Author's response to reviews

Title: Association between Helicobacter pylori cagA-related genes and clinical outcomes in Colombia and Japan

Authors:

Masahide Watada (m08d9036@oita-u.ac.jp)
Seiji Shiota (sshiota@oita-u.ac.jp)
Osamu Matsunari (narry@oita-u.ac.jp)
Rumiko Suzuki (rusuzuki@oita-u.ac.jp)
Kazunari Murakami (murakam@oita-u.ac.jp)
Toshio Fujioka (fujioka@oita-u.ac.jp)
Yoshio Yamaoka (yyamaoka@bcm.tmc.edu)

Version: 3 Date: 24 November 2011

Author's response to reviews: see over
Dear Dr. Damian Hussey,

Editor, *BMC Gastroenterology*:

November 24, 2011

Thank you for your letter. We were pleased to receive a positive evaluation of our manuscript. The manuscript was fully evaluated in consideration of the reviewer’s comments, and we believe that the changes they suggested led to an improvement of our manuscript. Especially, according to the comments by Reviewer 2, we examined the status of \( \text{vacA} \), which is another important virulence factor of *H. pylori* and which has been widely studied. Because 3 samples from Colombia were positive for both \( \text{vacA} \) m1 and m2 genotypes, which suggested mixed infections, we excluded those samples in the final analysis. Therefore, we recalculated all of the data and rewrote the text. In addition, our manuscript was edited by a native English speaker. The changes made during the revision are printed in blue letters in the revised manuscript. In addition, we formatted our Tables according to the Instructions for authors.

Finally, I am very appreciative that you could extend the deadline for the revision.

We hope that the revised manuscript is now acceptable for publication in *BMC Gastroenterology*.

Yours sincerely,

Yoshio Yamaoka, MD, PhD

Department of Environmental and Preventive Medicine, Oita University Faculty of Medicine, 1-1 Idaigaoka, Hasama-machi, Yufu-City, Oita 879-5593, Japan
Tel: +81-97-586-5740; Fax: +81-97-586-5749
E-mail: yyamaoka@oita-u.ac.jp
ITEMIZED RESPONSES
Responses to comments raised by Reviewer 1

- Major Compulsory Revisions
There is no need to major revision.

Response:
Thank you for your comment. We were pleased to receive a positive evaluation of our manuscript. Our responses to the issues raised by you under “Minor Essential Revisions” were incorporated into the revised manuscript.

Minor Essential Revisions
1- Page 6, line 6: each deoxynucleotide concentration must be written in millimolar not Micro liter (µL)
Response:
We corrected the unit of each deoxynucleotide concentration as “millimolar” as suggested (Page 6, line 10 in the revised version).

2- Page 6, line 8: regarding the product size of used primers are short (table 1), thereby the PCR program as 30 cycles; 5 min at 94°C, 30 sec and 5 min at 72°C seems uncommon and must be checked.
Response:
We are sorry for the confusion. We used 5 min at 94°C for the initial denaturation and 5 min at 72°C for the final extension. When we performed 30 cycles of amplification, we used 30 s for denaturation, annealing, and extension. We therefore rewrote the sentence as follows: “Each reaction mixture was amplified as follows: initial denaturation at 94°C for 5 min, which was followed by 30 cycles of denaturation at 94°C for 30 s, annealing at the indicated temperature in Table 1 for 30 s, extension at 72°C for 1 min, and then final extension at 72°C for 5 min.” (Page 6, lines 12-15 in the revised version)

3-page 6 and 7: In some cases, the standard material sources addressing (company, county) is not considered as the name of country is missed.
Response:
We corrected the material sources as suggested.

4- In case of table 2, since the references of selected genes have given in the context, the table2 is no longer mandatory.
Response:
We moved table 2 to the supplementary file as suggested.

5- In case of the table 4 and 5, P value should be calculated for each row.

Response:
We added P values for each row in the tables as suggested.

6- Page 9, paragraph 2, in result section under subtitle of “The association between candidate genes and clinical outcomes” only two genes (jhp0045, jhp0046) of four studied genes (hp0967, jhp0045, jhp0046 and jhp0951) have been considered.

