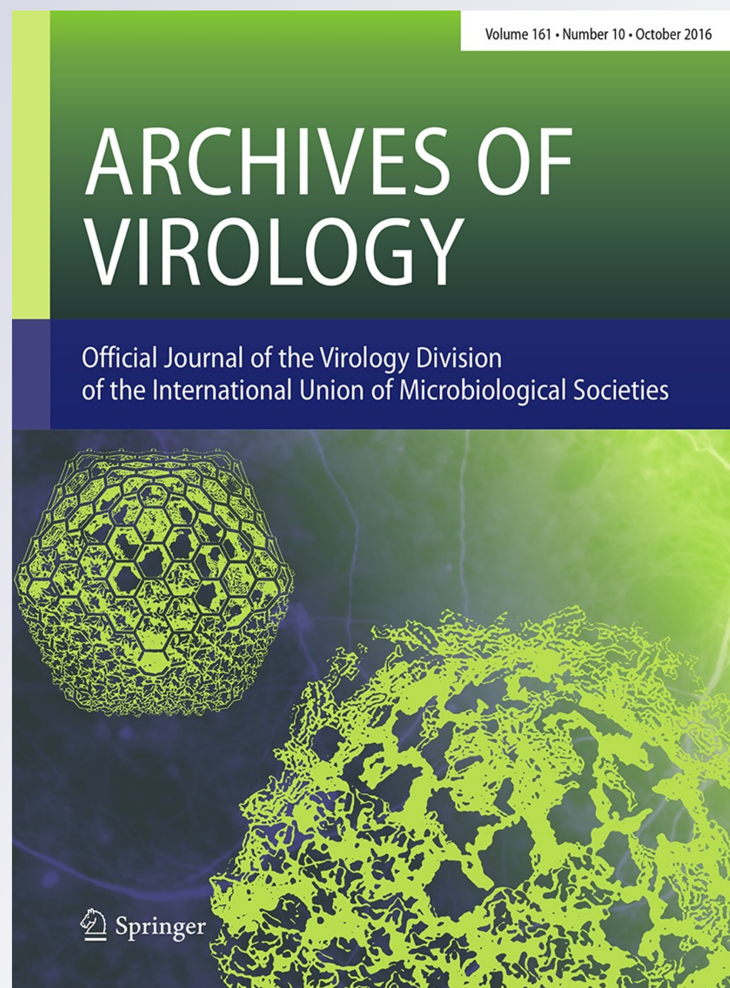
Hepatitis B virus resistance substitutions: long-term analysis by next-generation sequencing

Leandro R. Jones, Mariano Sede, Julieta M. Manrique & Jorge Quarleri

Archives of Virology
Official Journal of the Virology Division of the International Union of Microbiological Societies

ISSN 0304-8608
Volume 161
Number 10

Arch Virol (2016) 161:2885-2891
DOI 10.1007/s00705-016-2959-8



Springer

Your article is protected by copyright and all rights are held exclusively by Springer-Verlag Wien. This e-offprint is for personal use only and shall not be self-archived in electronic repositories. If you wish to self-archive your article, please use the accepted manuscript version for posting on your own website. You may further deposit the accepted manuscript version in any repository, provided it is only made publicly available 12 months after official publication or later and provided acknowledgement is given to the original source of publication and a link is inserted to the published article on Springer's website. The link must be accompanied by the following text: "The final publication is available at link.springer.com".

Springer

Author's personal copy

Arch Virol (2016) 161:2885–2891
DOI 10.1007/s00705-016-2959-8

BRIEF REPORT

Hepatitis B virus resistance substitutions: long-term analysis by next-generation sequencing

Leandro R. Jones[1,2] · Mariano Sede[1,3] · Julieta M. Manrique[1,2] · Jorge Quarleri[1,3]

Received: 11 March 2016 / Accepted: 28 June 2016 / Published online: 22 July 2016
© Springer-Verlag Wien 2016

Abstract HBV phylogenetics and resistance-associated mutations (RAMs) were surveyed by next-generation sequencing of 21 longitudinal samples from seven patients entering antiviral therapy. The virus populations were dominated by a few abundant lineages that coexisted with substantial numbers of low-frequency variants. A few low-frequency RAMs were observed before treatment, but new ones emerged, and their frequencies increased during therapy. Together, these results support the idea that chronic HBV infection is dominated by a few virus lineages and that an accompanying plethora of diverse, low-frequency variants may function as a reservoir that potentially contribute to viral genetic plasticity, potentially affecting patient outcome.

