THE ROYAL SOCIETY PUBLISHING

PROCEEDINGS B

Exceptional but vulnerable microbial diversity in coral reef animal surface microbiomes

Marlène Chiarello, Jean-Christophe Auguet, Nicholas A. J. Graham, Thomas Claverie, Elliott Sucré, Corinne Bouvier, Fabien Rieuvilleneuve, Claudia Ximena Restrepo-Ortiz, Yvan Bettarel, Sébastien Villéger and Thierry Bouvier

Article citation details
Proc. R. Soc. B **287**: 20200642.
http://dx.doi.org/10.1098/rspb.2020.0642

Review timeline
Original submission: 17 September 2019
1st revised submission: 23 March 2020
2nd revised submission: 16 April 2020
Final acceptance: 17 April 2020

Note: Reports are unedited and appear as submitted by the referee. The review history appears in chronological order.

Review History

RSPB-2019-2159.R0 (Original submission)

Review form: Reviewer 1

Recommendation
Major revision is needed (please make suggestions in comments)

Scientific importance: Is the manuscript an original and important contribution to its field?
Good

General interest: Is the paper of sufficient general interest?
Good

Quality of the paper: Is the overall quality of the paper suitable?
Acceptable

Is the length of the paper justified?
Yes

Should the paper be seen by a specialist statistical reviewer?
Yes

Do you have any concerns about statistical analyses in this paper? If so, please specify them explicitly in your report.
Yes

It is a condition of publication that authors make their supporting data, code and materials available - either as supplementary material or hosted in an external repository. Please rate, if applicable, the supporting data on the following criteria.

Is it accessible?
Yes

Is it clear?
Yes

Is it adequate?
No

Do you have any ethical concerns with this paper?
No

Comments to the Author
Summary
The authors analyzed the microbiotas over 300 reef animal (fish, invertebrate, and coral) species and seawater samples in order to understand the microbial diversity of coral reefs. This paper expanded the current knowledge on reef macroorganismal microbial diversity as fewer than 150 species of coral and fish have been sampled in the past (this study analyzed 74 taxa). The authors were successful in showing that fish, invertebrates, and coral surface microbiomes constitute a major component of reef microbial alpha and beta diversity, especially in comparison to the significantly lower diversity of the surrounding seawater. Furthermore, I found a major strength of the study was when the authors quantified the erosion of diversity given the vulnerability of specific fish and coral to future extinction. That component is relevant and important considering the global decline of coral reefs and mass extinction in the Anthropocene.

Major Concerns
The sequences were grouped into 97%-identity de novo OTUs (line 192), which is a method that is generally less favored due to the rather arbitrary 97% cutoff. Generally, many scientists are beginning to favor ASVs, which are more reproducible, tractable, and generally represent a finer-scale taxonomic identity than de novo OTUs (Callahan et al. 2016). While re-analyzing the complete dataset is likely unreasonable, I suggest re-analyzing a subset of the data to confirm the trends seen with the de novo OTU pipeline used. Alternatively, I suggest the authors add in a sentence justifying the use of de novo OTUs at a 97% cutoff over other, finer-scale and more biologically-informed methods.
Rarefaction down to 2,000 sequences was used in this paper (lines 189-191), which is a major concern given the demonstrated statistical issues with rarefaction, as outlined in McMurdie and Holmes (2014) and Willis (2019). The authors report that samples ranged from 2,443 to 43,504 sequences, which indicates that some samples had over 40,000 reads removed, which constitutes a major component of the data. Given that the majority of the paper focuses on the analysis of diversity, removing such a significant portion of the data may lead to underestimates of the total alpha and beta diversity. While this is a major concern, the authors do point out that "Chao's non-parametric coverage computed using entropart R-package averaged 93±5% across all samples testifying that sampling depth was sufficient [26]" (line 195). Another way to deal with rarefaction for Faith's PD, would be to produce rarefaction and extrapolation curves of both a sample size and coverage-based method to assess sample completeness and is implemented in

software: iNEXT-pd (Chao et al. 2015).

The majority of analyses in the body of the paper use Faith's phylogenetic diversity measure and unweighted and weighted UniFrac, which is not a direct measure of number of OTUs and is severely biased by different sequencing depths, making rarefaction important for these diversity metrics (Schloss 2008). To complement this phylogenetic diversity viewpoint that relies on removing a significant portion of sequences, I suggest complementing it with an analysis of the unrarefied sequences to add weight to the findings of the paper.

More specifically, the supplementary figures (Figure S3, Figure S5, Figure S8) that report number of OTUs based on random sampling of the rarefied samples could be reproduced using the un-rarefied samples (maintain the same number of original sequences). This would improve the robustness of the findings in the body of the paper. Additionally, it would be helpful to report the number of sequences per sample for transparency. For re-analysis, I advise using the R packages breakaway that account for differing sequence depths without removing data (For more details, see Willis and Bunge 2015). In addition, it provides error bars for the richness estimation. This package and analysis does not account for phylogeny and therefore will offer an added comparison/analysis of the reproducibility of the phylogenetic diversity analysis. If the findings are sound using two analytical pipelines that do and do not include rarefaction, then the robustness of the findings and impact of this paper would improve.

A strength of the paper was the vulnerability analysis and extinction scenario. I appreciated the use of location-relevant bleaching responses of the coral (W. Indian Ocean, line 303), and the use of the FishBase database for generating the vulnerability indices they calculated in the supplement. I also found the supplemental figure of vulnerability indices to be a nice addition (Figure S4). An area for improvement would be to provide the code for the analysis of prokaryotic diversity vulnerability. The explanation is sufficient; however, a formula or code would make it easier to assess the statistical rigor and validity of the analysis. Furthermore, as future scientists may be interested in reproducing a similar analysis given the extensive current extinctions in the Anthropocene, making the code available will allow them to build upon your work rather than beginning from scratch.

Minor Concerns

Microbiome data analysis pipelines are constantly changing and getting updated and as a result, I recommend adding version numbers and citations for each R package listed in the methods section.

In a similar vein, I would suggest explicitly stating the function used in each package in the methods section, as that is often a good indication of the purpose of using the package. That said, I appreciated how well the authors explained why each package was used. Beyond this, I encourage the authors to publish the code used for the analyses beyond simply the novel code. Knowing exact parameters used helps maintain the reproducibility in microbiome science, which is currently a major issue in the field.

Including a supplementary table with each OTU and the representative sequence will help future scientists interested in comparing OTUs found on reef animal surfaces to their study.

Line 363 – Figure 3C does not seem to fit in with the statement. The statement explains how animal surface microbiomes are distinct from planktonic microbiomes, but Figure 3C does not include planktonic microbiomes. It would be a nice addition to include the planktonic community dissimilarity values as an added bar chart in Figure 3C.

Figure 3A – I appreciate the consistent color themes that identify the Fish, Anthozoa, and Other invertebrates categories. It might be helpful to also write out these categories on the axis of Figure 3A for ease of interpretation.

Figure 2 – I found the discussion of this figure to be limited in the body of the paper, and I feel it would be well-suited as a supplementary figure.

Citations

Callahan, Benjamin J, Paul J McMurdie, Michael J Rosen, Andrew W Han, Amy Jo A Johnson, and Susan P Holmes. 2016. "DADA2: High-Resolution Sample Inference from Illumina Amplicon Data." Nature Methods 13 (7): 581–83. https://doi.org/10.1038/nmeth.3869.

Chao, Anne, Chun-Huo Chiu, T. C. Hsieh, Thomas Davis, David A. Nipperess, and Daniel P. Faith. 2015. "Rarefaction and Extrapolation of Phylogenetic Diversity." Edited by Robert B. O'Hara. Methods in Ecology and Evolution 6 (4): 380–88. https://doi.org/10.1111/2041-210X.12247.

McMurdie, Paul J., and Susan Holmes. 2014. "Waste Not, Want Not: Why Rarefying Microbiome Data Is Inadmissible." Edited by Alice Carolyn McHardy. PLoS Computational Biology 10 (4): e1003531. https://doi.org/10.1371/journal.pcbi.1003531.

Schloss, Patrick D. 2008. "Evaluating Different Approaches That Test Whether Microbial Communities Have the Same Structure." The ISME Journal 2 (3): 265–75. https://doi.org/10.1038/ismej.2008.5.

Willis, Amy. 2019. "Rarefaction, Alpha Diversity, and Statistics." Frontiers in Microbiology 10 (2407): 1–5. https://doi.org/10.3389/fmicb.2019.02407.
Willis, Amy, and John Bunge. 2015. "Estimating Diversity via Frequency Ratios: Estimating Diversity via Ratios." Biometrics 71 (4): 1042–49. https://doi.org/10.1111/biom.12332.

Review form: Reviewer 2 (Jesse R. Zaneveld)

Recommendation
Accept with minor revision (please list in comments)

Scientific importance: Is the manuscript an original and important contribution to its field?
Excellent

General interest: Is the paper of sufficient general interest?
Excellent

Quality of the paper: Is the overall quality of the paper suitable?
Good

Is the length of the paper justified?
Yes

Should the paper be seen by a specialist statistical reviewer?
No

Do you have any concerns about statistical analyses in this paper? If so, please specify them explicitly in your report.
No

It is a condition of publication that authors make their supporting data, code and materials available - either as supplementary material or hosted in an external repository. Please rate, if applicable, the supporting data on the following criteria.

Is it accessible?
Yes

Is it clear?
Yes

Is it adequate?
No

Do you have any ethical concerns with this paper?
No

Comments to the Author
I was pleased to read the manuscript "Exceptional but Vulnerable Microbial Diversity in Coral Reef Animal Surface Microbiomes" from Chiarello et al.

While many studies have addressed planktonic and coral microbiome diversity, it has been less common for studies to fill in a picture of the interrelations between the microbial communities of the various types of animals that live on coral reefs.

This is important both for understanding the consequences of animal extinction on the overall reef microbial meta-community, but also because microbes can transmit between animals (either directly or via the plankton).

Overall I thought the paper was well-considered, the experiments and methods appropriately straightforward and clearly designed, and most of the interpretations justified.

The estimates of which species are vulnerable to extinction are necessarily approximate, but I think the choices are reasonable given the aims of the study and the available data.

In total, I think the paper is a good fit for this journal and would recommend publication following inclusion of some additional supplementary data files (e.g. sample metadata) and minor revision to the text to address a few important points of interpretation laid out below.

Major Issues –

Clarification needed in discussion of phylosymbiosis. The results line 397 states that the observation of similar beta-diversity within the fish and coral samples vs. between them indicates that phylosymbiosis is not present, and that this contradicts previous studies. I think the author's observation is interesting, but I think the discussion of it needs some clarification:

1. Phylosymbiosis can occur at multiple levels of phylogeny. So for example, there can easily be phylosymbiosis near the tips of the tree, but saturation deeper in the tree. Without explicitly testing for phylosymbiosis across all distances the authors cannot rule it out based on a single distance class (e.g. fish vs. corals). Similarly, I do not see these results as directly contradicting the author's previous findings of phylosymbiosis within reef fish. It can simultaneously be true that phylosymbiosis is present, but has signal only for certain distance classes.

