Author’s response to reviews

Title: Analysis of the Relationship between MIR155HG Variants and Gastric Cancer Susceptibility

Authors:

Wenjing Zou (zouwenjing84@163.com)
Xu Li (lixu0204@sohu.com)
Cheng Li (licheng1125@163.com)
Dan Liu (danliu600@126.com)
Yanyan Lv (lyywjlxj@163.com)
Ying Yang (273786660@qq.com)
Nan Ye (276000612@qq.com)
Dan Guo (1184795810@qq.com)
Shuixiang He (shuixiang__he@163.com)

Version: 1 Date: 27 Nov 2019

Author’s response to reviews:

Dear Editors and Reviewers:

Thank you very much for your constructive and positive comments. Those comments are all valuable and very helpful for revising and improving our paper, as well as the important guiding significance to our researches. We appreciate the seriousness of the reviewers, which is worth learning in the future. I am very sorry for our lack of seriousness. Here we have studied comments carefully and tried our best to revise the manuscript to strengthen our paper. The edits are marked in red in the revised manuscript. I really hope the revised version can be approved.

Our responses to these comments and the details of these point-by-point responses are enumerated below.

Sincerely,
Professor Shuixiang He
Address: #227 West Yanta Road, Xi’an 710061, Shaanxi Province, China
Tel/Fax: +86-18991232026
E-mail: shuixiang__he@163.com
Reviewer reports:
Kimberly J. Perez, MD (Reviewer 1): This is a well constructed endeavor. Your methods section was very clear.

I would recommend that you tighten up your discussion. Your data suggests a correlation between this particular SNP and gastric cancer risk by the genotype and recessive models in men in your cohort. You don't address this specifically. I anticipate that is because there is no corroborating data available by your group or others but I think you should address this.

You also make the assertion that genetic variation in MIR155HG influences the transcriptional product. What is this based on?

Response: Thank you very much for your comments.
We revised the discussion section. In the discussion section, we explain the above problems. However, inevitably, we considered that there are some limitations in this study. This study lacks functional validation, and we will elaborate on these questions in subsequent studies.

In the discussion section, we supplement and modify as follows:“For many years, intron sequences have been considered essentially non-functional. However, subsequent study showed that intron-containing genes presented higher levels of transcription when compared to intron-less genes in mammalian cells(22), suggesting that introns may be enhancers of transcription. In addition, we hypothesized that the genetic variation of MIR155HG gene may influence the expression of its transcription product miR-155, thus affecting its function in tumors. Stratified analysis found that MIR155HG-rs11911469 was associated with an increased risk of gastric cancer only in the male group. This result may be due to the presence of some gender-specific genes in the target genes of miR-155. But these internal connections need to be verified by subsequent functional experiments.”

Jose Saenz, M.D., Ph.D. (Reviewer 2): In this manuscript, Zou et al. describe a case-control study looking at the correlation between genetic variants of a miRNA, MIR155HG, and gastric cancer. In particular, they analyzed single nucleotide polymorphisms (SNPs) within this gene using a MassARRAY platform to identify particular SNPs that correlated with either increased or decreased gastric cancer risk. While I found that the overall approach of attempting to identify gastric cancer susceptibility loci is worthwhile, I believe that, in general, the authors did not adequately describe their methods and seemed to overstate certain conclusions. Some of their association analyses failed to show correlations between certain SNPs, which decreased the overall impact of the manuscript. In addition, it remains unclear whether some of their conclusions can be applied to different populations aside from the Han population in China. Perhaps the biggest weakness is that it remains to be seen how clinically and physiologically relevant the SNPs identified in this study are to gastric cancer. The authors make little to no mention of the possible functions or genetic targets of MIR155HG and how these might contribute to or prevent the progression to gastric cancer. Some additional comments are listed below:
* The introduction of the various inheritance models seems problematic to me. That the χ2 test showed no statistical difference in the frequencies of any of the SNPs in the MIR155HG gene between the gastric cancer cases and controls makes it seem that the authors needed to turn to another statistical test to find an association. The authors rely on a previously published inheritance model but do not introduce the model or explain why they chose this model. What are the limitations of this model, and why would it better explain the association than the χ2 test? This model appears to be central to
validating the association between the SNPs and gastric cancer, so it needs to be discussed in more detail, and a more rigorous explanation of its utility needs to be made.
Response: Thank you very much for your comments.

