Author's response to reviews

**Title:** Genome Wide Association Scan for Chronic Periodontitis Implicates Novel Locus

**Authors:**

Ping Feng (psxffp@gmail.com)
Xiaojing Wang (timanwang@gmail.com)
Priscila L Casado (plcasado@hotmail.com)
Erika C Kuchler (eck51@pitt.edu)
Kathleen Deeley (kbd1@pitt.edu)
Jacqueline Noel (jbn9@pitt.edu)
Hyongsup Kimm (hskimm88@gmail.com)
Ji-Hye Kim (vitamin8030@hotmail.com)
Alex N Haas (alexnhaas@gmail.com)
Valquiria Quinelato (valquiriaquinelato@yahoo.com.br)
Leticia L Bonato (atm.leticia@gmail.com)
Jose M Granjeiro (jmgranjeiro@gmail.com)
Cristiano Susin (csusin@gru.edu)
Alexandre R Vieira (arv11@pitt.edu)

**Version:** 3  
**Date:** 25 June 2014

**Author's response to reviews:** see over
June 25, 2014

Editor BMC Oral Health

Dear editor:

Here we are resubmitting the original work entitled "Genome Wide Association Scan for Chronic Periodontitis Implicates Novel Locus" for further consideration for possible publication in your highly regarded journal. In this manuscript, we report a genome-wide association study, which provides new insight into loci possibly contributing to chronic periodontitis.

We believe these findings are of interest to the field and we are looking forward to hearing your assessment and possible the results of the peer-review process. The manuscript in its submitted form has been read and approved by all authors.

Our research is original, not under publication consideration elsewhere, and free of conflict of interest. Below are point-by-point responses to the reviewers’ critiques. Changes in the text are marked in yellow.

My best regards, Alex

Alexandre R. Vieira, DDS, MS, PhD
Department of Oral Biology
614 Salk Hall
School of Dental Medicine
University of Pittsburgh
3501 Terrace Street
Pittsburgh, PA 15261
Phone # 412 383-8972
FAX # 412 624-3080
Email: arv11@pitt.edu
Reviewer: FLAVIA M MARTINEZ DE CARVALHO

Reviewer's report:

I think the table 2 could be reorganized in an ascending order by locus, because in my opinion it is the best way to analyze the data. (page 15 of 16)

RESPONSE: We reordered Table 2 as requested.

Reviewer: Gareth Griffiths

Reviewer's report:

This is an innovative and extensive piece of work on the genetic risk factors for chronic periodontitis. It is a well written manuscript and has lots to commend the research conducted. There are some omissions in the descriptions which worry me slightly in terms of the methods but I am sure further descriptive narrative should help overcome these problems.

Page 5

The authors need to clarify better the inclusion and exclusion criteria for disease and health. “Participants with 30% or more teeth with sites of clinical attachment loss of five millimeters or more were defined as having chronic periodontitis [11,12].” This is the definition for generalised chronic periodontitis which is a severe form of periodontal disease. What happened to any subjects who had attachment loss of 5 mm or more on less than 30% of sites?

RESPONSE: We added the requested clarification.

Ninety-nine patients were diagnosed with chronic periodontitis (42 women and 57 men with ages ranging from 30 to 83 years). As indicated by the authors this is a
low % of disease, is this because they have been selected on generalised chronic periodontitis? Whilst the follow up studies have 430 and 180 affected individuals the pilot only has 99 affected individuals which is quite a small number in a polygenic disease to start forming hypotheses. This in part is overcome by testing in the other populations. Nevertheless, some other more important polymorphisms may not have shown up in the “pilot” due to the small numbers.

RESPONSE: This is true for all studies of this nature; there is the possibility that statistical power was not enough to detect true associations. We included originally some discussion on that limitation.

“All research records were reviewed to exclude the possibility that cases may have aggressive periodontitis”. A better description needs to be given of how this was performed. As the authors point out the current view is that this represents a different genetic disease. It is therefore considered important to have clearly defined groups. Currently the manuscript is inadequate in this respect.

RESPONSE: We added additional clarification in the Methods Section.

Rio de Janeiro has a very high % of affected individuals almost 50/50% whereas in the other 2 groups 11/89% and 30/70 %. This requires some discussion and explanation.

RESPONSE: The two studies with lower frequency of affected individuals are prospect cohorts in comparison to the case-control study from Rio de Janeiro. We added a clarification in the Discussion section.

The % of affected individuals is described along with the ethnicity of the groups overall, but what is the ethnicity of the affected individuals?

RESPONSE: We added this information on Table 1.
Reviewer: Dahlia Nielsen

Reviewer's report:

This manuscript describes a genome-wide association study of Chronic Periodontitis. Analyses were first performed on a sample of individuals collected from the US, then follow-up studies were performed on two populations collected from Brazil. The authors find several suggestive results, which they present as providing the basis of a specific hypothesis test for future studies.

General comments (Compulsory)

1) The background section is very brief, and really does not provide much background information. Most of the description provided in this section is about the current study, and even this is terse. For instance, in the second paragraph, the authors note that "the two lists of associated SNPs did not obviously overlap."

Which two lists they are referring to is unclear at this point.
RESPONSE: We revised the sentence to improve clarity. The descriptions on this section are mostly from the samples from the ARIC study (reference 9 and 10 of our manuscript).

