid error rate $<0.1\%$ with the exception of rtS202G, which had a median plasmid error rate of 0.15%.

Inferential statistics. Standard statistical tests, including the Mann-Whitney test and Friedman rank sum test for comparing medians, the Fisher exact test for comparing proportions, and the Pearson (parametric) and Spearman (nonparametric) correlation coefficients, were performed using the core package of libraries in the R statistics program (19). The Jaccard correlation coefficient was used to detect statistically significant correlations between pairs of mutations within the same sample. This coefficient examines sequences containing pairs of mutations within the context of sequences containing at least one of the mutations, thereby eliminating the heavy weight given to the double negative category in analyses such as the Fisher exact test (20). The details of the method are described in the supplemental material accompanying the analysis.

RESULTS

Subject characteristics. Twenty-one of the LAM-experienced and 19 of the LAM-naive subjects had a history of receiving IFN- α prior to study enrollment. The median plasma HBV DNA levels were 3.7×10^6 IU/ml (range, 5.3×10^3 to 8.9×10^8 IU/ml) in the LAM-naive samples and 3.9×10^6 IU/ml (range, 7.6×10^3 to 3.4×10^9 IU/ml) in the LAM-experienced samples (Table 1 and see Table S1 in the supplemental material). The median duration of LAM treatment was 21 months (interquartile range: 12 to 36 months, Fig. 1A). The median duration of time since LAM discontinuation was 22 months (interquartile range [IQR], 13 to 40 months, Fig. 1B). Of the 45 LAM-naive subjects, 3 were infected with genotype A strains, 6 were infected with genotype B strains, 6 were infected with genotype C strains, and 30 were infected with genotype D strains. Of the 46 LAM-experienced subjects, 9 were infected with genotype A strains, 1 was infected with genotype B strains, 6 were infected with genotype C strains, 29 were infected with genotype D strains, and 1 was infected with a genotype E strain.

Sanger sequencing results. Direct PCR Sanger sequencing detected a median of 7.5 amino acid mutations per sample in both the LAM-naive and LAM-experienced subjects. NRTI resistance mutations were detected in 0 of the 45 LAM-naive subjects versus 5 of the 46 LAM-experienced subjects ($P = 0.06$; Fisher exact test) (Table 2). The five LAM-experienced subjects with LAM resistance mutations had the following mutations: (i) rtL80I/V + rtM204I, (ii) rtL80I + rtM204I, (iii) rtA181T, (iv) rtL180M, and (v) rtM204I. The GenBank accession numbers for each of the Sanger sequences are listed in Table S1 in the supplemental material.

UDPS coverage. UDPS produced 3.7 million reads of 200 or more nucleotides in length. The median read coverage prior to filtering was 4,200 reads per primer (IQR, 2,840 to 6,160). However, following removal of reads with low quality scores or read lengths shorter than 200 nucleotides (consisting primarily of reads that began upstream of RT but were truncated at the start of RT or reads that ended downstream of RT but which were truncated at the end of RT), the median coverage per sample was 2,829 reads

TABLE 1 Virological characteristics of the HBV samples from 91 study subjects^a

| Parameter | Subject treatment category | | | |
|--|------------------------------------|------------------------|------------------------------|----------------|
| | Naive | LAM | LAM+IFN | IFN |
| No. of samples | 26 | 25 | 21 | 19 |
| Study subject gender, no. (%) | | | | |
| Females | 7 (26.9) | 7 (28) | 7 (33.3) | 4 (21) |
| Males | 19 (73.1) | 18 (72) | 14 (66.7) | 15 (79) |
| Median age in yrs (IQR) | 45 (37–52) | 45 (35–55) | 50 (38–56) | 47 (36–54) |
| Ethnicity, no. (%) | | | | |
| White | 14 (53.8) | 18 (72) | 18 (85.7) | 17 (89.5) |
| Black | | 1 (4) | 2 (9.5) | |
| Asian | 10 (38.5) | 6 (24) | 1 (4.8) | 2 (10.5) |
| Other | 2 (7.7) | | | |
| Median virus level ^b (IQR) | 6.3 (5.4–7.2) | 6.6 (5.2–7.3) | 6.6 (5.7–8.0) | 7.1 (6.1–7.7) |
| LAM treatment duration, median no. of mos. (IQR) | | | | |
| On LAM | | 20 (9–33) | 30 (19–43) | |
| Off LAM | | 26 (12–36) | 13 (12–24) | |
| HBeAg ⁺ (%) | 19 | 20 | 14 | 11 |
| Genotypes (%) | D (54), B (23), A (11.5), C (11.5) | D (58), C (22), A (20) | D (71), A (19), B (5), E (5) | D (89), C (11) |

