Dear Dr. Liam Messin

We would like to thank you and the reviewers for your insightful comments in response to our manuscript entitled “Investigation of molecular biomarker candidates for diagnosis and prognosis of chronic periodontitis by bioinformatics analysis of pooled microarray gene expression datasets in Gene Expression Omnibus (GEO)” (OHEA-D-18-00592). We have further revised the manuscript based on these comments and have addressed each of the issues raised, both major and minor, in point-by-point responses that can be found below. The modified text is indicated by red colored font in the revised manuscript (Supplementary Manuscript file). The revised manuscript was also checked by a native English-speaking editor.

We believe that the revised manuscript adequately addresses all of the reviewers’ concerns. We respectfully resubmit this revised and improved manuscript for your consideration for publication in BMC Oral Health.

Dear Yuichiro Kikuchi (Reviewer 1):

Major Compulsory Revisions

1. (Abstract: page4, line55-56)

Why did you choose only CSF3, CXCL12, IL1B, TAGLN, CD19, IL8, 55CD79A, TNF, and FGF2 as potential biomarkers for CP diagnosis? Are FCGR3B, SELL, FCGR2B, C3, CD53 and
IL10RA not so important? Please explain this reason to readers in the results or discussion section.

-> FCGR3B, SELL, FCGR2B, C3, CD53 and IL10RA are also important.

CSF3, CXCL12, IL1B and TAGLN are dominant biomarker candidates for CP diagnosis. CD19, IL8 and CD79A (<= 55CD79A) are the top 3 hub genes. And FCGR3B, SELL, FCGR2B, C3, CD53 and IL10RA are the following hub genes. TNF and FGF2 are upstream regulators of dominant biomarker candidates. Then we chose only CSF3, CXCL12, IL1B, TAGLN, CD19, IL8, 55CD79A, TNF, and FGF2 as potential biomarkers for CP diagnosis.

We rewrote sentences, as FCGR3B, SELL, FCGR2B, 3C, CD53 and IL10RA were hub genes. (Abstract: lines 51–56, pages 2-3; Discussion section: lines 277–279, page 13).

2. (Methods: page9, line104 and Table 1)

The authors should provide information about the clinical diagnosis of the patients (healthy control or chronic periodontitis) with the GSE 10334, GSE16134 and GSE23586.

[The sampling time, aged, probing depth, clinical attachment loss, smoking history, and so on.]

-> We added clinical information, such as probing depth, clinical attachment loss, smoking status, bleeding on probing, diabetes status, pregnancy status, and medication history (antibiotics and anti-inflammatory drugs) to Table 1 (Table 1, page 24).

Unfortunately, we do not have information about the age of patients with chronic periodontitis in GEO10334 and GSE16134 datasets.

3. (Discussion)

The discussion is very succinct. Since the explanations of this results for the majority, please increase the consideration of the results and restructure the discussion.

For example,

1. how about comparing to the results of other studies (especially reference No. 17 and 24)?

-> We have revised the discussion with consideration of the results of other studies (e.g., refs. 5 and 12), (Discussion section: lines 240–255, pages 11–12). (Reference No.17 and No.24 were changed to reference No.5 and No.12)
The authors state 'Among biomarker candidates and hub genes, the association of CD53, CD79A, MS4A1, PECAM1, and TAGLN with CP has not been previously reported.' Readers might want some additional explanations of this reasons.

->We added the reasons to the Discussion section (Discussion section: lines 268–270, pages 12-13).

Minor Essential Revisions

4. (Abstract)

Line 37. Change "Method" to "Methods".

->This has been corrected (Abstract: line 36, page 2).

Line 55. Change "Conclusion" to "Conclusions".

->This has been corrected (Abstract: line 51, page 2).

5. (Introduction)

Line 62. Change "Introduction" to "Background".

->This has been corrected (Background section: line 61, page 4).

6. (Methods)

Page 10, Line 125-127. What parameters (combined score) were used in STRING to construct the PPI network?

->We used ‘default’ in STRING to construct PPI network, did not used ‘Advanced function’. Because in this study, the PPI network is not huge. (Methods section: line 122, page 7)

7. (Table 4, 5, 6, 7)

It is difficult to understand the Table 4-7 at a glance. Like a Table 8, depending on the item to the left side, please separate a line space between right each item and delimit it.