2. The results presented here do not, in my opinion, contradict some of the studies cited. Specifically, Pollock et al (ref 55) showed phylosymbiosis in coral tissue and skeleton *but did not find a significant correlation signal of phylosymbiosis in mucus* when using Bray-Curtis dissimilarity, and saw significant phylosymbiosis only in skeleton (not mucus) when using weighted UniFrac. (Perversely, as a phylogenetically-weighted algorithm, Weighted UniFrac may be somewhat bad at picking up phylosymbiosis as fine-scale differences in the microbial phylogeny – precisely the place where associations with animal hosts should be strongest – are downweighted!). Since this study sampled coral mucus, the result of no phylosymbiosis (at least based on the within vs. between comparison of fish vs. coral mucus samples) would seem consistent with those findings.

I agree with the author's interpretation on line 397, that a likely explanation is that the phylogenetic signal of the host on the microbiome may saturate at very large distances. A simple fix would be to reword the paragraph to highlight this observation, while making it clear that

phylosymbiosis may still be happening at lower distance classes (as I think the authors have shown within reef fish in previous work), or in other tissue compartments.

Indeed, when combining these results with the author's previous work, I think there may be some additional evidence to bolster the author's idea that microbiome change may saturate at large distance classes. Pollock et al (Supplemental Note 10) write:

"Across all compartments, three dissimilarity measures (Weighted UniFrac distance, Unweighted UniFrac distance, and Bray Curtis dissimilarity) and two rarefaction depths (1000 sequences/sample or 10000 sequences/sample), more specific taxonomic levels always explained microbiome beta-diversity better than more general ones. The largest fall in explanatory power typically occurred between genus and family-level group, and between family-level group and Complex vs. Robust clade membership. This analysis is largely consistent with signals of phylosymbiosis in Mantel test results, and also shows that the relative effect of coral taxonomy on microbiome membership is robust to common choices for rarefaction depth and dissimilarity measure."

This seems pretty consistent with what the authors see.

Alternatively, if the authors wanted to emphasize the findings on phylosymbiosis and test them rigorously, a Mantel test based on the overall tree of all sampled species would help resolve whether the whole sample set shows phylosymbiosis when considered holistically across all distance classes.

Potential confounding effects of mitochondria. It is common in animal studies to pick up mitochondrial SSU rRNA sequences alongside the intended bacterial and archaeal 16S rRNA sequences. This has upsides – authors have used these to detect the presence of animals that couldn't be seen directly or to confirm host taxonomic annotations. Unfortunately, most standard microbiome reference taxonomies do not adequately annotate mitochondria for certain taxa. Sceleractinian corals are particularly notorious in this regard. It is common for a very high propotion of 'Unclassified' sequences according to SILVA or greengenes_13_8 to be mitochondrial sequences. Such sequences will be taxon specific and may therefore inflate within-group beta-diversity (since each e.g. fish or coral genus may have distinct mitochondrial SSU rRNA gene sequences). Moreover, the combination of misannotated mitochondria with sequencing error can result in many OTUs (not just one) that derive from host mitochondria (or the mitochondria of epiphytes etc).

The authors report unclassified OTUs made up 0 – 77% of total microbial abundance, and did a good job trying to refine these using parsimony insertion in SILVA. It may be worth checking whether these match mitochondrial sequences not in SILVA or greengenes, esp. for the corals.

Caution should be used when discussing absolute abundance of OTUs or ASVs.

Around line 410, the authors discuss the total OTUs present in the lagoon of Mayotte. If it is critical to discuss the absolute abundance of microbes, I think probably an improved denoising strategy is needed. For relative comparisons I think OTUs and the current pipeline are OK (see note about inflation of OTU counts seen in mock community studies below).

Minor issues –

Use of OTUs rather than ASVs.

The paper uses somewhat older OTU-clustering methods rather than more recent ASV methods. While other reviewers may take a harsher view, my own perspective is that it is not necessary to redo the whole analysis to update this aspect of the methodology. However, it will be necessary to update the discussion of microbial diversity, especially when absolute (rather than relative)

quantities are discussed, since this choice of methodology is predicted to consistently over-estimate diversity across all samples.

Specifically, the authors should address findings from mock community studies that show that OTU-based approaches can dramatically overestimate microbial diversity relative to ASVs (e.g. say deblur in QIIME2 or DADA2). Nearing et al., 2018 "Denoising the Denoisers" provide a nice summary of this issue in an independent analysis of the situation conducted by a group that didn't develop any of the methods testing (link: https://www.ncbi.nlm.nih.gov/pmc/articles/PMC6087418/). The authors cite some of the previous literature on this issue.

The reason I list this as a minor rather than major issue is that we should probably expect this overestimation to be roughly similar across taxa, so I wouldn't expect that this is likely to change the estimates for e.g. the proportion of microbial phylogenetic diversity that is lost with species extinction. Other methodological choices the authors made (iike focusing on phylogenetic diversity and/or Allen's metric rather than e.g. Chao1 estimates) will help insulate their diversity results against inflation of the number of OTUS. In most cases, erroneous OTUs that appear due to sequencing noise will be phylogenetically clustered with the true sequence. Since the authors use phylogenetic measures for both alpha and beta diversity, their richness estimates will down-weight these spurious OTUs.

DNA extraction method (line 156). Different DNA extraction methods can preferentially favor certain microbial taxa over others. It was unclear to me whether the kit used was designed for microbiome work? Many of the kits that are commonly used across microbiome studies have been tested on mock microbial communities to assess their taxonomic coverage (e.g. how badly will using any given kit alter observed microbial taxonomic ratios?). I know the authors have used this kit before, and I don't doubt that it was able to get out microbial DNA given the results. However, I was unable to tell from the manufacturer's website whether it has been tested on mock communities to see if it *evenly* extracts DNA from across the microbial tree? No method is likely to be perfect in this regard, but it would be good to discuss a bit more what we know or don't know about the properties of this DNA extraction approach.

Diversity of coral microbiomes vs. seawater: These results seem to contradict those summarized in the McDevitt-Irwin review (https://www.frontiersin.org/files/Articles/286253/fmars-04-00262-HTML/image_m/fmars-04-00262-g004.jpg), where seawater is presented as much richer than coral microbiomes in terms of # of OTUs. It may be worth a bit more discussion of the previous literature on seawater vs. coral and fish microbiome richness. I suspect the differences could be due to a) failure of some other studies or the comparison between them to account for sequencing depth (which is accounted for in this study) or b) differences in sampling method (e.g. for amount of water sampled), c) differences in DNA extraction method.

Other comments:

Line 167: The choice of primers is standard and reasonable.

Line 189: I felt that the 2000 sequences/sample rarefaction depth, and the use of rarefaction to avoid the possibility of false positives due to unequal sequencing depth was reasonable in the context of this study. Some reviewers may argue that parametric methods should be used exclusively, but I would suggest that the authors made a reasonable choice given the study questions (we should be more scared of inducing a false-positive result using parametric methods and an incorrect model than using the conservative but reliable rarefaction approach).

Line 422: Expand 'within the skeleton' to 'within the tissue and skeleton' (since we know the tissue is a key habitat for coral-associated microbes from microscopy and sequencing).

Line 424: samplings -> sampling

Line 423: ‘The reported estimate of animal surface microbiome diversity is conservative’. I understand the authors point, but given the OTU clustering methods used I think this statement should be revised.

Line 433: Of course, strictly speaking this should really be a phylogenetic independent contrasts (PIC) analysis, not a plain Spearman since we are comparing across species and the structure of the tree will effectively cause the data to be pseudo-replicated.

Decision letter (RSPB-2019-2159.R0)

15-Nov-2019

Dear Dr Chiarello:

I am writing to inform you that your manuscript RSPB-2019-2159 entitled "Exceptional but vulnerable microbial diversity in coral reef animal surface microbiomes" has, in its current form, been rejected for publication in Proceedings B.

This action has been taken on the advice of referees, who have recommended that substantial revisions are necessary. With this in mind we would be happy to consider a resubmission, provided the comments of the referees are fully addressed. However please note that this is not a provisional acceptance.

The resubmission will be treated as a new manuscript. However, we will approach the same reviewers if they are available and it is deemed appropriate to do so by the Editor. Please note that resubmissions must be submitted within six months of the date of this email. In exceptional circumstances, extensions may be possible if agreed with the Editorial Office. Manuscripts submitted after this date will be automatically rejected.

Please find below the comments made by the referees, not including confidential reports to the Editor, which I hope you will find useful. If you do choose to resubmit your manuscript, please upload the following:

1) A ‘response to referees’ document including details of how you have responded to the comments, and the adjustments you have made.
2) A clean copy of the manuscript and one with 'tracked changes' indicating your 'response to referees' comments document.
3) Line numbers in your main document.

To upload a resubmitted manuscript, log into http://mc.manuscriptcentral.com/prsb and enter your Author Centre, where you will find your manuscript title listed under "Manuscripts with Decisions." Under "Actions," click on "Create a Resubmission." Please be sure to indicate in your cover letter that it is a resubmission, and supply the previous reference number.

Sincerely,
Dr Daniel Costa
mailto: proceedingsb@royalsociety.org

Reviewer(s)' Comments to Author:
Referee: 1

Comments to the Author(s)
Summary

The authors analyzed the microbiotas over 300 reef animal (fish, invertebrate, and coral) species and seawater samples in order to understand the microbial diversity of coral reefs. This paper expanded the current knowledge on reef macroorganismal microbial diversity as fewer than 150 species of coral and fish have been sampled in the past (this study analyzed 74 taxa). The authors were successful in showing that fish, invertebrates, and coral surface microbiomes constitute a major component of reef microbial alpha and beta diversity, especially in comparison to the significantly lower diversity of the surrounding seawater. Furthermore, I found a major strength of the study was when the authors quantified the erosion of diversity given the vulnerability of specific fish and coral to future extinction. That component is relevant and important considering the global decline of coral reefs and mass extinction in the Anthropocene.

Major Concerns

The sequences were grouped into 97%-identity de novo OTUs (line 192), which is a method that is generally less favored due to the rather arbitrary 97% cutoff. Generally, many scientists are beginning to favor ASVs, which are more reproducible, tractable, and generally represent a finer-scale taxonomic identity than de novo OTUs (Callahan et al. 2016). While re-analyzing the complete dataset is likely unreasonable, I suggest re-analyzing a subset of the data to confirm the trends seen with the de novo OTU pipeline used. Alternatively, I suggest the authors add in a sentence justifying the use of de novo OTUs at a 97% cutoff over other, finer-scale and more biologically-informed methods.

Rarefaction down to 2,000 sequences was used in this paper (lines 189-191), which is a major concern given the demonstrated statistical issues with rarefaction, as outlined in McMurdie and Holmes (2014) and Willis (2019). The authors report that samples ranged from 2,443 to 43,504 sequences, which indicates that some samples had over 40,000 reads removed, which constitutes a major component of the data. Given that the majority of the paper focuses on the analysis of diversity, removing such a significant portion of the data may lead to underestimates of the total alpha and beta diversity. While this is a major concern, the authors do point out that "Chao's non-parametric coverage computed using entropart R-package averaged 93±5% across all samples testifying that sampling depth was sufficient [26]" (line 195). Another way to deal with rarefaction for Faith's PD, would be to produce rarefaction and extrapolation curves of both a sample size and coverage-based method to assess sample completeness and is implemented in software: iNEXT-pd (Chao et al. 2015).

