Dear Editor,

Please find enclosed the edited manuscript in Word format (file name: 16083-review.doc).

Title: Distribution and Character of Chinese Hepatitis B Virus

Author: Hong-Mei Li, Jian-Qiong Wang, Rui Wang, Qian Zhao, Li Li, Jin-Ping Zhang, Tao Shen.

Name of Journal: *World Journal of Gastroenterology*

ESPS Manuscript NO: 16083

The manuscript has been improved according to the suggestions of reviewers:

1 Format has been updated

2 Revision has been made according to the suggestions of the reviewer

Reviewer 1

(1). I suggest to short the Discussion section.

Revised: Delete the following sentences

For instance, genotype A has been shown to be primarily distributed in Northern Europe and Africa; genotype B and C in Southeastern Asia; genotype D in the Middle East, North Africa, and Europe; genotype E in Africa; and genotype F in South America. (page 7 line 14)

Delete: Moreover, it has been reported that, in human liver cancer cell lines (HepG2 and Hep3B) the wild-type HBX protein could activate the tumor suppressor P^{21}, and consequently inhibit HBV DNA replication and cell proliferation [50-55]. The BCP region double mutation diminished this function, resulting in the aggravation of HBV. (page 9 line 6)

Delete: The pre-C mutation has been shown to lead to the mutation of the 28th amino acid residue in hepatitis B e Antigen (HBeAg) from tryptophan (TGG) to a stop codon (TAG), in a human liver cancer cell line (Huh7) the HBeAg protein synthesis is consequently terminated, leading to highly replication HBV strain [59-61]. (page 9 line 21)

Delete: For instance, E. coli has a start codon in the forms of ATG, GTG, TTG, ATT and CTG [65], while a mitochondrial start codon can be ATG, ATA or GTG [66-68].

Additionally, Kadowaki et al [69] suggested that a cytochrome oxidase subunit from tomato mitochondria I employed ACG as the start codon for RNA editing. (page 10 line 16)

Reviewer 2

(1) The title should be more specific, e. g.: Hepatitis B Virus Genotypes and Genome Characteristics in China.

Revised: Hepatitis B Virus Genotypes and Genome Characteristics in China (page 1 line 1)

(2) Key-words should also include: Hepatitis B virus (HBV) and genotypes?

Revised: Hepatitis B virus (HBV), genotypes, (page 3 line 2)

(3) In the first sentence, where the name of the viral family in Latin is mentioned, it should be written with the capital letter and in italic?

Revised: *Hepadnaviridae* (page 3 line 6)

(4) In the sentence where first discovered genotypes were discussed ("HBV can be divided into 4 genotypes (A, B, C, and D)"), it should be emphasized that this was the distribution at the given moment, since 6 more genotypes were discovered later?

Revised: HBV genotype A has been shown to be primarily distributed in Northern Europe and Africa; genotype B and C in Southeastern Asia; genotype D in the Middle East, North Africa, and Europe. With technological development, more HBV genotypes have been found, genotype E in Africa; genotype F in South America; genotype G in USA and France; and genotype H in Europe and North America. (page 3 line 21)

(5) "Previous studies have indicated that HBV genotypes might associate with serotype, liver disease progression and mutation" – should not be with just any mutation but with more than one and specific mutations?

Revised: Previous studies have indicated that HBV genotypes might associate with serotype, liver disease progression and mutation like BCP and pre-C region. (page 3 line 30)

(6) The aims of the study should be defined more clearly. Numbering the aims in order of importance would be helpful.

Revised: Therefore, in this study, the distribution of genotypes and subgenotypes in China were analysed firstly. Then, the characters of HBV sub-genotype B2 and C2 were analysed. Finally, the correlation between BCP double mutation/pre-C mutation and clinical symptoms were also determined. (page 3 line 30)

(7) It is said that sequences from patients who were just HBsAg positive were excluded but then, at the end, HBsAg-positive group is listed. Also, what would be the difference between patients who are just HBsAg positive and those who have asymptomatic hepatitis? Patients without symptoms are most often just carriers and can be recognized only by presence of HBsAg and anti-HBc IgG. It is also possible that patients without symptoms have active chronic hepatitis (verified by liver biopsy and presence of HBV DNA) but then that should be clearly indicated.

Revised: HBsAg positive patients may have a variety of clinical symptoms, such as CHB and ASC. That is to say, HBsAg positive no clear annotation of clinical symptoms, therefore, we excluded it in clinical symptoms analysis. (page 5 line 1)

(8) What is “chronic to acute hepatitis B” defined in the paper as ACHB? It should be defined more clearly or excluded from the study.

Revised: Based on the NCBI databases, chronic to acute hepatitis B should be annotated as acute-on-chronic liver failure. Therefore, we suggested that acute-on-chronic liver failure was abbreviated as ACLF. (page 4 line 28)

(9) Recombinant genotype should be marked C/D and not C/d?

Revised: C/d has been amended as C/D in manuscript. (page 2 line 21)

(10) In Table 1, two different genotyping methods were compared to genotypes annotated in NCBI. These two methods are defined in the Table’s legend but should be clarified in the Table too?

Revised: We have had add these two methods to Table 1. (page 20 line 25)

(11) In the sentence where it is explained that genetic distance between isolate FJ386674 and other sub-genotypes was more than 4%, the brackets should contain the genetic distance of this isolate compared to other subgenotypes and not the overall distance between B1-B9. ?

Revised: About this question, we have specified in table 2. In text, furthermore, we

revised as the genetic distance between FJ386674 and genotypes (A,C-J) were more than 8%，but the genetic distance between FJ386674 and other B sub-genotypes were more than 4% (0.05±0.00-0.07±0.01) and less than 8% (Table 2). (page 6 line 4)

(12) The certain mutations can be more frequently present but not “elevated” in some clinical conditions?

