Biogeosciences Discuss., 9, C174–C179, 2012
www.biogeosciences-discuss.net/9/C174/2012/

Biogeosciences Discussions

Interactive comment on "Microbial community diversity of the eastern Atlantic Ocean reveals geographic differences" *by* C. J. Friedline et al.

A. Ramette (Referee)

aramette@mpi-bremen.de

Received and published: 27 February 2012

GENERAL COMMENTS

Overall, the study is of good quality and a better understanding of the nature of bacterial communities and of their drivers of diversity in the environment was achieved. In this sense, the study represents an interesting contribution to the scope of Biogeosciences by revealing the interplay of multiple factors on the diversity of specific pelagic bacterial communities in the Atlantic.

The work is clearly presented. Yet, the abstract is rather descriptive of the experimental strategy and results, and the main ecological hypotheses or research questions need to be stated more clearly. There were only few typos, which are indicated below. Tables

Full Screen / Esc

Printer-friendly Version

Interactive Discussion

Discussion Paper

and figures are appropriate.

The two important points of the study were: 1) Detailed description of vertical and spatial variation in community structure of bacterial assemblages in the water column of the eastern Atlantic ocean. 2) Beyond traditional community ecology statistical methods, using Bayesian inference to investigate OTU patterns, as a way to complement the classical community ecology approach.

The molecular methods that were used have become the gold standard in the field, and the main questions and comments I have mostly concern the statistical treatment of the data, as this is one of the main points of the study. In particular the use of Bayesian inference, which offers an interesting option to look at the data, would need to be better justified and explained.

The results could also incorporate some recent work that has been published and which compare pelagic and benthic bacterial communities based on the same molecular approach (Zinger et al. 2011 Global patterns of bacterial beta-diversity in seafloor and seawater ecosystems PLoS ONE), and also work that investigate the impact of rarity or dominance on ecological interpretation of community patterns (e.g. Gobet et al. 2010 Nucleic Acids Research).

SPECIFIC TECHNICAL COMMENTS

I. Ecological statistics.

A) L8-9P117: Using ACE or Chao1 estimators make use of singletons in their formulae. Therefore if they are removed (L18 P116) from the data, what kind of effects this may have on your reported diversity estimates?

B) If the authors' goal was to explain community variation as a function of environmental vs. spatial variables, the use of indirect gradient analysis (i.e. PCoA axes calculated on the community data, which are in a second step further explained by environmental and spatial variables) or (partial or simple) Mantel tests might not be appropriate. Those

BGD
9, C174–C179, 2012

methods are not recommended when one wants to determine the respective effects of space vs. environment in explaining changes in beta diversity (see Legendre et al. 2005 Ecol Monographs; Peres-Neto et al. 2006 Ecology; Ramette and Tiedje 2007 PNAS for an example in microbial ecology). If biotic and environmental variations share a common spatial structure, spatial processes must generally be considered when examining the effects of environment on biotic variation (Borcard et al. 1992 Ecology). Instead, variation partitioning using constrained ordination (e.g. redundancy analysis) should be preferred because the amount of covariation between groups of factors is also quantified, and this is not the case when using (simple or partial) Mantel tests. The second reason is that the variance of a dissimilarity matrix among sites is not the variance of the community composition table nor a measure of beta diversity (Legendre et al. 2005 Ecol Monographs), so what is represented by PCoA is the "variation of the variation" in community structure, while the variance of the community composition table is a measure of beta diversity. The Mantel approach is, however, still appropriate for testing other hypotheses, such as the variation in beta diversity among groups of sites. It is not clear if the authors were aware of the implications the choice of the numerical strategies has on their ecological interpretation.

C) Before computing the environmental dissimilarity matrix needed for the Mantel test, it seemed that standardization of the variables to the same variance was not performed. Standardization to a variance of unity gives the same weight to each variable in the matrix. If not, variables measured on larger scales will have more impact on the resulting ecological interpretation.

