Diversity and mineral substrate preference in endolithic microbial communities from marine intertidal outcrops (Isla de Mona, Puerto Rico).

Garcia-Pichel et al.

We would like to thank the referee #1 for the time spent in a thorough review and his/her comments. We take this opportunity to address his/her concerns regarding the approach and methodology choices that we made. Before that we would like to set for the record straight that Garcia-Pichel is the senior author, not the lead, of this article, therefore usage would indicate to refer to this contribution as Couradeau et al.

This study investigated substrate specificity of endolithic communities in phosphate rock, limestone and dolostone outcrops from Isla de Mona, Puerto Rico. Authors implemented a high-throughput 16SrDNA genetic diversity approach that reveled the dominance of euendolithic cyanobacteria associated to a high community diversity. Results did not support the hypothesis that community composition would relate to mineral substrate but particular euendolithic cyanobacteria seemed to be specialized at the mineral substrate level. Also, the question regarding the existence of a specialized community associated to dolostone vs limestone could not be resolved.

Since authors used a very short region of the 16S rDNA and then used a culture-collection of euendoliths to ex- tract the euendolith sequences, there seems to be a lot diversity that was not included, thus they conclude that only a small fraction of the community (3.5%) is influenced by substrate. The data analysis need to be redone. The method used provides short frag- ments and does not allow for thorough phylogenetic analysis. This study would greatly benefit from longer reads and maybe a metagenomic approach.

We, as referee #1, do principally worry about technical aspects, but have to disagree with the criticisms leveraged at our approach and analyses.

Being aware that the length of the sequences is critical for phylogenetic reconstruction, we took advantage of the recent progress of the Illumina chemistry and use general primers that amplify 465bp of the 16S (V3-V4 regions) instead of classical set of primers centered on the V3 region only (291bp) (Caporaso et al., 2012). To reconstruct the phylogeny (Figure S3) we manually selected 736 well aligned positions from our multiple alignment. The ends of our Illumina reads were filled up with the "?" character treated as an unknown position by the evolutionary model as defined in the Treefinder manual http://www.treefinder.de/tf-march2011-manual.pdf. The obtained topology is well supported and allows us to resolve the position of the OTUs of interest compared to reference sequences. We made the careful assumption that only OTUs that fell within clades of proven euendolithic strains could be deemed possibly euendolithic themselves. To do so we decided to put effort into increasing the number of reference sequences of proven euendolithic strains through targeted cultivation.

We note that our sequencing and bioinformatics approach is currently standard in microbial ecology. Here are a few examples of recent papers that used the same technique to detect microbial dynamics and distribution:

Angelakis, E., Yasir, M., Bachar, D., Azhar, E.I., Lagier, J.-C., Bibi, F., et al. (2016). Gut microbiome and dietary patterns in different Saudi populations and monkeys. Sci. Rep., 6, 32191

Boetius, A., Anesio, A.M., Deming, J.W., Mikucki, J.A. & Rapp, J.Z. (2015). Microbial ecology of the cryosphere: sea ice and glacial habitats. Nat Rev Micro, 13, 677–690

Clayton, J.B., Vangay, P., Huang, H., Ward, T., Hillmann, B.M., Al-Ghalith, G.A., et al. (2016). Captivity humanizes the primate microbiome. Proc. Natl. Acad. Sci., 113, 201521835

Hu, J., Raikhel, V., Gopalakrishnan, K., Fernandez-Hernandez, H., Lambertini, L., Manservisi, F., et al. (2016). Effect of postnatal low-dose exposure to environmental chemicals on the gut microbiome in a rodent model. Microbiome, 4, 26

Lal, C.V., Travers, C., Aghai, Z.H., Eipers, P., Jilling, T., Halloran, B., et al. (2016). The Airway Microbiome at Birth. Sci. Rep., 6, 31023

Props, R., Kerckhof, F.-M., Rubbens, P., De Vrieze, J., Sanabria, E.H., Waegeman, W., et al. (2016). Absolute quantification of microbial taxon abundances. ISME J. Adv. online Publ., 1–4

The main question addressed was if there is a highly adapted endolithic flora to spe- cific mineral substrates, yet in lines 299-301, authors state "At this level of taxonomic resolution, we did not detect any significant association of substrate mineralogy and community composition". To answer the main question authors used high-throughput sequencing to describe the microbial diversity and test the effect of different substrates on community composition.