^a No subject was on therapy at the time of virological testing. Naive, naive subjects that had never received antiviral therapy; LAM, subjects with a history of receiving lamivudine (LAM); LAM+IFN, subjects with a history of receiving LAM + interferon (IFN); IFN, subjects with a history of receiving IFN but not LAM.

^b Virus levels were measured as the log₁₀ IU/ml.

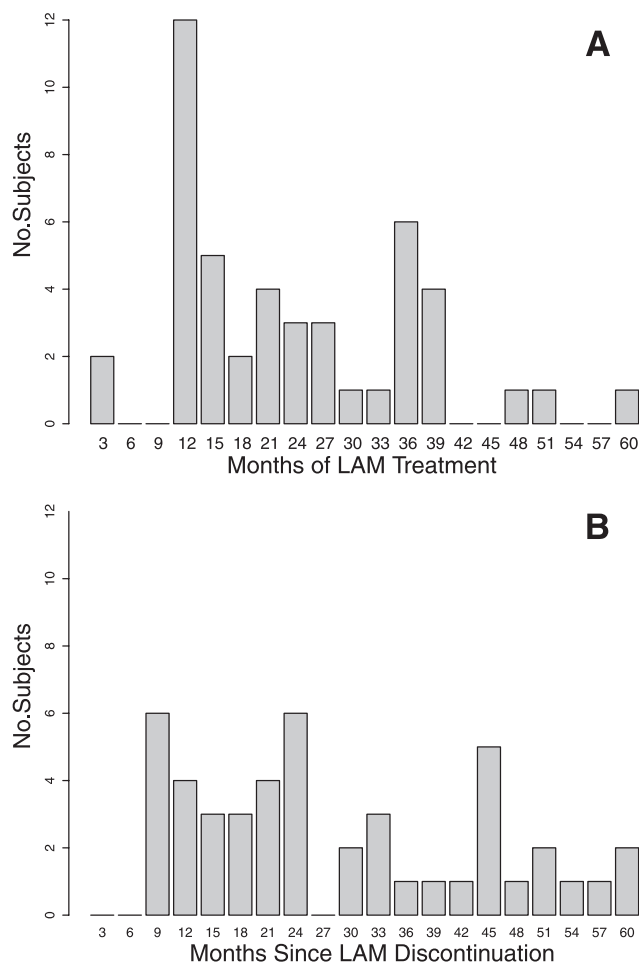


FIG 1 Distribution of months of past lamivudine (LAM) treatment (A) and of months since LAM discontinuation (B) among the 46 LAM-experienced subjects.

(range, 1192 to 5497). The median coverage per amino acid position was 2,637 (range, 1,423 to 4,620). Overall, UDPS detected a median of 15 minority-variant RT amino acid (nonsynonymous) mutations per sample (IQR, 10 to 23) and a median of 22 minority-variant RT synonymous mutations per sample (IQR, 13 to 34).

UDPS detection of minority variant LAM resistance mutations. UDPS detected minority-variant LAM resistance mutations in three of the five samples with LAM resistance mutations detected by Sanger sequencing (Table 2). UDPS also detected one primary LAM resistance mutation (either rtM204V or rtM204I) plus one to two secondary minority variant LAM resistance mutations (rtL180M, rtL80I, and/or rtV173L) per sample in five additional samples from the remaining 41 LAM-experienced subjects. Therefore, overall 10 (21.7%) of 46 samples from LAM-experienced subjects had LAM resistance mutations.