The majority of analyses in the body of the paper use Faith's phylogenetic diversity measure and unweighted and weighted UniFrac, which is not a direct measure of number of OTUs and is severely biased by different sequencing depths, making rarefaction important for these diversity metrics (Schloss 2008). To complement this phylogenetic diversity viewpoint that relies on removing a significant portion of sequences, I suggest complementing it with an analysis of the unrarefied sequences to add weight to the findings of the paper.

More specifically, the supplementary figures (Figure S3, Figure S5, Figure S8) that report number of OTUs based on random sampling of the rarefied samples could be reproduced using the un-rarefied samples (maintain the same number of original sequences). This would improve the robustness of the findings in the body of the paper. Additionally, it would be helpful to report the number of sequences per sample for transparency. For re-analysis, I advise using the R packages breakaway that account for differing sequence depths without removing data (For more details, see Willis and Bunge 2015). In addition, it provides error bars for the richness estimation. This package and analysis does not account for phylogeny and therefore will offer an added comparison/analysis of the reproducibility of the phylogenetic diversity analysis. If the findings are sound using two analytical pipelines that do and do not include rarefaction, then the robustness of the findings and impact of this paper would improve.

A strength of the paper was the vulnerability analysis and extinction scenario. I appreciated the use of location-relevant bleaching responses of the coral (W. Indian Ocean, line 303), and the

use of the FishBase database for generating the vulnerability indices they calculated in the supplement. I also found the supplemental figure of vulnerability indices to be a nice addition (Figure S4). An area for improvement would be to provide the code for the analysis of prokaryotic diversity vulnerability. The explanation is sufficient; however, a formula or code would make it easier to assess the statistical rigor and validity of the analysis. Furthermore, as future scientists may be interested in reproducing a similar analysis given the extensive current extinctions in the Anthropocene, making the code available will allow them to build upon your work rather than beginning from scratch.

Minor Concerns

Microbiome data analysis pipelines are constantly changing and getting updated and as a result, I recommend adding version numbers and citations for each R package listed in the methods section.

In a similar vein, I would suggest explicitly stating the function used in each package in the methods section, as that is often a good indication of the purpose of using the package. That said, I appreciated how well the authors explained why each package was used. Beyond this, I encourage the authors to publish the code used for the analyses beyond simply the novel code. Knowing exact parameters used helps maintain the reproducibility in microbiome science, which is currently a major issue in the field.

Including a supplementary table with each OTU and the representative sequence will help future scientists interested in comparing OTUs found on reef animal surfaces to their study.

Line 363 – Figure 3C does not seem to fit in with the statement. The statement explains how animal surface microbiomes are distinct from planktonic microbiomes, but Figure 3C does not include planktonic microbiomes. It would be a nice addition to include the planktonic community dissimilarity values as an added bar chart in Figure 3C.

Figure 3A – I appreciate the consistent color themes that identify the Fish, Anthozoa, and Other invertebrates categories. It might be helpful to also write out these categories on the axis of Figure 3A for ease of interpretation.

Figure 2 – I found the discussion of this figure to be limited in the body of the paper, and I feel it would be well-suited as a supplementary figure.

Citations

Callahan, Benjamin J, Paul J McMurdie, Michael J Rosen, Andrew W Han, Amy Jo A Johnson, and Susan P Holmes. 2016. “DADA2: High-Resolution Sample Inference from Illumina Amplicon Data.” Nature Methods 13 (7): 581–83. https://doi.org/10.1038/nmeth.3869.

Chao, Anne, Chun-Huo Chiu, T. C. Hsieh, Thomas Davis, David A. Nipperess, and Daniel P. Faith. 2015. “Rarefaction and Extrapolation of Phylogenetic Diversity.” Edited by Robert B. O’Hara. Methods in Ecology and Evolution 6 (4): 380–88. https://doi.org/10.1111/2041-210X.12247.

McMurdie, Paul J., and Susan Holmes. 2014. “Waste Not, Want Not: Why Rarefying Microbiome Data Is Inadmissible.” Edited by Alice Carolyn McHardy. PLoS Computational Biology 10 (4): e1003531. https://doi.org/10.1371/journal.pcbi.1003531.

Schloss, Patrick D. 2008. “Evaluating Different Approaches That Test Whether Microbial Communities Have the Same Structure.” The ISME Journal 2 (3): 265–75. https://doi.org/10.1038/ismej.2008.5.

Willis, Amy. 2019. “Rarefaction, Alpha Diversity, and Statistics.” Frontiers in Microbiology 10 (2407): 1–5. https://doi.org/10.3389/fmicb.2019.02407.

Willis, Amy, and John Bunge. 2015. “Estimating Diversity via Frequency Ratios: Estimating Diversity via Ratios.” Biometrics 71 (4): 1042–49. https://doi.org/10.1111/biom.12332.

Referee: 2

Comments to the Author(s)
I was pleased to read the manuscript “Exceptional but Vulnerable Microbial Diversity in Coral Reef Animal Surface Microbiomes" from Chiarello et al.

While many studies have addressed planktonic and coral microbiome diversity, it has been less common for studies to fill in a picture of the interrelations between the microbial communities of the various types of animals that live on coral reefs.

This is important both for understanding the consequences of animal extinction on the overall reef microbial meta-community, but also because microbes can transmit between animals (either directly or via the plankton).

Overall I thought the paper was well-considered, the experiments and methods appropriately straightforward and clearly designed, and most of the interpretations justified.

The estimates of which species are vulnerable to extinction are necessarily approximate, but I think the choices are reasonable given the aims of the study and the available data.

In total, I think the paper is a good fit for this journal and would recommend publication following inclusion of some additional supplementary data files (e.g. sample metadata) and minor revision to the text to address a few important points of interpretation laid out below.

Major Issues –

Clarification needed in discussion of phylosymbiosis. The results line 397 states that the observation of similar beta-diversity within the fish and coral samples vs. between them indicates that phylosymbiosis is not present, and that this contradicts previous studies. I think the author’s observation is interesting, but I think the discussion of it needs some clarification:

1. Phylosymbiosis can occur at multiple levels of phylogeny. So for example, there can easily be phylosymbiosis near the tips of the tree, but saturation deeper in the tree. Without explicitly testing for phylosymbiosis across all distances the authors cannot rule it out based on a single distance class (e.g. fish vs. corals). Similarly, I do not see these results as directly contradicting the author’s previous findings of phylosymbiosis within reef fish. It can simultaneously be true that phylosymbiosis is present, but has signal only for certain distance classes.

2. The results presented here do not, in my opinion, contradict some of the studies cited. Specifically, Pollock et al (ref 55) showed phylosymbiosis in coral tissue and skeleton *but did not find a significant correlation signal of phylosymbiosis in mucus* when using Bray-Curtis dissimilarity, and saw significant phylosymbiosis only in skeleton (not mucus) when using weighted UniFrac. (Perversely, as a phylogenetically-weighted algorithm, Weighted UniFrac may be somewhat bad at picking up phylosymbiosis as fine-scale differences in the microbial phylogeny – precisely the place where associations with animal hosts should be strongest – are downweighted!). Since this study sampled coral mucus, the result of no phylosymbiosis (at least based on the within vs. between comparison of fish vs. coral mucus samples) would seem consistent with those findings.

I agree with the author's interpretation on line 397, that a likely explanation is that the phylogenetic signal of the host on the microbiome may saturate at very large distances. A simple fix would be to reword the paragraph to highlight this observation, while making it clear that phylosymbiosis may still be happening at lower distance classes (as I think the authors have shown within reef fish in previous work), or in other tissue compartments.

Indeed, when combining these results with the author's previous work, I think there may be some additional evidence to bolster the author's idea that microbiome change may saturate at large distance classes. Pollock et al (Supplemental Note 10) write:

"Across all compartments, three dissimilarity measures (Weighted UniFrac distance, Unweighted UniFrac distance, and Bray Curtis dissimilarity) and two rarefaction depths (1000 sequences/sample or 10000 sequences/sample), more specific taxonomic levels always explained microbiome beta-diversity better than more general ones. The largest fall in explanatory power typically occurred between genus and family-level group, and between family-level group and Complex vs. Robust clade membership. This analysis is largely consistent with signals of phylosymbiosis in Mantel test results, and also shows that the relative effect of coral taxonomy on microbiome membership is robust to common choices for rarefaction depth and dissimilarity measure."

This seems pretty consistent with what the authors see.

Alternatively, if the authors wanted to emphasize the findings on phylosymbiosis and test them rigorously, a Mantel test based on the overall tree of all sampled species would help resolve whether the whole sample set shows phylosymbiosis when considered holistically across all distance classes.

Potential confounding effects of mitochondria. It is common in animal studies to pick up mitochondrial SSU rRNA sequences alongside the intended bacterial and archaeal 16S rRNA sequences. This has upsides – authors have used these to detect the presence of animals that couldn't be seen directly or to confirm host taxonomic annotations. Unfortunately, most standard microbiome reference taxonomies do not adequately annotate mitochondria for certain taxa. Sceleractinian corals are particularly notorious in this regard. It is common for a very high propotion of 'Unclassified' sequences according to SILVA or greengenes_13_8 to be mitochondrial sequences. Such sequences will be taxon specific and may therefore inflate within-group beta-diversity (since each e.g. fish or coral genus may have distinct mitochondrial SSU rRNA gene sequences). Moreover, the combination of misannotated mitochondria with sequencing error can result in many OTUs (not just one) that derive from host mitochondria (or the mitochondria of epiphytes etc).

The authors report unclassified OTUs made up 0 – 77% of total microbial abundance, and did a good job trying to refine these using parsimony insertion in SILVA. It may be worth checking whether these match mitochondrial sequences not in SILVA or greengenes, esp. for the corals.

Caution should be used when discussing absolute abundance of OTUs or ASVs.

Around line 410, the authors discuss the total OTUs present in the lagoon of Mayotte. If it is critical to discuss the absolute abundance of microbes, I think probably an improved denoising strategy is needed. For relative comparisons I think OTUs and the current pipeline are OK (see note about inflation of OTU counts seen in mock community studies below).

Minor issues –

Use of OTUs rather than ASVs.

The paper uses somewhat older OTU-clustering methods rather than more recent ASV methods. While other reviewers may take a harsher view, my own perspective is that it is not necessary to redo the whole analysis to update this aspect of the methodology. However, it will be necessary to update the discussion of microbial diversity, especially when absolute (rather than relative)

quantities are discussed, since this choice of methodology is predicted to consistently over-estimate diversity across all samples.

Specifically, the authors should address findings from mock community studies that show that OTU-based approaches can dramatically overestimate microbial diversity relative to ASVs (e.g. say deblur in QIIME2 or DADA2). Nearing et al., 2018 "Denoising the Denoisers" provide a nice summary of this issue in an independent analysis of the situation conducted by a group that didn't develop any of the methods testing (link: https://www.ncbi.nlm.nih.gov/pmc/articles/PMC6087418/). The authors cite some of the previous literature on this issue.

