

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

Anonymous Referee #1

Received and published: 1 May 2013

General comments: Tai and coauthors studied the community of nitrogen-fixing bacteria in two different sites of the upper reaches of Heihe River, China. Three biological replicates per site obtained from 5 sub-replicates that were pooled were used for 1) diversity analysis based on *nifH* environmental clone libraries, 2) *nifH* copy numbers based on qPCR, 3) diversity and number of CFUs, 4) soil physicochemical analysis, and 5) constrained multivariate statistics (RDA). The two sites studied varied in altitude and vegetation cover with the lower shrub soil being dominated by *Potentilla* and the higher meadow soil dominated by *Carex*. Based on the Shannon diversity index and in comparison to related studies, the two sites, shrub and meadow soil, showed a high diversity. The authors found pronounced differences between the two sites in basically all properties targeted (community structure, diversity, *nifH* abundance etc.) and also with respect to the environmental parameters influencing them. While, for example, in

C1423

the shrub soil diversity and *nifH* copy numbers were both correlated with aboveground biomass, no such correlation was observed in the meadow soil. The study by Tai and coauthors adds some valuable data to existing environmental studies on nitrogen fixing bacteria. Accordingly, in their manuscript the authors account for this by putting their study into the context of related studies (e.g. Table 3). Nevertheless, the study suffers from an often sloppy performance and poor English. The statistical analysis is not well explained and interpreted and the figures lack appropriate figure captions. For publication, the manuscript needs to be significantly improved.

Specific comments: 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). Give some explanation on why an RDA was chosen and not a CCA? How was the RDA plot done? 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. 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? Finally, some of the environmental data (those with obviously no information) could be deleted so the plots are easier to be read. 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 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. 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? 4. The molecular data are

C1424

presented without any rate measurements. Do you have any data on nitrogen fixation activity supporting the qPCR results? 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. 6. The objectives could include the purpose of linking community data with environmental parameters (or what was the reason for the statistical analysis?). 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? 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. 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. 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? 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?

Technical comments: 1. Italics should be used throughout the manuscript for taxa and genes (e.g. section 3.2, and *nifH* throughout the manuscript). 2. p: 5022: give references for the equations 3. Table 2: add assignment to the study sites (shrub and meadow) for site 1-6 4. Table 3: Replace "Cites" with "Reference", "Phylum" with "Dominant phyla" (if this is what is meant), "In this study" with "This study" 5. Figure 2: "Copy numbers" instead of "Copies" 6. p: 5016, l: 5: Replace "In present study" with "In the present study" 7. p: 5016, l: 13: Replace "than shrub soil " with "than in shrub soils" 8. p:5016, l: 13: Why "contrarily"? Does this refer to the shrub soil, where copy numbers are higher? Also "copies" should be replaced by "copy numbers" 9. p:5017, l:1: Omit "rather" 10. same page, line 5: "of *nifH* gene" should be replaced with "of the *nifH* gene"

C1425

11. same page, line 10: "is still poorly described" instead of "was still poorly described" 12. same page line 23: replace "and they are" with "which is" 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 . . ." 14. page 5018 heading 2.1: "Study site and soil sampling" instead of "soils samples" 15. same page line 5: omit "latitude" and "longitude" 16. same page line 10: replace "vegetation types are varied" with "vegetation types varied" 17. same page line 23: "." missing after the reference" 18. page 5019, line 22: Omit the last sentence of the section or describe in more detail. 19: same page section 2.6: rephrase this and give more information here (see above). Was ClustalX 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" 21. same page line 16: Start new sentence after copies with "This was used for a standard curve.." 22. same page line 22: add bracket after "25 ng" 23. page 5021 line 4: omit "cultured", line 8: start new sentence after "Eq(1)" with "Thereby, ..", line 9 insert "and" before "N" 24. same paragraph: explain qPCR at least when using the first time ("quantitative or real time PCR – qPCR") 25. page 5022: "Gen-Bank" instead of "Gen Bank" 26. same page line 5: Replace "Cultured nitrogen-fixing.." with "Colony forming units (CFUs) of nitrogen-fixing bacterial communities.." 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? 28. same page line 8: replace "than shrub soil" with "than in shrub soil" 29. same page line 12: replace "abundant" with "higher" and insert an "in" before "meadow soil" 30. same page line 16: replace "from meadow soil clustered" with "from meadow soil also clustered" 31. page 5023 line 9: replace "reveals" with "shows", line 11: what means "principally" here? Rephrase 32. same page line 20: Which conclusion is supported? Rephrase this. Also, qPCR data cannot support diversity data. 33. page 5024 line 7: replace "from same" with "from the same" 34. same page line 14: what means "a lower OTU"? Rephrase this and use either "lower Shannon diversity index" or "lower

C1426

number of OTUs". 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? 36. page 5025 line 13: Replace "e.g. *Potentilla parvifolia* or *Carex alrofusca*" with "e.g. *Potentilla parvifolia* and *Carex alrofusca*, respectively" 37. same page line 25: omit "principally" or explain in more detail 38. same page line 26: replace "The result was" with "This result is"

Interactive comment on Biogeosciences Discuss., 10, 5015, 2013.

C1427