In contrast, minority-variant LAM resistance mutations were not detected in any sample from the 45 LAM-naive subjects. Thus, LAM resistance mutations were more likely to be detected in samples from the 46 LAM-experienced (10/46) than LAM-naive subjects (0/45; $P = 0.001$; Fisher exact test). The number of subjects with LAM resistance mutations detectable only by UDPS was more frequent among the LAM-experienced subjects (5/46) compared to the LAM-naive subjects (0/45), but the difference did not reach statistical significance ($P = 0.06$; Fisher exact test).

For five subjects, it was possible to assess how frequently minority-variant LAM resistance mutation were present on the same HBV genome as another LAM resistance mutation. Subjects 36, 82, 72, 73, and 75 each had two or more of mutations at positions 173 (rtV173L), 180 (rtL180M), and 204 (rtM204I/V), close enough to be detected on sequence reads generated from the same amplicon. In all five subjects, each of these mutations occurred both alone and in combination. In four of five subjects, the Jaccard

TABLE 2 HBV drug resistance mutations detected by Sanger sequencing and ultra-deep pyrosequencing

| Patient no. | Treatment history (no. of months) | | Sanger sequence DRMs | UDPS DRM(s) (%) ^a |
|-------------|-----------------------------------|---------|----------------------|---|
| | On LAM | Off LAM | | |
| 42 | 13.5 | 7 | M204M/I, L80L/I | M204I (41.2), L80I (31.3), L80V (4.1) , M204V (0.5) |
| 81 | 33 | 4 | M204I, L80I/V | M204I (99.8), L80V (67.2), L80I (32.3) |
| 73 | 12 | 13.5 | M204M/L/I | M204I (15.8), L180M (13.2) , M204V (2.2) |
| 85 | 36.5 | 10 | L180L/M | L180 M (80.7), L80V (4.8) |
| 5 | 12 | 7.5 | A181A/T | A181T (51.5) |
| 82 | 36.5 | 12 | None | L180 M (69.8) , V173L (69.6) , <u>M204V (1.0)</u> |
| 36 | 10 | 7 | None | <u>M204I (9.2)</u> , L80I (8.2) , L180M (1.8) |
| 75 | 11.5 | 14 | None | <u>M204V (2.0)</u> , L180M (1.7) |
| 87 | 36 | 19.5 | None | <u>M204I (0.8)</u> , L80I (0.7) |
| 72 | 3 | 13.5 | None | <u>M204I (0.6)</u> , L180M (0.5) |

^a Mutations in boldface type were detected solely by ultra-deep pyrosequencing (UDPS). Underlined mutations (either rtM204V or rtM204I) are primary LAM resistance mutations that were not detected by Sanger sequencing. The parenthetical percentages indicate the proportion of UDPS reads at an amino position encoded by the preceding drug resistance mutation. rtM204L, a mutation of unknown significance was present in 41.4 and 2.8% of sequence reads of the plasma samples from subjects 73 and 72, respectively. rtA181S, another mutation of unknown significance, was present in 0.7 and 1.9% of sequence reads of the plasma samples from subjects 5 and 75, respectively.

correlation coefficient (20) showed that mutations were significantly much more likely to occur in combinations than alone (see Table S2 in the supplemental material).

Two atypical mutations of uncertain biological and clinical significance were detected at known LAM resistance position: rtM204L and rtA181S (Table 2). rtM204L was detected as part of a mixture by Sanger sequencing and was present in 41.4% of UDPS reads in the sample from subject 73; it was detected only by UDPS in subject 72 (2.8% of sequence reads); rtA181S was present in 0.7 and 1.9% of sequence reads of the plasma samples from subjects 5 and 75, each of which also contained known LAM resistance mutations. rtM204L was negatively correlated with rtL180M. rtA181S occurred in a proportion of reads that was too low to assess its correlation with rtL180M and rtM204V, which were present in the same sample but never on the same read as rtA181S (see Table S2 in the supplemental material).

