Author’s response to reviews

Title: A cis-eQTL allele regulating reduced expression of CHI3L1 is associated with late-onset adult asthma in Japanese cohorts

Authors:

Jun Kanazawa (jk-822@khaki.plala.or.jp)
Haruna Kitazawa (kita_masa_haru@yahoo.co.jp)
Hironori Masuko (hmasuko@md.tsukuba.ac.jp)
Yohei Yatagai (yohei820yatagai@yahoo.co.jp)
Tohru Sakamoto (t-saka@md.tsukuba.ac.jp)
Yoshiko Kaneko (yoshipy0429@yahoo.co.jp)
Hiroaki Iijima (iijima@tmch.or.jp)
Takashi Naito (naitota@tmch.or.jp)
Takefumi Saito (takefumisaito@yahoo.co.jp)
Emiko Noguchi (enoguchi@md.tsukuba.ac.jp)
Satoshi Konno (satkonno@pop.med.hokudai.ac.jp)
Masaharu Nishimura (ma-nishi@med.hokudai.ac.jp)
Tomomitsu Hirota (tomomitsu.hirota@riken.jp)
Mayumi Tamari (mayumi.tamari@riken.jp)
Nobuyuki Hizawa (hironori816@yahoo.co.jp)

Version: 2 Date: 07 Mar 2019

Author’s response to reviews:

Re: MGTC-D-18-00495 A cis-eQTL allele regulating reduced expression of CHI3L1 is associated with late-onset adult asthma in Japanese cohorts

Dear Dr Li:
We appreciate the detailed review of our manuscript. The comments of the three reviewers have been very helpful for us in revising the manuscript. We have addressed their comments and provide a point-by-point response to the issues they raised below. We believe that the current version of the manuscript is much improved, and we hope that it is now acceptable for publication in BMC Medical Genetics.

Sincerely yours,

Hironori Masuko, MD

Point-by-point responses to comments

Reviewer reports:

Erik Fransen (Reviewer 1): However, as a statistician and genetic epidemiologist, I noticed several methodological issues, which make reported associations not convincing.

Major comment:

Comment #1

My main concern is the total lack of correction for multiple testing. The genotyping was carried out through a GWAS, which gives data on about 4 million markers. One SNP was picked out - rs946261 in the promoter of CHI3L1, with a p-value for association with asthma of 0.018. There are by definition 1.8 percent of all SNPs reaching that p-value under the null hypothesis of no association. Even in the total absence of any genuine association, hundred thousands of SNPs are expected to reach a p-value of 0.018. The fact that the SNP is a known e-QTL, does not add to the argumentation that the SNP has an effect on disease risk. Neither does the possible functional role of CHI3L1 provide any additional proof - many genes are involved in the asthma pathophysiology.

Response #1

Thank you very much for the valuable comment. We did not apply any corrections for multiple testing at the GWAS level because this study is basically a replication of the previous study that showed that an eQTL at CHI3L1 was associated with the risk of asthma in 3 populations of European ancestry with mild asthma (P = 1.2 × 10^-5). We previously conducted a GWAS for adult asthma [E1] using Japanese populations and identified some novel genes for specific phenotypes of asthma [E2, E3]. Using this Japanese GWAS data, we also had the opportunity to evaluate some previously reported asthma associations in non-Japanese populations [E4, E5, E6, E7], assess the reproducibility of these associations, test additional common variants in these loci, and ask the question: “Are these genes truly associated with asthma or specific phenotypes
of asthma especially in Japanese?“ Although this approach is unlikely to identify novel genes for asthma, we believe this is still very important given that asthma is a complex and heterogeneous disease and a single study in a single population cannot cover all potential phenotypes, all possible gene-environment interactions, and all different genetic models. In fact, in our previous studies, we successfully identified Japanese-specific SNPs that had not been previously tested in non-Japanese populations, yet were associated with asthma in Japanese, demonstrating the capacity of GWAS surveys to identify novel susceptibility SNPs within the candidate genes [E5, E6, E7].

In the current work, we initially used our own GWAS data to replicate the genetic effects of CHI3L1 as one of the promising asthma candidate genes (ie, those previously associated in two or more non-Japanese cohorts) in Japanese populations. Also, given the ethnic difference in the LD pattern in the CHI3L1 region, we conducted gene-level replication but not SNP-level replication. We selected 6 eQTLs for the initial association; however, we did not apply the Bonferroni correction here because we consider it possibly to be too conservative. Considering that functional SNPs in the biological candidate may not yield significant p-values after correction for multiple testing but may replicate consistently over many populations tested, we used the liberal standard of “any nominal p < 0.05” to suggest gene-level replication, and we believe our finding may support some roles of YKL-40 in the pathogenesis of asthma, especially late-onset adult asthma. Meanwhile, it is certainly possible that our finding just represents a false positive, and we understand that our finding cannot stand alone but should be seen as a contribution to the discussion about the genetic role of CHI3L1 in asthma. Our findings need to be treated with caution until a sufficient level of replication has been achieved in other populations. In the revised manuscript, we have acknowledged these points in lines 159-161, 214-220, and 331-336.

