

Biogeosciences Discuss., author comment AC1
<https://doi.org/10.5194/bg-2021-103-AC1>, 2021
© Author(s) 2021. This work is distributed under
the Creative Commons Attribution 4.0 License.

Reply on RC1

Tzu-Hsuan Tu et al.

Author comment on "The biogeographic pattern of microbial communities inhabiting terrestrial mud volcanoes across the Eurasian continent" by Tzu-Hsuan Tu et al.,
Biogeosciences Discuss., <https://doi.org/10.5194/bg-2021-103-AC1>, 2021

- The analysis of the article is not sufficient to support the conclusion drawn by the author. The authors emphasized stochastic process in the title, but they told nothing related to stochastic processes in either introduction or result or M&M section. How will readers know stochasticity and determinism if you do not mention it in the introduction? The author only analyzed the influence of environmental factors on microbial diversity and the distance attenuation relationship. How can you link these analyses to community assembly mechanisms? So, I don't know how the author concluded that random factors dominated the variation of microbial community.

Thanks for the comment. We decide to change the title to a more general form as "Biogeographic pattern of microbial communities inhabiting in mud volcanoes across the Eurasian continent" to clear the inconsistency raised by the reviewer. Intrinsically, drift, dispersal, and mutation could all be categorized as stochastic processes even though they are non-exclusive and interact with deterministic processes. These processes would generate community diversity pattern indistinguishable from random chance. In contrast, "selection" is regarded as deterministic process which involves nonrandom, niche-based mechanisms, including environmental filtering (e.g., pH, temperature, salinity, and geochemical or redox variations) and various biological interactions (e.g., competition, mutualisms, predation, and tradeoffs). Such selection factors could have led to the establishment of a variety of core microbiomes inhabiting distinct environments or organized spatially as in a gradient and, thus, spatially autocorrelated. We have revised the first paragraph of the "Introduction" to provide readers with more description about deterministic and stochastic processes. As being suggested by reviewer 2 and to provide the assessment of the relative stochasticity of community diversity variation, we have introduced a new calculation of the NST index ("NST" stands for normalized stochastic ratio; Ning et al., 2019). We have categorized all samples into "bubbling fluid", "surface sediments", and "within site sediment", each representing source communities, source communities susceptible to minor surface impact, and communities potentially altered by localized geochemical context, respectively. The results demonstrate a high stochastic control on "bubbling fluid" (NST = 1.00) and "surface sediment" (NST = 0.70) and a relatively strong control of deterministic process on "within site sediment" (12 out of 15 sites with NSTs < 0.5 and the remaining 3 with NSTs between 0.54 and 0.61), a data pattern consistent with the results obtained from existing DDR, NMDS, and CCA analyses. Similar approach has been applied to the calculation of pNST, which is a measure of stochasticity based on phylogenetic distance between communities and represents an analogy of β NTI. The results demonstrate a pattern similar to that from NST. High

stochastic control has been observed for “bubbling fluid” (pNST = 0.65) and “surface sediment” (pNST = 0.73), whereas mild deterministic control has been inferred for “within site sediment” (8 out of 15 sites with pNSTs < 0.5 and the remaining 7 with pNSTs between 0.51 and 0.92). Compared with ASV abundance dissimilarity, phylogenetic dissimilarity leads to a pattern with the stochastic control (pNST > 0.5) on more sites. This could be exemplified with sites SI and SH where relatively small numbers of ASVs were recovered. The pNSTs for these two sites are much higher than the corresponding NSTs (0.44-0.51 versus ~0.01), suggesting that contrast proportions of phylogenetically distant ASVs between communities lead to a higher degree of random dissimilarity. While all of our existing analyses are based on the ASV abundance dissimilarity (e.g., Bray-Curtis distance for NMDS and DDR), we will present the NST results only in order to maintain the consistency.

- The article lacks clear goal and scientific issues, which prevents readers to understanding the article quickly.

Thanks for the comment. To make the goal even clearer, we have revised the objective of the study as “This study aims to determine prokaryotic community compositions and structures associated with terrestrial MVs and the underlying mechanisms by examining the control of deterministic and stochastic processes on community variations over a horizontal scale across the Eurasian continent and a vertical scale over a redox transition”.

