Author's response to reviews

Title: Association of common variants identified by recent genome-wide association studies with obesity in Chinese children: a case-control study

Authors:

Hai-Jun Wang (whjun1@bjmu.edu.cn)
Anke Hinney (anke.hinney@uni-due.de)
Jie-Yun Song (songijeyun1983@126.com)
André Scherag (andre.scherag@med.uni-jena.de)
Xiang-Rui Meng (meng_xiangrui@163.com)
Harald Grallert (harald.grallert@helmholtz-muenchen.de)
Thomas Illig (Illig.Thomas@mh-hannover.de)
Johannes Hebebrand (johannes.hebebrand@uni-due.de)
Yan Wang (wangyan@bjmu.edu.cn)
Jun Ma (majunt@bjmu.edu.cn)

Version: 5
Date: 26 December 2015

Author's response to reviews: see over
We greatly appreciated the valuable comments and suggestion of two reviewers, which helped us to improve our paper. We have considered the reviewers’ comments carefully and made a point-by-point response to the comments as below.

Reviewer’s report
Title: Association of common variants identified by recent genome-wide association studies with obesity in Chinese children: a case control study
Version: 1
Date: 6 May 2015
Reviewer: Annique Claringbould

Reviewer’s report:
Wang and colleagues describe an association study of 40 SNPs previously found associated with obesity. The association of 32/40 SNPs is checked in a sample of 2030 Chinese children, of which 718 are overweight, and 705 obese. The authors use logistic regression to test for association with overweight or obesity, and linear regression to test the effect of each SNP on the BMI standard deviation score. The study is executed well and shows some interesting associations. However, low power of the study (as indicated by appropriate power calculations) results in the detection of only two associations after p-value correction for multiple testing.

Minor Essential Revisions
The correction for multiple testing could be avoided, given the prior hypothesis of the common origin of the variants leading to obesity in major ethnic groups, therefore the expected transferability between ethnicities (several T2D and obesity papers could be cited). Therefore, any variant associated at 0.05 significance could be referred to as associated in this relatively small sample.
Response: We appreciated your comment. There were several T2D and obesity papers without correction for multiple testing. But some statisticians believed the correction for multiple testing is reliable because it is more conservative. So we think it would be better to provide the results without and with correction for multiple testing. Please see Page 3 Line 17-21, Page 4 Line 6, Page 13 Line 5-10, Page 17 Line 15-17.

It would be good to expand on the comment on different effect sizes between ethnicities for one locus. The differences found at rs543874 of SEC16B and rs2241423 of MAP2K5 are interesting, but a statistical evaluation along the lines of “Transethnic meta-analysis of genomewide association studies.” Morris, A.P. (DOI: 10.1002/gepi.20630) would provide more insight into this effect. It could be worthwhile to look into other loci with less pronounced evidence of heterogeneity as well to find more potentially interesting differences in effect sizes between populations.
Response: Thank you for helpful suggestion. We used the statistical evaluation of “Transethnic meta-analysis of genomewide association studies.” Morris, A.P. (DOI: 10.1002/gepi.20630). The results are the rs543874 of SEC16B and rs2241423 of MAP2K5 having evidence for heterogeneity ($P$(heterogeneity)$>50\%$). We did not find more SNPs with
differences in effect sizes between populations. We added the description of traditional meta-analysis in “Statistical analyses”. Then we replaced the results of traditional meta-analysis with the results of transethnic meta-analysis developed by Morris AP et al. Please see Page 11 Line 11-16, Page 13 Line 12-15 and Figure 1.

Discretionary Revisions
Mention the participants’ average age in the text on p.7, and refer to table 1 with characteristics of participants here, rather than in the results section.
Response: According to your suggestion, we have added the participants’ average age in the subjects section and refer to Table 1, and deleted the previous text in the results section. Please see Page 8 Line 5-9.

Could the authors state what value the BMI in the 95th percentile is (p.7)? Maybe a graph of the BMI distribution or a table of what BMI constitutes obesity and overweight at different ages could be added.
Response: According to your suggestion, we added a table to state the 85th and 95th percentiles of BMI, as the reference for screening overweight and obesity, respectively. Please see Page 7 Line 19-20 and Table 1.