The reason I list this as a minor rather than major issue is that we should probably expect this overestimation to be roughly similar across taxa, so I wouldn't expect that this is likely to change the estimates for e.g. the proportion of microbial phylogenetic diversity that is lost with species extinction. Other methodological choices the authors made (iike focusing on phylogenetic diversity and/or Allen's metric rather than e.g. Chao1 estimates) will help insulate their diversity results against inflation of the number of OTUS. In most cases, erroneous OTUs that appear due to sequencing noise will be phylogenetically clustered with the true sequence. Since the authors use phylogenetic measures for both alpha and beta diversity, their richness estimates will down-weight these spurious OTUs.

DNA extraction method (line 156). Different DNA extraction methods can preferentially favor certain microbial taxa over others. It was unclear to me whether the kit used was designed for microbiome work? Many of the kits that are commonly used across microbiome studies have been tested on mock microbial communities to assess their taxonomic coverage (e.g. how badly will using any given kit alter observed microbial taxonomic ratios?). I know the authors have used this kit before, and I don't doubt that it was able to get out microbial DNA given the results. However, I was unable to tell from the manufacturer's website whether it has been tested on mock communities to see if it *evenly* extracts DNA from across the microbial tree? No method is likely to be perfect in this regard, but it would be good to discuss a bit more what we know or don't know about the properties of this DNA extraction approach.

Diversity of coral microbiomes vs. seawater: These results seem to contradict those summarized in the McDevitt-Irwin review (https://www.frontiersin.org/files/Articles/286253/fmars-04-00262-HTML/image_m/fmars-04-00262-g004.jpg
), where seawater is presented as much richer than coral microbiomes in terms of # of OTUs. It may be worth a bit more discussion of the previous literature on seawater vs. coral and fish microbiome richness. I suspect the differences could be due to a) failure of some other studies or the comparison between them to account for sequencing depth (which is accounted for in this study) or b) differences in sampling method (e.g. for amount of water sampled), c) differences in DNA extraction method.

Other comments:

Line 167: The choice of primers is standard and reasonable.

Line 189: I felt that the 2000 sequences/sample rarefaction depth, and the use of rarefaction to avoid the possibility of false positives due to unequal sequencing depth was reasonable in the context of this study. Some reviewers may argue that parametric methods should be used exclusively, but I would suggest that the authors made a reasonable choice given the study questions (we should be more scared of inducing a false-positive result using parametric methods and an incorrect model than using the conservative but reliable rarefaction approach).

Line 422: Expand 'within the skeleton' to 'within the tissue and skeleton' (since we know the tissue is a key habitat for coral-associated microbes from microscopy and sequencing).

Line 424: samplings -> sampling

Line 423: 'The reported estimate of animal surface microbiome diversity is conservative'. I understand the authors point, but given the OTU clustering methods used I think this statement should be revised.

Line 433: Of course, strictly speaking this should really be a phylogenetic independent contrasts (PIC) analysis, not a plain Spearman since we are comparing across species and the structure of the tree will effectively cause the data to be pseudo-replicated.

Author's Response to Decision Letter for (RSPB-2019-2159.R0)

See Appendix A.

RSPB-2020-0642.R0

Review form: Reviewer 1

Recommendation
Accept with minor revision (please list in comments)

Scientific importance: Is the manuscript an original and important contribution to its field?
Excellent

General interest: Is the paper of sufficient general interest?
Good

Quality of the paper: Is the overall quality of the paper suitable?
Excellent

Is the length of the paper justified?
Yes

Should the paper be seen by a specialist statistical reviewer?
No

Do you have any concerns about statistical analyses in this paper? If so, please specify them explicitly in your report.
No

It is a condition of publication that authors make their supporting data, code and materials available - either as supplementary material or hosted in an external repository. Please rate, if applicable, the supporting data on the following criteria.

Is it accessible?
Yes

Is it clear?
Yes

Is it adequate?
Yes

Do you have any ethical concerns with this paper?
No

Comments to the Author
--Summary.
--The authors used an extensive dataset of reef animal surface and seawater microbiome samples (almost 300 samples) to estimate the phylogenetic and taxonomic diversity of coral reefs. This paper successfully expands the current baseline knowledge on coral reef macro-organismal microbial diversity because they sampled species of coral, fish, and other invertebrates that have never been sampled previously. This allowed them to significantly contribute to the coral reef microbiome literature. A particular strength of the paper is the thorough robust and careful analyses. They show that diversity of animal microbiomes in reef environments are strikingly more diverse than the surrounding seawater, a finding that is even more interesting considering that the surfaces of each of these organisms is in constant contact with surrounding seawater. Finally, the major novel component of this paper is the estimation of decline in prokaryotic diversity given relevant extinction scenarios of fish and coral in the Anthropocene. Given the importance, yet global decline of coral reefs I feel this extinction scenario is a timely, relevant, and unique aspect of this study.

--Major Comments.
--I have no strong major changes to the manuscript. I feel the study was careful and thorough. I especially appreciate the extensive supplementary figures using un-rarefied data to explicitly show that their rarefaction procedure did not introduce significant bias into the data to alter the conclusions. Furthermore, inclusion of DADA2 code, R packages with version numbers and relevant functions, novel code, and functions support the reproducibility of this study, which is important in analyses with significant analytical and coding components.

--Minor Comments.
--Overall, I have only minor comments and changes to the manuscript. The following comments are mostly questions regarding figures and statements made in the text.

--Line 55 in abstract vs. 462: "would induce an erosion of 28% of the prokaryotic diversity" vs. "would induce a loss of ca. 28 to 29% of ASVs, based on rarefied and un-rarefied data". I think the intent is the same, but it seems like the extinction scenario is only modeled with richness data. Maybe change the abstract to say prokaryotic richness? Though I recognize richness is a diversity metric, so if you do not change it I understand.

--In general, I think you should check the figure labels in the body of the text. I think you reference incorrect figures at times. Here are the examples I found where I believe you meant a different figure or figure part:
--Line 115: Figure S1 should be Figure SM1-1.
--Line 343: Figure 3c should be Figure 3a.
--Line 394: Figure 3a should be Figure 3b or 3c.
--Line 403: Figure SM2-6 should be SM2-7.
--Line 413: Figure 3b and SM2-4 should be Figure 3c and SM2-5.
--Line 456: SM1-6 should be SM1-7, the vulnerability values. (SM1-6 is DADA2 script in my document).

--In the methods I appreciate the inclusion of all R packages and version numbers. In addition, the inclusion of specific functions from each package will help improve the reproducibility of this paper. Along those lines, please make sure the formatting of each package name and function are

consistent throughout the methods. It seems they are formatted as package and 'function', but in Line 150 and 163 there is different formatting.

--Figure 1 and associated supplementary figures: Figure 1 is generally inconsistent with the figure legend and with the body of the text in the paper associated with the figure. Figure 1a is labeled as "Taxonomic richness (Nb ASVs)". The legend states that this should be phylogenetic richness. Indeed, the associated supplementary figure, SM2-1 has figure a as "Phylogenetic richness (% total Faith's PD)" and a different scale. I think you added the wrong figure here. SM2-3a is supposed to be the figure for taxonomic richness, so maybe that was added to Figure 1a accidentally.
For figure 1b, the numbers are not reported as percentages as indicated in the legend. In SM2-3b, the raw number of ASVs are reported for taxonomic richness, but they are different than the values in Figure 1b, so it is unclear which numbers reported are correct.
For Figure 1 c and d, they are labeled as "Taxonomic richness" and "Taxonomic diversity", respectively. In the legend, they are supposed to be "Phylogenetic richness" and "Phylogenetic diversity", respectively. If they were phylogenetic metrics, they would match the associated Figure SM2-1c,d. It is unclear what the correctly reported figure should be.

--Line 344: Change "clades" to classes. It seems you are reporting bacterial classes in the figure. It might be worthwhile to change "clades" in the body text to classes to be more clear about the taxonomic hierarchy you are considering.

--Line 351-59: Your finding of Arcobacter in the coral microbiome is indeed interesting considering the pristine nature of the sites you sampled. You discuss how this genus is found in thermally stressed and fish effluent stressed corals. It might be important to note that Arcobacter is a common microbe (based on ASVs and OTUs) found associated with coral disease, though your corals were not in any diseased states. Here are some citations that found Arcobacter in coral diseased lesions:
Gignoux-Wolfsohn SA, Vollmer SV (2015) Identification of Candidate Coral Pathogens on White Band Disease-Infected Staghorn Coral. PLoS ONE 10:e0134416.
Sunagawa S, DeSantis TZ, Piceno YM, Brodie EL, DeSalvo MK, Voolstra CR, Weil E, Andersen GL, Medina M (2009) Bacterial diversity and White Plague Disease-associated community changes in the Caribbean coral Montastraea faveolata. ISME J 3:512–521.
Meyer JL, Castellanos-Gell J, Aeby GS, Häse CC, Ushijima B, Paul VJ (2019) Microbial Community Shifts Associated With the Ongoing Stony Coral Tissue Loss Disease Outbreak on the Florida Reef Tract. Front Microbiol 10:2244.

--Line 435: "cured" should be "curated"

--Line 463: "such extinction would induce 23% of phylogenetic richness in the ecosystem"
Do you mean that it would induce a "loss of" 23% richness? I found this sentence confusing after the previous sentence discussing loss of ASV-based richness in the community.

--Line 466-8: "slope of microbial diversity extinction, which increased only slightly with increasing loss of macroscopic species"
I could be mis-interpreting Figure 4b, but it seems that with increasing loss of species, the slope decreased from -0.5 to closer to -0.75. It increased in the negative direction, if that is what is meant by that sentence. I could be interpreting the graph incorrectly, but I found that sentence misleading.

--Figure 2 Legend: The legend does not seem to fully fit with the figure and I think you may need to edit it. Specifically, I interpreted the following fixes: "(a)" at the beginning might actually be "(c)"? I think c in the following sentence, "(b and c) Average rank-abundance curve obtained from surface microbiomes of all animal taxa (b) and all 35 planktonic communities (c)," is supposed to be "a".

Decision letter (RSPB-2020-0642.R0)

14-Apr-2020

Dear Dr Chiarello

I am pleased to inform you that your manuscript RSPB-2020-0642 entitled "Exceptional but vulnerable microbial diversity in coral reef animal surface microbiomes" has been accepted for publication in Proceedings B.

The referee(s) have recommended publication, but also suggest some minor revisions to your manuscript. Therefore, I invite you to respond to the referee(s)' comments and revise your manuscript. Because the schedule for publication is very tight, it is a condition of publication that you submit the revised version of your manuscript within 7 days. If you do not think you will be able to meet this date please let us know.

To revise your manuscript, log into https://mc.manuscriptcentral.com/prsb and enter your Author Centre, where you will find your manuscript title listed under "Manuscripts with Decisions." Under "Actions," click on "Create a Revision." Your manuscript number has been appended to denote a revision. You will be unable to make your revisions on the originally submitted version of the manuscript. Instead, revise your manuscript and upload a new version through your Author Centre.

When submitting your revised manuscript, you will be able to respond to the comments made by the referee(s) and upload a file "Response to Referees". You can use this to document any changes you make to the original manuscript. We require a copy of the manuscript with revisions made since the previous version marked as ‘tracked changes’ to be included in the ‘response to referees’ document.

Before uploading your revised files please make sure that you have:

1) A text file of the manuscript (doc, txt, rtf or tex), including the references, tables (including captions) and figure captions. Please remove any tracked changes from the text before submission. PDF files are not an accepted format for the "Main Document".