We disagree on the reading of the results by the reviewer. This particular study had a double aim, (i) we wanted to apply the widely used 16 rRNA gene high throughput sequencing tool to describe intertidal endolithic communities and (ii) to test whether there exists a specialized community associated to the type of mineral they colonize. The motivation of our first aim was the lack of such dataset for these globally relevant microbial communities. The work presented here definitely contributes that part. The second aim was driven by the hypothesis that, if there exists a substrate preference of the pioneer euendolithic cyanobacteria, this preference could drive the total microbial community towards different climax communities.

Our dataset revealed that endolithic habitat hosts a large variety of microbial species, a lot wider that could have been foreseen from classical literature. As noted by the referee #1 we did not observe a correlation between the proportion of prokaryotic phyla and the mineralogy of their substrate. In other words, if there is a substrate preference it does not reflect into the proportion of prokaryotic phyla.

However, the proportion of a given phylum does not indicate the nature of the microbes that constitute it, therefore is not sufficient to reject our hypothesis. We demonstrated that there is substrate preference among the Cyanobacteria, so even if their proportion of the total community does not vary significantly with the substrate, their composition does. This clearly supports our hypothesis in a statistically robust and significant way. We could identify several cyanobacterial OTUs that were differentially abundant on limestone compared to dolostone (Figure 4). It is correct, as referee#1 pointed out, that only a small fraction of the cyanobacterial OTUs diversity (3.5%) showed a significant change in abundance with substrate. However, these very OTUs account for 16 ±4% of the total number of cyanobacterial sequences analyzed here (some of these OTUs, especially the possible euendoliths, being very abundant). Thus, they are not only differentially distributed, but also a significant proportion of the community.

We demonstrated that the effect of the substrate is not dramatic enough to change the proportion of prokaryotic phyla but still affects the abundance of some keystones species such as pioneer euendolithic Cyanobacteria.

We do agree with referee #1 that we could not bring a "yes / no" answer to the hypothesis; we argue that we enhanced our hypothesis by showing that the answer depends on the taxonomic level set for the analysis. Yes at fine resolution, no at coarse resolution.

We are grateful that referee #1 pointed out the reference sequences of the newly cultured euendolithic strains, as we realized that the details regarding the amplification and sequencing of their 16S rRNA genes were missing. We used the primers and PCR conditions recommended by Nübel to retrieve these sequences (Nübel et al., 1997). These details will be added to the methods section of the revised manuscript.

We agree with the referee #1 that a metagenomics study could be a nice follow-up step to the present piece of work. This contribution constitutes a pioneer study that explored the endolithic microbial diversity associated to various substrates. For that purpose, we used the 16S rRNA gene as a proxy that allowed us to genetically sample a large variety of locations with the appropriate sequencing depth. This is a required first step to ask relevant functional questions that could justify a new study involving metagenomics or other relevant methods such as in-situ biogeochemistry measurement, fluorescent labelling, and metabolomics.

Specific comments: Abstract- The last claim "The cationic mineral component was. . .. existence in nature of alternatives to the boring mechanism. . .. based. . .on transcellu- lar calcium transport" is not sustained from the results presented.

In their recent contribution, (Guida and Garcia-Pichel, 2016) showed that the boring mechanism in the model strain *Mastigocoleus testarum* BC008 was based on vectorial transcellular transport of calcium from the boring front to the boring hole. Here we show that some possible euendoliths, including close relatives to *Mastigocoleus testarum* BC008, do not show exclusive preference for ca-carbonate substrate. This indicates that the *Mastigocoleus testarum* BC008 vectorial transport of calcium ions to bore cannot be the sole mechanism, and that there might exist alternative mechanisms.

Methods- Authors justify the need to re-assess the diversity of euendolithic cyanobac- teria yet only include a high-throughput sequencing approach that produced very short reads, which are not informative for pylogenetic analysis. Itag sequencing is not the best platform to analyze deep phylogenetic affiliations and to resolve the mentioned issues on euendolithic cyanobacteria for this study model.