2) The discussion section is also a bit rough, especially the first part of it. It might read better if it opened with the third sentence rather than as it does. It may be beneficial to reorganize it somewhat.
RESPONSE: We revised the initial statement of our Discussion section as suggested. This point was also suggested by reviewer 2.

Specific points (Compulsory)

3) Methods, p6: The authors state "In the analysis of the complete dataset, we also adjusted for the principal components ...." The information that the principal
components are from an analysis of population structure needs to be given, not just in the parenthetical comment.

RESPONSE: We included the information as requested.

4) Methods, Follow-up samples, p6: "... we selected the most consistent findings of both analyses." It would be worth repeating here very briefly what "both analyses" refers to. Also, how were "consistent findings" quantified? Lastly, my understanding from the previous section is that the two analyses were both performed on subsets of the full dataset, but that these subsets are not independent. How does the lack of independence affect the decisions that are made regarding downstream analyses?

RESPONSE: We included the information requested. Regarding the analyses there were done twice, one included the whole dataset, and the second just data from Whites. These analyses aimed to help identify associations that were driven by the White dataset, despite the loss of statistical power due to the analysis of less samples.

5) Methods, Meta-analysis, p6: the description "follow up sample analyses across populations" needs to be clarified.

RESPONSE: We added the requested clarification.

6) Results, Follow-up studies, p7: the authors state that three SNPs "show a trend" for association in the Rio de Janeiro sample, but don't two SNPs show statistical significance? In the methods the authors state they are using a Bonferroni adjustment, so that with 20 tests, the significance threshold is 0.05/20 = 0.0025. Two of the results listed in Table 2 have p-values = 0.001.

RESPONSE: We revised the Results section to clarify this issue.

7) Results: What does LD look like between the 20 SNPs chosen for the follow-up analyses? (Particularly the three SNPs on 21q22.11)

RESPONSE: They are in strong LD. We added this information in the Results section.
8a) Results, Meta-analysis, p8, a minor question and a more substantive one: the authors state that only one SNP shows "an association that is consistent in direction for the three studied populations." Presumably, the three populations include the two sets of samples from Brazil and the original set of samples: for the original samples, is this the full set, or just those individuals who self-reported as white? Or was agreement between the two one of the criteria for consistency (per Methods, Follow-up samples, p6)?

**RESPONSE:** The meta-analysis included the full set of samples from Pittsburgh. We added a clarification in the Methods section.

8b) More substantive: if one population produces a large p-value for a given SNP when two other populations produced small ones for that SNP, it seems there are several possible reasons. One would be that the SNP is truly associated with the trait in the populations conferring the signal, but the SNP is not associated with the trait in the third population. Another possibility is that the SNP is associated with the trait in all populations, but the sample of individuals collected from one of the populations by chance happened to provide low power. A third possibility of course is that the SNP is not associated with the trait, and the two populations that showed a signal were both false positives. One question is this: are the two Brazilian populations similar in ethnic composition? Would a true association in one population imply an association (whether or not it is detected) in the other one? It would be useful for the authors to provide this kind of detail. Are the allele frequencies between populations similar? Does a principal components analysis pull the three populations apart from each other? Or does it group the two
Brazilian populations? Going back to the reasons two populations might provide a signal when a third one does not, if it is hypothesized that the signal is real, then it is not clear if it should be expected that the direction of the "effect" seen in the population with the high p-value to be in the same direction. It definitely wouldn't be expected if there was no association in the third population (as in this case there is no real effect, so the direction of any measured effect in a given sample would be arbitrary). If there is, in truth, association in the third population, but the sample happened to display low power, then while the direction of any effect seen in the sample would be expected to be the same, it also seems not unlikely that by chance it might actually be opposite (low p-values mean effect sizes near zero in a given sample, and as such, the "effect" could be in either direction). It seems this needs to be modeled more carefully.

**RESPONSE:** We added a comment in the Discussion to address this issue inspired by this very insightful comment. Yes, the two Brazilian groups are of similar composition, a principal component analysis does not separate them. We agree that not necessarily identifying association in one group means the other is associated, however, since one group is a cohort study and the other a case-control design, differences may be due to the distinct study designs and statistical power.

**Minor Essential Revisions**

9) 2nd paragraph of Methods, Discovery Sample: authors note that "DNA was extracted according to the manufacturer's instructions", but do not note who the manufacturer is (or what the kit is).

**RESPONSE:** We originally added this information as a footnote. Oragene kits were used.

10) Methods, p6: authors should note what software/method was used for their meta-analysis.

**RESPONSE:** We added this information in the Methods section.
11) p10: "Even though we carefully took into consideration this factor, we cannot exclude the possibility that the suggestive associations we found are not influenced by variation in ethnic background of the samples." Are influenced by? or are not influenced by? **RESPONSE: We corrected the sentence as requested.**

**English Language Issues:**

p5: "Nonaffected individuals were then 767"  
**RESPONSE: We revised the sentence.**

p5: "the study group was comprise of individuals"  
**RESPONSE: We made this correction.**

p8: "only one marker showed association when the samples were pulled in a meta-analysis" (pooled in a meta-analysis?)  
**RESPONSE: We revised the sentence.**

p8: "These results are interesting if considered that ..."  
**RESPONSE: We revised the sentence.**

p9: "We have previously shown that aggressive periodontitis aggregates in families and its most parsimonious mode of inheritance is a semi-general transmission model that allows the heterozygote transmission probably varies."  
**RESPONSE: We made this correction.**