Minor comments:

Comment #2

The association test between the genotype and the disease status was reported with a one-sided p-value. A one-sided test starts from the null hypothesis that one allele (the putatively deleterious allele) leads to an improvement of the disease status. This does not make sense in the present situation - here one should start from the null hypothesis of no association between the allele and the disease status. The association test should be carried out using a two-sided test.

Response #2

According to the suggestion, in the revised manuscript, the results described are those obtained using a two-sided test.

Comment #3
The partitioning of the patients into clusters is poorly explained and argued. On line 180-181 the authors mention previously determined subgroups of patients, but no reference is provided. A model to partition patients into subgroups is built using a classification and regression tree (CART) model, using a set of 880 adult patients of which no details are provided. They seem to be not part of the current genotyping study, but no descriptive statistics are given. It is highly unclear if this cohort is matched to the collection 971 genotyped individuals. The CART model is then used to classify the 971 individuals into the six groups.

Response #3

We apologize for not clearly providing the reference about the cluster analysis. We previously established a model to partition patients into subgroups using a classification and regression tree (CART) model, using a set of 880 adult patients [E8]. For the multinomial regression model in the current study, we used 968 patients with adult asthma, which included 605 patients examined in the previous study and an additional 363 patients with adult asthma. They were collected in the same geographic regions, and were appropriately divided into the clusters using the CART model. No differences between the 605 patients and the additional 363 patients were found in age, sex, age at disease onset, smoking status, total levels of serum IgE, percent of predicted forced expiratory volume in the first second (%FEV1), ratio of FEV1 to forced vital capacity (FVC), and specific IgE responsiveness to common inhaled allergens.

Comment #4

The output of the multinomial regression is confusing. In a multinomial regression with a six-level outcome, one of the outcome levels has to be the reference level. The effect sizes (odds ratios) are then referred to that reference level - the odds ratio of a non-reference level equals the odds of the non-reference level, divided by the odds of the reference level. Here, we see six odds ratios, each of which with a p-value. Was the control group included in the analysis as a seventh group, and used a reference? Assuming this latter scenario do the p-values represent the null hypothesis that the odds ratio (with the control group) equals 1? In this latter case, you need to adjust for multiple testing.

Response #4

We sincerely apologize for having caused any confusion. The control group was included in the analysis as a seventh group, and used as a reference. In the revised manuscript, we have carefully revised the sentence stating how the multinomial regression with 6 asthma clusters plus a control group was conducted (lines 202-212 of the revised manuscript). Here, given a stronger association with late-onset asthma that developed at 41 years of age or older, we used the multinomial regression model to predict the probabilities of the different possible outcomes of specific phenotypes. We decided not to correct the analysis by 6 because this analysis was rather a hypothesis-generating analysis to examine if any specific phenotype(s) can be predicted by the CHI3L1 eQTL. In addition, we believe that examining a single multinomial regression model using all the individuals already adjusted the p-value inflation, which is different from running the multiple logistic regression analysis for each group separately, each analysis on a different
smaller number of individuals. Finally, the larger odds ratio with a specific asthma phenotype characterized by late onset, less atopy, and mild airflow obstruction may indicate a potential role of CHI3L1 in susceptibility to a specific phenotype of asthma, contributing to the discussion about the genetic role of CHI3L1 in asthma with some insight into the mechanisms underlying the complexity of the gene and development of asthma or specific phenotypes of asthma.

Xingnan Li, Ph.D. (Reviewer 2): In this study, Dr. Jun et al. performed a candidate gene study of CHI3L1, and indicated that the C allele of rs946261 may be associated with late-onset adult asthma through down-regulation of YKL-40 expression. Some of my concerns are:

Major comments:

Comment #1

The organization of the manuscript is not clear, sometimes confusing. For example, (a) In the Materials and methods section (Page 9, Lines 140-144): 'Thus, we first searched the GTEx Portal database,......, (Table 1: Figure 1)'. The results of these 34 SNPs and/or 6 SNPs (Table 1 and Figure 1) should belong to Results section. (b) In the Materials and methods section (Page 11-12, Lines 182-199): 'We performed a two-step cluster analysis,......". Is this cluster analysis performed in this study or previously? If it is done previously, this part should go to Introduction section. If it is done currently, then it should go to Results section.

Response #1

We sincerely apologize for having caused any confusion. According to the comments, we have revised the organization of the manuscript. We have moved the results of the GTEx portal database to the Results section (lines 223-227 of the revised manuscript). We also apologize for not clearly providing the reference about the cluster analysis [E8]. We previously established a model to partition patients into subgroups using a classification and regression tree (CART) model, using a set of 880 adult patients. The multinomial regression model in the current study included 605 of the 880 adult patients in the previous study, and the additional 363 patients with asthma, who were collected in the same geographic region as the original 880 patients, were appropriately divided into the clusters using the data for age at asthma onset and %FEV1 in the CART model.

