Interactive comment on "High diversity of nitrogen-fixing bacteria in upper reaches of Heihe River, Northwestern China" *by* X. S. Tai et al.

X. S. Tai et al.

leafturnred@163.com

Received and published: 16 May 2013

Dear Tina Treude,

Interactive comment on "High diversity of nitrogen-fixing bacteria in upper reaches of Heihe River, Northwestern China" by X. S. Tai et al.

Thanks a million for your suggestions and the reviewers' comments. They are really helpful to improve our manuscript! I've discussed with all the other authors seriously, which cost such long time, and carefully revised the manuscript by using blue colored text. We would like to address the changes of the manuscript and our response to the comments of the Reviewers as follows (We keep the original text of the Reviewers' comments and responded them seriatim with blue colored text.):

Reviewer(s)' General comments to Author:

1. The statistical analysis (RDA) needs more explanation. Firstly, no interpretation of the two components PC1 and PC2 is given. This would be important since PC1 explains more than 90% of the variance while PC2 explains only less than 10%. In this context, correlations with PC1 (e.g. CFUs and underground biomass in the shrub soils) should be specifically accounted for (e.g. in comparison with diversity, copy numbers and aboveground biomass correlating with PC2 in the shrub soil).

Thanks for your suggestion. PC1 explains more than 90% of the variance while PC2 explains only less than 10%. Thus, the first axis account for a valuable part of the variance while the second axis do not (Fig. 6 in p:5039). Underground biomass has a large positive correlation with the first axis and a strong influence on bacterial communities in shrub soil while available K dose in meadow soil. You can approximate correlation (depending on the standardization option used in the analysis) between species and environmental variables by a perpendicular projection of one arrow-tip onto the line overlaying the other arrow. If the projection point lies far from the coordinate origin, in the direction indicated by the arrow, the two variables are predicted to have a positive correlation. If, on the other hand, the projection point lies in an opposite direction, the two variables are predicted to have a negative correlation. A projection point near the coordinate origin (zero point) suggests that the two variables have a low correlation. The detailed explanation of correlations between bacterial data and environmental variables has been given in 3.3 (p:5023) of the ms.

Give some explanation on why an RDA was chosen and not a CCA?

Redundancy analysis (RDA) was performed to analyze the nitrogen-fixing bacterial data combined with the environmental parameters by using the CANOCO program for Windows (Version 4.5) (Braak and Smilauer, 1998; Zumsteg et al., 2012). The reasons we chose RDA other than CCA were because the data of nitrogen-fixing bacteria exhibited a linear, rather than unimodal, response to the environmental variables.

(Braak CJFt & Smilauer P (1998) CANOCO Reference Manual and User's Guide to Canoco for Windows: Software for Canonical Community Ordination (Version 4). Centre for Biometry, Wageningen.; Zumsteg A, Luster J, Göransson H, et al. (2012) Bacterial, Archaeal and Fungal Succession in the Forefield of a Receding Glacier. Microbial Ecology 63: 552-564.)

How was the RDA plot done?

The RDA plot was done with biplots method and joint plots between species and environmental variables by CanoDraw based on linear model

How confident is the analysis -did you do some ANOVA? As I see it, the whole statistical analysis is based on six samples only so please test for significance.

A Monte Carlo test was performed in CANOCO within the RDA to assess the significance of the correlations of nitrogen-fixing bacterial data with the environmental parameters. In addition, ANOVA with SPSS (version 16.0) was used to verify the results of RDA (data not shown).

Also, to my knowledge this kind of constrained analysis should result in a triplot rather than a biplot. Is there a reason that the sampling sites are missing?

Grouping pattern of sampling sites was well supported by the hierarchical cluster analysis by using SPSS (version 16.0) (Fig. 3, p:5036). So that the sampling sites are missing and a triplot is taken place by a biplot between species and environmental variables.

Finally, some of the environmental data (those with obviously no information) could be deleted so the plots are easier to be read.

We make changes as being suggested.

2. There is no discussion on the coverage/specificity of the nifH primers. Please also give a reference for them. Was it to be expected that they would mainly target

C1905

Proteobacteria? There are various universal (and usually highly degenerated) nifH primers. Give some explanation on the choice of your primers. Also, for the review process, please provide some gel pictures and melting curve data from the qPCR analysis to verify specific amplification of the targeted nifH gene fragment.

The nifH-F/nifH-R primers were chosen according to the previous studies (Rösch et al., 2002; Kizilova et al., 2012), which target a wide range of nitrogen-fixing Bacteria including Proteobacteria. The gel picture and melting curve data from the q-PCR analysis will be provided in revised ms.