(For example, Table 6. Separate a line space between CD19 and IL8, IL8 and CD79A, CD79A and FCGR3B, --------------)}

->We revised Tables 4–8 as suggested (Tables 4–8: pages 29–36).
8. (List of abbreviations)

Please add the lists of abbreviations.

->We added a list of abbreviations as suggested (Abbreviations: page 15, lines 296–308). We also added gene descriptions to Tables 2 and 3 (Tables 2 and 3: pages 25–28) and defined abbreviations in Table 4 caption (Table 4: lines 483–484, page 30).

9. (References)

17. Change "2015;11:2541-7.." to "2015;11:2541-7.".

->This has been changed (References: line 360, page 18).

(Reference number were changed No.17 to No.5)

Dear Ulvi Gursoy (Reviewer 2):

- Abstract: Authors state that "for the diagnosis of CP and the associated systemic diseases...", the present study protocol does not allow them to evaluate anything from associated systemic diseases. Same statement is also in introduction, which gives a false expectation.

->We agreed with the reviewer’s comment. We removed those description and rewrote. (Abstract: lines 28–30, page 2).

(‘Introduction’ was changed to ‘Background’)

- Abstract: Conclusion is the repetition of results

->We agree with the reviewer’s comment and rewrote the conclusion (Abstract: lines 51–56, pages 2-3).

- Main text: Use of abbreviations is too much. This is a gene expression study and the need to use abbreviations is understandable. However, it is getting very difficult when authors use abbreviations also for chronic periodontitis (CP), molecular function (MF), biological process (BP, etc.).

->While we agree with the reviewer, we would like to use an abbreviation for chronic periodontitis because it is a long term. Additionally, as BP, CC, and MF are commonly used terms in Gene Ontology, we would like to use these abbreviations as well. We added gene descriptions to Tables 2 and 3 (Tables 2 and 3: pages 25–28) and defined abbreviations in Table 4 caption (Table 4: lines 483–484, page 30). We also added a list of abbreviations (Abbreviations: lines 296–308, page 15).
- Data sets: GSE10334 included 63 GP and 27 AP patients, GSE 16134 included 120 periodontitis patients, and GSE23586 included 3 periodontitis and 3 periodontally healthy participants. Number of healthy control are very low. I see that the first dataset also took samples from periodontally healthy sites, but at the end those were periodontitis patients. Is the outcome reliable?

- At first we analyzed each dataset to identify DEGs to reduce bias on individual sampling and experimental condition. Then we compared respective DEGs of each dataset.

- There are unclear expressions, such as: "...positive regulation of the biosynthetic process, the biosynthetic process, and... (page 14 line 197)"

- We apologize for the oversight; “biosynthetic process” was accidentally included twice. This sentence has been revised (Results section: lines 190–191, page 9).

- Discussion is a repetition of results and needs to be rewritten. There are several studies that can be discussed: PMID:26793622, PMID: 27924602, PMID: 26334995. Basically I would expect to see a wider description of the candidate biomarkers and their interactions in disease pathogenesis. Limitations of this approach is also missing.

- We rewrote the discussion as the reviewer suggested. (Discussion section: lines 282-293, page 14)

  The aims of the studies suggested by the reviewer were to elucidate biomarkers and/or pathology of periodontitis. The groups suggest that biomarkers should be investigated using microarray mRNA analyses to further understand global health.

  PMID: 26793622 was reported importance of the clinical application of biomarkers. (Discussion section: lines 222-226, page 11)

  PMID: 27924602 was reported omics analysis such as mRNA expression was useful to elucidate pathobiology of periodontitis. (Discussion section: lines 222-225, page 11)

  And PMID: 26334995 also analyzed GEO dataset GSE16134, we discussed our results in comparison with their results (Discussion section: lines 246–255, page 12).

  We also added discussion of other studies (refs. 5) (Discussion section: lines 240–245, pages 11-12).

  We added descriptions of biomarker candidates and interactions in disease pathogenesis, and limitation of this approach. (Conclusions section: lines 284-293, page 14)