Comment #2

The association results of 6 SNPs should be shown in the Table 1.

Response #2
We have now revised the manuscript so that the association results of the 6 SNPs are shown in Table 1.

Comment #3

The association of rs946261 with asthma is not convincing without other evidence, such as the findings from previously published large sample size GWAS. On the basis of previous studies, the association of SNPs in CHI3L1 with asthma has never been confirmed, at most the association is very weak and inconsistent. It needs to be carefully discussed.

Response #3

As the reviewer pointed out, to date, studies are not conclusive regarding the genetic impacts of the SNPs at CHI3L1. Initially, a GWAS of serum YKL-40 levels in a founder population of European descent by Ober et al [E9] identified an association between the rs4950928 and circulating YKL-40 levels. This SNP was also associated with the risk of asthma in 3 populations of European ancestry with mild asthma \( (P = 1.2 \times 10^{-5}) \). Subsequent reports have shown an association with rs4950928 and asthma [E10, E11], although the risk allele was opposite to that reported by Ober et al. Additional publications have been conflicting with respect to the effect of genetic variation in the CHI3L1 gene and asthma [E12, E13]. Such inconsistencies are particularly the case for asthma where a single study in a single population cannot cover all potential phenotypes, all possible gene-environment interactions, and all different genetic models. Therefore, we believe that replication in as many populations as possible is still important and that our replication study in Japanese populations may indeed give some insight into the mechanisms underlying the complexity of the gene and development of asthma or specific phenotypes of asthma.

Comment #4

Reference #5 and #6 both indicated higher levels of YKL-40 were correlated with asthma and asthma severity. In this study, lower CHI3L1 mRNA was correlated with asthma. How to explain this contradiction?

Response #4

Thank you for the comment. In some other studies as well as in the current study, the original genetic association of CHI3L1 with asthma was also replicated but with the risk allele reversed from the original report [E10, E11]. In other words, lower CHI3L1 mRNA has been correlated with asthma in some populations including ours. This inconsistency may be caused by the small sample sizes of the previous studies, the presence of multiple causal variants at the loci, and differences in ethnicity, smoking exposure, and age. The populations studied in the original report included one founder population consisting of genetically related individuals and other populations consisting of many children with asthma [E9], which were totally different from the populations used in the replication studies consisting of mainly adult patients with asthma in
outbred populations. In fact, in the current study, stronger association was found in asthma patients who developed the disease at 41 years of age or older, indicating that some CHI3L1 SNPs have an opposite or independent effect on the risk of asthma later in life as compared with early-onset asthma. We have discussed this issue in the revised manuscript (lines 299-309).

Comment #5

The function of CHI3L1 is not clear or it has multiple potential function. The conclusion of inflammasome activation is far-reaching and needs to be weakened.

Response #5

According to the suggestion, we have weakened our conclusion of inflammasome activation throughout the manuscript including the last paragraph of the Discussion in the revised manuscript (lines 343-347).

Martha Patricia Gallegos-Arreola, PhD (Reviewer 3):

Comment #1

The article entitles "A cis-eQTL allele regulating reduced expression of CHL3L1 is associated with late-onset adult asthma in Japanese cohorts", is an interesting study in which is notorious the importance of genetic heterogeneity that exists between populations. Minimal suggestions are enlisted below:

1. On line 117. The reference 20 is cited before that 19.

2. On line 234. Only clarify that the genotype to which they refer is the risk.

Response #1

Thank you for the comment. In complex diseases such as asthma, single-locus association studies are always complicated by interaction between an investigated locus and other genetic and environmental risk factors. As the reviewer correctly pointed out, genetic heterogeneity among different ethnic populations remains an important obstacle. We have made the suggested corrections in the revised manuscript.

Editor Comments:

#1. Please kindly remove the watermark from your submission.

Response #1
Thank you for the comment. We removed the watermark from our revised manuscript.

Editor Comments:

#2. In the “Funding” section of your declarations, please clarify the role of the funding body in the design of the study and collection, analysis, and interpretation of data and in writing the manuscript.

Response #2

Thank you for the comment. We described the role of the funders in the design of the study and collection, analysis, and interpretation of data and in writing the manuscript.

Editor Comments:

#3. Please confirm whether informed consent to publish was written or verbal, and clearly state this in your "Consent to Publish" section. If verbal, please state the reason and whether the ethics committee approved this procedure.

Response #3

Thank you for the comment. We stated this in the Consent to Publish section. All participants have given their written informed consents for publication of this study.

Editor Comments:

#4. Please remove the funding information from the Acknowledgements and include it in the Funding section instead. If you have no further acknowledgements please put “Not Applicable” in the Acknowledgements section.

Response #4

Thank you for the comment. We removed the funding information from Acknowledgement section and included it in the Funding section.

Editor Comments:

#5. Thank you for providing your marked revised manuscript. However, at this stage it is not required and so we kindly ask that you remove it from your manuscript.

Response #5
Thank you for the comment. We removed the marked revised manuscript.

References


