Author's response to reviews

Title:A genome-wide association study of copy-number variation identifies putative candidate loci associated with osteoarthritis in Koreans

Authors:

Sanghoon Moon (jiap72@hanmail.net)
Bhumsuk Keam (bhumsuk@snu.ac.kr)
Mi Yeong Hwang (myhwang85@gmail.com)
Young Lee (lyou7688@gmail.com)
Suyeon Park (sooyeon1002@hanmail.net)
Ji Hee Oh (ohjh83@gmail.com)
Yeon-Jung Kim (cocona35@hanmail.net)
Nam Hee Kim (shnhkim@hanmail.net)
Young Jin Kim (anwltlarkr@gmail.com)
Dong-Hyun Kim (dhkims@hallym.ac.kr)
Bok-Ghee Han (bokghee@korea.kr)
Bong-Jo Kim (kbj6181@cdc.go.kr)
Juyoung Lee (jylee@cdc.go.kr)

Version: 2 Date: 22 November 2014

Author's response to reviews: see over
Response to the reviewers:

Thanks for the reviewer’s thorough review of our manuscript, which greatly improved the quality of the manuscript. The reviewer’s very helpful comments were addressed, point by point, in our response to each comment or revision request.

Reviewer #1: James Robinson

MINOR ESSENTIAL REVISIONS

1) The authors should describe the cohorts in enough detail to enable meta-analysis of their data in the future, paying particular attention to the guidelines in Kerkhof et al. (Osteoarthritis Cartilage. 2011 19;3:254-64).

We agreed with reviewer’s comment. According to the guidelines in Kerkhof’s paper, we described properties of our study to enable meta-analysis in the future. Table 1 and table 2 provide exact definition of the phenotype to reduce heterogeneity of OA study.

Table 1. Basic characteristics of study subjects.

<table>
<thead>
<tr>
<th>OA phenotype</th>
<th>Number</th>
<th>Gender (%women)</th>
<th>Age (mean)</th>
<th>BMI (mean)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Cases</td>
<td>Controls</td>
<td>Cases</td>
<td>Controls</td>
</tr>
<tr>
<td>ROA</td>
<td>3/11</td>
<td>40/1</td>
<td>84%</td>
<td>80%</td>
</tr>
<tr>
<td>Osteoarthritis (K&gt;2)</td>
<td>36</td>
<td>42</td>
<td>84%</td>
<td>80%</td>
</tr>
</tbody>
</table>

BMI: body mass index

Table 2. Number of cases and controls in each OA type

<table>
<thead>
<tr>
<th>Study</th>
<th>Classification System</th>
<th>Cut-off value for OA</th>
<th>Exact OA definition</th>
<th>Kace Ort. Cases</th>
<th>Kace Ort. Controls</th>
<th>Weiss OA Cases</th>
<th>Weiss OA Controls</th>
</tr>
</thead>
<tbody>
<tr>
<td>KARE</td>
<td>K/L score</td>
<td>2</td>
<td>One definite osteophyte (original K&gt;2)</td>
<td>167</td>
<td>467</td>
<td>204</td>
<td>467</td>
</tr>
</tbody>
</table>

Moreover, we added a sentence on page 5, line 13~15 to describe our data more details as follows.

“OA phenotype was diagnosed based on clinical and radiographic findings without assessment of symptoms. Alternative methods of assessment such as self-reported and symptomatic OA could not be used in this study.”
2) The use of CNVtools to analyse the aCGH data should be described in more detail. CNVtools allows for association testing without applying copy number calls. Reporting of the parameter estimates for each of the identified CNVs would enhance the paper.

We thank the reviewer for pointing out ambiguous expression and we added several sentences for clarity. Detailed description of CNVtools was reworded as follows.

Page 6, line 16~23

“We used the software package, CNVtools with default parameters to assign individuals to each CNV genotype [18]. A mixture-based model used in CNVtools can separate samples into more exact CNV genotype group. To do this, all individuals of each CNV regions from the CNV detection with GADA were clustered according to the log₂ ratio between test sample and reference sample. Additional figure 1 shows type of CNV class. It is not easy to discriminant exact CNV genotype in the single-class CNVs, in which all individuals of the CNV region belonged to one cluster. On the contrary, multi-class CNVs that consisted of more than two clusters can assign highly confident CNV genotype.”

And we also added two sentences about parameter selection for CNV discovery.

Page 6, line 11~13

“To overcome the limitation of single algorithm detection, we tested CNV discovery with several parameters to find best parameter using known CNV region. Consequently, we select best parameter with highly concordant with known CNV region.”

3) Figure 1 AND Figure 3 panel A - Axes and labels should be clearly visible. What are the colored lines? Presuming they represent CNVtools posterior probability of copy number there should be a secondary Y axis labelled appropriately. Figure 3 panel B - the labels are unreadable.

We agreed with the reviewer’s suggestion. We added labels to the figures as follows.
Moreover, we also added description of the colored lines in the Figure 3 panel A: 0 copies, 1copy and 2copies.

Figure 1
DISCRETIONARY REVISIONS

4) The authors should indicate what efforts were made to detect batch effects, a common feature of CNV studies.

According to the reviewer’s comment, we added a sentence on page 6, line 1~4 as follows.

“To adjust systemic biases that are present in microarray experiment, all of samples passed experimental control metrics such as chromosome X shift and mad.1dr with NimbleScan v.2.5.”
5) While discussing limitations of the study the authors could consider estimates of false discovery rates in the discovery phase.