(Rösch C, Mergel A & Bothe H (2002) Biodiversity of Denitrifying and Dinitrogen-Fixing Bacteria in an Acid Forest Soil. Applied and Environmental Microbiology 68: 3818-3829.; Kizilova AK, Titova LV, Kravchenko IK & Iutinskaya GA (2012) Evaluation of the diversity of nitrogen-fixing bacteria in soybean rhizosphere by nifH gene analysis. Microbiology 81: 621-629.)

3. Diversity analysis (Shannon diversity index) was done based upon OTU calculation. However, no explanation or reference is given on the cutoff for nifH OTU calculation. In equation 2 (calculation of H), what does p stand for?

The nifH OTU calculation was based on RFLP patterns. Clones with same RFLP pattern were cluster into one OUT (Jungblut and Neilan, 2010; Strassert et al., 2012). The "pi" in equation 2 means how much the clone numbers of i-th OTU account for in the clone library of a sampling site.

(Jungblut AD & Neilan BA (2010) NifH gene diversity and expression in a microbial mat community on the McMurdo Ice Shelf, Antarctica. Antarctic Science 22: 117-122. Strassert JFH, Köhler T, Wienemann THG, et al. (2012) 'CandidatusAncillula trichonymphae', a novel lineage of endosymbiotic Actinobacteria in termite gut flagellates of the genus Trichonympha. Environmental Microbiology 14: 3259-3270.)

4. The molecular data are presented without any rate measurements. Do you have

C1906

any data on nitrogen fixation activity supporting the qPCR results?

Thanks for the suggestion. This article mainly discussed nitrogen-fixing bacterial diversity and the correlations of nitrogen-fixing bacterial data with the environmental parameters, so that the rate measurements were not taken into consideration.

5. Some general statement or concluding remark referring to the objectives of the study mentioned in the Introduction should be added at the end of the abstract.

It is a good suggestion and we will add the information in abstract in revised ms.

6. The objectives could include the purpose of linking community data with environmental parameters (or what was the reason for the statistical analysis?).

Soil quality is determined by its chemical, physical and biological components and how they interact. The biological components of the soil is mainly represented by microorganisms, which carry out important functions and play a key role in the food web chain. Microorganisms can be affected by abiotic factors or by biotic factors, it is important to understand how the environment affects communities of microorganisms. In this study, discussion on the correlations between species variations and environmental gradients is one of our objectives. This kind of constrained ordination including RDA (Redundance analysis) target for finding out species variations responding to environmental gradients, and looking for potential or indirect environmental gradients which can explain changes of species data.

7. Equation 3 (p: 5022): Does this mean that there was only a single qPCR run? Or was the same standard curve used for several runs? If so, why are there different values for efficiencies?

The q-PCR run was done in triplicate for both standard and unknown samples. The standard curve was based on the average of those triplicate data. "R and efficiency were 0.996 and 105.1% respectively." (p:5021, l: 14-15) showed that the PCR efficiency is 105.1% and R2 value of the standard curve is 0.996.

C1907

8. The last paragraph of the discussion section (p: 5026, l: 3-13) is redundant to the conclusion section. Also, the conclusion rather reads like a summary, not like conclusions. Either just omit this part or rewrite in a concluding manner.

It's a helpful suggestion. We will rewrite the conclusion in revised ms as suggestion.

9. If possible, include some discussion on what is known about nitrogen fixation in association with Potentilla and Carex as the two main plant taxa.

It is a good suggestion. We will add some discussion in revised ms as suggestion.

10. The figure captions are insufficient and need more information. For example in figure 1 and 2: what do the numbers stand for? Is this reflecting significance?

We will add some new information for the figure captions. In figure 1 and 2, the numbers stand for confidence levels above 95%.

11. Figure 3 and 4 should be combined into one tree including both sampling sites. Replace 'phylogeny analysis' with 'phylogenetic tree' and give information on tree construction (algorithm, filters, outgroup, bootstrap calculation, is the tree based on nucleotides or amino acids?). What was the amount of sequences used for tree calculation? How was the alignment done?

The tree is based on nucleotide sequences. Phylogenetic dendrogram based on a comparison of the nifH sequences determined from the samples of shrub soil (18 sequences) and meadow soil (44 sequences) and their closest phylogenetic relatives (13 and 25 sequences respectively). The reference sequences were obtained by BLAST. The alignment was performed by ClustalX (version 1.8). The tree was created by using the neighbor-joining method. The numbers on the tree indicate the percentage of bootstrap sampling derived from 1000 replications.

Reviewer(s)' Technical comments to Author: 1. Italics should be used throughout the manuscript for taxa and genes (e.g. section 3.2, and nifH throughout the manuscript).

C1908

Thanks! We italicized it.

2. p: 5022: give references for the equations

The reference is provided now.

3. Table 2: add assignment to the study sites (shrub and meadow) for site 1-6

We added the assignment.