2) A separate electronic file of each figure (tiff, EPS or print-quality PDF preferred). The format should be produced directly from original creation package, or original software format. PowerPoint files are not accepted.

3) Electronic supplementary material: this should be contained in a separate file and where possible, all ESM should be combined into a single file. All supplementary materials accompanying an accepted article will be treated as in their final form. They will be published alongside the paper on the journal website and posted on the online figshare repository. Files on figshare will be made available approximately one week before the accompanying article so that the supplementary material can be attributed a unique DOI.

Online supplementary material will also carry the title and description provided during submission, so please ensure these are accurate and informative. Note that the Royal Society will not edit or typeset supplementary material and it will be hosted as provided. Please ensure that the supplementary material includes the paper details (authors, title, journal name, article DOI). Your article DOI will be 10.1098/rspb.[paper ID in form xxxx.xxxx e.g. 10.1098/rspb.2016.0049].

4) A media summary: a short non-technical summary (up to 100 words) of the key findings/importance of your manuscript.

5) Data accessibility section and data citation
It is a condition of publication that data supporting your paper are made available either in the electronic supplementary material or through an appropriate repository.

In order to ensure effective and robust dissemination and appropriate credit to authors the dataset(s) used should be fully cited. To ensure archived data are available to readers, authors should include a 'data accessibility' section immediately after the acknowledgements section. This should list the database and accession number for all data from the article that has been made publicly available, for instance:
• DNA sequences: Genbank accessions F234391-F234402
• Phylogenetic data: TreeBASE accession number S9123
• Final DNA sequence assembly uploaded as online supplemental material
• Climate data and MaxEnt input files: Dryad doi:10.5521/dryad.12311
NB. From April 1 2013, peer reviewed articles based on research funded wholly or partly by RCUK must include, if applicable, a statement on how the underlying research materials – such as data, samples or models – can be accessed. This statement should be included in the data accessibility section.

If you wish to submit your data to Dryad (http://datadryad.org/) and have not already done so you can submit your data via this link
http://datadryad.org/submit?journalID=RSPB&manu=(Document not available) which will take you to your unique entry in the Dryad repository. If you have already submitted your data to dryad you can make any necessary revisions to your dataset by following the above link.
Please see https://royalsociety.org/journals/ethics-policies/data-sharing-mining/ for more details.

6) For more information on our Licence to Publish, Open Access, Cover images and Media summaries, please visit https://royalsociety.org/journals/authors/author-guidelines/.

Once again, thank you for submitting your manuscript to Proceedings B and I look forward to receiving your revision. If you have any questions at all, please do not hesitate to get in touch.

Sincerely,
Dr Daniel Costa
mailto: proceedingsb@royalsociety.org

Reviewer(s)' Comments to Author:

Referee: 1

Comments to the Author(s).
--Summary.
--The authors used an extensive dataset of reef animal surface and seawater microbiome samples (almost 300 samples) to estimate the phylogenetic and taxonomic diversity of coral reefs. This paper successfully expands the current baseline knowledge on coral reef macro-organismal microbial diversity because they sampled species of coral, fish, and other invertebrates that have never been sampled previously. This allowed them to significantly contribute to the coral reef microbiome literature. A particular strength of the paper is the thorough robust and careful analyses. They show that diversity of animal microbiomes in reef environments are strikingly more diverse than the surrounding seawater, a finding that is even more interesting considering that the surfaces of each of these organisms is in constant contact with surrounding seawater. Finally, the major novel component of this paper is the estimation of decline in prokaryotic diversity given relevant extinction scenarios of fish and coral in the Anthropocene. Given the importance, yet global decline of coral reefs I feel this extinction scenario is a timely, relevant, and unique aspect of this study.

--Major Comments.
--I have no strong major changes to the manuscript. I feel the study was careful and thorough. I especially appreciate the extensive supplementary figures using un-rarefied data to explicitly show that their rarefaction procedure did not introduce significant bias into the data to alter the conclusions. Furthermore, inclusion of DADA2 code, R packages with version numbers and relevant functions, novel code, and functions support the reproducibility of this study, which is important in analyses with significant analytical and coding components.

--Minor Comments.
--Overall, I have only minor comments and changes to the manuscript. The following comments are mostly questions regarding figures and statements made in the text.

--Line 55 in abstract vs. 462: "would induce an erosion of 28% of the prokaryotic diversity" vs. "would induce a loss of ca. 28 to 29% of ASVs, based on rarefied and un-rarefied data". I think the intent is the same, but it seems like the extinction scenario is only modeled with richness data. Maybe change the abstract to say prokaryotic richness? Though I recognize richness is a diversity metric, so if you do not change it I understand.

--In general, I think you should check the figure labels in the body of the text. I think you reference incorrect figures at times. Here are the examples I found where I believe you meant a different figure or figure part:
--Line 115: Figure S1 should be Figure SM1-1.
--Line 343: Figure 3c should be Figure 3a.
--Line 394: Figure 3a should be Figure 3b or 3c.
--Line 403: Figure SM2-6 should be SM2-7.
--Line 413: Figure 3b and SM2-4 should be Figure 3c and SM2-5.
--Line 456: SM1-6 should be SM1-7, the vulnerability values. (SM1-6 is DADA2 script in my document).

--In the methods I appreciate the inclusion of all R packages and version numbers. In addition, the inclusion of specific functions from each package will help improve the reproducibility of this paper. Along those lines, please make sure the formatting of each package name and function are consistent throughout the methods. It seems they are formatted as package and 'function', but in Line 150 and 163 there is different formatting.

--Figure 1 and associated supplementary figures: Figure 1 is generally inconsistent with the figure legend and with the body of the text in the paper associated with the figure. Figure 1a is labeled as "Taxonomic richness (Nb ASVs)". The legend states that this should be phylogenetic richness. Indeed, the associated supplementary figure, SM2-1 has figure a as "Phylogenetic richness (% total Faith's PD)" and a different scale. I think you added the wrong figure here. SM2-3a is supposed to be the figure for taxonomic richness, so maybe that was added to Figure 1a accidentally.
For figure 1b, the numbers are not reported as percentages as indicated in the legend. In SM2-3b, the raw number of ASVs are reported for taxonomic richness, but they are different than the values in Figure 1b, so it is unclear which numbers reported are correct.
For Figure 1 c and d, they are labeled as "Taxonomic richness" and "Taxonomic diversity", respectively. In the legend, they are supposed to be "Phylogenetic richness" and "Phylogenetic diversity", respectively. If they were phylogenetic metrics, they would match the associated Figure SM2-1c,d. It is unclear what the correctly reported figure should be.

--Line 344: Change "clades" to classes. It seems you are reporting bacterial classes in the figure. It might be worthwhile to change "clades" in the body text to classes to be more clear about the taxonomic hierarchy you are considering.

--Line 351-59: Your finding of Arcobacter in the coral microbiome is indeed interesting considering the pristine nature of the sites you sampled. You discuss how this genus is found in thermally stressed and fish effluent stressed corals. It might be important to note that Arcobacter is a common microbe (based on ASVs and OTUs) found associated with coral disease, though your corals were not in any diseased states. Here are some citations that found Arcobacter in coral diseased lesions:
Gignoux-Wolfsohn SA, Vollmer SV (2015) Identification of Candidate Coral Pathogens on White Band Disease-Infected Staghorn Coral. PLoS ONE 10:e0134416.
Sunagawa S, DeSantis TZ, Piceno YM, Brodie EL, DeSalvo MK, Voolstra CR, Weil E, Andersen GL, Medina M (2009) Bacterial diversity and White Plague Disease-associated community changes in the Caribbean coral Montastraea faveolata. ISME J 3:512–521.
Meyer JL, Castellanos-Gell J, Aeby GS, Häse CC, Ushijima B, Paul VJ (2019) Microbial Community Shifts Associated With the Ongoing Stony Coral Tissue Loss Disease Outbreak on the Florida Reef Tract. Front Microbiol 10:2244.

--Line 435: “cured” should be “curated”

--Line 463: “such extinction would induce 23% of phylogenetic richness in the ecosystem”
Do you mean that it would induce a “loss of” 23% richness? I found this sentence confusing after the previous sentence discussing loss of ASV-based richness in the community.

--Line 466-8: “slope of microbial diversity extinction, which increased only slightly with increasing loss of macroscopic species”
I could be mis-interpreting Figure 4b, but it seems that with increasing loss of species, the slope decreased from -0.5 to closer to -0.75. It increased in the negative direction, if that is what is meant by that sentence. I could be interpreting the graph incorrectly, but I found that sentence misleading.

--Figure 2 Legend: The legend does not seem to fully fit with the figure and I think you may need to edit it. Specifically, I interpreted the following fixes: “(a)” at the beginning might actually be “(c)”? I think c in the following sentence, “(b and c) Average rank-abundance curve obtained from surface microbiomes of all animal taxa (b) and all 35 planktonic communities (c),” is supposed to be “a”.

Author's Response to Decision Letter for (RSPB-2020-0642.R0)

See Appendix B.

Decision letter (RSPB-2020-0642.R1)

17-Apr-2020

Dear Dr Chiarello

I am pleased to inform you that your manuscript entitled "Exceptional but vulnerable microbial diversity in coral reef animal surface microbiomes" has been accepted for publication in Proceedings B.

You can expect to receive a proof of your article from our Production office in due course, please check your spam filter if you do not receive it. PLEASE NOTE: you will be given the exact page

length of your paper which may be different from the estimation from Editorial and you may be asked to reduce your paper if it goes over the 10 page limit.

If you are likely to be away from e-mail contact please let us know. Due to rapid publication and an extremely tight schedule, if comments are not received, we may publish the paper as it stands.

If you have any queries regarding the production of your final article or the publication date please contact procb_proofs@royalsociety.org

Your article has been estimated as being 10 pages long. Our Production Office will be able to confirm the exact length at proof stage.

Open Access
You are invited to opt for Open Access, making your freely available to all as soon as it is ready for publication under a CCBY licence. Our article processing charge for Open Access is £1700. Corresponding authors from member institutions (http://royalsocietypublishing.org/site/librarians/allmembers.xhtml) receive a 25% discount to these charges. For more information please visit http://royalsocietypublishing.org/open-access.

Paper charges
An e-mail request for payment of any related charges will be sent out shortly. The preferred payment method is by credit card; however, other payment options are available.

Electronic supplementary material:
All supplementary materials accompanying an accepted article will be treated as in their final form. They will be published alongside the paper on the journal website and posted on the online figshare repository. Files on figshare will be made available approximately one week before the accompanying article so that the supplementary material can be attributed a unique DOI.

You are allowed to post any version of your manuscript on a personal website, repository or preprint server. However, the work remains under media embargo and you should not discuss it with the press until the date of publication. Please visit https://royalsociety.org/journals/ethics-policies/media-embargo for more information.

Thank you for your fine contribution. On behalf of the Editors of the Proceedings B, we look forward to your continued contributions to the Journal.