Approximately 240 million individuals worldwide are chronically infected with hepatitis B virus (HBV) [15]. Chronic hepatitis B (CHB) can be treated with either pegylated interferon alpha or nucleoside/nucleotide analogs (NAs) with the primary goal of achieving sustained virological suppression. Interferon-based therapy remains a benchmark therapy for CHB, with the absence of resistance and the possibility of immune-mediated clearance of hepatitis B as its main advantages over NAs. Treatments including entecavir (ETV) and tenofovir (TDF) are recommended as first-line therapies by major international guidelines [22]; however, these therapies have been difficult to implement in several developing countries due to economic concerns. Therefore, the low-genetic-barrier NAs are still used extensively in those places, especially in countries where generic lamivudine (LMV) and adefovir dipivoxil (ADV) are available, resulting in a high risk of partial virological response and drug resistance [23].

HBV is a circular partially double-stranded DNA virus containing a reverse-transcriptase enzyme. HBV replicates via an RNA intermediate that is responsible for the generation of several related virus variants called quasispecies, favoring the emergence of HBV drug resistance. Such resistance to antiviral therapy can lead to HBV treatment failure and progression of liver disease. In general, this is gradually acquired through the selection of pre-existing variants with resistance-conferring mutations in the polymerase and the accumulation of new amino acid substitutions [2]. Clinical resistance based on amino acid substitutions resulting from genotypic mutations causes viral population fluctuations [24].

A better understanding of the dynamics of the resistance-associated mutations (RAMs) would improve clinical outcomes and help to elucidate the mechanism of resistance to nucleoside/nucleotide analogs resistance mechanism [9]. To date, the most widely used method to characterize HBV variants is direct PCR and sequencing, but considering that this method only detects variants representing ≥20 % of the total population, we used next-generation sequencing (NGS) techniques to produce DNA sequences in much greater numbers than direct PCR and

✉ Jorge Quarleri
quarleri@fmed.uba.ar

[1] Consejo Nacional de Investigaciones Científicas y Técnicas, Av. Rivadavia 1917 (C1083ACA), Buenos Aires, Argentina

[2] Laboratorio de Virología y Genética Molecular, Facultad de Ciencias Naturales sede Trelew, Universidad Nacional de la Patagonia San Juan Bosco, 9 de Julio y Begrano S/N (9100), Trelew, Chubut, Argentina

[3] Instituto de Investigaciones Biomédicas en Retrovirus y Sida (INBIRS), Facultad de Medicina, Universidad de Buenos Aires, Paraguay 2155-Piso 11, C1121ABG Buenos Aires, Argentina

Author's personal copy

clone-based sequencing, also detecting minor variants [5, 18]. Therefore, the aim of the present retrospective observational study was to analyze the long-term dynamics of HBV resistance mutations in chronically infected patients under different treatment schemes, including pegylated interferon alpha and nucleoside/nucleotide analogs, using NGS. For this goal, samples from seven patients were analyzed during the follow-up. Inter-host phylogenetic relationships among the HBV isolates were also investigated.

Peripheral blood samples were collected from seven patients who were chronically infected with HBV (3 females; mean age ± SD, 34.5 ± 8.8 years: HBsAg and HBeAg positive). Serum levels of HBV DNA were quantified using a real-time PCR assay on a Cobas TaqMan 48 analyzer (lower detection limit 60 copies/mL, Roche Molecular Systems, Branchburg, NJ). The dynamics of the HBV viral load and therapy schemes at different time points are presented in Figure 1. Baseline samples (prior to NA antiviral therapy) were collected for six of seven patients, and for the remaining patient, samples were collected only during LMV-based therapy. For the present analysis, the first samples taken for two patients (P3, P6) under IFN-based therapy were considered to be the baseline, taking into account that interferon-based therapy does not induce resistance-associated mutations. Serum ALT and AST levels were measured using a commercially available automatic analyzer.

NGS was applied to serum samples taken at baseline and consequently during therapy, representing a total of 21 serial samples from seven patients. HBV-DNA was extracted from 200 μL of plasma by using a QIAamp DNA Blood Kit (QIAGEN AG, Basel, Switzerland). The HBV polymerase domain was amplified by nested PCR [7, 21]. Both reactions were run with high-fidelity Platinum Taq DNA polymerase (Life Technologies, Gaithersburg, Md.) according to the manufacturer's instructions. Primers used in the second round were modified by 5' tag extensions, which provided binding sites for multiplex identifiers (MIDs) that allow the identification of the samples after the pyrosequencing procedure is completed. Ultra-high-throughput next-generation sequencing (NGS) was carried out on a Roche/454 Life Sciences "Genome Sequencer-FLX" (GS-FLX) pyrosequencer.