Factors associated with LAM resistance mutations detected by Sanger sequencing and or UDPS. The only factor significantly associated with the presence of LAM resistance mutations by Sanger sequencing and/or UDPS was the time since LAM discontinuation (Fig. 2). Among the 10 LAM-experienced subjects with detectable LAM resistance mutations, the median time since LAM discontinuation was 11.2 months (range, 4.1 to 19.7). In contrast, among the 36 LAM-experienced subjects without detectable LAM resistance mutations, the median time since LAM discontinuation was 30.5 months (range, 7.1 to 58.5) ($P = 0.0001$; Mann-Whitney test). There was a significant inverse relationship between the time since LAM discontinuations and the log-transformed value of the most prevalent LAM resistance mutations within each subject ($r = 0.69$; $P < 0.001$; Spearman's rank correlation coefficient; Fig. 3). In an analysis confined to the 21 subjects who discontinued LAM within 20 months prior to study enrollment, 10 (48%) had detectable LAM resistance mutations.

The duration of LAM therapy, the plasma virus level, and the number of minority variants were not significantly associated with the presence of LAM resistance mutations (see Fig. S1 in the supplemental material). LAM resistance mutations were detected in 4/9 genotype A samples, 6/29 genotype D samples, 0/6 genotype C samples, 0/1 genotype B samples, and 0/1 genotype E samples ($P = \text{NS}$ for differences between all pairs of genotypes; Fisher exact test).

G-to-A hypermutation. G-to-A hypermutation was present in a median of 1.2% of reads per subject (range, 0.1 to 6.8%). The Friedman rank sum test (friedman.test; R Statistical Package) showed that the sample of origin (degrees of freedom [df] = 87) made a more significant contribution to the level of hypermutation in a read than the primer pair used for PCR amplification (df = 3), demonstrating that hypermutation was a characteristic of individual plasma samples rather than a PCR artifact. A G-to-A hypermutation was associated with the number of mutations detectable by Sanger sequencing (Pearson correlation coefficient, $r = 0.35$; $P = 6 \times 10^{-4}$) and the number of minor variants detected by UDPS (Pearson $r = 0.48$; $P = 1 \times 10^{-6}$). However, G-to-A hypermutation was not more common in the 10 subjects with LAM resistance mutations detected either by Sanger or UDPS. It was also not associated with the plasma HBV DNA level, HBV genotype, history of receiving LAM, or history of receiving IFN.

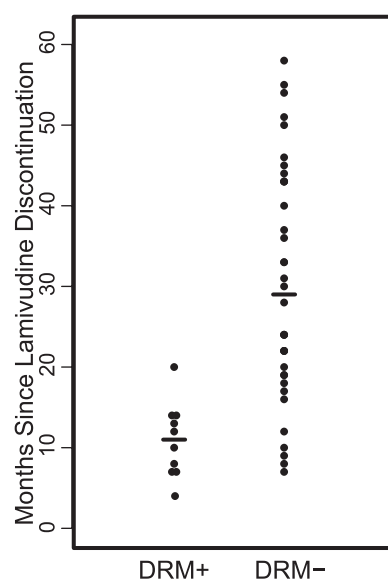


FIG 2 Months since discontinuation of lamivudine (LAM) for the 10 subjects with and 36 subjects without viruses containing LAM resistance mutations (median: 11.2 versus 30.5 months; $P = 0.0001$, Mann-Whitney test).

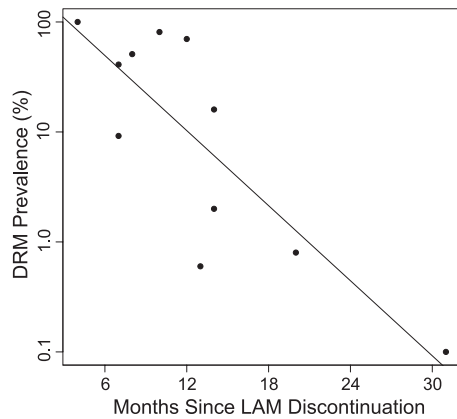


FIG 3 Inverse relationship between the time since lamivudine (LAM) discontinuation and the log-transformed value of the most prevalent LAM resistance mutations within each of the 46 LAM-experienced subjects ($r = 0.69$, $P < 0.001$; Spearman's rank correlation coefficient). A single point ($x = 30.5$ months and $y = 0.1\%$) is used to represent the 36 subjects who had discontinued LAM a median of 30.5 months prior to sampling and who did not have detectable LAM resistance mutations.