Sincerely,
Proceedings B
mailto: proceedingsb@royalsociety.org

Appendix A

Dear Dr Chiarello:

I am writing to inform you that your manuscript RSPB-2019-2159 entitled "Exceptional but vulnerable microbial diversity in coral reef animal surface microbiomes" has, in its current form, been rejected for publication in Proceedings B.
This action has been taken on the advice of referees, who have recommended that substantial revisions are necessary. With this in mind we would be happy to consider a resubmission, provided the comments of the referees are fully addressed. However please note that this is not a provisional acceptance.
The resubmission will be treated as a new manuscript. However, we will approach the same reviewers if they are available and it is deemed appropriate to do so by the Editor. Please note that resubmissions must be submitted within six months of the date of this email. In exceptional circumstances, extensions may be possible if agreed with the Editorial Office. Manuscripts submitted after this date will be automatically rejected.
Please find below the comments made by the referees, not including confidential reports to the Editor, which I hope you will find useful. If you do choose to resubmit your manuscript, please upload the following:
1) A 'response to referees' document including details of how you have responded to the comments, and the adjustments you have made.
2) A clean copy of the manuscript and one with 'tracked changes' indicating your 'response to referees' comments document.
3) Line numbers in your main document.
To upload a resubmitted manuscript, log into http://mc.manuscriptcentral.com/prsb and enter your Author Centre, where you will find your manuscript title listed under "Manuscripts with Decisions." Under "Actions," click on "Create a Resubmission." Please be sure to indicate in your cover letter that it is a resubmission, and supply the previous reference number.

Sincerely,

Dr Daniel Costa
mailto: proceedingsb@royalsociety.org

Dear Dr Costa,

Thank you for considering our article and offering us the opportunity to revise our manuscript. In agreement with points raised by both referees, we entirely reanalyzed our sequencing data using a method based on Amplicon Sequence Variants (ASVs) using DADA2-R package. We agree that ASVs are more accurate to assess microbial diversity than Operational Taxonomic Units (OTUs) defined using the arbitrary 97%-identity criterion that we used previously. Our results are now entirely based on ASVs and they re-enforced the conclusion that animal surfaces host a much higher diversity than surrounding bacterioplankton.

Using DADA2, we also improved the quality of the taxonomic classification, as only 4% of sequences (to compare to the 13% of unclassified left by Mothur in the previous version of this paper) remained unclassified at class level. We then removed any mitochondrial sequences that could remain as "unclassified", especially in *Scleratinian* corals as stated by referee #2, by mapping unclassified sequences onto a database. The final dataset contains as little as 3% unclassified sequences after this two-step procedure. We mentioned such improvement hereafter in our detailed answer to referees.

Classification and removal of non-bacterial and mitochondrial sequences yielded at least 7,000 sequences per sample, which allow us to run rarefaction analysis with 3.5 more sequences than in the previous version of the manuscript (2,000 sequences per sample). As requested by referee #1, we also provided the number of sequences per sample in Supplementary Material SM1-4.

According to the points raised by both reviewers, we now address the effect of such rarefaction on diversity assessment by comparing the patterns obtained using both rarefied and non-rarefied data. Such results are now included as supplementary results SM2. We found that rarefaction neither altered the high contrast of diversity between plankton and animal surfaces microbiomes, nor modified the dissimilarity among animal microbiomes.

As requested by referee #1 we now provide a *phyloseq* R-object containing the ASV table, metadata and ASVs representative sequences and classification in SM3 to facilitate further use of our data. As requested by referee #2, we also edited the discussion, to clarify our paragraphs about phylosymbiosis (L427-432) and bacterial composition (L363-381). We also clarified our methods and provided the R-code allowing replicating the erosion simulation.

Please find below in blue our detailed answers to reviewers' comments. Changes in main text are also in blue.

Thank you for your time,

Marlène Chiarello on behalf of all authors

Referee: 1

Summary

The authors analyzed the microbiotas over 300 reef animal (fish, invertebrate, and coral) species and seawater samples in order to understand the microbial diversity of coral reefs. This paper expanded the current knowledge on reef macro-organismal microbial diversity as fewer than 150 species of coral and fish have been sampled in the past (this study analyzed 74 taxa). The authors were successful in showing that fish, invertebrates, and coral surface microbiomes constitute a major component of reef microbial alpha and beta diversity, especially in comparison to the significantly lower diversity of the surrounding seawater. Furthermore, I found a major strength of the study was when the authors quantified the erosion of diversity given the vulnerability of specific fish and coral to future extinction. That component is relevant and important considering the global decline of coral reefs and mass extinction in the Anthropocene.

We thank Referee #1 for these positive comments and relevant suggestions that helped us to improve our manuscript. Please find hereafter our detailed answers to reviewers' comments.

Major Concerns

The sequences were grouped into 97%-identity de novo OTUs (line 192), which is a method that is generally less favored due to the rather arbitrary 97% cutoff. Generally, many scientists are beginning to favor ASVs, which are more reproducible, tractable, and generally represent a finer-scale taxonomic identity than de novo OTUs (Callahan et al. 2016). While re-analyzing the complete dataset is likely unreasonable, I suggest re-analyzing a subset of the data to confirm the trends seen with the de novo OTU pipeline used. Alternatively, I suggest the authors add in a sentence justifying the use of de novo OTUs at a 97% cutoff over other, finer-scale and more biologically-informed methods.

Thank you for you comments. We agree that ASVs-based analyses are biologically more relevant than OTUs defined using the arbitrary 97%-identity cutoff. We re-analyzed our data using the software DADA2, permitting to define ASVs instead of OTUs. ASV-based analysis increased the quality of bacterial taxonomic classification. Therefore, we replaced all previous results based on OTUs to replace them by the ones based on ASVs. We modified the methodology section accordingly, L176-204, indicating the parameters that were used in DADA2 in the script provided in Supplementary Material SM1-5. All findings reported with the OTUs analyzes were confirmed and even re-inforced by ASV-based analyzes.

Rarefaction down to 2,000 sequences was used in this paper (lines 189-191), which is a major concern given the demonstrated statistical issues with rarefaction, as outlined in McMurdie and Holmes (2014) and Willis (2019). The authors report that samples ranged from 2,443 to 43,504 sequences, which indicates that some samples had over 40,000 reads removed, which constitutes a major component of the data. Given that the majority of the paper focuses on the analysis of diversity, removing such a significant portion of the data may lead to underestimates of the total alpha and beta diversity. While this is a major concern, the authors do point out that "Chao's non-parametric coverage computed using entropart R-package averaged 93±5% across all samples testifying that sampling depth was sufficient [26]" (line 195). Another way to deal with rarefaction for Faith's PD, would be to produce rarefaction and extrapolation curves of both a sample size and coverage-based method to assess sample completeness and is implemented in software: iNEXT-pd (Chao et al. 2015).

The majority of analyses in the body of the paper use Faith's phylogenetic diversity measure and unweighted and weighted UniFrac, which is not a direct measure of number of OTUs and is severely biased by different sequencing depths, making rarefaction important for these diversity metrics (Schloss 2008). To complement this phylogenetic diversity viewpoint that relies on removing a significant portion of sequences, I suggest complementing it with an analysis of the unrarefied

sequences to add weight to the findings of the paper.

More specifically, the supplementary figures (Figure S3, Figure S5, Figure S8) that report number of OTUs based on random sampling of the rarefied samples could be reproduced using the un-rarefied samples (maintain the same number of original sequences). This would improve the robustness of the findings in the body of the paper. Additionally, it would be helpful to report the number of sequences per sample for transparency. For re-analysis, I advise using the R packages breakaway that account for differing sequence depths without removing data (For more details, see Willis and Bunge 2015). In addition, it provides error bars for the richness estimation. This package and analysis does not account for phylogeny and therefore will offer an added comparison/analysis of the reproducibility of the phylogenetic diversity analysis. If the findings are sound using two analytical pipelines that do and do not include rarefaction, then the robustness of the findings and impact of this paper would improve.

Indeed, rarefaction to 2,000 sequences as in the first version of the manuscript might have induced an underestimation of the diversity in some samples. DADA2 processing of raw sequences yielded, more curated sequences per sample 7,074 and 56, 927 (number of sequences per sample are now in Supplementary Information SM1-4). The analyses included in the new version of the manuscript are based on rarefaction to 7,000 sequences per sample, raising coverage values of 98.0±1.5% across samples after rarefaction (L201). To assess the effect of rarefaction, we also re-analyzed our dataset without any rarefaction, *i.e.* using all curated sequences. While increasing coverage (99.8±0.1%), the absence of rarefaction did not substantially modify the results presented in the manuscript, but were included as Supplementary Information SM2.
As underlined by referee #2, phylogenetic diversity indices are complementary to the information accounting only for units (ASVs), and also permit to reduce overestimation of alpha-diversity, as spurious sequences are usually phylogenetically close to more abundant ones (Callahan et al. 2016). In the previous version of our manuscript, we compared the results obtained using taxonomic diversity (measured by ASVs richness) and phylogenetic diversity (measured by Faith's PD). We kept such comparison in the present paper, with figures based on taxonomic diversity provided in SM2, and the ones based on phylogenetic diversity provided in the main text and in SM2-1. We found overall similar results between both types of biodiversity measures, both on rarefied and un-rarefied data.

A strength of the paper was the vulnerability analysis and extinction scenario. I appreciated the use of location-relevant bleaching responses of the coral (W. Indian Ocean, line 303), and the use of the FishBase database for generating the vulnerability indices they calculated in the supplement. I also found the supplemental figure of vulnerability indices to be a nice addition (Figure S4). An area for improvement would be to provide the code for the analysis of prokaryotic diversity vulnerability. The explanation is sufficient; however, a formula or code would make it easier to assess the statistical rigor and validity of the analysis. Furthermore, as future scientists may be interested in reproducing a similar analysis given the extensive current extinctions in the Anthropocene, making the code available will allow them to build upon your work rather than beginning from scratch.

Thank you. We now provide the code for the extinction scenario, as a ready-to-use R function available on GitHub (https://github.com/marlenec/MicroErosion).

Minor Concerns

Microbiome data analysis pipelines are constantly changing and getting updated and as a result, I recommend adding version numbers and citations for each R package listed in the methods section.
We now indicate the version of each package throughout the Methods section and citations in the references list.

In a similar vein, I would suggest explicitly stating the function used in each package in the methods section, as that is often a good indication of the purpose of using the package. That said, I appreciated how well the authors explained why each package was used. Beyond this, I encourage the authors to publish the code used for the analyses beyond simply the novel code. Knowing exact parameters used helps maintain the reproducibility in microbiome science, which is currently a major issue in the field.

The script for DADA2 sequence processing is now provided in SM1-5 as well as the R code for the extinction scenario. We now also indicate parameters used within each function whenever it is necessary in the Methods section.

Including a supplementary table with each OTU and the representative sequence will help future scientists interested in comparing OTUs found on reef animal surfaces to their study.

The *phyloseq* object containing the ASVs abundance table, their classification and reference sequences, and corresponding metadata, is now available as a Supplementary Material SM3.

Line 363 – Figure 3C does not seem to fit in with the statement. The statement explains how animal surface microbiomes are distinct from planktonic microbiomes, but Figure 3C does not include planktonic microbiomes. It would be a nice addition to include the planktonic community dissimilarity values as an added bar chart in Figure 3C.

The figure 3C and similar figures in SM2-4 have been modified as requested to include planktonic communities.

Figure 3A – I appreciate the consistent color themes that identify the Fish, Anthozoa, and Other invertebrates categories. It might be helpful to also write out these categories on the axis of Figure 3A for ease of interpretation.