For the sake of minimizing potential PCR and sequencing errors, we filtered the data obtained based on sequence length and quality of base calls. Low-quality read ends were trimmed out. Then, reads containing at least one base with a quality of ≤10 and/or Ns were dismissed. In addition, an average (mean) quality value ≥30 was

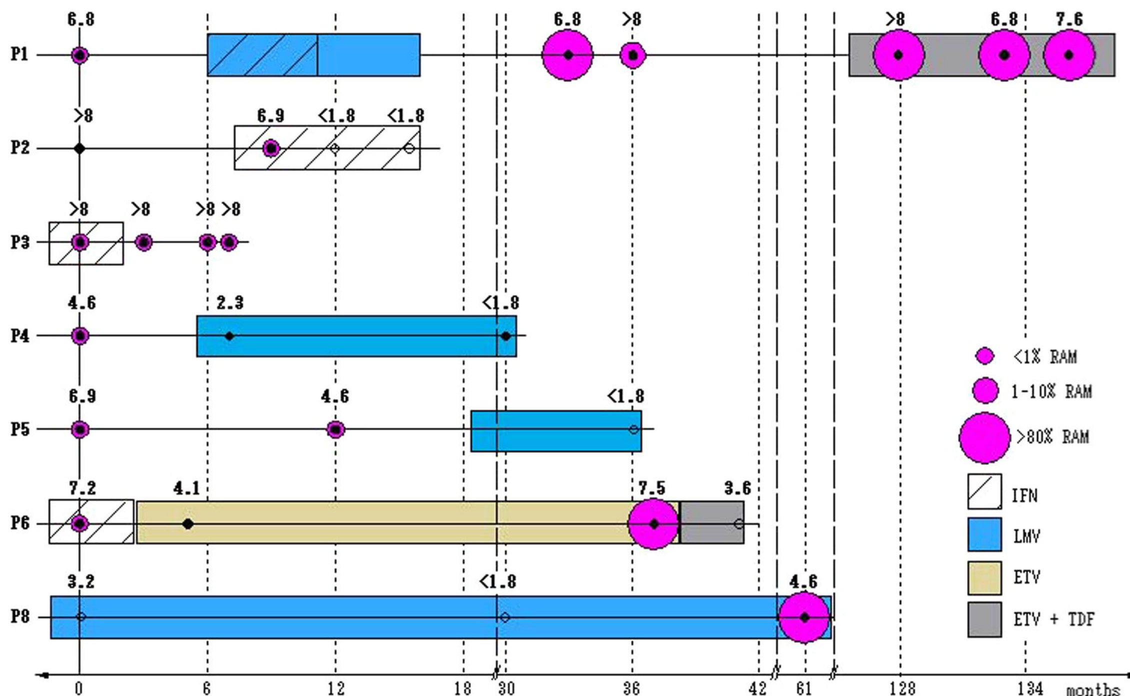


Fig. 1 Timeline representation (in months) of antiviral therapy (by rectangles) for each patient (P). HBV viral load levels (log copies/mL) are indicated by black and white dots. The black dots represent samples with successful HBV pol gene NGS analysis, while those represented as white dots were not available for NGS analysis. The sizes of purple circles sizes represent the relative abundance (as percentages) of HBV viral variants exhibiting RAMs. The vertical shorter dashed lines represent timelines separated by 6 months separated, and longer ones with parallel lines on the x-axis indicate a break in the timeline (color figure online)

Author's personal copy

required for each read in order to pass the quality control. All of these pre-processing steps were done using PRINSEQ v.0.15 [17]. Then, the data containing frameshifts were further filtered out, and low-frequency k-mers were corrected using Kec [19]. After that, we merged the sequences that were within 2 bp of a more abundant sequence using Mothur to further reduce potential PCR and sequencing errors [16]. Finally, the absence of chimeric sequences was confirmed with the *UCHIME* algorithm implemented in Mothur. After these quality controls, we obtained 8,121 high-quality sequence reads (P1: 2116, P2:514, P3:1763, P4:848, P5:844, P6:1645, P8:391) representing 1,241 different sequences (here after *variants*).

An appropriate model of nucleotide substitution (JC) was selected according to Akaike's information criterion using the program Mr.AIC [14]. Phylogenetic trees were constructed by the maximum-likelihood method [6] using the PhyML v3.0 software, and the branch support was assessed by nonparametric bootstrapping (100 pseudoreplicates). HBV sequences were aligned and visualized using the Mafft online application and the BioEdit v7.0.5.3 program (T. Hall Ibis Biosciences, Carlsbad, CA, USA), respectively.