On p.9, line 8: could you also select proxy SNPs based on the CHB HapMap sample, rather than CEU? It does not seem logical to use CEU with the Chinese participants.
Response: We select 6 proxy SNPs in total. One proxy SNP (rs261966) was selected for rs261967based on CHB HapMap sample, because the discovery study for rs261967 was done in East Asians and there are linkage disequilibrium (LD) data for the SNP in CHB HapMap database. In addition, we could not select the other proxy SNPs based on the CHB HapMap sample, because the discovery study for these SNPs were done in Europeans and there are no LD data for rs10150332 and rs12444979 in CHB HapMap database. In order to do the validation study, we could only select the CEU proxy SNPs.

Use the term GWA studies, rather than GWAS. The authors often write ‘GWAS studies’, which translates to ‘genome-wide association studies studies’.
The abbreviation BMI is introduced twice (p. 5 in background and p. 7 in subjects)
Response: According to your suggestion, we used the term GWA studies, rather than GWAS. We deleted the second introduction of the abbreviation BMI.

In table 2, indicate the proxy SNPs.
Response: We indicated the proxy SNPs in Table S1 (Additional file 1) and Table 3, which is previously Table 2.

Reviewer's report
Overall, this is a conservative, thorough and clearly-written study of the impact of genetic variation on obesity in Chinese children. I have suggested some modifications to the manuscript (points 4 and 7) that the authors could make at their discretion; and have identified some typographical and grammatical errors that should be corrected (point 10).

Detailed comments follow:

1. Is the question posed by the authors well defined?
   Yes. The authors wished to determine whether SNPs associated with BMI or obesity in large GWAS of Caucasians and Asians were individually associated with, or had collective value for predicting obesity in Chinese children.
   Response: Thank you for the comment.

2. Are the methods appropriate and well described?
   Yes. The authors have identified 40 obesity related loci, based on a review of genome-wide association studies that were conducted mainly using Caucasian adults. In six cases, genetically linked proxy SNPs were chosen to allow a multiplexed assay design. Association was tested in 2030 unrelated Chinese children, who were previously recruited in two cross-sectional studies performed in Beijing. Obese or overweight participants were identified using age- and sex-specific BMI percentiles. Logistic and linear regression was used to identify SNPs associated with the weight classification and BMI standard deviation score, respectively. Of the 32 SNPs that passed quality control, 6 showed nominal association (P<0.05) with childhood obesity, of which 2 showed association at a more stringent threshold reflecting multiple testing (P<0.00156).
   My major concern with the methods used in this study surrounds the SNP selection, which is based on studies published in 2010 and 2012. Since then, the GIANT consortium has gone on to identify 56 additional loci associated with obesity (Locke, Nature. 2015 Feb 12;518(7538):197-206). Since this more than doubles the number of obesity-related loci, the paper would be stronger if these loci had been tested.
   Response: Thank you for the comment. We did not perform the genotyping for the 56 additional loci associated with obesity (Locke, Nature. 2015 Feb 12;518(7538):197-206), because of our limited research fund. Those SNPs would be further tested if we could apply more funds.

3. Are the data sound?
Yes. The methods used for data cleaning are clearly described as are statistical approaches used to detect confounders (including study group), heterogeneity and bias. The rate of genotyping failure seems appropriate for the genotyping technology (Sequenom MassARRAY) and data cleaning procedure.

Response: Thank you for the comment.

4. Do the figures appear to be genuine, i.e. without evidence of manipulation? Yes. The figures, tables and supplemental tables are clear and provide a full description of the study results (including results for the SNPs excluded from the analysis for technical reasons). As minor points, the authors should consider redrawing Figure 1 to better separate the confidence intervals seen in the initial and present study (either by using colors that are more easily distinguishable or displacing the overlapping CI whiskers and centroids). It’s also hard to identify the SNPs that were used to assess directional consistency of effects. Finally, it would also help to add summary statistics to the body or legend of Figure 2.

