

Interactive comment on “Characterization of particle-associated and free-living bacterial and archaeal communities along the water columns of the South China Sea” by Jiangtao Li et al.

Anonymous Referee #2

Received and published: 7 July 2020

General comments

This manuscript by Li et al. is an examination of the PA and FL microbial communities found throughout depth at two stations in the South China Sea, and is an interesting addition to the body of literature on particle association of ocean microbes. The general patterns found in the microbial community composition data are reasonable. However, it is unclear whether the authors performed specific important transformations of the count table data before statistical analyses. Without that point being clarified, I would be very cautious to interpret anything from the ordinations and diversity calculations. The authors also included an analysis of seawater age, which is a unique aspect of

C1

this dataset. I would like to see a bit more exploration of that in relation to specific microbial taxa. Generally, I think this is an interesting and publishable dataset, but some refinement of the statistical methods are necessary.

Specific Comments

L107- A little context on the stations would be nice. They seem to be part of a larger study. What is their significance and why were these two chosen?

L172 This is a very outdated version of SILVA. I'm not going to argue that the classification should be redone, but there are likely implications that can be discussed (eg. it may explain the large amount of unidentified archaeal taxa. Also, another example: the Nitrospinaceae are no longer considered part of the delta-proteobacteria, but in their own Nitrospinae phylum, L348).

Section 2.4- There is no mention of transformation/normalization of count tables or removing singletons. Removing singletons is absolutely vital for analyzing OTU data because 97% clustering introduces lots of singleton artifacts (Edgar RC. 2017. Accuracy of microbial community diversity estimated by closed- and open-reference OTUs. PeerJ 5:e3889. DOI: 10.7717/peerj.3889), and this could greatly skew estimates of diversity and ordination results. Removal of singletons will also change the results for Figure 9 and the diversity estimates. Transformation is absolutely necessary for ordinations (see Legendre and Gallagher. 2001. Ecologically meaningful transformations for ordination of species data. Oecologia 129:271–280. DOI: 10.1007/s004420100716 and Gloor GB, Macklaim JM, Pawlowsky-Glahn V, Egozcue JJ. 2017. Microbiome Datasets Are Compositional: And This Is Not Optional. Front Microbiol. 8(NOV):1–6. doi:10.3389/fmicb.2017.02224), so it needs to be made clear if this was done or not. Tables S1 and S2 appear to be raw count data with no transformation or normalization

Section 2.5- No quality parameters of the qPCR assays are reported (eg. R² of the standard curve or efficiency of the reaction). Also what standard was used for qPCR? A PCR product? Genomic DNA from cultured organism with a known 16S rDNA copy

C2

number? This should be briefly mentioned.

L249-256 & L377 It's very interesting that diversity decreased mid-water column and then increased again below that. Can the authors speculate what's going on here? Could they relate it to their DOC/ POC data or age of seawater?

L257- 259 & Fig. 3 I see the separation of the 3 identified groups in the ordination but it is unclear which test was used to statistically distinguish these groups or if the circles were just drawn based on looking at the figure.

L410-432 Since the authors analyzed the age of seawater, it would be nice to interpret this more directly with respect to DOC/POC quality and microbial community composition. What is the precise impact on microbial community composition based on age of seawater (which groups were important and why?). I like that this part of the discussion begins to interpret the impact of silicate (which is really an indirect correlate and likely a signal of diatom biomass impacting microbial community, as the authors begin to suggest). But I think this can go deeper given the high-resolution community composition data that is available here (similar to the detailed discussion on PA/ FL preference).

Figure 9 is not introduced in the results but heavily discussed in the discussion. The results reported for Fig. 9 in the Discussion should be moved to the Results.

L602- Bchl a is introduced for the first time with no context on what this is or what it is short for.

Technical Corrections

L37- A high proportion "of" overlap

L140- What is CTAB?

L151- "each DNA was" should be each "DNA pellet was"?

L259- I am not sure what is meant by incompact.

C3

L388: "were supposed to" is a misleading phrase. It sounds like an expectation of a result. Perhaps this would better be "several environmental parameters played a pivotal role. . ."

L403 impaction should be impact

L412, "It is considered. . ." I am not sure what the 'subset' is and I think this can be better phrased.

L414-415- should be 'A recent study' (not 'A most recent study')

L425 – should be 'unexpected' rather than 'out of our expectation'

L425-426 – should be 'generally exhibits N- or P-limited phytoplankton production'

L436- 'niches' is not the correct word here. Maybe habitats? Locations?

L445- The phrase 'significantly divergent' implies statistical significance, but no such test was done to prove that PA and FL communities were significantly different (also in lines 641, 27, and 103). I think just 'divergent' would be acceptable unless a test is incorporated.

L463- 'dominantly govern' should just be 'dominate'

L498 – I don't understand the meaning of this phrase: 'nothing is available to elaborate the selection better PA and. . .' I think it needs to be reworded.

L580- The phrase 'intelligibly convinced' is unclear. Also the entire sentence L580-583 is a run-on sentence with some unclear phrasing and I'm not sure what the intended meaning is.

Interactive comment on Biogeosciences Discuss., <https://doi.org/10.5194/bg-2020-115>, 2020.

C4