Phylogenetic analysis showed that patients P1, P2, and P8 were infected with sub-genotype A2 while the remaining ones were infected with sub-genotype F1. The frequency distributions of the variants were highly skewed in all the patients, with a few variants dominating the virus populations and substantial numbers of low-frequency variants. To graphically summarize and further assess these

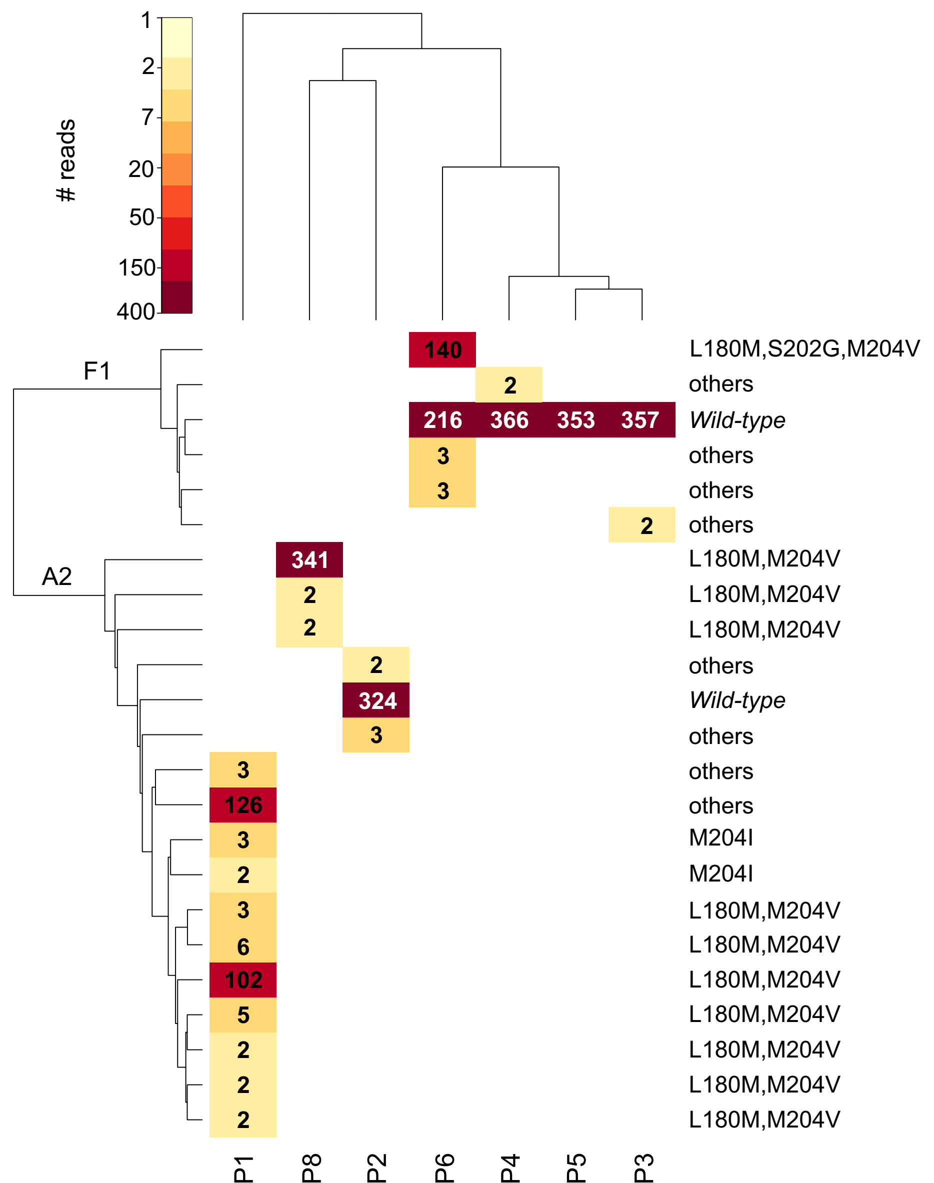


Fig. 2 Heat map displaying the frequencies of 23 variants (2372 sequences or reads) after equalizing the number of sequences from eight patients (bottom: P1-P8). Wild-type residues and RAMs associated with each HBV variant are shown at the right, ("others" include variants without primary and/or compensatory RAMs). The tree shown at the left corresponds to the phylogeny of the viral variant, with branch lengths set arbitrarily for better presentation. The dendrogram at the top corresponds to a hierarchical cluster analysis of the patients' virus populations performed by the complete-linkage method as implemented in the *gplots* R package. *F1,* subgenotype F1; *A2,* subgenotype A2

Author's personal copy

results, we first equalized the number of sequences per patient in order to compensate for potential biases due to sequence depth differences. To this aim, we randomly sampled, with replacement, 391 reads per patient. Then, we took all of the sequences from each patient that were represented by two or more reads (2372 sequences belonging to 23 variants) and inferred the virus phylogeny, constructed a phenogram grouping the patients according to the genetic background of the corresponding virus population, and produced a heat map depicting the frequencies of the variants using the R package gplots (https://cran.r-project.org/web/packages/gplots/index.html). The results are shown in Figure 2. The observed patterns agree with our previous analyses showing that chronically infected patients have virus populations with a few dominant variants and a significant number of low-frequency variants [8]. In patients infected with HBV sub-genotype F1 (P3, P4, P5, and P6), the wild-type virus was the most abundant variant. Among those infected with HBV sub-genotype A2, the wild-type variant only predominated in P2, while in P1 and P8, the most frequent variants exhibited RAMs, most likely due to shifts in response to the selective effect of antiviral therapy.

