Salem Medical Center • Dept. of Medicine • Zeppelinstr. 11-33 • D-69121 Heidelberg

SCIENTIFIC EDITOR
WJH
Ze-Mao Gong
8226 Regency Drive
Pleasanton, CA 94588,
USA
Fax: +1 925 223 8243

Prof. Dr. med. Sebastian Mueller, MD, PhD
Professor of Medicine
Vice Head and Research Director
Department of Internal Medicine, Salem Medical Center and
Center for Alcohol Research, University of Heidelberg
Zeppelinstraße 11 – 33
69121 Heidelberg, Germany

Tel.: 06221 483 210
Fax.: 06221 483 494
sebastian.mueller@urz.uni-heidelberg.de
www.sebmueller.com

Heidelberg, 12.07.2016

Dear Prof. Gong,

In response to our earlier correspondence, we hereby send you our revised manuscript entitled

'Primary liver injury and delayed resolution of liver stiffness after alcohol detoxification in heavy drinkers with the PNPLA3 variant I148M'

accepted for publication in World Journal of Hepatology as BASIC STUDY.

As requested by you, we have incorporated all revisions suggested by the peer-reviewers that are cited in the following point-by-point response. Changes are highlighted in red in the updated version of the manuscript. Specifally, the following points have been addressed:

1. According to the reviewer's request and for better readability, we have marked significance levels in the tables by asterisk and highlighted in bold.
2. In addition, Fig. 2 has been improved and completed.
3. The major key messages are better visible both in title and abstract.

Sincerely yours,

Sebastian Mueller

Sebastian Mueller, MD, PhD
Salem Medical Center, University of Heidelberg
Zeppelinstraße 11 – 33, 69121 Heidelberg, Germany
Tel.: 06221 483 210
Fax.: 06221 484 494
Email: Sebastian.Mueller@urz.uni-heidelberg.de

Point-by-point response:

Reviewer 1:

1. In the results, the authors stated "Mean LS was lowest in CC carriers (13.1 kPa) as compared to CG and GG carriers (both 17.6 kPa)". But in table 2, mean LS if CC, CG and GG carriers were 13.1 kPa, 17.6 kPa and 17.2 kPa repectively. The authors did not show any statistical difference between CC and GG group. The authors need to make the statement clear and precise.

We thank for this important comment. The typo in the abstract has been corrected. We also improved table 2 for better readability that now clearly shows the statistical significance of CG versus CC and GG versus CC. In order to improve the readability we now use asterisks and bold highlighting for levels of significance.

2. In the results, the authors stated "After alcohol withdrawal, LS significantly decreased in CG-carriers from 17.6 to 12.7 kPa but not significantly in CC carriers. Moreover, LS decreased to a lesser extent in GG carriers from 17.6 to 14.5 kPa due to prolonged resolution of inflammation". But in fig. 2A, the LS decrease in GG group was not statistically significant. The statement is not adequate.

The reviewer correctly cites our formulation. LS decreased indeed to a lower extent and non-significantly. However, AST levels were significantly higher after alcohol detoxification in GG carriers (see. Tab. 2 and Fig. 2B). Together with the histological findings this suggests to us that the prolonged inflammation/ballooning in GG carriers contributes to LS elevation. This is now better demonstrated in Fig. 2A and B.

3. In both the results and discussion the authors speculated that LS decreased to a lesser extent in GG carriers was due to "prolonged resolution of inflammation", but they did not explain why LS did not significantly decreased in CC carriers, whose transaminase level significantly reduced after stop alcohol drinking.

We would like to thank for this important comment. Indeed, a striking feature of the CC carrier is the fast resolution of the transaminase levels while almost no changes of LS are observed in response to alcohol withdrawal. This is in line with recent observations (Mueller et al, Liver international, 2016). According to this study, only a minority of ca. 30% with elevated transaminase levels shows a decrease of LS after alcohol withdrawal. It suggests to us that inflammation not necessarily causes LS elevation in all patients. Tab. 4 may offer an explanation with regard to the histological findings. Here, the most significant histological finding with the GG carrier was ballooning/steatohepatitis. As cited in the main text, however, AST did not correlate with GG status. In other words, ballooning seems not to translate directly and tightly into higher transaminase levels. This discrepancy between serum markers and

histology has been known for many years. Moreover, and as shown in this paper, liver stiffness does not correlate with steatosis. We therefore have taken up this important comment and we have revised the discussion section accordingly (highlighted in red).

4. In the conclusion, the authors stated "In heavy drinkers, PNPLA3 GG primarily correlates with liver damage but not steatosis". However, this may not be true for all heavy drinkers. Other factors more than PNPLA3 genotype may cause liver damage in heavy drinkers. This statement is only true in particular group of patient. The definition of this particular group should be well defined.

As shown in the patient section, we have indeed a quite homogenous cohort of German heavy drinkers and the major goal of the present study was to find associations with the PNPLA3 genotype. We certainly agree that there are additional unknown (genetic) factors but our sound statistical approach clearly reveals a tighter association of GG with ballooning/steatohepatitis but not steatosis. Moreover, no significant differences with regard to steatosis as measured by ultrasound were observed between the genotypes (table 2).

5. In the conclusion, the authors stated "sustained LS elevation could be a major risk factor in PNPLA3 GG carriers". Dose they mean "GG carrier is a major risk factor for sustained elevated LS value".

We refer indeed to the recently introduced pressure hypothesis (reference 24) that predicts pressure (inflammation-associated) LS elevation as important causative condition for fibrosis. The sentence has been modified accordingly.