DISCUSSION

This study shows that LAM resistance mutations may persist in plasma for 20 or more months following LAM discontinuation. Among the 46 LAM-treated subjects, 10 (21.7%) had LAM resistance mutations detected either by Sanger sequencing (8 mutations in 5 subjects) or by UDPS (26 mutations in 10 subjects). Among the subjects who discontinued LAM within the 20 months prior to study enrollment, nearly one-half (10 of 21) had detectable LAM resistance mutations.

HBV RT has a high mutation rate and generates a highly heterogeneous population of viral variants that diversify during the course of infection. Nonetheless, NRTI resistance emerges much more slowly for HBV than for other chronic viruses such as HIV and HCV (1). Moreover, LAM resistance, which is often caused by just one mutation in HIV RT, is often caused by two or three mutations in combination in HBV RT. The slower evolution of HBV resistance may result from the evolutionary constraints imposed by HBV's overlapping reading frames and by the reservoir of drug-susceptible cccDNA genomes (21, 22).

The slow reversion of the LAM resistance mutations despite their lower levels of replication fitness parallels their slow emergence in subjects receiving LAM (23). The reversion to wild type may be slowed in viruses that developed beneficial viral mutations elsewhere on the HBV genome during the period of LAM therapy. In addition, the presence of linked mutations suggests that some reversions to wild type might disrupt epistatic combinations of mutations resulting in less fit intermediately resistant genomes (24).

Nine subjects had two or more of the canonical LAM resistance mutations: rtM204I/V ($n = 10$), rtL80I/V ($n = 7$), rtL180M ($n = 6$), and rtV173L ($n = 1$). The mutations at positions rt204, rt180, and rt173 were significantly more likely to co-occur on the same genomes than would be expected if they had arisen independently. One subject had the multi-NRTI-associated mutation rtA181T. The ADV resistance mutation rtN236T and the ETV resistance mutations at positions rt169, rt184, rt202, and rt250 were not detected. The detection of minor variant mutations was not asso-

ciated with the genotype of the sample sequence nor with the specific codon present in the Sanger sequence. Indeed, it has been shown by others that in all genotypes, most HBV resistance mutations arise from single nucleotide transitions (25).

Two novel mutations at known NRTI-associated positions (rtA181S and rtM204L) each occurred in two subjects. The combination of rtA181S + rtM204I has been reported to emerge in two patients receiving LAM monotherapy and has been shown to reduce susceptibility to LAM, FTC, ADV, and TDF (26, 27). rtM204L has been reported in a sequence from one of ~4,000 individuals with HBV RT sequences in GenBank (18). In contrast to rtM204V and rtM204I, rtM204L was negatively correlated with the secondary LAM resistance mutation rtL180M.

Although the combination of PCR and pyrosequencing can introduce technical artifact, multiple lines of evidence suggest that the minority LAM resistance variants that we detected were not artifacts. First, only those variants detected in 0.5% or more sequence reads and at a proportion more than five times higher the median plasma error rate for that mutation (usually $< 0.1\%$) were considered valid minority variants. Second, each of the minority LAM resistance variants occurred in a sample that also contained one or more additional LAM resistance variants either in the Sanger sequence or as another minority variant. Finally, the linkage among many of the mutations would be unlikely if these mutations had resulted from PCR or pyrosequencing errors.

Our analysis of the evolution of LAM resistance mutations in the subjects who discontinued therapy, however, presents an incomplete picture of HBV population dynamics *in vivo*. First, genotypic resistance data were not available on samples from the subjects at the time of LAM discontinuation. Second, the combined technical error resulting from PCR plus pyrosequencing decreases the reliability of variants as they decrease in proportions below 0.5%. Finally, ascertaining the ultimate fate of LAM-resistant variants would require examining the proportion of LAM resistance mutations within the hepatic cccDNA reservoir.