D) The study uses the 3D location of the samples (latitude, longitude, depth). Why log-transforming variable depth and not the other ones (L4 P118)? Keeping those spatial data on the same scale would enable a more homogeneous description of spatial patterns.

II. Bayesian inference.

BGD
9, C174–C179, 2012

Bayesian inference was used to find the best relationships (as visualized by an unrooted dendrogram) among samples, by using the recoded relative sequence abundance of each OTU. A) Hypothesis ii (L24, P118) “the distribution and abundance pattern for any OTU is independent of the pattern of any other OTU” is difficult to assume, given the way data were collected and the fact that bacteria in communities do interact. If this hypothesis is true, it would correspond to bacterial cells inactively floating in the water column, without any function or interactions with their neighbors.

B) L8-9P119: traditional approaches loose some information, but by transforming the data into ranks, is it not the same that was achieved?

C) L6-7P120: How does the recoding affect the result? In fact, by rank-transforming the data and recoding them, the initial raw data are therefore very different from this new set. It would be good to apply the same approaches to both sets to see how different the results might be.

D) Why was Bayesian inference required here? For instance, could the authors make use of prior knowledge? L10-11P120: what does “across-site rate variation following a gamma distribution” have to do with the current problem? There is no rate of evolution here, so are those parameters and model adequate for the data at hand? Generally MrBayes can be used to infer phylogenies based on sequence data so phylogenetic or evolutionary models need to be supplied. If the authors used the default settings, one should ask if this is really appropriate given the type of data (relative sequence abundance) that is used in the study.

E) L9P130: PCoA is designed to summarize the variation in a data matrix onto few axes (generally up to three components) to facilitate the interpretation of the main patterns of variation in a dataset. The Mantel test is designed to correlate two matrices and test for the significance of the correlation coefficient by considering the fact that the data originates from matrices (i.e. data not freely exchangeable in the permutation scheme). Therefore, with both approaches the idea is clearly not to visualize individual OTU

BGD

9, C174–C179, 2012

variation. Yet, the information from individual OTU variations is considered in the final matrix. Those approaches are not dissimilar to what the authors acknowledged (L19-20P130) as being their goal: “to identify the optimal tree topology that best explains the relationships based on overall patterns of OTU abundance”. One could also obtain the same idea by applying a clustering algorithm on the Bray-Curtis dissimilarity matrix, and use a bootstrapping approach to test for the reliability and support of the branches in the dendrogram. So again, it is not clear why and how the Bayesian approach is better than the “classical” approach to deal with such dataset.

F) L3-5P127: The hypothesis that the less (sequence) abundant members drive community change can be tested with the data. Considering only parts of the data to investigate the role of dominance or rarity has already been explored elsewhere (e.g. Gobet et al. 2010 Multivariate Cutoff Level Analysis (MultiCoLA) of Large Community Datasets. Nucl Acids Research).

MINOR COMMENTS

- The title should probably indicate “Bacterial” instead of “Microbial” community diversity to be more in line with what has been done in the study.

- Results: Many values are reported with some indication of variation. It would be useful to know if these are standard deviation, and also how many observations were used in each case to calculate the sd.

- L10P110: “16s rDNA” should be changed to “16S rRNA gene”.

- L25P110: “suggest”

- L9P115: “PCR reactions” change to “PCR”

- L3P116: Why using 95% bootstrap support? Provide a reference or experimental justification.

- L10P116: reformulate “sequence containing “N””.

BGD

9, C174–C179, 2012

Interactive Comment

- L3P117: why not using “OTU richness” instead of “species richness” as the reason for using “OTU” is to avoid dealing with the microbial species concept in microbes?

- L28P118: define what M and N are here.

- L3-5P127: Make sure here that your readership understands that you are talking about relative sequence abundance and not actual counts.

- L14-15P127: Chlorophyll-a is written in two different ways.

- L6P130: “experimental observations” sounds like different experiments were carried. Maybe use “values” instead?

Interactive comment on Biogeosciences Discuss., 9, 109, 2012.