Revised: “Elevated” has been amended as more frequently present in text. (page 6 line 16)

(13) What does GP in the phylogenetic tree stand for? It should be defined in the legend.

Revised: GP is abbreviation of group, we have annotated in figure2. (page 24 line 3)

(14) The following paragraph is unclear: “This study showed BCP double mutation was significant differences in several hepatitis symptoms (ASHB vs. CHB, and ACHB vs. HCC). The results suggested that severe liver disease had a lower mutation rate in the BCP region as compared with that of CHB, which is consistent with previous report [32].”

Revised: This study showed BCP double mutation was significant differences in ACLF and HCC. The results suggested that HCC had a lower mutation rate in the BCP region as compared with that of ACLF, which is consistent with previous report. (page 9 line 3)

(15) It was not defined what was meant by term “severe liver disease”.

(1) Revised: Severe liver disease represented some diseases, it stand for ACLF, HCC and LC in this study, we have modified according to the text. Such as: the mutation rate of pre-C in ACLF, HCC and LC were higher than that in CHB. (page 9 line 19); this may explain why patients infected with genotype C HBV are more susceptible to the development of ACLF, HCC or LC. (page 9 line 30)

(16) It would be useful to repeat, in short, the major findings of the study in the last paragraph.

Revised: C2 and B2 were identified as the two major sub-genotypes in China. FJ386674 might be a new sub-genotype as B10. The major stop codon of S-ORF were TAA (92.2%) and TGA (79.65%) in B2 and C2 subtype, respectively. (page 11 line

19)

3 References and typesetting were corrected

Thank you again for publishing our manuscript in the *World Journal of Gastroenterology.*

Yours sincerely

Li Hongmei

Corresponding author:

Name: Shen Tao

E-mail: ts902@126.com

Dear Editor

Please find enclosed the edited manuscript in Word format (file name: 16083-review.doc).

Title: Hepatitis B Virus Genotypes and Genome Characteristics in China

Author: Hong-Mei Li, Jian-Qiong Wang, Rui Wang, Qian Zhao, Li Li, Jin-Ping Zhang, Tao Shen.

Name of Journal: *World Journal of Gastroenterology*

ESPS Manuscript NO: 16083

The manuscript has been improved according to the suggestions of Journal Editor-in-Chief:

1. Add page and line numbers.

Revised: We have added page and line number in manuscript.

2. Hepadnaviridae should be in italics.

Revised: *Hepadnaviridae.* (page 3 line 2)

3. The lates figures for CHB is 240 million not 400 million see WHO fact sheet: WHO. Hepatitis B Fact sheet no 204. 2014 February 2015]; Available from: http://www.who.int/mediacentre/factsheets/fs204/en/.

Revised: We have modified the number of CHB according to the WHO fact sheet. (page 3 and line9)

4. Subtype refers to serological subtypes and not to subgenotypes. Please correct throughout the manuscript. Please refer to Kramvis, A., M. Kew, and G. Francois, Hepatitis B virus genotypes. Vaccine, 2005. 23(19): p. 2409-23 for definitions.

Revised: Subtype correct to subgenotypes according to the reference (Kramvis, A., M. Kew, and G. Francois, Hepatitis B virus genotypes. Vaccine, 2005. 23(19): p. 2409-2423).

5. The first sentence under paragraph on start and stop codons does not make sense. Abandoned the incomplete sequences (< 3215 bp), the start and stop codons from 197 B2 genotype sequences and 439 C2 genotype sequences were analyzed. Please correct.

Revised: Abandoned the incomplete sequences (< 3215 bp), 197 B2 subgenotype

sequences and 439 C2 subgenotype sequences were analyzed with the start and stop codons. (page 5 and line 23)

6. More than one isolate is required in order to classify a strain as a new subgenotype. Refer to Kramvis, A., Genotypes and genetic variability of hepatitis B virus. Intervirology, 2014. 57(3-4): p. 141-50.for latest genotype/subgenotype classification criteria and guidelines.

Revised: Based on the reference (Kramvis, A., Genotypes and genetic variability of hepatitis B virus. Intervirology, 2014. 57(3-4): p. 141-150. page 148 line 12) and previous reports of classification of subgenotype (C9: AP011108, C10: AB540583, C15: AB644286, C16: AB644287), we changed "FJ386674 as a new subgenotype B10" to" FJ386674 as a *putative* subgenotype B10" (page 7 line 19),

7. What is meant by the following sentence: Furthermore, some study suggested that the correlation between the PreS2 start codon mutation and hepatocirrhosis was either liver cancer progression or active DNA replication [65].

Revised: Furthermore, some study suggested that the PreS2 start codon mutation might be related with liver cancer progression or active DNA replication [65]. (page 9 line 17)

8. Correct the following sentences: We further analysis the termination codon usage bias in genotype B2 and C2 by RNAdraw software to predict RNA secondary structure of the HBsAg protein.

Revised: We further analysis the termination codon of subgenotype B2 and C2 by RNAdraw software to predict RNA secondary structure of the HBsAg protein. (page 10 line 5)

9. Subtype B2 with TAA termination codon might be more stable than subtype C2 with TGA termination codon because of base paired. Codon usage bias is the disequilibrium phenomenon of synonymous codon usage, which encoding of the same kinds of amino acids in biological.

Revised: The RNA secondary structure of subgenotype B2 with TAA termination codon might be more stable than subgenotype C2 with TGA termination codon because of base paired. (page 10 and line 12)