G-to-A hypermutation results from an innate antiviral defense mechanism mediated by host enzymes belonging to the APOBEC3 family of cytidine deaminases, which cause extensive deamination of cytidine bases to uridine in negative-stranded DNA, resulting in G-to-A hypermutation in positive-stranded DNA (15). APOBEC-mediated G-to-A hypermutation has been reported in HIV, HBV (16, 28, 29), other retroviruses, and retrotransposons (15). Because the effects of G-to-A hypermutation on a complete viral genome almost certainly lead to mutational impairment, it is essential to identify hypermutated reads because hypermutated reads containing rtA181T (GCN→ACN) and/or rtM204I (ATG→ATA) would otherwise be spuriously interpreted as being NRTI resistant. Therefore, we applied a classification model developed in a recent study to exclude sequence reads with a high probability of having G-to-A hypermutation (17).

Because little is known about the correlates of APOBEC-mediated hypermutation, we sought to determine whether this phenomenon was associated with the subjects' treatment histories and with other aspects of the subjects' HBV sequences. First, we showed that the proportion of hypermutated reads was an intrinsic property of each sample rather than sequencing artifact. Second, APOBEC-mediated G-to-A hypermutation was associated with the number of mutations detectable by Sanger sequencing and of minority variants detected by UDPS. These associations

have not previously been reported but are consistent with the concept that APOBEC3G proteins are encoded by IFN response genes and that periods of increased host immunity may be associated with both increased viral immune selection pressure and G-to-A hypermutation.

Most of the minority variant mutations detected here were the canonical LAM resistance mutations rtM204VI, rtL180M, rtV173L, and rtL80VI. Subjects with these mutations would be expected to respond to treatment with TDF. One patient had rtA181T, and this mutation, which partially reduces TDF susceptibility, might be an indication for combination therapy (e.g., TDF + ETV) (1, 3).

In conclusion, we show here that LAM-experienced patients who discontinue LAM are often likely to harbor LAM resistant variants, particularly if LAM was discontinued within the 20 months prior to deep sequencing. However, the absence of mutations in plasma at a level of 0.5% does not exclude the possibility that clinically significant levels of drug resistance mutations at lower levels or within hepatic cells. Case-control studies of UDPS sequencing on baseline samples from subjects failing and responding to salvage therapy would be needed to define the utility of UDPS in clinical practice.

ACKNOWLEDGMENT

The study of hepatitis B minority variants in subjects treated with lamivudine and/or IFN was sponsored by Gilead Sciences.