In order to evaluate phenotypic resistance of HBV to antiviral drugs, reads that passed the quality controls described above were input to the Geno2pheno(hbv) application at http://www.geno2pheno.org. In a previous study, amino acid substitutions at 42 potential RAMs in the HBV full-length RT sequence were analyzed and classified into five categories [10]. Those classified as primary RAMs

Fig. 3 Temporal dynamics of HBV variants found in the studied patients. Sampling times are given in the first horizontal axis (x-axis, in months). Viral loads (log copies/mL) at each sampling time are given in the vertical axis (y-axis), with sampling times connected by a solid line. The frequencies of RAMs during the follow-up are depicted using cones with color codes as indicated in the insets. The heights of cones are proportional to each variant frequency at each sampling time. The gray cones labeled as "others" include wild-type sequences and variants without RAMs (color figure online)

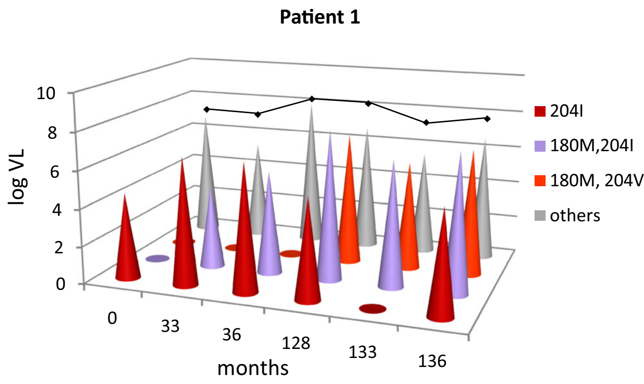
(category 1: I169T, A181T/V, T184A/C/F/G/I/L/M/S, A194T, S202C/G/I, M204I/V/S, N236T, M250I/L/V), and those classified as compensatory mutations (category 2: V173L, L180M) were dynamically monitored. Taking into account the HBV viral load, wild-type and viral variants exhibiting RAMs were quantified and expressed as copies/mL. We observed several amino acid substitutions known to confer HBV resistance among the characterized variants. Single or multiple RAMs were detected in the seven patients but with different frequencies during the course of the study. Samples from six patients (P1, P3, P4, P5, P6, and P8) harbored M204V/I substitutions. The substitution L180M was also found multiple times in three patients (P1, P6, and P8), only during NA therapy and concomitantly with M204V/I. V173L was only found in P2 as a minor population (<1 %) at baseline. R202G was detected at a low level (<1 %) at baseline (P1, P6) and during therapy (P3) as a single mutation. The frequency increased dramatically (>99 %) during entecavir monotherapy in a patient with virological failure (P6), but it appeared in

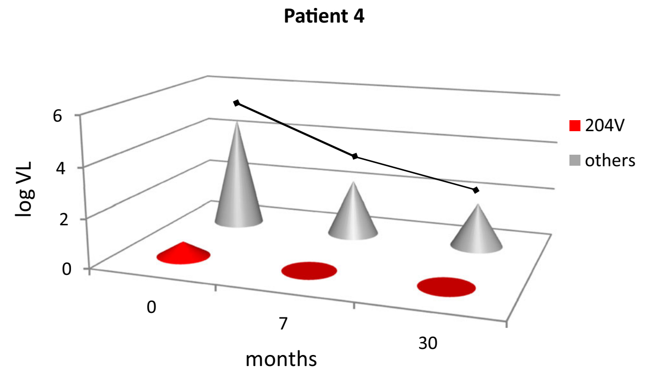
Table 1 Patient-by-patient dynamics of primary and compensatory RAMs. At sampling times when the HBV variants were successfully analyzed by NGS, the relative abundance of each RAM (as a percentage) is shown