Our aim was to and compare the microbial diversity among 34 samples, for which we used 16S rRNA gene genetic sampling. Saying that a) read are shorts and b) uninformative, is simply incorrect. Again we point the reviewer to the fact that this is a standard methodology with the power to show differences (see some other examples above). The reads produced here were 465bp. Using 16S rDNA to assess the microbial diversity allowed us to both reach enough sequencing depth to get appropriate coverage (see Table S2) and to compare our sequences with the largest library of taxonomically assigned sequences (Greengenes 13-8).

Lines 196-201, repeated phrase.

This will be fixed.

Line2 209-212, I don't understand why mention an- other site, and sequences that are afterwards not discussed in this analysis.

We regret that our point was missed by the reviewer. We mentioned these samples because we included them in the meta-analysis, figure 4. We included them for comparative purposes, processing them in parallel to the Mona samples. They came from a different type of environment (alkaline lake) and therefore represent an internal control of our analysis to supports the fact that the difference that we see is due to the environment, rather than analytical. Differences due to analytical aspects are principally possible when comparisons are done on dataset retrieved from the Qiita database.

Lines 237- 238, by using a dataset with proven boring cultured strains and using that to assess which of the cyanobacteria OTUs could be euendolyths, this study is losing the power to identify other euendoliths. Why compare only to the known euendolith dataset?

How does one know the metabolic activity of any one particular organism identified based solely on the presence of its 16S rDNA? The best approach one can take is to compare this particular 16S rDNA to the reference sequences of organisms with proven activity. There is no other way we could have identified euendoliths, and one can never identify a new euendolith (or any other putative metabolic activity) based on a sequence only. This also justifies why we put effort into increasing the number of references sequences through targeted cultivation effort.

Again this approach, which is rather commonplace and the basis of most functional bioinformatics, will indeed miss absolute novelty, but will secure identification of a large part of the community.

Results and Discussion Lines 274-275, please give information on coverage.

Please see Table S2 column 2 for coverage information.

Lines 307-314, in this paragrapha, authors mention that the sequences obtained in this study clustered together and discuss that 1) this could happen since euendolithic assemblages are distinct in a microbiological and adaptive way, or that alternatively 2) the clustering pattern reflects a biogeographical island effect, since all samples come from a small area. Authors discuss the second is unlikely given the cosmopolitan na- ture of marine cyanobacteria. Nonetheless the references cited are for a cosmopolitan, nonmarine cyanobacterium, M. vaginatus and for M. chthonoplastes. This discussion should be revised; there are different methods to proof for biogeography in communi- ties, and to analyze diversity patterns related to biogeography. Also, it is possible to do analysis to disentangle which environmental variables, in this case including mineral composition, are relevant and explain community composition.

We would like to thank referee #1 for pointing us to a follow-up hypothesis that could be tested in the framework of this experiment. However, the point of the present contribution is not to discuss the biogeography of Cyanobacteria in general, but to look at substrate preference among endolithic communities. The meta-analysis that was performed here used an aggregated dataset from various studies looking at marine and lake sediments, intertidal mollusks shells, microbialites and hot springs.

We agree with referee #1 that it would be great to be able to correlate the pattern that we observed with some environmental parameters, however this type of data is not consistently available for the chosen datasets. We agree with the reviewer that this point of discussion being speculative it would be best to let our readers make their own opinion, we will therefore present the two hypothesis as equally valuable in the revised version of the manuscript.

The references associated to that paragraph point both to *Microcoleus chtonoplastes* as an example of marine cosmopolitan cyanobacterium and not to *Microcoleus vaginatus*, a terrestrial biocrust forming cyanobacterium.

Line 323, are there cyanobacterial communities? Or populations that interact with others to form communities? Line 327, The cyanobacterial community (diversity?) ap- peared quire diverse (elevated?) with a specific. . ..