Previously, Kim and Jeon reported that the prevalence of OA (aged 45 to 64) was assessed 20.1% (Kim and Jeon, Korean J Rehabil Nurs (2011) 14:111-117). Based on this knowledge, we estimated sample size with 80% power for OA study using Quanto software. At least 3,996 case samples are required to satisfy acceptable significance (significance level: $2 \times 10^{-5}$, allele frequency: 0.1, genetic effect: 1.3). On the contrary, our sample size (371 cases) is quite small.

The sample size required for satisfying good power and low FDR is usually large. Therefore we considered that estimation of FDR value in our study cannot guarantee of statistical significance. Moreover, because we considered well-clustered CNV regions according to the CNV class (i.e. model-specific), CNV region did not distributed in genome-wide. Therefore, we change our manuscript title to “Model-based analysis of copy-number variation identifies putative candidate loci associated with osteoarthritis in Koreans” rather than considering false discovery rate.
Reviewer #2: F. David Carmona

Moon and colleagues have performed a large-scale analysis of CNVs in a Korean cohort of osteoarthritis patients and matched controls. The study aimed to shed light into the genetic background underlying this complex disease.

- Major Compulsory Revisions:

1) The study has important limitations that are acknowledged by the authors. The most relevant one in my opinion is the low statistical power that may rise reasonable doubts about the consistency of the results. The P-values obtained are distant from an acceptable significance threshold in this kind of studies. Therefore, I would recommend the authors to gather additional replication cohorts in order to discard the possibility that the proposed associations are a consequence of type-I errors due to an underpowered analysis.

We agreed with the reviewer’s comments.

As mentioned above, we estimated sample size with 80% power for OA study using Quanto software. At least 3,996 case samples are required to satisfy acceptable significance (significance level: $2 \times 10^{-5}$, allele frequency: 0.1, genetic effect: 1.3). On the contrary, our sample size (371 cases) is quite small. As reviewer’s comment, small samples size may result in the inconsistency of the results.

However, gathering independent replication samples within a month was not easy for us. According to the decision of the Associate editor, we did not deal with this point in our manuscript. We respected the valuable comment of the reviewer and we also appreciated the decision of the Associate editor.

Instead of conducting replication study, we addressed our limitations in the discussion section as follows.

Page 13, line 13~22

“The present study has some limitations. First, diagnosis of OA was radiologically based (i.e.,
using X-ray) without assessment of symptoms. Different methods of assessment such as self-reported and symptomatic OA could not be available in population-based study samples. This limitation has resulted in small sample size of discovery stage and a difficulty of collecting samples for replication analysis. Second, because of first limitation which was mentioned above, we did not success to find radiographic case samples in independent general population samples. Eventually, we did not conduct a replication study, which is highly desirable with association studies. This limitation can be solved through a meta-analysis with other data in the future. To reduce heterogeneity of meta-analysis, we provide our cohort information such as exact phenotype definition, cut-off, and number of cases of each OA type.”

2) I also have major concerns regarding the structure of the manuscript. The methods and results sections are poor and need a more detailed description.

We also agreed with reviewer’s comment regarding the structure of the manuscript. Thus we added more detailed descriptions in both methods and results on page 6~8. We addressed the quality control process in the method section and also added a detailed description of genotype estimation. Moreover, detailed description of CNVtools was reworded. For the results, we newly added study of population section.

3) In addition, the discussion is too descriptive and speculative, and lacks continuity in the argumentation.

According to the reviewer’s comment, to reduce speculatively described sentence, we changed second phrase in discussion on page 10, line 1~16 as follows.

“The canonical Wnt/β-catenin signaling pathway has been implicated in the pathogenesis of osteoarthritis (Corr 2008). Increased β-catenin accumulations have been observed in degenerative cartilage, suggesting that activated Wnt signaling might contribute to progression of osteoarthritis by downstream target genes such as matrix metalloproteinases (MMPs) (Monfort et al. 2006). For instance, the secreted Wnt antagonist Dickkopf-1 is related with slowed progression of hip OA in elderly women (Lane et al. 2007) and the small molecule XAV939, which selectively inhibits β-catenin-mediated transcription, is associated with protecting cartilage degradation in a rat osteoarthritis model (Zeng et al. 2014).
Tankyrase, which is encoded by TNKS, bind directly to axin, a negative regulator of the canonical Wnt/β-catenin signaling pathway, forming a destruction complex with glycogen synthase kinase 3β (GSK-3β) and adenomatous polyposis coli (APC) to degrade β-catenin (Huang et al. 2009). Inhibition of TNKS by treating the compound XAV939 or small interfering mediated silencing of TNKS can cause an increase in levels of axin, leading to phosphorylation and degradation of β-catenin, and inhibition of target gene transcription by Wnt signaling. Thus, although further studies are needed to establish the role of TNKS in osteoarthritis through axin-mediated inhibition of Wnt signaling, it is quite possible that TNKS could be risk factor for osteoarthritis.”

4) The authors should focus more on developing the main message of their paper rather than just stating the known role of the putative risk loci.

We agreed with the reviewer’s suggestion. We changed confusable expressions and speculations to more clear expression. And to reduce discursiveness of the discussion section, we also removed description of the known role of the putative risk loci such as MAMDC2, KCND3, and ME3 in the manuscript.

References