(1) First, I explain the statistical methods used in this study. In case/control association studies, the \( \chi^2 \) test was used to assess allelic associations, traits and allelic associations. The odds ratios (ORs), 95% confidence intervals (CIs) and p-values were calculated with the logistic regression model to assess the associations between genotype and risk of gastric cancer.

(2) I explain the model analysis in this study. Four models (genotype, dominant, recessive, and additive) analysis is the most commonly used method in association analysis to analyze disease risk in individuals with different genotypes and different genetic models. The four genetic models are based on Mendel's law of heredity, which deal with the pathogenicity of heterozygotes and homozygotes in isolation or in combination. On the basis of this study, we were able to know which genetic model the individual had a higher risk of gastric cancer. Therefore, it can provide a certain theoretical basis for the advance prevention and individualized diagnosis of gastric cancer.

(3) Inevitably, this study has some limitations. We also discuss the limitations of the study at the end of the article (p.12, lines 2-8). Cases are often not representative of all cases, nor are they as usual representative of the target population and thus prone to selection bias. We did our best to avoid crowd bias in choosing participants. All the individuals were genetically unrelated.

(4) Finally, as you mentioned, I refer to similar association analysis studies published earlier. I mainly refer to the following articles:


* The abstract presents too many loci and SNPs and no clear explanation of how these might contribute to gastric cancer. Along these lines, on p.8, lines 45-53, how are the functions of these SNPs predicted and/or validated? Are there any references to substantiate these hypothetical functions? In addition, it seems like the functions listed are fairly broad and non-specific (e.g., protein binding).
Response: Thank you very much for your comments.
We modified the abstract to highlight significant SNP sites.
HaploReg v4.1 (https://pubs.broadinstitute.org/mammals/haplogreg/haplogreg.php) was employed to predict the potential functions of the candidate SNPs.
HaploReg v4.1 is an online software. We used HaploRegv4.1 to predict the potential functions of the candidate SNPs. However, the prediction information we obtained from the analysis of this software is limited and only includes some general SNP potential functions. This information provides new ideas for future functional research.
In order to clarify the exact functions of candidate SNPs, future functional experiments will be the direction of our efforts.

* In Table 1, what does "absence" mean for staging and LNM?
Response: Thank you very much for your comments.
"Absence" means that we were unable to collect information about these patients. That is, missing information. It may be due to the patient's own memory deviation or case record confidentiality.
* On p.8, lines 17-23, the authors claim that "the gender distribution was statistically different between the case and control groups (p<0.05)." Is this accurate? They look nearly identical, certainly not different enough for a p value of 0.001, as shown in Table 1.
Response: Thank you very much for your comments.
We recalculated the gender distribution difference between case and control.
I am very sorry for our lack of seriousness.
The updated p value is 0.970.
We modified the statements in the manuscript (p.8, lines 17-23) and the p values in the Table 1.

* Why are only two SNPs shown in Table 3? How did the inheritance models fare for the other SNPs shown in Table 2?
Response: Thank you very much for your comments.
In the current study, there were no significant results for the remaining SNPs in the four genetic models. We are trying to collect more samples to verify these results.