Response:
According to this comment, we added the following sentence: “On the other hand, the prevalences of hp0967 and jhp0951 in cagA-positive cases were not associated with clinical outcomes.” (Page 11, lines 6-7 in the revised version)

7- Page 9, under subtitle of “The association between candidate genes and clinical outcomes”: it is better to present each country result separately and more clearly.

Response:
We presented each country result separately in the revised manuscript as suggested.

Responses to comments raised by Reviewer 2

Major compulsory revisions
1. This addresses virulence factors in the bacteria but the complex interplay between host and pathogen are not really addressed in this analysis. Were the populations studied tested for established host polymorphisms that increase the risk of GC (eg IL-1R)?

Response:
Thank you very much for your comments. Unfortunately, in our study population, we did not obtain informed consent for measuring host polymorphisms. According to your comments, we added the following sentences in the last part of the discussion: “Our study had several limitations. First, not only the *H. pylori* virulence factors, but also environmental factors (eg, diet) and host factors have been demonstrated to be predictors of severe clinical outcomes. Especially, inflammatory cytokine gene polymorphisms (*IL-1* gene cluster, *TNF-α*, *IL-10*, and *IL-8*) have been reported to be correlated with gastric cancer. Further study will be necessary in order to elucidate the role of our candidate genes of *H. pylori.*” (Page 14, lines 1-6 in the revised version)
2. I have an issue with the methodology used for selection of genes. Microarray analyses involving correlations require correction for multiple testing. The p values reported here do not appear to be corrected for multiple testing. This would reduce some of the significance of the genes reported as correlated to cagA status. If you perform appropriate multiple testing correction do these genes still reach significance?

Response:
We appreciate your helpful comment. In this study, we used the microarray data from our previous report by Gressmann et al., PLoS Genet 2005. The methods for the microarray are detailed in that paper. Based on a cut-off value of our analyses, we calculated the accuracy as 93-96%, the sensitivity as 96-98%, and the specificity as 80-83%, as described in that article. Therefore, we think that our microarray data can be trusted. The gene expressions were recorded as positive or negative. We used that data. In addition, we calculated the correlations between cagA and candidate genes by Spearman rank coefficients ($r$). We believe that this process is sufficient for determining the candidate genes.

3. I think it is insufficient to perform a multivariate analysis of age and gender for risk of GC in these populations. There are other risk factors that could be incorporated that are more relevant for risk of GC (see question 1).

Response:
We appreciate your helpful comment. As we described above, we did not obtain data for host polymorphisms. Therefore, we described this limitation in the Discussion section, as described above.

4. It would be relevant to ascertain whether other established H. pylori genetic markers of risk as mentioned in the manuscript are also relevant to risk in these cohorts. The results would serve as a positive control if they indeed correlated as the authors suggest they would.

Response:
We agree with your comment. Because of your comment, we examined the status of vacA, which is another important virulence factor of H. pylori that has been widely studied. Because 3 samples from Colombia were positive for both vacA m1 and m2 genotypes, which suggested mixed infections, we excluded the samples in the final analysis. Therefore, we recalculated all of the data and rewrote the text. As a result, the vacA m1 genotype was significantly associated with gastric cancer (GC) in the
samples from Colombia. However, \textit{vacA} status was not associated with GC among the \textit{cagA}-positive cases in Colombia. We added these findings in the Results section on page 10. However, as you commented, it is still possible that our candidate genes may correlate with other virulence factors. Therefore, we added the following limitation to the Discussion section: “Second, we did not examine known virulence factors other than \textit{cagA} and \textit{vacA} of \textit{H. pylori}. It is possible that our candidate genes might correlate with other known virulence factors, even in \textit{cagA}-positive cases. It is better to examine host factors and other known virulence factors in order to clarify the role of our candidate genes in future studies.” (Page 14, lines 6-10 in the revised version)

5. This study seems to investigate presence of genes, however gene expression would be more relevant. Does the level of expression of the genes being studied have biologically significant expression differences? There is no data presented about fold change of gene expression only that there was negative or positive association with \textit{cagA} status. Whereas a correlation may be statistically significant between genes the expression difference may be marginally different and hence have little biological meaning.