The figure 3A and all other similar plots in SM2-6 have been modified as requested.

Figure 2 – I found the discussion of this figure to be limited in the body of the paper, and I feel it would be well-suited as a supplementary figure.

This figure is important to us, because it shows that planktonic communities not only host only a small fraction of the microbial diversity, but also that their “unique” biodiversity is in low relative abundance in these communities compared to the “shared” fraction of biodiversity. Therefore, we chose to keep this figure in the main body of the article.

Citations

Callahan, Benjamin J, Paul J McMurdie, Michael J Rosen, Andrew W Han, Amy Jo A Johnson, and Susan P Holmes. 2016. “DADA2: High-Resolution Sample Inference from Illumina Amplicon Data.” Nature Methods 13 (7): 581–83. https://doi.org/10.1038/nmeth.3869.

Chao, Anne, Chun-Huo Chiu, T. C. Hsieh, Thomas Davis, David A. Nipperess, and Daniel P. Faith. 2015. “Rarefaction and Extrapolation of Phylogenetic Diversity.” Edited by Robert B. O’Hara. Methods in Ecology and Evolution 6 (4): 380–88. https://doi.org/10.1111/2041-210X.12247.

McMurdie, Paul J., and Susan Holmes. 2014. “Waste Not, Want Not: Why Rarefying Microbiome Data Is Inadmissible.” Edited by Alice Carolyn McHardy. PLoS Computational Biology 10 (4): e1003531. https://doi.org/10.1371/journal.pcbi.1003531.

Schloss, Patrick D. 2008. “Evaluating Different Approaches That Test Whether Microbial Communities Have the Same Structure.” The ISME Journal 2 (3): 265–75. https://doi.org/10.1038/ismej.2008.5.

Willis, Amy. 2019. “Rarefaction, Alpha Diversity, and Statistics.” Frontiers in Microbiology 10 (2407): 1–5. https://doi.org/10.3389/fmicb.2019.02407.

Willis, Amy, and John Bunge. 2015. “Estimating Diversity via Frequency Ratios: Estimating Diversity via Ratios.” Biometrics 71 (4): 1042–49. https://doi.org/10.1111/biom.12332.

Referee: 2

Comments to the Author(s)
I was pleased to read the manuscript “Exceptional but Vulnerable Microbial Diversity in Coral Reef Animal Surface Microbiomes" from Chiarello et al.
While many studies have addressed planktonic and coral microbiome diversity, it has been less common for studies to fill in a picture of the interrelations between the microbial communities of the various types of animals that live on coral reefs.
This is important both for understanding the consequences of animal extinction on the overall reef microbial meta-community, but also because microbes can transmit between animals (either directly or via the plankton).
Overall I thought the paper was well-considered, the experiments and methods appropriately straightforward and clearly designed, and most of the interpretations justified.
The estimates of which species are vulnerable to extinction are necessarily approximate, but I think the choices are reasonable given the aims of the study and the available data.
In total, I think the paper is a good fit for this journal and would recommend publication following inclusion of some additional supplementary data files (e.g. sample metadata) and minor revision to the text to address a few important points of interpretation laid out below.

We thank the referee #2 for these positive comments and relevant suggestions that helped us to improve our manuscript. Please find hereafter our detailed answers to specific comments.

Major Issues
Clarification needed in discussion of phylosymbiosis. The results line 397 states that the observation of similar beta-diversity within the fish and coral samples vs. between them indicates that phylosymbiosis is not present, and that this contradicts previous studies. I think the author’s observation is interesting, but I think the discussion of it needs some clarification:

1. Phylosymbiosis can occur at multiple levels of phylogeny. So for example, there can easily be phylosymbiosis near the tips of the tree, but saturation deeper in the tree. Without explicitly testing for phylosymbiosis across all distances the authors cannot rule it out based on a single distance class (e.g. fish vs. corals). Similarly, I do not see these results as directly contradicting the author’s previous findings of phylosymbiosis within reef fish. It can simultaneously be true that phylosymbiosis is present, but has signal only for certain distance classes.

2. The results presented here do not, in my opinion, contradict some of the studies cited. Specifically, Pollock et al (ref 55) showed phylosymbiosis in coral tissue and skeleton *but did not find a significant correlation signal of phylosymbiosis in mucus* when using Bray-Curtis dissimilarity, and saw significant phylosymbiosis only in skeleton (not mucus) when using weighted UniFrac. (Perversely, as a phylogenetically-weighted algorithm, Weighted UniFrac may be somewhat bad at picking up phylosymbiosis as fine-scale differences in the microbial phylogeny – precisely the place where associations with animal hosts should be strongest – are downweighted!). Since this study sampled coral mucus, the result of no phylosymbiosis (at least based on the within vs. between comparison of fish vs. coral mucus samples) would seem consistent with those findings.

I agree with the author's interpretation on line 397, that a likely explanation is that the phylogenetic signal of the host on the microbiome may saturate at very large distances. A simple fix would be to

reword the paragraph to highlight this observation, while making it clear that phylosymbiosis may still be happening at lower distance classes (as I think the authors have shown within reef fish in previous work), or in other tissue compartments.

Indeed, when combining these results with the author's previous work, I think there may be some additional evidence to bolster the author's idea that microbiome change may saturate at large distance classes. Pollock et al (Supplemental Note 10) write:

"Across all compartments, three dissimilarity measures (Weighted UniFrac distance, Unweighted UniFrac distance, and Bray Curtis dissimilarity) and two rarefaction depths (1000 sequences/sample or 10000 sequences/sample), more specific taxonomic levels always explained microbiome beta-diversity better than more general ones. The largest fall in explanatory power typically occurred between genus and family-level group, and between family-level group and Complex vs. Robust clade membership. This analysis is largely consistent with signals of phylosymbiosis in Mantel test results, and also shows that the relative effect of coral taxonomy on microbiome membership is robust to common choices for rarefaction depth and dissimilarity measure."

This seems pretty consistent with what the authors see.

Alternatively, if the authors wanted to emphasize the findings on phylosymbiosis and test them rigorously, a Mantel test based on the overall tree of all sampled species would help resolve whether the whole sample set shows phylosymbiosis when considered holistically across all distance classes.

We agree and rephrased the sentences mentioning phylosymbiosis L432-437. We now emphasize the very likely saturation of phylogenetic signal at high phylogenetic distances that could explain why pairs of coral species have as distinct microbiomes as coral and fish species. We also tested the correlation between microbiome dissimilarities and phylogenetic distances between all host species included in our dataset, using a phylogenetic tree obtained using the tree of life, and found no significant signal. However such result could be driven by biases in phylogenetic distances due the absence of resolution for most recent speciation events in this large tree, as well as biases in long branches estimation (e.g. between corals and fish). Therefore, we prefer to not include this rough-analysis in the manuscript and to restrain ourselves to writing a few lines on the phylosymbiosis in the discussion.

Potential confounding effects of mitochondria. It is common in animal studies to pick up mitochondrial SSU rRNA sequences alongside the intended bacterial and archaeal 16S rRNA sequences. This has upsides – authors have used these to detect the presence of animals that couldn't be seen directly or to confirm host taxonomic annotations. Unfortunately, most standard microbiome reference taxonomies do not adequately annotate mitochondria for certain taxa. Sceleractinian corals are particularly notorious in this regard. It is common for a very high propotion of 'Unclassified' sequences according to SILVA or greengenes_13_8 to be mitochondrial sequences. Such sequences will be taxon specific and may therefore inflate within-group beta-diversity (since each e.g. fish or coral genus may have distinct mitochondrial SSU rRNA gene sequences). Moreover, the combination of misannotated mitochondria with sequencing error can result in many OTUs (not just one) that derive from host mitochondria (or the mitochondria of epiphytes etc).

The authors report unclassified OTUs made up 0 – 77% of total microbial abundance, and did a good job trying to refine these using parsimony insertion in SILVA. It may be worth checking whether these match mitochondrial sequences not in SILVA or greengenes, esp. for the corals.

After DADA2 processing, 4% of sequences remained unclassified at class level (while the proportion of unclassified reached 13% in the previous version of the manuscript). Whether such difference may

be due to a better filtering of non-bacterial, and/or chimeric mitochondrial-bacterial sequences by DADA2 than Mothur, or to differences in merging reads, or in ASV vs. OTU reconstruction methods, remains undetermined.
Still, in order to identify potentially remaining mitochondrial sequences in our new dataset, we mapped all the 3,677 ASVs that were unclassified at class level onto the Genbank database (NCBI) using BLASTn alignment tool. 54% of such sequences could not be identified, *i.e.* showed more than 3% dissimilarity to the entire database. 39% showed more than 97% identity with unidentified prokaryotes, and 7% (279 ASVs) showed more than 97% identity with a mitochondrial sequence. Such ASVs totalized 0 to 15% of sequences in animal microbiomes, and were therefore removed before computing further relative comparisons between plankton and animal microbiomes (L177-204).

Caution should be used when discussing absolute abundance of OTUs or ASVs.
Around line 410, the authors discuss the total OTUs present in the lagoon of Mayotte. If it is critical to discuss the absolute abundance of microbes, I think probably an improved denoising strategy is needed. For relative comparisons I think OTUs and the current pipeline are OK (see note about inflation of OTU counts seen in mock community studies below).

We agree that further post-processing was needed. We thus have applied two additional steps to our curated data (see above) before making such absolute abundance estimates. First, we removed all unclassified ASVs that could not be matched against NCBI's Genbank database (i.e. that showed >3% dissimilarity to all sequences in the database). Then, to fairly compare our microbial species estimates to previously published estimates, we grouped the remaining 37 758 ASVs into 97%-identity OTUs using QIIME software, and removed any OTU that was found in only one animal taxon or planktonic community.
This additional curation step allowed us to make very conservative estimates of the unsampled microbial diversity. We also removed any estimation at lagoon scale, because such estimation might in fact be biased by the unknown shared amount of diversity between unsampled animal taxa (Louca et al. 2019). We made modifications in the methods and discussion sections L277-283 and L451-465 accordingly.

Minor issues –

Use of OTUs rather than ASVs.

The paper uses somewhat older OTU-clustering methods rather than more recent ASV methods. While other reviewers may take a harsher view, my own perspective is that it is not necessary to redo the whole analysis to update this aspect of the methodology. However, it will be necessary to update the discussion of microbial diversity, especially when absolute (rather than relative) quantities are discussed, since this choice of methodology is predicted to consistently over-estimate diversity across all samples.

Specifically, the authors should address findings from mock community studies that show that OTU-based approaches can dramatically overestimate microbial diversity relative to ASVs (e.g. say deblur in QIIME2 or DADA2). Nearing et al., 2018 "Denoising the Denoisers" provide a nice summary of this issue in an independent analysis of the situation conducted by a group that didn't develop any of the methods testing (link: https://www.ncbi.nlm.nih.gov/pmc/articles/PMC6087418/). The authors cite some of the previous literature on this issue.

We agree and now based all our results based on an ASV method. We made changes accordingly in the methods section (L177-205) and all throughout the results and discussion section, as well as in the figures, that are now all based on ASV data.

