Biogeosciences Discuss., 10, C2173–C2177, 2013
www.biogeosciences-discuss.net/10/C2173/2013/

Biogeosciences Discussions
Open Access

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 #2

Received and published: 22 May 2013

In their paper "High diversity of nitrogen-fixing bacteria in upper reaches of Heihe River, Northwestern China", Tai and colleagues studied soil N-fixing bacterial communities in two habitats, namely Potentilla-dominated vs. Carex-dominated grasslands, along an altitudinal gradient. To this end, they determined the abundance of both N-fixing culturable bacteria and nifH genes (by using qPCR). They also characterised N-fixing bacterial community structure and – to a certain extent – community composition by combining RFLP and cloning/sequencing approaches. All parameters studied were found different between the two habitats.

This study has some potential and may provide insights into this bacterial functional group in alpine environments. Authors also made a significant effort in integrating their observations in a broader context by comparing their results with previous studies.

Full Screen / Esc
Printer-friendly Version
Interactive Discussion
Discussion Paper

However, I strongly struggle with the manuscript form and content, as well as with the results reliability.

Main comments: 1/ One of the main weaknesses of the study is that it does not address clear questions and does not have hypotheses. It therefore remains too descriptive and lacks of structure. It's a pity when one look at the sampling design, which I've found quite interesting: the sampling sites are located along an altitudinal gradient, the two habitat types studied are dominated by plant species that may differ in their growth strategy/ecophysiology... And all this may affect nitrogen cycling (see e.g. Chapman et al 2006 New Phytol). Why did the authors choose to compare these two environments? Why did they work on an altitudinal gradient? What are the challenges behind this? What were the expected results? What conclusions can be drawn about the diversity and/or functioning of these systems? Reorganizing/rewriting the introduction / discussion with these questions in mind may significantly improve the manuscript quality.

2/Statistical analyses: 2.1/ The number of samples/replicates used for the statistical analyses is really unclear: 5 samples/quadrats were pooled, rendering 3 composite soils / sites (one per quadrat). Then, authors performed 3 DNA extractions per "samples" (composite samples?) and 3 PCR replicates per extractions (or pools of extractions?). But at the end, there is not any information on the number of clone libraries, qPCRs, cultivation libraries, or sets of RFLP profiles obtained for each site and used for statistical analyses (ideally one per quadrat, 3 per sites). This needs to be clarified. Besides, I've noticed that SD values were absent in most tables/figures, and that no statistical tests of significance (e.g. mean comparison tests or permutations tests) were used (or at least reported: pvalues must be provided) to prove that shrub vs. meadow bacterial communities characteristics are indeed different or that they indeed co-vary with environmental variables. 2.2/ The analyses themselves are poorly described and not always correctly performed: 2.2.1/ Phylogenetic trees: First, authors should indicate which model of DNA evolution was used to construct the trees

BGD
10, C2173–C2177, 2013

(amongst e.g. Jukes-Cantor, Kimura, Tamura models). Second, many branches in Fig. 4 and 5 are not very well supported (bootstrap value < 95%) to make reliable taxonomic assignments. Furthermore, these two trees were not constructed with the same references (e.g. no γ-proteobacteria references in Fig. 4). Making inferences on the differences between meadow and shrub communities (in terms of composition) based on such trees is not reliable (as done p. 5022 l.20ff). I would suggest constructing the phylogenetic tree (i) by using a larger number of references, encompassing a larger number nifH groups to make it more robust, and (ii) by including both meadow and shrub OTUs in the same phylogenetic tree. Authors might be interested, for instance, in the UNIFRAC website (http://bmf.colorado.edu/unifrac/), which provide statistical tools for testing differences in bacterial community composition between samples based on phylogenetic trees. 2.2.2/ Clustering analysis: Which clustering method was used? (average/complete/single linkage?). Authors should also be more careful with their interpretations regarding this analysis: it shows that shrub and meadow communities display different characteristics (in terms of evenness, nifH genes abundances), not that the communities are different (as stated/suggested e.g. p.5022 l.18, p. 5024 l.8), the latter rather evoking that they harbour different taxa, an observation that is not supported by the analyses performed here (cf. 2.2.1). 2.2.3/ The "Correlation" analysis, which appears to be an RDA analysis (i.e. based on linear regressions). First, authors should be careful: correlations differ from linear regressions. Second, it is unclear to me what has been really done: authors' reply to reviewer 1 suggests that RDA was chosen to find out what are the environmental parameters that may be responsible for species variations (that's sound ok for me, it is how I use it). But what I understand from the Material&Method section and Fig 6 is that community characteristics, not community composition, were used as response variables. This should be clarified.

3/ Results interpretations are not always clear: p.5024 l.8ff: Awkward: "samples spatially closer to each other, regardless of their location in the geographic range"? Ramette & Tiedje actually reported that bacterial community composition varies at small spatial scales due to environmental heterogeneity. It is not the point here: the fact that

BGD
10, C2173–C2177, 2013

similar soils harbour communities that display similar characteristics (again, only in terms of evenness and nifH gene abundance) is to be related to their abiotic and biotic contexts, which tend to be the same within the same habitat, and which also happen to be the same for close samples (environmental context and geographic location are cofounding factors with the sampling design used by the authors). I don't believe that isolation by distance may be at play at this spatial scale. p.5024 l.15ff: What "disturbance" stands for exactly? Freezing? Was it true during the sampling campaign? The sampling period is not indicated in the Material & Methods section. Furthermore, most of these statements are not supported by Table 1: meadow and shrub soils seem to display the same TN content (both $\sim$ 5 g.kg -1 on average), and one cannot have any idea of moisture and temperature fluctuations from the data presented here. p. 5025 l.16: Do the authors really talk about natural selection (related to evolutionary processes, which is very unlikely: again, I don't believe that isolation by distance may be at play at this spatial scale) or about habitat filtering, i.e. the fact that the abiotic/biotic conditions may enhance the fitness of certain taxa?

4/ Finally, the literature is often cited inappropriately, mostly because the cited references do actually already cite other references in their introduction. For instance: p. 5016 l. 26: wrong ref. The one really discussing that point is Lynch, JM, Hobbie, JE (1998) Microorganisms in Action. Blackwell Scientific Publications, Oxford p. 5017 l. 5: idem. Results from this ref do not support this statement. p. 5024 l. 5: Kizilova et al., 2012 is not a review and does not really test for primer pair taxonomic coverage. See rather Gaby and Buckley 2012 PLoS One for an in silico evalutation of nifH primer pairs.

Minor comments p. 5016 l. 23: "particularly in those without any chemical fertilizer": did the authors mean low nutrient availability? p.5020: As GenBank is constantly evolving, authors should indicate the date at which they made their BLAST analyses (or the GenBank database version) p.5021 l.6: why did the authors used the 16S gene here and not the nifH gene instead? Results obtained from those data cannot be crossed with

BGD
10, C2173–C2177, 2013

those obtained from the clone library. p. 5021 l.10: redundant with p.5019 l.23 p.5023 l.12: is it copies or number of copies? To be corrected throughout the manuscript p. 5023 l.20: qPCR instead of Q-PCR. Table 2: What are the percentages provided in the first row? Fig. 6: What are the different colours for?

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