REFERENCES

- Locarnini S, Zoulim F. 2010. Molecular genetics of HBV infection. *Antivir. Ther.* 15(Suppl 3):3–14.
- Lok AS, Zoulim F, Locarnini S, Bartholomeusz A, Ghany MG, Pawlotsky JM, Liaw YF, Mizokami M, Kuiken C. 2007. Antiviral drug-resistant HBV: standardization of nomenclature and assays and recommendations for management. *Hepatology* 46:254–265.
- European Association for the Study of the Liver. 2012. EASL clinical practice guidelines: management of chronic hepatitis B virus infection. *J. Hepatol* 57:167–185.
- Wiersma ST, McMahon B, Pawlotsky JM, Thio CL, Thursz M, Lim SG, Ocama P, Esmat G, Maimuna M, Bell D, Vitoria M, Eramova I, Lavanchy D, Dusheiko G. 2011. Treatment of chronic hepatitis B virus infection in resource-constrained settings: expert panel consensus. *Liver Int.* 31:755–761.
- Margeridon-Thermet S, Shulman NS, Ahmed A, Shahriar R, Liu T, Wang C, Holmes SP, Babrzadeh F, Gharizadeh B, Hanczaruk B, Simen BB, Egholm M, Shafer RW. 2009. Ultra-deep pyrosequencing of hepatitis B virus quasispecies from nucleoside and nucleotide reverse-transcriptase inhibitor (NRTI)-treated patients and NRTI-naïve patients. *J. Infect. Dis.* 199:1275–1285.
- Powdrill MH, Tchesnokov EP, Kozak RA, Russell RS, Martin R, Svarovskaia ES, Mo H, Kouyos RD, Gotte M. 2011. Contribution of a mutational bias in hepatitis C virus replication to the genetic barrier in the development of drug resistance. *Proc. Natl. Acad. Sci. U. S. A.* 108:20509–20513.
- Solmone M, Vincenti D, Prosperi MC, Bruselles A, Ippolito G, Capobianchi MR. 2009. Use of massively parallel ultradeep pyrosequencing to characterize the genetic diversity of hepatitis B virus in drug-resistant and drug-naïve patients and to detect minor variants in reverse transcriptase and hepatitis B S antigen. *J. Virol.* 83:1718–1726.
- Wang C, Mitsuya Y, Gharizadeh B, Ronaghi M, Shafer RW. 2007. Characterization of mutation spectra with ultra-deep pyrosequencing: application to HIV-1 drug resistance. *Genome Res.* 17:1195–1201.
- Margulies M, Egholm M, Altman WE, Attiya S, Bader JS, Bemben LA, Berka J, Braverman MS, Chen YJ, Chen Z, Dewell SB, Du L, Fierro JM, Gomes XV, Godwin BC, He W, Helgesen S, Ho CH, Irzyk GP, Jando SC, Alenquer ML, Jarvie TP, Jirage KB, Kim JB, Knight JR, Lanza JR, Leamon JH, Lefkowitz SM, Lei M, Li J, Lohman KL, Lu H, Makhijani VB, McDade KE, McKenna MP, Myers EW, Nickerson E, Nobile JR, Plant R, Puc BP, Ronan MT, Roth GT, Sarkis GJ, Simons JF, Simpson JW, Srinivasan M, Tartaro KR, Tomasz A, Vogt KA, Volkmer GA, Wang SH, Wang Y, Weiner MP, Yu P, Begley RF, Rothberg JM. 2005. Genome sequencing in microfabricated high-density picolitre reactors. *Nature* 437:376–380.
- Heathcote EJ, Marcellin P, Buti M, Gane E, De Man RA, Krastev Z, Germanidis G, Lee SS, Flisiak R, Kaita K, Manns M, Kotzev I, Tchernev K, Buggisch P, Weibert F, Kurdas OO, Shiffman ML, Trinh H, Gurel S, Snow-Lampart A, Borroto-Esoda K, Mondou E, Anderson J, Sorbel J, Rousseau F. 2011. Three-year efficacy and safety of tenofovir disoproxil fumarate treatment for chronic hepatitis B. *Gastroenterology* 140:132–143.
- Marcellin P, Heathcote EJ, Buti M, Gane E, de Man RA, Krastev Z, Germanidis G, Lee SS, Flisiak R, Kaita K, Manns M, Kotzev I, Tchernev K, Buggisch P, Weibert F, Kurdas OO, Shiffman ML, Trinh H, Washington MK, Sorbel J, Anderson J, Snow-Lampart A, Mondou E, Quinn J, Rousseau F. 2008. Tenofovir disoproxil fumarate versus adefovir dipivoxil for chronic hepatitis B. *N. Engl. J. Med.* 359:2442–2455.
- Snow-Lampart A, Chappell B, Curtis M, Zhu Y, Myrick F, Schawalter J, Kitrinis K, Svarovskaia ES, Miller MD, Sorbel J, Heathcote J, Marcellin P, Borroto-Esoda K. 2010. No resistance to tenofovir disoproxil fumarate detected after up to 144 weeks of therapy in patients mono-infected with chronic hepatitis B virus. *Hepatology* 53:763–773.