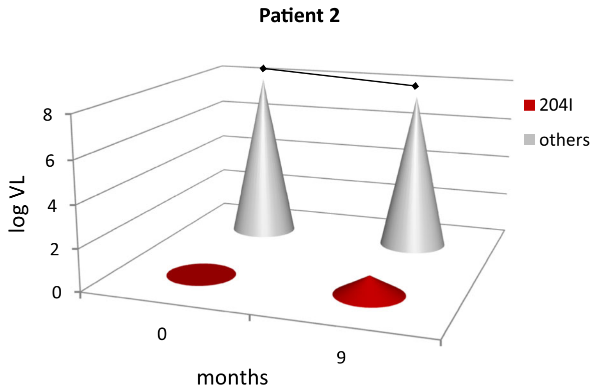
Patient	RAM	0	3	5	6	7	9	12	30	33	36	37	61	128	133	136
P1	M204I	<1								96.1	6.5			<1	0	<1
	L180M,M204I	0								<1	<1			83.4	84.1	70.4
	L180M,M204V	0								0	0			10.5	10.2	15.6
P2	M204I	0					1.0									
P3	A181V	0	<1		0	0										
	M204V	<1	0		<1	0										
	M204I	0	0		0	<1										
P4	M204V	<1				0			0							
P5	M204I	<1						0								
	M204V	0						<1								
P6	A181T	<1		0								0				
	M204V	<1		0								0				
	S202G	<1		0								0				
	S202G,M204V,L180M	0		0								99.0				
	M204V,L180M	0		0								<1				
	S202G,M204V	0		0								<1				
P8	L180M,M204V												99.2			

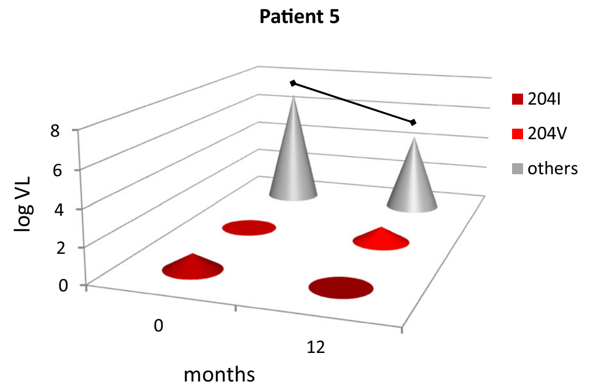
Shaded cells represent the period of time in which the patient received antiviral therapy

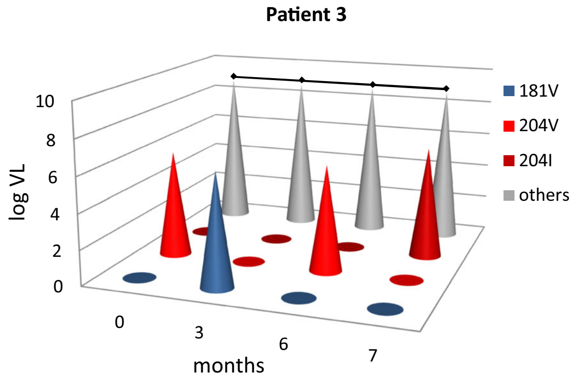
Author's personal copy

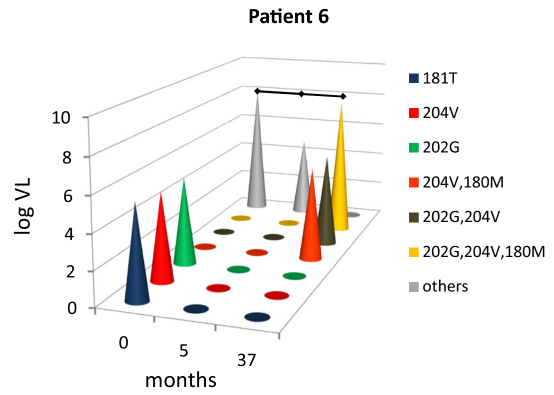


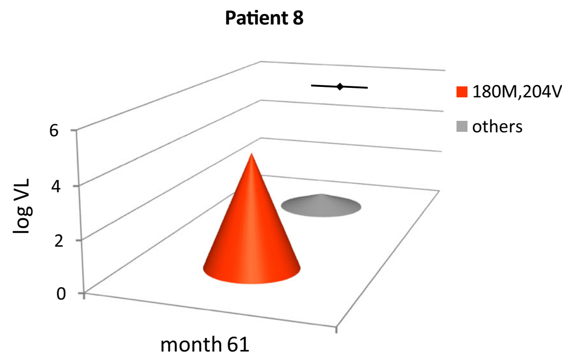












Author's personal copy

association with M204V and L180M/V. The substitution A181V/T was found in a minor population in two patients (P3 and P6). The I169T and T184G/S substitutions were detected at a very low frequency (<1 %), either alone or together with other mutations, at baseline (P1) and during therapy (P1, P6) (Table 1). None of the samples were found to harbor A194T, N236T, or M250V substitutions.