* On p. 9, lines 9-23, the authors claim that the identified SNP genotypes either increase or decrease the risk of gastric cancer. This statement, as it is worded, is not accurate. The SNP genotypes are associated with, or correlate with, an increased/decreased risk of gastric cancer, but the SNPs themselves do not increase or decrease the risk of gastric cancer. No functional validation of these SNPs was done in this study, so the authors cannot conclude that these SNPs increase or decrease the cancer risk.
Response: Thank you very much for your comments.
I am very sorry for our lack of seriousness. We modified the inappropriate wording.
On the p. 9, lines 9-23, we modify as follows: "The result indicated that the “CC” genotype of rs4143370 was associated with a decreased risk of gastric cancer in genotype model and recessive model. Inversely, the “CC” genotype of rs1893650 was associated with an increased risk of gastric cancer in genotype model and recessive model."

* To appeal to a broader audience, the authors may want to briefly discuss the differences between the dominant, recessive, and additive models. What does it mean for some genotypes to be correlated to gastric cancer in one model but not the other? Does this have clinical significance?
Response: Thank you very much for your comments.
The models used in PLINK said if A is minor allele, G is wild-type allele: dominant: GG Vs AG+AA; recessive: AA Vs AG+GG; genotypic: AA vs AG vs GG; additive: AA VS GG.
Four models (genotype, dominant, recessive, and additive) analysis is the most commonly used method in association analysis to analyze disease risk in individuals with different genotypes and different genetic models. The four genetic models are based on Mendel's law of heredity, which deal with the pathogenicity of heterozygotes and homozygotes in isolation or in combination. On the basis of this study, we were able to know which genetic model the individual had a higher risk of gastric cancer. Therefore, it can provide a certain theoretical basis for the advance prevention and individualized diagnosis of gastric cancer.

* Haplotype association analysis showed no association between specific haplotypes and risk of gastric cancer, significantly weakening the authors' conclusions.
Response: Thank you very much for your comments.
Our results show that there is no significant association between these haplotypes and gastric cancer susceptibility. However, haplotype results did not affect the association between a single SNP site and
gastric cancer risk. Genotype and haplotype are different concepts. Haplotype is a combination of two or more polymorphism sites on a chromosome.

* Are these SNPs specific to the Han population? Were other Chinese ethnicities studied or used as controls?
Response: Thank you very much for your comments. These SNPs are not unique to Han Chinese. The genetic pattern of each race is different, and this study was conducted only in Han Chinese. These SNPs have not been studied in other ethnic groups. We will try to collect more abundant samples to verify the conclusion in subsequent studies.

* No validation of the effects of these SNPs on miR-155 levels was demonstrated.
Response: Thank you very much for your comments. The focus of this study was to clarify the association between gene polymorphism and gastric cancer risk. Inevitably, this study had some limitations, and we did not confirm the level of mir-155. The levels of mir-155 in the blood and tissues were different. In this study, only blood samples were collected from volunteers, and we are actively collecting a large number of tissue samples. In subsequent studies, we will identify the level of mir-155 in tissue samples and further study the function of MRI155HG gene.

* Minor edit: On p.11, line 50, it should read "gastric cancer," not "CRC (colorectal cancer)."
Response: Thank you very much for your comments. I am very sorry for our lack of seriousness. “CRC” is changed to “gastric cancer” in Page 11 line 50.

Ulrich Peitz, M.D. (Reviewer 3): This is an important study on the relationship between gastric cancer and SNPs of a certain microRNA involved in gene Regulation. One strength of the study is its large sample of gastric cancer patients and controls.

I have some suggestions concerning wording:
Page 5 line 26
Tending
Tend to or are tending to
Page 8 line 17
The gender distribution was statistically difference
different
Page 8 line 33
was showed
shown
Page 10 line 28
were found significantly affect
were found to significantly….
Page 11 line 50
CRC
What does this stand for?
Response: Thank you very much for your comments. I am very sorry for our lack of seriousness. Here we tried our best to revise the above problems in manuscript.
Lastly, in Page 11 line 50 “CRC” is changed to “gastric cancer”.
The edits are marked in red in the revised manuscript.