- Rozanov M, Plikat U, Chappey C, Kochergin A, Tatusova T. 2004. A web-based genotyping resource for viral sequences. *Nucleic Acids Res.* 32:W654–W659.
- Huse SM, Huber JA, Morrison HG, Sogin ML, Welch DM. 2007. Accuracy and quality of massively parallel DNA pyrosequencing. *Genome Biol.* 8:R143.
- Cullen BR. 2006. Role and mechanism of action of the APOBEC3 family of antiretroviral resistance factors. *J. Virol.* 80:1067–1076.
- Suspene R, Guetard D, Henry M, Sommer P, Wain-Hobson S, Vartanian JP. 2005. Extensive editing of both hepatitis B virus DNA strands by APOBEC3 cytidine deaminases in vitro and in vivo. *Proc. Natl. Acad. Sci. U. S. A.* 102:8321–8326.
- Reuman EC, Margeridon-Thermet S, Caudill HB, Liu T, Borroto-Esoda K, Svarovskaia ES, Holmes SP, Shafer RW. 2010. A classification model for G-to-A hypermutation in hepatitis B virus ultra-deep pyrosequencing reads. *Bioinformatics* 26:2929–2932.
- Rhee SY, Margeridon-Thermet S, Nguyen MH, Liu TF, Kagan RM, Beggel B, Verheyen J, Kaiser R, Shafer RW. 2010. Hepatitis B virus reverse transcriptase sequence variant database for sequence analysis and mutation discovery. *Antiviral research.* 88:269–275.
- Development Core Team R. 2011. R: a language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria.
- Rhee SY, Liu TF, Holmes SP, Shafer RW. 2007. HIV-1 subtype B protease and reverse transcriptase amino acid covariation. *PLoS Comput. Biol.* 3:e87. doi:10.1371/journal.pcbi.0030087.
- Mizokami M, Orito E, Ohba K, Ikeo K, Lau JY, Gojobori T. 1997. Constrained evolution with respect to gene overlap of hepatitis B virus. *J. Mol. Evol.* 44(Suppl 1):S83–S90.
- Soriano V, Perelson AS, Zoulim F. 2008. Why are there different dynamics in the selection of drug resistance in HIV and hepatitis B and C viruses? *J. Antimicrob. Chemother.* 62:1–4.
- Pallier C, Castera L, Soulier A, Hezode C, Nordmann P, Dhumeaux D, Pawlotsky JM. 2006. Dynamics of hepatitis B virus resistance to lamivudine. *J. Virol.* 80:643–653.
- Khudyakov Y. 2010. Coevolution and HBV drug resistance. *Antivir. Ther.* 15:505–515.
- Svicher V, Cento V, Salpini R, Mercurio F, Fraune M, Beggel B, Han Y, Gori C, Wittkop L, Bertoli A, Micheli V, Gubertini G, Longo R, Romano S, Visca M, Gallinaro V, Marino N, Mazzotta F, De Sanctis GM, Fleury H, Trimoulet P, Angelico M, Cappiello G, Zhang XX, Verheyen J, Ceccherini-Silberstein F, Perno CF. 2011. Role of hepatitis B virus genetic barrier in drug-resistance and immune-escape development. *Digestive Liver Dis.* 43:975–983.
- Karatayli E, Karayalcin S, Karaaslan H, Kayhan H, Turkyilmaz AR, Sahin F, Yurdaydin C, Bozdayi AM. 2007. A novel mutation pattern emerging during lamivudine treatment shows cross-resistance to adefovir dipivoxil treatment. *Antivir. Ther.* 12:761–768.
- Moriconi F, Colombatto P, Coco B, Ciccorossi P, Oliveri F, Flichman

- D, Maina AM, Sacco R, Bonino F, Brunetto MR. 2007. Emergence of hepatitis B virus quasispecies with lower susceptibility to nucleos(t)ide analogues during lamivudine treatment. *J. Antimicrob. Chemother.* **60**: 341–349.
28. Noguchi C, Imamura M, Tsuge M, Hiraga N, Mori N, Miki D, Kimura T, Takahashi S, Fujimoto Y, Ochi H, Abe H, Maekawa T, Tateno C, Yoshizato K, Chayama K. 2009. G-to-A hypermutation in hepatitis B virus (HBV) and clinical course of patients with chronic HBV infection. *J. Infect. Dis.* **199**:1599–1607.
29. Noguchi C, Ishino H, Tsuge M, Fujimoto Y, Imamura M, Takahashi S, Chayama K. 2005. G-to-A hypermutation of hepatitis B virus. *Hepatology* **41**:626–633.