Author’s response to reviews

Title: Interaction between obesity and the Hypoxia Inducible Factor 3 Alpha Subunit rs3826795 polymorphism in relation with plasma alanine aminotransferase

Authors:

Shuo Wang (wangshuo_20080512@126.com)
Jieyun Song (songjieyun1983@126.com)
Yide Yang (yangyide2007@126.com)
Yining Zhang (zyn1221medical@163.com)
Nitesh Chawla (nchawla@nd.edu)
Jun Ma (majunt@bjmu.edu.cn)
Haijun Wang (whjun1@bjmu.edu.cn)

Version: 1 Date: 08 Jun 2017

Author’s response to reviews:

The response to reviewers' comments can also be found in the cover letter.

Reviewer reports:

Andrew Walley (Reviewer 1): I have a few points about the manuscript that need to be sorted out before it should be considered acceptable for publication:

* The only new laboratory work in this manuscript appears to be the SNP typing. However, the last paragraph of the introduction suggests the methylation data is new as well when it looks like it is from the authors' previous publication in PLoS ONE 2015. This should be much clearer in all parts of the manuscript. For example, the DNA methylation method should not be included. The methods do point the reader to the previous paper but only for the study design when it looks like this is exactly the same dataset.

Thank you for your suggestion. For the HIF3A DNA methylation data, we did use the same dataset as the previous paper. As you suggested, we changed the text in the last paragraph of the introduction to ‘In this paper, we perform genotyping for the HIF3A rs3826795 polymorphism among 2030 Chinese children aged 7-18 years old was performed, and also detected the HIF3A methylation data of among 220 children detected in our previous study [15] was also used.’. In the Method section- DNA methylation detection, we pointed to the previous study in the first
sentence ‘Details of the methylation examination were described previously [15]’. In the first paragraph of the Result section, we also used the following content to point to the previous study: ‘Details of the characteristics of the 110 extremely obese children and 110 matched controls with methylation examination were described previously [15].’

* Equally, in this paper the DNA methylation data is not described in sufficient detail. For example, it is not possible to know what the distribution of methylation is for the cases and controls. This is presumably because it has been described before and, if so, the authors should cite the previous work in the results to ensure the reader can investigate this dataset further. If it is not the same dataset then it needs more detail in the results, such as distribution, success rates, etc.

Thank you for your suggestion. For the HIF3A methylation, we used the same dataset as the previous work. We added the following content to help readers find our previous work, and also describe the characteristics of samples with methylation.

‘Details of the characteristics of the 110 extremely obese children and 110 matched controls with methylation examination were described previously [15]. Briefly, there was no difference between the two groups in age (P=0.934) or gender (P=0.946), and the 110 extremely obese children had a higher level of ALT as compared with controls (P<0.001).’

* Abbreviations in the title (HIF3A and ALT) need to be spelled out in full.

We changed the spelling in the title using full words instead of abbreviations. Thank you for your correction.

* Lg is used in the statistical analysis section. This is not a normal abbreviation, it looks like it is being used as an abbreviation for logarithm, in which case it should be logALT

We changed Lg to log throughout the text, and also figures and tables. Thank you for your correction.

* I am not a statistical expert so it is impossible for me to confirm that the mediation analysis is appropriate. It would be very helpful if the authors included a reference to mediation analysis.

Thank you for your comments. The mediation model was only described briefly in the original version of the manuscript. In the current version, we added reference and described the mediation analysis in more detail. The following content was added in the manuscript, Methods section - statistical analyses.

The mediation analysis was based on the model brought forward by Baron, et al. [21], and was used by other literatures, especially in the domain of psychology [31]. Three multivariate linear regression models were conducted, all adjusting for age, age2 and gender: (1) \( Y = cX + p_{1age} + q_{1age2} + r_{1gender} + e_1 \); (2) \( M = aX + p_{2age}+q_{2age2} + r_{2gender} + e_2 \); (3) \( Y = cX + bM + p_{3age}+ q_{3age2} + r_{3gender} + e_3 \). The statistical test of the mediation effect included several steps: (1) the association between the independent variable and the dependent variable was tested
(the coefficient c); (2) the association between the independent variable and the potential mediator was tested (the coefficient a); (3) both of the independent variable and the potential mediator were entered simultaneously as predictors of the dependent variable, and the coefficient b was tested to establish the significance of the mediation effect; (4) if the mediation effect was significant, the type of mediation effect could be determined by testing the coefficient cr, indicating either full mediation effect (cr not significant) or partial mediation effect (cr still significant).


