Dear Jia-Ping Yan,

Along with this letter we include the revision of our manuscript entitled "Characterization of Hepatitis B Virus X Gene Quasispecies Complexity in Monoinfection and Hepatitis Delta Virus Superinfection" (Manuscript ID 45308), which we submitted as an invited manuscript (ID: 00225318). We would like to thank you and the reviewers for the comments, which gave us the opportunity to clarify and improve the paper. We have answered the comments you made within the same text boxes you created in the version of the manuscript that you edited, and we provide the point-by-point response to the reviewers below.

We hope that you will find the manuscript suitable for publication in World Journal of Gastroenterology in its present form.

Sincerely,

Francisco Rodríguez-Frías, PhD

Reviewer #1 (Number ID 00722050)

Specific Comments To Authors:

Hepatitis B is a chronic infection which affects 257 million people worldwide with 15 to 20 million affected with hepatitis delta virus. The hepatitis delta virus strongly suppresses hepatitis B virus replication. The mechanism of their interaction is unknown. These viruses show a dynamic distribution of mutants that results in viral quasispecies. This study included 24 patients of which 7/24 (29.2%) with HBeAg-negative chronic HBV infection (CI, previously termed inactive carriers), 8/24 (33.3%) with HBeAg-negative chronic hepatitis B (CHB) and 9/24 (37.5%) with chronic hepatitis delta (CHD). The aim of this study was to compare the HBV quasispecies complexity in the

HBX 5′ region between chronic hepatitis delta and chronic HBV monoinfected patients. The researchers also evaluated the pattern of nucleotide changes to investigate which nucleotides could be the cause of the quasispecies complexity. The researchers used a variety of serological and virological tests, in addition to amplification of hepatitis B and hepatitis D viruses regions of interest by next generation sequencing. The complexity of quasispecies and nucleotides were also analyzed.

Novelty /Originality This article is sufficiently novel and interesting to warrant publication. No previous studies were found for the characterization of hepatitis B virus X gene quasispecies complexity in monoinfection and hepatitis delta virus superinfection. This article can contribute to the advancement of science and the delivery of healthcare as it has the potential to improve the management of hepatitis B infections.

Presentation This article was clearly laid out with all the key elements present. The title clearly described the content of the article, while the abstract provided a good summary of the content of the manuscript. In the introduction the authors clearly stated their objectives and the aim of their investigation. The methodology used and results obtained were described by the authors. The study design was suitable for the aim of the study with adequate statistical analysis conducted on the results obtained. Appropriate graphs and pictures which were both clear and informative were included in the manuscript. In the discussion, the authors summarized their findings with these findings being relevant to previous studies. The results obtained supported the claims of the researchers with the speculations and extrapolations being reasonable. The article used language that was scientific. While the article provided a lot of information, the authors could have presented the information in a more organised and reader friendly manner.

Importance The study demonstrated that the lower replication chronic hepatitis D and hepatitis B virus infection groups show a trend to higher quasispecies complexity than the higher replication chronic hepatitis B group. The findings of this study have the potential to improve the understanding of the mechanism of quasispecies in hepatitis B.

References The references used in this manuscript were sufficient, appropriate and recent.

Scientific Merit The researchers were able to come up with two possible hypothesis in an attempt to possibly explain the mechanism by which hepatitis delta virus enhances the

hepatitis B virus quasispecies. The first hypothesis suggests that the activation of the host innate immune response under the effect of hepatitis delta virus. While the second hypothesis postulates a possible interaction between HDAg and RNA pol II, which could affect the replicative capacity and functionality of the enzyme. This study provided data on the influence of hepatitis delta virus on hepatitis B virus genetic diversity in the HBX gene. Results showed that in the hepatitis B stages with lower replication (CHD and CI), the hepatitis B virus quasispecies in the 5'end of HBX exhibited a trend toward higher complexity than in chronic hepatitis B. This study was warranted and the findings of this study can provide a better understanding of the mechanism of interaction between hepatitis B virus and hepatitis delta virus with regards to the complexity of hepatitis B virus quasispecies. Given the prevalence of Hepatitis B infections, a larger study population could have been used, which would have further validated the results obtained by the researchers. The total study population consisted of 24 patients. While the study population used to compare the HBV quasispecies complexity indices between HBeAg-negative and HBeAg-positive chronic hepatitis D patients consisted of ONLY 4 patients in each group. Recommendations such as the need for further research were addressed by the researchers. The section of limitations with comparison with other studies should be expanded.

Ethical Issues There was neither plagiarism nor fraud in this manuscript.

We would like to thank the reviewer for the time and effort spent reviewing our manuscript and the positive comments made.