Response: According to your suggestion, we redrew Figure 1 to better separate the confidence intervals seen in the initial and present study. We added summary statistics ($\chi^2$ tests overall $P=0.001$) to the body of Figure 2.

5. Does the manuscript adhere to the relevant standards for reporting and data deposition? Yes. The summary level data is fully reported.

Response: Thank you for the comment.

6. Are the discussion and conclusions well balanced and adequately supported by the data? Yes. The authors have taken a conservative approach in their discussion and conclusions.

Response: Thank you for the comment.

7. Are limitations of the work clearly stated? The main limitation of this work is the sample size used for validation. As the authors point out, only 6 of the 42 previously associated SNPs showed association in their population sample of 2030 individuals. Differences in ethnic origin could account for the discrepancy; and, in fact, 8 of the SNPs were monomorphic or low-frequency (MAF<1%) in this Chinese population and 3 SNPs showed differences in minor allele frequency compared to Europeans (FST#0.10). However, none of the four SNPs that were initially associated with in East Asians achieved significance in this study; and 23 of the remaining 26 SNPs showed the same direction of effect on obesity as reported in the original study (as well as a dosing effect on obesity), suggesting that the failure to individually validate these SNPs could be better explained by either the younger age of the study sample or lower statistical power. The authors discuss this well in the paper; but I think that the discussion should be expanded to note that none of the
SNPs previously associated with obesity in East Asians, showed association in this study (this information can be confirmed by digging around in the supplemental tables, but should be mentioned in the text). I would also add a sentence to explain the implications of the heterogeneity of effect sizes that was seen between the current and the discovery studies (Page 16, line 9). It's probably also worth noting that none of the proxy SNPs showed significant association.

Response: Thank you for the comment.

One of the four SNPs (rs2206734 of CDKAL1) that were initially associated with in East Asians achieved significance in this study (see Table 4). We added a sentence in “Discussion” to note it. Please see Page 16 Line 11-14.

According to one comment of another reviewer, we used the statistical evaluation of “Transethnic meta-analysis of genome wide association studies.” Morris, A.P. (DOI: 10.1002/gepi.20630). The results are the rs543874 of SEC16B and rs2241423 of MAP2K5 having evidence for heterogeneity ($P$(heterogeneity)>50%). None of the 26 SNPs without statistical significance showed heterogeneity of effect sizes between the present and discovery studies. So we deleted the previous sentence in Page 16 line 21-22.

We added a sentence to note that none of the proxy SNPs showed significant association. Please see Page 17 Line 1-3.

8. Do the authors clearly acknowledge any work upon which they are building, both published and unpublished?
Yes.
Response: Thank you for the comment.

9. Do the title and abstract accurately convey what has been found?
Yes. The abstract clearly defines the source of the candidate SNPs, the study's sample population, the major results of the paper and their implications.
Response: Thank you for the comment.

10. Is the writing acceptable?
Yes. The paper is well-organized and clearly written.
Response: Thank you for the comment.

There are a few typographical errors:
Response: We greatly appreciated your helps. We revised them accordingly.

page 5, line 12 "identified 18 new loci associated with BMI at a ..."
Response: We revised it.

page 7, line 12 "were approved by the ethics ..."
Response: We revised it.

page 9, line 3 "Multiplex SNP assays designs failed for 6 out of 40 SNPs (rs10968576, rs4771122, rs10150332, rs12444979, rs29941 and rs261967); these were replaced by ..."
Response: We revised it.

page 9, line 9 "(CEU) for the other 5 SNPs)."
Response: We revised it.

page 9, line 14 "group, 31 oof the 32 SNPs showed ..."
Response: We revised it.

page 9, line 21 "representing the effect of population ..."
Response: We revised it.

page 10, line 14 "weight the risk alleles ..."
Response: We revised it.

page 10, line 16 "that weighting the risk alleles ..."
Response: We revised it.

page 15, line 22: "our study would be underpowered to detect ..."
Response: We revised it.

page 16, line 18 "consequently, reduced statistical power."
Response: We revised it.

page 16, line 18 "by the number of loci tested, which is growing as new genome-wide meta-analyses are conducted ..."
Response: We revised it.