* It is not clear why the authors are adjusting for age-squared when this is not commonly done in statistical genetics analysis. Some statement in the methods section of why they are doing this would be helpful.

Thank you for your comments. Age and age-squared were adjusted for simultaneously because phenotypes could be highly influenced by age, especially in the children population, and the association between the phenotype and age could be non-linear. The method was also used in other genetic studies. Musunuru et al. performed association analyses of SNPs with lipid traits, adjusted for age and age-squared. Hernaez et al. investigated association between SNP and steatosis controlling for age, age-squared, sex, and alcohol consumption. We added the following content in the Method section.

In order to fully control the effect of age on ALT, age and age2 were adjusted for simultaneously, since phenotypes could be highly influenced by age especially in the children population, and the association between the phenotype and age could be non-linear. The method of age-squared adjustment was also used in previous studies [32, 33].


* It is also not clear why the authors’ selection process resulted in obese and non-obese groups with significantly different proportions of male subjects. The fact that the controls are not gender-matched may be the source of the weakly positive associations they are seeing. Some discussion of this point is needed.
Thank you for your comments. For the samples with methylation data, controls were gender-matched. However, for the SNP data, there was a higher proportion of boys in the obese group than in the non-obese group. Although we adjusted for gender in the regression model, the different proportion of male subjects may still influence the result of association. We added the following content in last paragraph of the Discussion section.

‘Besides, for the SNP data, there was a higher proportion of boys in the obese group than in the non-obese group. Although we adjusted for gender in the regression model, the different proportion of male subjects may still influence the result of association.’

* No mention of multiple testing correction is made and this as a clear weakness of the study. Any correction of the reported p-values reported is likely to render them non-significant and the discussion should mention this.

Thank you for your suggestion. Gene-environment interaction analysis and mediation analysis usually need a large sample size. Although only one polymorphism was tested in the current study, multiple tests in the interaction analysis and the mediation analysis have been performed, and the result should be interpreted with caution. We have added the following content in the last paragraph of the Discussion section.

‘Gene-environment interaction analysis and mediation analysis usually need a large sample size. Although only one polymorphism was tested in the current study, multiple tests in the interaction analysis and the mediation analysis have been performed, therefore the result should be interpreted with caution.’

Yvonne Böttcher, PhD (Reviewer 2): Interaction between obesity and HIF3A rs3826795 polymorphism in relation with plasma ALT

The study was conducted to investigate a potential association of the HIF3A polymorphism with ALT in Chinese children (obese cases (705) and non-obese controls (1325)). Moreover, the authors aimed to analyze whether obesity interacts with the single variant on ALT.

The authors describe a significant interaction between obesity and rs3826795 polymorphism on ALT, expressed by association in obese children only, while the G risk allele increases ALT levels. Further, the authors performed methylation analyses in 110 severely obese children and 110 non-obese children and find that these effects are mediated via DNA methylation at the CpG site 46801699 within first HIF3A intron.

This is a nicely written, concise and clear manuscript. The authors discussed several limitations of their study. I raise here a few issues that should be considered:

1. The authors may want to further elaborate on what the potential mechanism of the meQTL (rs3826795 on methylation at 46801699) might be.
Thank you for your suggestion. The mechanism of meQTL is an interesting topic. Although many obesity-associated SNPs have been found to be related to DNA methylation alteration at proximal promoters and enhancers, most of the studies were correlation analysis in statistics, and we did not find literatures that can provide evidence on the mechanism of rs3826795 on HIF3A methylation. We did not test or cytokines or gene expression data other than DNA methylation, which is one of the limitations of the current study. Further studies were needed for understanding the mechanism of HIF3A meQTL.