Regarding the concerns about the organization of the manuscript: "*While the article provided a lot of information, the authors could have presented the information in a more organised and reader friendly manner.*" In the previous version of the manuscript we tried to facilitate comprehension of the main results of the study by presenting them as figures and tables, and organizing the text into several subsections. In the discussion section we then discussed the results in the same order as they were presented in the manuscript. Nonetheless, if you believe that some parts may require better organization, we would be grateful if you would point them out, as we would be pleased to clarify them in a subsequent revision.

Regarding the concerns about the number of patients: "*Given the prevalence of Hepatitis B infections, a larger study population could have been used, which would have further validated the results obtained by the researchers. The total study population consisted of 24 patients. While the study population used to compare the HBV quasispecies complexity indices between HBeAg-negative and HBeAg-positive chronic hepatitis D patients consisted of ONLY 4 patients in each group.*" It is true that the patient groups are relatively small, but we must mention that it was difficult to find patients for the groups with lower HBV replication (i.e. chronic hepatitis delta [CHD] and HBV chronic infection [CI]) having high enough HBV DNA levels for sample processing with next-generation sequencing. To increase the number of CHD and CI patients included, we improved the PCR protocol described in our previous study (González C et al World J Gastroenterol 2018;24:2095-107) (please see the second paragraph of the subsection *Amplification of HBV and HDV regions of interest by next-generation sequencing*, pages 10 and 11) and we decided to include both HBeAg negative and positive CHD patients (CI patients are HBeAg negative by definition [EASL guidelines on the management of hepatitis B infection, J Hepatol 2017;67:370-98]). Because of that, we were able to include 9 patients in the CHD group. We finally included 7 patients in the CI group, and to maintain a balanced number of patients between groups, we included 8 chronic hepatitis B (CHB) patients in the CHB group. However, inclusion of both HBeAg negative and positive patients in the CHD group led us to question whether their different HBeAg status could be an additional factor affecting their HBV X gene quasispecies complexity, as we had found a more complex viral population in HBeAg-negative than HBeAg-positive CHB patients in the preCore/Core region of the HBV genome in our previous study (Homs M et al. PLoS ONE 2014;9:e112306). Thus, we compared the complexity of the HBV quasispecies in 4 HBeAg negative vs. 5 HBeAg positive CHD patients, which showed no statistically significant differences (please see Table 2, page 37). However, we acknowledge that although the HBeAg status of CHD patients, and the other clinical characteristics of our patients such as progression to cirrhosis or HCC did not

seem to affect HBV quasispecies complexity, the effect of these factors should be confirmed in larger groups of patients, to improve the statistical power of the analysis. We have added this comment to the seventh paragraph in Discussion section, please see page 21.

Regarding the concerns about the limitations of the study *"The section of limitations with comparison with other studies should be expanded."* We were unable to find other studies analyzing HBV quasispecies complexity in HBV+HDV superinfection. However, we found 2 interesting recent studies comparing HBV sequence variation in the surface (S) open reading frame (ORF), encoding HBsAg, between HBV/HDV and HBV infection (Baig S et al. J Med Virol 2018;90:1328-36, and Colagrossi L et al. Viruses 2018;10:E363). Both studies included a larger number of patients than ours, but their patients were recruited from larger populations: in the study by Baig et al. patients were recruited in the highly populated city of Karachi (Pakistan), an area with a higher prevalence of HBV infection among adults than ours (Guidelines for the prevention, care and treatment of persons with chronic hepatitis B infection, WHO, March 2015), and the authors did not specify the patients' clinical stage (i.e. chronic or acute infection). In the study by Colagrossi et al, chronic HDV patients were recruited from 4 different hospitals in Italy and France. The results of these 2 studies are not directly comparable to ours, as the authors did not analyze the HBV quasispecies of each patient, but instead, compared their consensus sequences. However, it is interesting to note that both studies concluded that HDV can exert selective pressure over certain positions of the S ORF, constraining HBV evolution. In the light of these findings, it would be interesting to assess the effect of HDV over the HBV quasispecies in the S ORF and compare it with the effect in other regions of the viral genome, such as that analyzed in the present study. In our view, this illustrates the importance of future studies investigating the HBV quasispecies in HDV superinfected or coinfected patients to deepen the current knowledge about the interference between HDV and HBV. We have added a comment about these studies in the sixth paragraph of the Discussion section, page 20.