- The author should check the article carefully before submitting it. There are many confusing descriptions and mistakes across the article.

Thanks for the comment. We will revise the confusing descriptions more carefully.

- Lines 80: There are 17 samples in Fig.S1, but 16 are written in the material.

Thanks for the comment. Indeed, 16 cores were investigated in this study. We have revised the manuscript and supplementary information.

- Lines 105: Clarify the number of ASVs.

Thanks for the comment. Different sequencing depths could lead to the overestimate or underestimate of diversity (both alpha and beta diversities). To circumvent the shortcoming, ASV normalization using the function cumNorm embedded within metagenomeSeq in R was adopted in this study. The method considers the ASV quantile distribution for each sample by adjusting the sequence number for individual ASVs while keeping the total ASVs the same before and after the normalization. The method does not sacrifice sample diversity contributed from rare ASVs, thereby providing a better assessment on the community variation across different spatial scales or controlled by localized environmental factors. Nevertheless, the number of ASVs in each sample has been added to Table S1.

- Lines 110: Why not consider pH, which is generally believed to be the most important

factor affecting the microbial community.

Thanks for the comment. The measurement of pH was only conducted for bubbling fluid primarily because the volume of porewater is often not sufficient. In addition, the pH of bubbling fluid could not be directly extrapolated to that of subsurface porewater because microbial metabolisms along a redox transition could divert pH into a range distinct from its starting value. For example, anaerobic methane oxidation, the key metabolism in mud volcano, converts methane into carbon dioxide that could lead to the enhanced pH. Under the context that the in situ pH could be constrained only with the real measurement, we did not incorporate pH into CCA.

- Lines 115: How to distinguish between stochastic and deterministic.

Please see response 1 in which we have performed the calculation of NST and pNST in addition to the existing DDR, NMDS and CCA. The combination of the methods described above provides qualitative and quantitative assessments on the mechanisms controlling biogeographic pattern.

- Lines 160: There is no information about bacteria orders in Fig.2b.

Thanks for the correction. We have removed the description in Fig. 2b.

- Lines 170: How did the 136 samples come from? The x-axis in Fig.3 seems to be only 15.

We did not plot the occurrence of Proteobacteria in all 136 samples in Fig 3 because the sample size is too big to be properly shown in the figure. Instead, the occurrence was condensed into the site level for the simplicity of presentation. Detailed data regarding the occurrence of individual ASVs (so called ASV table) would be available upon request.

- Lines 200: The author calculated shannon, chao1 and richness earlier, but why only shannon is studied here.

The three diversity indices represent different estimates of community richness. Our calculations yielded that their variation patterns were essentially comparable to each other. Therefore, we only chose the Shannon index for further presentation.

- Lines 210: Why is geographic distance not included in the CCA analysis. I would suggest authors to consider both physicochemical and geographical factors to reveal the contribution of environmental and spatial factors to community variation. Based on this, you may state stochastic and deterministic processes in this paper.

CCA attempts to elucidate the relationships between biological assemblages and environmental characteristics intrinsically attributed to the corresponding niches, thereby facilitating to identify environmental variables that are vital in determining community compositions. In essence, the data required for CCA are all inherited from the attributes of

individual samples, such as abundances of specific taxonomic units (e.g., ASV table) or physio-chemical metadata (e.g., temperature, pH). Geographic distance represents a distance gauge between samples, not an intrinsic attribute associated with individual samples. Therefore, it could not be considered as an environmental variable for CCA. Instead, it could be used to constrain the effect of dispersal on community (dis)similarity by relating beta diversity with distance, which is what we did in the original manuscript (shown in Fig. 5). To provide a more quantitative assessment on the stochastic control, we have additionally calculated the NST and pNST for three sample categories, including "bubbling fluid", "surface sediment", and "within site sediment". Please see response 1 for more details.

- Lines 220: Which graph has a slope of 0.210, I can't find it. Fig.5b: One missing point in the formula.

Thanks for the correction. The slope and the figure have been revised.