This part of limitation was stated in the last paragraph of the Discussion section, as following:

‘Finally, we did not test cytokines or gene expression data other than DNA methylation. Further studies were needed for understanding the mechanism of the relation among obesity, ALT and the HIF3A gene, and also the mechanism of HIF3A meQTL.’

2. The authors describe an additive genetic model to assess the association between the polymorphism and ALT which is non-significant in the entire cohort. What about the dominant mode of inheritance for the G-allele?

Thank you for your suggestion. We implemented different genetic models (additive model, dominant model and recessive model), and as the following table shows, the association between ALT and rs3826795 was not significant in any of the three genetic models.

Table. Association between ALT and rs3826795.

<table>
<thead>
<tr>
<th></th>
<th>Additive model</th>
<th>Dominant model</th>
<th>Recessive model</th>
</tr>
</thead>
<tbody>
<tr>
<td>β’</td>
<td>0.020</td>
<td>0.017</td>
<td>0.016</td>
</tr>
<tr>
<td>p</td>
<td>0.340</td>
<td>0.417</td>
<td>0.455</td>
</tr>
</tbody>
</table>

All three models were adjusted for gender, age and age-square.

Besides, other studies that focused on rs3826795 used the additive model, including the EWAS study that examined the SNP-methylation association (Dick et al.[13]), and the gene-nutrition interaction study (Huang et al.[25]). Therefore, we assumed an additive allele effect, in order to compare with other studies.


3. The authors should describe the statistics behind the mediation analysis in more detail.

Thank you for your comments. The mediation model was only described briefly in the original version of the manuscript. In the current version, we added reference and described the mediation analysis in more detail. The following content was also added in the manuscript, Methods section - statistical analyses.

The mediation analysis was based on the model brought forward by Baron, et al. [21], and was used by other literatures, especially in the domain of psychology [31]. Three multivariate linear regression models were conducted, all adjusting for age, age2 and gender: (1) \( Y = cX + p1age + q1age2 + r1gender + e1 \); (2) \( M = ax + p2age+q2age2 + r2gender + e2 \); (3) \( Y = crX + bM + p3age+ q3age2 + r3gender + e3 \). The statistical test of the mediation effect included several steps: (1) the association between the independent variable and the dependent variable was tested (the coefficient c); (2) the association between the independent variable and the potential mediator was tested (the coefficient a); (3) both of the independent variable and the potential mediator were entered simultaneously as predictors of the dependent variable, and the coefficient b was tested to establish the significance of the mediation effect; (4) if the mediation effect was significant, the type of mediation effect could be determined by testing the coefficient cr, indicating either full mediation effect (cr not significant) or partial mediation effect (cr still significant).


4. The sample characteristics for the 110 extremely obese and 110 matched controls should be given in another table.

Thank you for your suggestion. For the HIF3A methylation, we used the same dataset as the previous work. We added the following content to help readers find our previous work, and also describe the characteristics of samples with methylation.

‘Details of the characteristics of the 110 extremely obese children and 110 matched controls with methylation examination were described previously [15]. Briefly, there was no difference between the two groups in age (P=0.934) or gender (P=0.946), and the 110 extremely obese children had a higher level of ALT as compared with controls (P<0.001).’

5. There is some redundancy between the introduction and the discussion which should be removed or re-written.
Thank you for your suggestion. We have removed the redundancy content in the introduction and the discussion. The reference [10-12] (about meQTL) and [13] (about the EWAS study which identified the HIF3A-obesity association) were still used in both the introduction part and the discussion part, but we used different writing to introduce their findings, in order to avoid redundancy.

6. As the authors acknowledge in their limitations they have analyzed peripheral blood samples while they had obviously no access to e.g. adipose tissue etc. Nevertheless, it would be good to describe such studies in the introduction or discussion. For example, the authors may want to cite the following reference in their introduction:


Thank you for your suggestion. The recommended reference is important for understanding the association between HIF3A and obesity, or adiposity. We cited the reference in the Discussion section, and added the following content in the last paragraph of Discussion.

As reported by Pfeiffer et al., HIF3A mRNA expression was higher in subcutaneous adipose tissue as compared with visceral adipose tissue, and correlated with parameters of adipose tissue dysfunction [34].