Reviewer #2 (Number ID 03020625)

Specific Comments To Authors:

The aim of this study was to analyze the 5' end of the hepatitis B X gene (HBX) coding region and its upstream non-coding region (nt 1255-1611) by next-generation sequencing (NGS) to evaluate HBV quasispecies complexity between chronic hepatitis delta (CHD)-infected patients and chronic HBV-monoinfected patients (CHB and CI). The HBV quasispecies showed a trend to higher complexity in groups with lower viral replication (CHD and CI) than in the higher-replicating CHB patients. The authors proposed two possible mechanisms to explain how HDV can change the HBV quasispecies. There are few questions:

1. The authors performed liver histology examination, the results showed that 2/9 patients with CHD and 3/8 with CHB had liver cirrhosis, and 1/9 patients with CHD had HCC. It would be better to analyze whether the liver histology results could affect the complexity of HBV

First, we must mention that in reviewing the previous version of the manuscript we realized that we expressed the number of patients with liver cirrhosis in an imprecise way. The CHD patient with HCC was one of the 2 CHD patients with liver cirrhosis. Thus, including the 3 CHB patients, a total of 5 patients had liver cirrhosis and 1 of these 5 also had HCC. We apologize if the previous statement caused any confusion, and we have clarified the explanation in the present version of the manuscript please see the *Patient Characteristics* subsection on page 15.

Having clarified this point, we used a Mann-Whitney test to compare the 6 parameters used to determine quasispecies complexity between the 5 cirrhotic patients and the 19 patients without liver cirrhosis, as was described in the first paragraph of the *Statistical Analysis* subsection (please see page 14). This

comparison showed no significant differences, as was reported in the second paragraph of the *Characterization of HBV Quasispecies Complexity* subsection on page 16, and in Table 3, page 38. However, because of the small number of cirrhotic patients included, the relationship between the severity of liver disease and HBV quasispecies complexity should be confirmed in larger groups of patients to improve the statistical power of the analysis. We have added this observation to the seventh paragraph in Discussion section, please see page 21.

2. There was no subgenotype C1 reference sequences in Supplementary Figure 1

The fragment of the HBV genome analyzed in the study (from nt 1255 to nt 1611, 357 nt in length) was extracted from a set of full-length HBV genomes with viral genotype and subgenotype well characterized, and the resulting fragments were used to determine HBV genotype. While this fragment is suitable to establish HBV genotypes A to H (González C et al World J Gastroenterol 2018;24:2095-107) it is too short to determine HBV subgenotypes, for which the ideal approach would be to sequence the entire HBV genome (Pourkarim MR et al. World J Gastroenterol 2014;20:7152-68). Therefore, we aimed at determining only the HBV genotype of the nucleotide haplotypes obtained.

On this basis, we included a sample of genetically diverse HBV genotyping patterns within each of the main HBV genotypes (A to H) in the HBV genomic region analyzed (shown in supplementary Figure 1, please see page 40). With these genotyping patterns we determined the maximum and minimum genetic distances between sequences from the same HBV genotype and between sequences from different genotypes. Thus, the genetic distance of the subgenotype C1 fragment between nt 1255 to 1611 will be closer to other genotype C patterns than to other HBV genotypes, and it will be correctly classified as genotype C even though no subgenotype C1 pattern was included. With the use of this method, the lack of a subgenotype C1 reference would not be a limitation to determine the HBV genotype, and for that reason inclusion of this reference was not contemplated.

3. It would be better to marke the significant difference of nt changes in Fig 3.

As the reviewer suggested, statistically significant changes within each patient group (CHB, CI, and CHD) between proportions of G-to-A and A-to-G, and between proportions of C-to-T and T-to-C nucleotide changes for genotype A and D haplotypes have been marked in Figure 3, please see page 35.

Reviewer #4 (Number ID 00722122)

Specific Comments To Authors:

The manuscript "Characterization of Hepatitis B Virus X Gene Quasispecies Complexity in Monoinfection and Hepatitis Delta Virus Superinfection" is nicely presenting a good scientific work. However it requires some minor corrections or queries to be answered as follows

1. The authors have divided their patients into three groups as HBeAg-negative chronic HBV infection (CI, previously termed inactive carriers), HBeAg-negative chronic hepatitis B (CHB) and with chronic hepatitis delta. The first two groups are essentially the same as with chronic HBV infection as both are HBeAg negative. Based on "previously termed inactive carriers" is insufficient to put them in a separate group. Authors need to elaborate on it.

It is true that patients in the chronic HBV infection (CI) and chronic hepatitis B (CHB) groups are all HBeAg-negative with chronic HBV monoinfection, and this shared characteristic could be used as a criterion to place them in a single group. In fact, we did compare quasispecies complexity in CHB and CI patients grouped together into a single group (HBV monoinfected patients) with the chronic hepatitis delta (CHD) group and no statistically significant differences were found, as commented in the third paragraph of *Characterization of HBV quasispecies complexity* subsection in page 16.