6. In the materials and methods, the authors included 521 alcoholic liver disease patients in the study. Subgroup analysis were performed according to the method of evaluation (non-invasive e.g. transient elastography, n=503 and histological, n=80). It means 62 patients received both examinations. The correlation between the evaluation results may be performed in these 62 patients. The comparison made between the whole cohorts is complex and may be misleading.

We have corrected the typos. Non-invasive assessment was done in 'all' 521 patients. Comparison between histology and TE was performed in 80 patients.

7. In the results, line 10, the authors stated "CC carriers represented 42.1% of the F0 cohort but 35.5% of the F4 cohort". They did not describe by which evaluation method were the results obtained.

This comment is not clear to us. We simply describe the distribution of genotypes in percentage as is routinely done in comparable genetic studies.

8. GC carriers and GG carriers are sometimes discussed together, and sometimes separately. The analysis makes the result complex and not easy to understand. The authors did not explain why they had to perform the analysis in this way.

We understand this helpful comment by this reviewer. However, PNPLA3 studies typically analyze the various genotypes independently since CG may lead to less severe phenotypes but is more abundant. In contrast, the GG carrier is rather rare but shows a more pronounced phenotype. As shown in Table 2, the independent analysis reveals additional information that could be missed otherwise.

9. In the results, the authors stated "In summary, GG-associated liver damage results in a reversible, inflammation-associated increase of liver stiffness. In addition, GG carriers show a slower resolution of liver damage and LS after withdrawal from alcohol". This statement is not exactly true according to figure 2B, which did not showed significant resolution of LS in GG group.

We would like to thanks for this important comment. As stated above we have added the statistical findings on LS and AST levels after detoxification in figure 2. Indeed, after detox, LS and AST are highest in GG carriers.

10. The results showed LS is highly associated with fibrosis stage, but the genotype distribution of F4 cirrhosis evaluated with fibroscan and histology are different. The authors did not explain the difference clear enough. The conclusion is not fully supported by the results. The statement should be revised.

The mentioned observation is interesting but not surprising. It simply demonstrates the bias of most histological studies that, for ethic reasons, require the patient consent for the invasive and potentially complicating liver biopsy. We clearly believe that our study demonstrates that recruitment of liver biopsy studies is biased by recruiting more severely ill patients. With the non-invasive approach this bias disappears showing a much higher percentage of F0 population. This important point is now taken up in the discussion section.

Reviewer 2:

Rausch et al. analyzed the influence of PNPLA3 genotype in heavy drinkers on serum markers and liver stiffness (LS) during all stages of alcoholic liver disease (steatosis, steatohepatitis, fibrosis) prior and after alcohol detoxification. This is a study of great interest that can help the researchers in evolving in this field. However, some minor points could be addressed.
1. The methods are incomplete and it needs to be implemented. It is not clear the study design and the inclusion and the exclusion criteria. How much time has passed after alcohol withdrawal?

Exclusion and inclusion criteria are more clearly described in the methods section. The time of alcohol withdrawal is mentioned and also shown in Fig. 2.

3. Please discuss limitations of the study, taking into account sources of potential imprecision.

Potential limitations are now taken up in the discussion section.

Reviewer 3:

1. Introduction: "However, twin studies, the enhanced sensitivity of female drinkers and the fact that only a minority of patients progress to cirrhosis despite heavy drinking clearly suggest a genetic predisposition". The enhanced sensitivity of female drinkers could also relate tot other (non-genetic) factors such as different gastric alcohol dehydrogenase, higher body fat, changes in alcohol absorption with menstrual cycle.

We revised the introduction accordingly.

2. Methods: “Other causes of liver diseases were ruled out serologically by screening for AMA, ANA, HCV and HBV.” How many cases were excluded? Were these tests performed in all patients?

These tests were performed in all patients and less than 20 patients had to excluded due e.g. presence of HCV.

3. Were all eligible patients attending the clinic consecutively included?

No, only patients who met inclusion criteria and who were compliant with the time schedule of measurements could be included.

4. 80 patients had liver biopsy. Which criteria determined the decision to perform this procedure? Probably this bias can explain significant part of the results and discrepancy between biopsy and Fibroscan.

Agreement was comparable to other centers. Patients with less pronounced disease were less willing to agree for liver biopsy. We discuss these issues now in the discussion section as mentioned above (reviewer 1).

5. Mean biopsy length was 15.6 mm. This could lead to significant underestimation of biopsy length. This limitation should be mentioned in the discussion (see ref. Bedossa, Hepatology 2003).

We do not completely understand this comment. Our biopsy length is typical for such studies and the sampling error of the biopsy is well known. But how should it underestimate the biopsy length? If the comment refers to the classically known rather high sampling error of the liver biopsy, we completely agree. However, the sampling error is especially relevant for the individual patients but levels out with increasing number of patients.

6. Methods: fibroscan was performed with M probe. How many patients were unsuccessfully measured? Was X-L probe available?

Successful Fibroscan measurement was an inclusion criterion for the study. The XL probe was also available at our medical center and performed in 5 patients. References have been added.

7. 80 patients apparently had fibroscan as well as liver biopsy. What was the time period between both investigations? What was the correlation of results with 2 methods in the same patients?

As mentioned above the time interval is provided in the method section and in Fig. 2.
Our analysis in the paper is provided for both methods (e.g. Table 4).

8. Fig 2: detailed info shoud be presented on the time periods elapsed between repeated fibroscan measurements resp repeated ast measurements in the patients of the various PNLPLA3 subgroups.

Fig. 2 provides the time intervals of alcohol withdrawal, which corresponds to the time of measurement.