In spite of differences in the dynamics of the RAMs among patients, at baseline, the substitutions appeared mainly as singletons with low frequency (<1 %) among studied patients. Five of them (P1, P3, P4, P5, and P6) harbored the M204I/V substitution. Other substitutions (I169T/V, V173A/M, T184A/I, S202G) were also present at low levels (<1 %) in two patients (P1, P6). During and after therapy, the number of resistance mutations was lower in those exposed only to IFN-based therapy (P2, P3), or in LMV virological responders (P4, P5) than in highly treatment-experienced patients (P1, P6) or those with failure of LMV treatment (P8). In the first group (P2, P3), IFN-therapy did not affect the RAMs pattern or frequency (<1 %) irrespective of the IFN therapy response. M204V/I was the only RAM detected at very low abundance (<1 %) prior to initiation of antiviral therapy in the patients who successfully responded to LMV, in sharp contrast to P8, who exhibited a high proportion of variants with the M204V+L180M double mutation (Fig. 3, Table 1).

The two patients exposed to several antiviral drugs (P1 and P6) showed similar dynamics of RAMs. As already mentioned, at baseline, they had single mutations (A181T, S202G, M204V/I), each at low frequency (<1 %). Under therapy pressure, multiple RAMs appeared (P1:L180M+M204I/V, P6: L180M+M204V+S202G), showing high predominance in the viral population (>99 %) (Fig. 3; Table 1). Recent studies have demonstrated that the HBV genotype contributes to drug-resistant mutant evolution and selection [10]. In accordance with previous reports, we observed that HBV sub-genotype A2 favors the M204V mutation in LMV-resistance mutation patterns [12]. However, in P1, the coexistence of M204I was observed after LMV failure. In addition, the L180M mutation is significantly associated with the M204V/I mutation in HBV-A [3]. Similarly, the HBV sub-genotype F1 appears to favor the M204V and L180M mutations. However, the number of patients studied was limited, and, therefore, further investigations are needed.

Drug resistance is the principal cause of antiviral treatment failure, which may result in clinical disease progression. NGS data analysis is a valuable tool for HBV resistance assessment that allowed us to find several amino acid substitutions, including both previously undescribed mutations (data not shown) and already described mutations that are associated with drug resistance. These were already present with variable frequencies in drug-naive patients, similar to what has been found in previous studies [11, 20]. Many mutations may be selectively neutral and thus subject to genetic drift, and their presence may enable the virus to respond more rapidly to changed environments [1]. Such variations may confer a selective advantage for generating fitter and thus more abundant virus variants under antiviral drug pressure. However, the kinetics of replacement of wild-type virus in liver cells by a dominant mutant is generally slow given the necessity of free replication space for the spread of the mutant virus. It requires the immune-mediated elimination of hepatocytes containing wild-type virus, and the generation of new susceptible cells. On the other hand, the assembly, secretion and/or infectivity of antiviral drug-resistant mutants may be impaired due to reduced fitness [4]. Among the RAMs, M204V/I was identified most frequently. It has been reported that G-to-A hypermutation at these codons contributes to the emergence of drug-resistant mutations. The mechanism underlying this phenomenon is not clear, and this could be analyzed further in future studies involving host factors such as APOBEC 3 [13].

In conclusion, NGS data indicate that the frequencies of primary antiviral-resistant mutants in HBV are low before the start of therapy, but can increase substantially as a response to the selective pressures exerted by the treatment. Furthermore, our analyses suggest that a reservoir of diverse, low-frequency virus variants confer genetic plasticity to the virus population along during chronic infection.

Compliance with ethical standards

Funding This work was supported by the Argentinean National Agency of Scientific and Technological Promotion (grant number PICT 2012-0422 to JQ and PICT PRH-14 120 to LRJ); the University of Buenos Aires (grant number 20020110100034); and the Argentinean National Scientific and Technical Research Council (grant number PIP-11220110101089).

Conflict of interest All authors declare that they have no conflict of interest.

Ethical approval All procedures performed in this study were in accordance with the ethical standards of the institution and with the 1964 Helsinki declaration and its later amendments or comparable ethical standards. The study was approved by the Huesped Foundation Ethics Committee, Buenos Aires, Argentina, and informed consent was obtained from each patient.

Informed consent Informed consent was obtained from all individual participants included in the study.