However, HBeAg-negative chronically HBV monoinfected patients from the CHB and CI groups had different levels of viral replication and severity of liver disease, as is shown in the subsection *Patient characteristics* (page 15), in Figure 1 (page 31), and in Table 1 (page 36). These characteristics coincided with those described in the definition of these groups by the European Association for the Study of the Liver (EASL) guidelines on the management of hepatitis B infection (J Hepatol 2017;67:370-98). Taking into account these differences, we thought that it would be interesting to study these patients separately at the virological level. Because of that, we observed that HBV quasispecies complexity in the HBV X gene (*HBX*) 5' region was more similar in CHD and CI than between either group and CHB patients. Considering the essential role of the hepatitis B X protein for HBV replication (please see the third paragraph from the introduction section, pages 8 and 9), these findings suggested to us that the greater *HBX* quasispecies complexity in CHD and CI patients was related to their lower HBV replication levels compared to CHB patients (please see figure 1 page 31 and Table 1 page 36). We believe that these findings are of value to improve our understanding of the mechanisms of interaction/interference between HBV and HDV, and that it was worthwhile to analyze HBeAg-negative chronically HBV monoinfected patients separately as CHB and CI.

2. In genotype section of Data treatment heading "The nt haplotypes aligned at 0.25% ……..1 to 8 obtained from GenBank" is a very long sentence. It should be broken down into smaller fragments for easy comprehensibility.

Thank you for this observation. Indeed, that sentence was very long and we have rewritten it according to the reviewer's suggestions, please see the *Genotyping* subsection (page 13).

3. In statistical analysis section, write "2-sample test" as 2-sample t test.

The name of the test the reviewer is referring to is "2-sample test for equality of proportions with continuity correction", a Chi-square test of equal proportions with correction of continuity. In the present study we used it to compare the proportions of nucleotide changes suggestive of deamination by APOBEC3G (G–to-A and C-to-T) vs. the opposite nucleotide changes (A-to-G and T-to-C respectively) within each different group of patients (CI, CHB and CHD). Thus, we believe the name of the test is correctly written in the manuscript.

4. In the result section, authors have mentioned that 6 patients of total had liver cirrhosis or HCC. This may have an impact on the results. The authors need to analyze/discuss these variable bias.

In fact 5 (not 6) patients showed liver cirrhosis or HCC. The CHD patient who showed HCC was one of the 2 CHD patients who showed liver cirrhosis. Thus, including the 3 CHB patients, a total of 5 patients showed liver cirrhosis. We regret that this information was imprecisely expressed, and we have clarified it in the present version of the manuscript please see *Patient characteristics* subsection in page 15.

To compare the 6 parameters used to determine quasispecies complexity between the 5 patients with cirrhosis and 19 patients without cirrhosis we used a Mann-Whitney test, as was mentioned in the first paragraph of *Statistical Analysis* subsection (please see page 14). This comparison showed no significant differences in quasispecies complexity between these groups, as reported in the second paragraph of the *Characterization of HBV Quasispecies Complexity* subsection on page 16, and in Table 3, page 38. In further reference to these findings, we have added sentence to the seventh paragraph of the Discussion (please see page 21), commenting that the relationship between the severity of liver disease and HBV quasispecies complexity should be confirmed in larger groups of patients to improve the statistical power of the analysis.

5. In discussion, write full form of "ADAR-1" and why in particular this enzyme mutation should be investigated. Any reference?

The reason why we thought that it would be interesting to investigate HBV edition by ADAR-1 was that according to recent studies (Suárez-Amarán L et al. J Hepatol 2017;37:669-79 and Giersch K J Hepatol 2015;63:346-53), ADAR-1 expression is more enhanced in HBV/HDV infected than in HBV monoinfected mice models. Despite these findings, we acknowledge that HBV-RNA edition by the double-stranded RNA-editing enzyme ADAR-1 has not been demonstrated; hence, it could be too speculative to suggest edition of HBV pregenomic RNA by this enzyme. Therefore, we have eliminated the suggestion to study ADAR-1 in the third paragraph of "Discussion" section (please see page 19).

6. HBsAg is not reported as log IU/mL. Therefore it is incorrectly written in table 1. Please revise or provide reference.

We determined the quantitative HBsAg using a commercial electrochemiluminescent immunoassay (ECLIA) on a COBAS 8000 instrument (Roche Diagnostics, Rotkreuz, Switzerland), as explained in the *Serological and Virological Determinations* subsection, page 10. According to the data sheet of this assay (Elecsys HBsAg II Quant assay), the results were standardized against the NIBSC standard (code number: 00/588; WHO Second International Standard for HBsAg, subtype adw2, genotype A). Therefore, the results are expressed in international units per milliliter (IU/mL). To simplify the presentation of the results in Table 1, we converted them to their decimal logarithm (ie, the decimal logarithm of IU/mL or logIU/mL). Expressing HBsAg quantification as logIU/mL is a common way to represent the results used in other scientific publications in viral hepatitis (Pfefferkorn M et al. Gut 2018;67:2045-2053, Zoulim F et al. J Hepatol 2015;62:56-63, Boglione L Liver Int 2013;33:580-5).