We would like to thank referee #1 for pointing out some terminology ambiguities. Here we use the word "community" as an aggregation of all the cyanobacteria sequences, there is therefore one community of Cyanobacteria in our dataset. We further give a quantification of the specific diversity/ richness of this community using the classical chao1 alpha-diversity metrics. We avoided the term "population" as this term might refer to population genetics and within species interaction which was not the subject of the present contribution.

Lines 348-349, What percentage of the community to euendolithis represent?

Good point. Euendoliths represent (based on the 7 OTUs that we could assign as putative euendolith based on their phylogenetic proximity to known microborers only) from 0.8% to 73% of the sequences depending of the sample considered. (Average value 29%). We will include this relevant information in the revised manuscript.

Lines 353-355, Authors could do microscopic observations to make sure the issue regarding the lack of P. terebrans.

Microscopy will be insufficient. Referee #1 will agree that it is particularly challenging to recognize *Plectonema terebrans*, this species being described based on very common morphological characteristics:

"Fila gracilia, elongata, flexuosa, vulgo parce pseudo-ramosa, pseudo-ramis saepius solitariis. Vaginae hyalinae tenuissimae , cylindraceae , chlorozincico iodurato non caerulescenles. Trichomata dilute aeruginea, non torulosa, 0,95 µ ad 1,5 µ crassa; arliculi diametro trichomatis longiores, 2 µ ad 6 µ longi; dissepimenta binis granulis protoplasmaticis nolata; cellula apicalis rotundata (v. v.)" Gomont, M. (1892 '1893'). Monographie des Oscillariées.

In fact, this description would even fit well members of the Chloroflexus bacteria, which are also present. Therefore, it is possible that several species, even non boring colonies that were secondary colonizers, have been called *P. terebrans* over the years, genetic tools would in that case help to resolve the abundance/presence of a particular boring *P.terebrans*, unfortunately there is no available reference sequences for this group. We hope that a cultured isolate will provide a reference sequence for *P.terebrans* in the future to help us overcome the limitations of microscopic observations for this group, similar to what we did for other euendolithic clades in the present contribution.

Lines 451-458, This discussion s very interesting bot out of place since this study did not focus on M. testarum but on the overall boarer diversity.

This part of the discussion focuses on how the current findings contrast with the knowledge that was gained from the only model strain of euendolithic cyanobacteria that exists, *M. testarum* BC008. We

regret that referee #1 judged it out of place as it seems important to tell our reader how these novel findings relate with the proposed boring mechanism deciphered from physiological studies of *M. testarum* BC008. So far, this strain is the single model one can compare with, so certainly not out of place.

Conclusion- Lines 462-466, please revise use of English. "These complex communities likely host. . . This phrase is stating the obvious.

"These complex communities likely host various microbial metabolic guilds beyond oxygenic phototrophs described during more than a century of naturalist's descriptions."

This sentence is recapitulating an important finding of this study which is that these communities are more diverse and likely hold more metabolic capabilities than one could have expect from previous literature. This was not obvious until the entire community (including heterotrophic members) was described using 16S rDNA based genetic sampling of the community here.

Lines 468-471, the claim regarding different boring mechanisms than those known for M. testarum is not sustained from these results.

This discussion point aims at casting the results presented here in the framework of the model developed for *M. testarum* BC008. See answer to the first specific comment above for more details.

References cited in the answer to referee #1

Caporaso, J. G., Lauber, C. L., Walters, W. a, Berg-Lyons, D., Huntley, J., Fierer, N., Owens, S. M., Betley, J., Fraser, L., Bauer, M., Gormley, N., Gilbert, J. A., Smith, G. and Knight, R.: Ultra-high-throughput microbial community analysis on the Illumina HiSeq and MiSeq platforms., ISME J., 6(8), 1621–4, 2012.

Guida, B. S. and Garcia-Pichel, F.: Extreme cellular adaptations and cell differentiation required by a cyanobacterium for carbonate excavation, Proc. Natl. Acad. Sci., in press, 2016.

Nubel, U., GarciaPichel, F., Muyzer, G. and Garcia-pichel, F.: PCR Primers To Amplify 16S rRNA Genes from Cyanobacteria, Appl. Environ. Microbiol., 63(8), 3327–3332, 1997.