References

1. Bimber BN, Dudley DM, Lauck M, Becker EA, Chin EN, Lank SM, Grunenwald HL, Caruccio NC, Maffitt M, Wilson NA, Reed JS, Sosman JM, Tarosso LF, Sanabani S, Kallas EG, Hughes AL,

Author's personal copy

O'Connor DH (2010) Whole-genome characterization of human and simian immunodeficiency virus intrahost diversity by ultra-deep pyrosequencing. J Virol 84:12087–12092

2. Chotiyaputta W, Lok AS (2009) Hepatitis B virus variants. Nat Rev Gastroenterol Hepatol 6:453–462
3. Damerow H, Yuen L, Wiegand J, Walker C, Bock CT, Locarnini S, Tillmann HL (2010) Mutation pattern of lamivudine resistance in relation to hepatitis B genotypes: hepatitis B genotypes differ in their lamivudine resistance associated mutation pattern. J Med Virol 82:1850–1858
4. Durantel D (2010) Fitness and infectivity of drug-resistant and cross-resistant hepatitis B virus mutants: why and how is it studied? Antivir Ther 15:521–527
5. Gong L, Han Y, Chen L, Liu F, Hao P, Sheng J, Li XH, Yu DM, Gong QM, Tian F, Guo XK, Zhang XX (2013) Comparison of next-generation sequencing and clone-based sequencing in analysis of hepatitis B virus reverse transcriptase quasispecies heterogeneity. J Clin Microbiol 51:4087–4094
6. Guindon S, Gascuel O (2003) A simple, fast, and accurate algorithm to estimate large phylogenies by maximum likelihood. Syst Biol 52:696–704
7. Jia S, Wang F, Li F, Chang K, Yang S, Zhang K, Jiang W, Shang Y, Deng S, Chen M (2014) Rapid detection of hepatitis B virus variants associated with lamivudine and adefovir resistance by multiplex ligation-dependent probe amplification combined with real-time PCR. J Clin Microbiol 52:460–466
8. Jones LR, Sede M, Manrique JM, Quarleri J (2016) Virus evolution during chronic hepatitis B virus infection as revealed by ultradeep sequencing data. J Gen Virol 97:435–444
9. Khudyakov Y (2010) Coevolution and HBV drug resistance. Antivir Ther 15:505–515
10. Liu BM, Li T, Xu J, Li XG, Dong JP, Yan P, Yang JX, Yan L, Gao ZY, Li WP, Sun XW, Wang YH, Jiao XJ, Hou CS, Zhuang H (2010) Characterization of potential antiviral resistance mutations in hepatitis B virus reverse transcriptase sequences in treatment-naive Chinese patients. Antiviral Res 85:512–519
11. 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-naive patients. J Infect Dis 199:1275–1285
12. Mirandola S, Sebastiani G, Rossi C, Velo E, Erne EM, Vario A, Tempesta D, Romualdi C, Campagnolo D, Alberti A (2012) Genotype-specific mutations in the polymerase gene of hepatitis B virus potentially associated with resistance to oral antiviral therapy. Antiviral Res 96:422–429
13. 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
14. Nylander JAA (2004) MrAIC.pl. Program distributed by the author. Evolutionary Biology Centre, Uppsala University
15. Rodriguez C, Chevaliez S, Bensadoun P, Pawlotsky JM (2013) Characterization of the dynamics of hepatitis B virus resistance to adefovir by ultra-deep pyrosequencing. Hepatology 58:890–901
16. Schloss PD, Gevers D, Westcott SL (2011) Reducing the effects of PCR amplification and sequencing artifacts on 16S rRNA-based studies. PLoS One 6:e27310
17. Schmieder R, Edwards R (2011) Quality control and preprocessing of metagenomic datasets. Bioinformatics 27:863–864
18. Shendure J, Ji H (2008) Next-generation DNA sequencing. Nat Biotechnol 26:1135–1145
19. Skums P, Dimitrova Z, Campo DS, Vaughan G, Rossi L, Forbi JC, Yokosawa J, Zelikovsky A, Khudyakov Y (2012) Efficient error correction for next-generation sequencing of viral amplicons. BMC Bioinform 13(Suppl 10):S6
20. 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-naive patients and to detect minor variants in reverse transcriptase and hepatitis B S antigen. J Virol 83:1718–1726
21. Stuyver LJ, Locarnini SA, Lok A, Richman DD, Carman WF, Dienstag JL, Schinazi RF (2001) Nomenclature for antiviral-resistant human hepatitis B virus mutations in the polymerase region. Hepatology 33:751–757
22. Wang YJ, Yang L, Zuo JP (2016) Recent developments in antivirals against hepatitis B virus. Virus Res 213:205–213
23. Zhou B, Dong H, He Y, Sun J, Jin W, Xie Q, Fan R, Wang M, Li R, Chen Y, Xie S, Shen Y, Huang X, Wang S, Lu F, Jia J, Zhuang H, Locarnini S, Zhao GP, Jin L, Hou J (2015) Composition and interactions of hepatitis B virus quasispecies defined the virological response during telbivudine therapy. Sci Rep 5:17123
24. Zoulim F, Locarnini S (2009) Hepatitis B virus resistance to nucleos(t)ide analogues. Gastroenterology 137(1593–1608): e1591–e1592