4. Table 3: Replace "Cites" with "Reference", "Phylum" with "Dominant phyla" (if this is what is meant), "In this study" with "This study"

We made changes as being suggested.

5. Figure 2: "Copy numbers" instead of "Copies"

We changed it.

6. p: 5016, l: 5: Replace "In present study" with "In the present study"

Thanks. We replaced it as suggestion.

7. p: 5016, l: 13: Replace "than shrub soil " with "than in shrub soils"

We replaced it.

8. p:5016, l: 13: Why "contrarily"? Does this refer to the shrub soil, where copy numbers are higher? Also " copies" should be replaces by "copy numbers"

Copy numbers of nifH in shrub soil are higher. We changed it as suggestion.

9. p:5017, l:1: Omit "rather"

We deleted it.

10. same page, line 5: "of nifH gene" should be replaced with "of the nifH gene"

Thanks. We replaced it as suggestion.

11. same page, line 10: "is still poorly described" instead of "was still poorly described"

We changed it.

12. same page line 23: replace "and they are" with "which is"

Thanks. We replaced it.

13. same page line 25: rephrase the sentence starting with "Therefore.." for example: "Therefore, environmental studies on nitrogen-fixing communities are needed in nitrogen limiting : : :"

Thanks. We rephrased the sentence as suggestion.

14. page 5018 heading 2.1: "Study site and soil sampling" instead of "soils samples"

We changed it.

15. same page line 5: omit "latitude" and "longitude"

We deleted them as suggestion.

16. same page line 10: replace "vegetation types are varied" with "vegetation types varied"

We changed it as suggestion.

17. same page line 23: "." missing after the reference"

Thanks! We added it.

18. page 5019, line 22: Omit the last sentence of the section or describe in more detail.

We deleted it.

19: same page section 2.6: rephrase this and give more information here (see above). Was ClustalX used for the alignment?

Thanks. We rephrased it and added some information. The ClustalX was used for the

alignment.

20. same page, line 10: replace "nifH gene" with "nifH gene copies" and "measured" with "determined", line12: "Standards" instead of "standard", line13: omit "typical" and replace with "a clone with correct insert"

We make changes as being suggestion.

21. same page line 16: Start new sentence after copies with "This was used for a standard curve.."

Thanks! We rephrased the sentence.

22. same page line 22: add bracket after "25 ng"

We changed it.

23. page 5021 line 4: omit "cultured", line 8: start new sentence after "Eq(1)" with "Thereby, ..", line 9 insert "and" before "N"

We make changes as being suggested.

24. same paragraph: explain qPCR at least when using the first time ("quantitative or real time PCR – qPCR")

The information is added now.

25. page 5022: "Gen-Bank" instead of "Gen Bank"

We changed it.

26. same page line 5: Replace "Cultured nitrogen-fixing.." with "Colony forming units (CFUs) of nitrogen-fixing bacterial communities.."

Thanks. We replaced it as suggestion.

27. in the same line: replace "are showed" with "are shown", line 6: Start new sentence after "Fig. 1" with "CFUs..", same line: are the copy numbers related to dry or wet soil?

We make changes as being suggestion. The copy numbers are related to wet soil.

28. same page line 8: replace "than shrub soil" with "than in shrub soil"

We replaced it.

29. same page line 12: replace "abundant" with "higher" and insert an "in" before "meadow soil"

Changed as suggestion.

30. same page line 16: replace "from meadow soil clustered" with "from meadow soil also clustered"

Thanks. We replaced it.

31. page 5023 line 9: replace "reveals" with "shows", line 11: what means "principally" here? Rephrase

Thanks. We changed it and the sentence is rephrased now.

32. same page line 20: Which conclusion is supported? Rephrase this. Also, qPCR data cannot support diversity data.

Thanks. We rephrased the sentence.

33. page 5024 line 7: replace "from same" with "from the same"

We replaced it.

34. same page line 14: what means " a lower OTU"? Rephrase this and use either "lower Shannon diversity index" or "lower number of OTUs".

Thanks. We rephrased it as suggestion.

35. in the same line: What means "Contrarily" here? Rewrite this and explain in more detail. In the same paragraph: What is meant by "disturbance"? Is there a definition for this?

We rewrote the sentence and added a definition for "disturbance" as suggestion.

36. page 5025 line 13: Replace "e.g. Potentilla parvivola or Carex alrofusca" with "e.g. Potentilla parvivola and Carex alrofusca, respectively"

We replaced it.

37. same page line 25: omit "principally" or explain in more detail

We deleted it.

38. same page line 26: replace "The result was" with "This result is"

Thanks. We changed it as suggestion.

Thanks again for your suggestions and referees' comments! I hope that the responses answer the Reviewers' comments, and that he and his colleague feel these are reflected in the revised manuscript. I look forward to hearing your decision soon.

Yours sincerely, Xi Sheng Tai