\textbf{Response:}

Thank you very much for your comments. As you commented, we examined gene expression as positive or negative. It is not clear whether these findings are biologically significant. However, \textit{jhp0045} and \textit{jhp0046} have been reported to work as a restriction-modification (R-M) system. Therefore, we believe these genes have biological significance, as described in the Discussion. In addition, not only gene expression, but also fold changes of genes can be important, as you mentioned. Therefore, we added the following limitation to the Discussion section: “Finally, we examined the status of the genes by only positivity or negativity. The levels of gene expression can be affected by clinical outcomes. In addition, gene expression is not always correlated with protein expression patterns. For example, the expression of the blood group antigen-binding adhesin (BabA) protein is not always correlated with \textit{babA} gene expression. Further analysis using real-time PCR or immunoblotting techniques is necessary to clarify the significance of our candidate genes.” (Page 14, lines 10-15 in the revised version)

Minor essential revisions
6. There are a number of grammatical errors throughout the manuscript that would need to be addressed

\textbf{Response:}
In order to address this comment, our manuscript was edited by a native English speaker.

Responses to comments raised by Reviewer 3
Response:
We really appreciate your helpful editing. We corrected all of our mistakes according to your editing. In addition, our manuscript was also edited by a native English speaker.

Responses to comments raised by Reviewer 4
Minor essential revisions:
1) (Throughout) The manuscript needs to be edited by a native English speaker.
   Response:
   We appreciate your advice. In order to address your comment, our manuscript was edited by a native English speaker.

2) (pg 8) The authors should give as much information as possible on the different genes/ORFs i.e. Are these conserved hypothetical genes? In which of the sequenced H. pylori genomes are these genes found? The authors also need to address why they investigated two of the ORFs (jhp0967 and jhp0951), which are associated with duodenal ulcer (DU) disease, when it is commonly accepted that DU and gastric cancer are mutually exclusive diseases.
   Response:
   Thank you for your comment. We apologize that we did not show information for several genes in the original table 2. We moved this table to the supplementary information in response to another reviewer’s comment. We showed that these genes are conserved hypothetical genes. Next, the purpose of this study was to find novel candidate genes for severe outcomes. In general, not only GC, but also DU are considered severe outcomes of *H. pylori*. In addition, most of the virulence genes, including *cagA*, were related to GC and DU. Therefore, we included *hp0967* and *jhp0951* in our project. Finally, we examined the presence of candidate genes in fully sequenced strains in the Genbank in response to your comment. We include the following sentences in the revised manuscript: “We examined the presence of these candidate genes in 28 full sequenced strains deposited in Genbank. The *hp0967*, *jhp0045*, *jhp0046*, and *jhp0951* were found in 18, 7, 7, and 13 strains, respectively.”
3) (pg 23) From “eye balling” of the numbers, it seems surprising that there is a significant association between the carriage of jhp0045 and jhp0046 in H. pylori strains and gastric cancer development (32.6% and 37%, respectively) when compared with gastritis (11.4% and 18.2%, respectively), yet there is no association with duodenal ulcer disease (28.6% and 28.6%, respectively).

**Response:**
Thank you for your comments. We calculated these values again, and our calculations were correct. There was a significant difference in these positive rates between patients with gastritis and GC but not those with DU.

Discretionary revisions:

4) (pg 20) What are PZ1 and PZ2?

**Response:**
PZ1 and PZ2 are the abbreviations of “Plasticity zone 1” and “2,” respectively. Although it is not so important to distinguish PZ1 from PZ2, we unified these terms into one (Plasticity zone) in the revised manuscript. In addition, table 2 was moved to a supplementary file in response to another reviewer’s comment.

5) (Tables 3 and 4) For consistency, the data in the two tables should be presented in the same order. i.e. Japanese strains then those from Columbia.

**Response:**
In response to this comment, we showed these tables in the same order.