The reason I list this as a minor rather than major issue is that we should probably expect this overestimation to be roughly similar across taxa, so I wouldn't expect that this is likely to change the estimates for e.g. the proportion of microbial phylogenetic diversity that is lost with species extinction. Other methodological choices the authors made (iike focusing on phylogenetic diversity and/or Allen's metric rather than e.g. Chao1 estimates) will help insulate their diversity results against inflation of the number of OTUS. In most cases, erroneous OTUs that appear due to sequencing noise will be phylogenetically clustered with the true sequence. Since the authors use phylogenetic measures for both alpha and beta diversity, their richness estimates will down-weight these spurious OTUs.

DNA extraction method (line 156). Different DNA extraction methods can preferentially favor certain microbial taxa over others. It was unclear to me whether the kit used was designed for microbiome work? Many of the kits that are commonly used across microbiome studies have been tested on mock microbial communities to assess their taxonomic coverage (e.g. how badly will using any given kit alter observed microbial taxonomic ratios?). I know the authors have used this kit before, and I don't doubt that it was able to get out microbial DNA given the results. However, I was unable to tell from the manufacturer's website whether it has been tested on mock communities to see if it *evenly* extracts DNA from across the microbial tree? No method is likely to be perfect in this regard, but it would be good to discuss a bit more what we know or don't know about the properties of this DNA extraction approach.

To our knowledge, there is unfortunately no published specific test of this kit on bacterial mock communities. In our study, we found that the Maxwell's 16 Buccal Swab LEV DNA extraction kit used with the extractor (Promega) gave better DNA yields and purity than two other commercial kits (Qiagen's PowerSoil and Blood & Tissue) when tested on our samples. Extraction from coral mucus, especially, was particularly challenging and we found that Maxwell's kit worked better.
This kit has been used for microbiome characterization of diverse types of samples, especially in complex matrices (driftwood (Kalenitchenko et al. 2018), human saliva, feces and bronchoalveolar lavage fluids (Lim et al. 2017; Liu et al. 2012); and bovine rumen (Golder et al. 2018)), where it has been notably shown to be efficient for assessing ultra-rare bacterial diversity and viral detection, and an absence of cross-contamination, when assessed. Another Promega kit based on the same technology for the same automated extractor gave better results on mock communities and human saliva samples than another commercial kit (Alessandrini et al. 2019). We added a statement about this point in Methods section L155-159. While we cannot rule out some biases related to the extraction method, Class-level bacterial composition found on animal surfaces and in bacterioplankton was similar to results from previous studies (Figure 3, L365-382) and bacterial clades abundant in animal surface microbiomes were similar to those reported previsouly (L396-410).

Diversity of coral microbiomes vs. seawater: These results seem to contradict those summarized in the McDevitt-Irwin review (https://www.frontiersin.org/files/Articles/286253/fmars-04-00262-HTML/image_m/fmars-04-00262-g004.jpg
), where seawater is presented as much richer than coral microbiomes in terms of # of OTUs. It may be worth a bit more discussion of the previous literature on seawater vs. coral and fish microbiome richness. I suspect the differences could be due to a) failure of some other studies or the comparison between them to account for sequencing depth (which is accounted for in this study) or b) differences in sampling method (e.g. for amount of water sampled), c) differences in DNA extraction method.

OTU/ASV richness is difficult to interpret *per se* as, as you pointed, since it is dependent of several methodological choices. In our study, we averaged the microbiomes *per* animal taxon (as stated in

Methods L207-209), therefore the number of ASVs *per* taxon is higher than the number of ASVs *per* sample. To provide more details, we now provide a table in SM2-2 displaying the number of observed ASVs and phylogenetic richness *per* sample based on rarefied data, with comparison with plankton ASVs and phylogenetic richness.
Samples of coral and fish microbiomes contained significantly more ASVs than planktonic samples (Mann-Whiney U tests, $P<0.05$).

This finding contrasts with the recent study from Pollock et al., which found higher OTU richness in seawater than in coral mucus. This difference may likely be due to the differences of volumes in coral mucus and seawater considered; as they considered both smaller coral surfaces and larger water volumes than we did (L327-336). However, as the observed OTU richness is highly dependent, not only to sampling, but also DNA extraction and amplification efficiencies and sequencing depth (1,000 sequences per sample in this study compared to 7,000 sequences per sample in ours), as well as subsequent data analysis (e.g. rarefaction or not, OTU picking method), we prefer to focus and phylogenetic richness and diversity, which are less biased by methodological choices. We edited the L323-350 in the results and discussion to detail this point.

Other comments:

Line 167: The choice of primers is standard and reasonable.

Line 189: I felt that the 2000 sequences/sample rarefaction depth, and the use of rarefaction to avoid the possibility of false positives due to unequal sequencing depth was reasonable in the context of this study. Some reviewers may argue that parametric methods should be used exclusively, but I would suggest that the authors made a reasonable choice given the study questions (we should be more scared of inducing a false-positive result using parametric methods and an incorrect model than using the conservative but reliable rarefaction approach).
Thank you. The sample rarefaction depth is now of 7000 sequences, based on curated ASVs data. To assess the possible bias due to rarefaction approach, we compared the results obtained with and without rarefaction, and obtained largely similar results, both on phylogenetic and taxonomic diversity measurements (Figures 1-4, SM2, and all along “Results and Discussion” section).

Line 422: Expand ‘within the skeleton’ to ‘within the tissue and skeleton’ (since we know the tissue is a key habitat for coral-associated microbes from microscopy and sequencing).
Done L461.

Line 424: samplings -> sampling
Corrected.

Line 423: ‘The reported estimate of animal surface microbiome diversity is conservative’. I understand the authors point, but given the OTU clustering methods used I think this statement should be revised.
We changed the method, to remove all non-described and mitochondrial ASVs, and grouped these ASVs into 97%-identity OTUs (26,589 OTUs). We also changed our method to estimate total diversity, which not longer calculated at lagoon’s scale, but at the scale of the animal taxa considered. We added more precision on L451-465.

Line 433: Of course, strictly speaking this should really be a phylogenetic independent contrasts (PIC) analysis, not a plain Spearman since we are comparing across species and the structure of the tree will effectively cause the data to be pseudo-replicated.
You are right. We ran the phylogenetic independent contrasts of the vulnerability and diversity

indices before assessing the correlation between them, and still got no signal. We made the modification in the methods L230-237.

References :

Alessandrini, Federica, Andrea Brenciani, Simona Fioriti, Filomena Melchionda, Marina Mingoia, Gianluca Morroni, and Adriano Tagliabracci. 2019. “Validation of a Universal DNA Extraction Method for Human and MicrobiAL DNA Analysis.” *Forensic Science International: Genetics Supplement Series*, The 28th Congress of the International Society for Forensic Genetics, 7 (1): 256–58. https://doi.org/10.1016/j.fsigss.2019.09.098.

Callahan, Benjamin J., Paul J. McMurdie, Michael J. Rosen, Andrew W. Han, Amy Jo A. Johnson, and Susan P. Holmes. 2016. “DADA2: High-Resolution Sample Inference from Illumina Amplicon Data.” *Nature Methods* 13 (7). https://doi.org/10.1038/nmeth.3869.

Golder, Helen M., Jennifer M. Thomson, Stuart E. Denman, Chris S. McSweeney, and Ian J. Lean. 2018. “Genetic Markers Are Associated with the Ruminal Microbiome and Metabolome in Grain and Sugar Challenged Dairy Heifers.” *Frontiers in Genetics* 9. https://doi.org/10.3389/fgene.2018.00062.

Kalenitchenko, Dimitri, Nadine Le Bris, Erwan Peru, and Pierre E. Galand. 2018. “Ultrarare Marine Microbes Contribute to Key Sulphur-Related Ecosystem Functions.” *Molecular Ecology* 27 (6): 1494–1504. https://doi.org/10.1111/mec.14513.

Lim, Yenkai, Makrina Totsika, Mark Morrison, and Chamindie Punyadeera. 2017. “The Saliva Microbiome Profiles Are Minimally Affected by Collection Method or DNA Extraction Protocols.” *Scientific Reports* 7 (August). https://doi.org/10.1038/s41598-017-07885-3.

Liu, Hongbo, Yan Gan, Bo Yang, Hui Weng, Chunmei Huang, Daofeng Yang, Ping Lei, and Guanxin Shen. 2012. “Performance Evaluation of the Maxwell 16 System for Extraction of Influenza Virus RNA from Diverse Samples.” *PLOS ONE* 7 (10): e48094. https://doi.org/10.1371/journal.pone.0048094.

Louca, Stilianos, Florent Mazel, Michael Doebeli, and Laura Wegener Parfrey. 2019. “A Census-Based Estimate of Earth’s Bacterial and Archaeal Diversity.” *PLOS Biology* 17 (2): e3000106. https://doi.org/10.1371/journal.pbio.3000106.

Appendix B

Dear Dr Costa,

Thank you very much for accepting your paper for publication in Proceedings B.
You will find hereafter in blue your answer to referee's comments.

Sincerely,

Marlène Chiarello on behalf of all authors

Reviewer(s)' Comments to Author:

Referee: 1

Comments to the Author(s).
--Summary.
--The authors used an extensive dataset of reef animal surface and seawater microbiome samples (almost 300 samples) to estimate the phylogenetic and taxonomic diversity of coral reefs. This paper successfully expands the current baseline knowledge on coral reef macro-organismal microbial diversity because they sampled species of coral, fish, and other invertebrates that have never been sampled previously. This allowed them to significantly contribute to the coral reef microbiome literature. A particular strength of the paper is the thorough robust and careful analyses. They show that diversity of animal microbiomes in reef environments are strikingly more diverse than the surrounding seawater, a finding that is even more interesting considering that the surfaces of each of these organisms is in constant contact with surrounding seawater. Finally, the major novel component of this paper is the estimation of decline in prokaryotic diversity given relevant extinction scenarios of fish and coral in the Anthropocene. Given the importance, yet global decline of coral reefs I feel this extinction scenario is a timely, relevant, and unique aspect of this study.
Thank you very much for your review and very positive comments on our study. We included all corrections you suggested in our manuscript. Please find our answers below in blue.

--Major Comments.
--I have no strong major changes to the manuscript. I feel the study was careful and thorough. I especially appreciate the extensive supplementary figures using un-rarefied data to explicitly show that their rarefaction procedure did not introduce significant bias into the data to alter the conclusions. Furthermore, inclusion of DADA2 code, R packages with version numbers and relevant functions, novel code, and functions support the reproducibility of this study, which is important in analyses with significant analytical and coding components.

--Minor Comments.
--Overall, I have only minor comments and changes to the manuscript. The following comments are mostly questions regarding figures and statements made in the text.

--Line 55 in abstract vs. 462: "would induce an erosion of 28% of the prokaryotic diversity" vs. "would induce a loss of ca. 28 to 29% of ASVs, based on rarefied and un-rarefied data". I think the intent is the same, but it seems like the extinction scenario is only modeled with richness data. Maybe change the abstract to say prokaryotic richness? Though I recognize richness is a diversity metric, so if you do not change it I understand.
We now mention richness instead of diversity in our abstract.

--In general, I think you should check the figure labels in the body of the text. I think you reference incorrect figures at times. Here are the examples I found where I believe you meant a different figure or figure part:
--Line 115: Figure S1 should be Figure SM1-1.
--Line 343: Figure 3c should be Figure 3a.
--Line 394: Figure 3a should be Figure 3b or 3c.
--Line 403: Figure SM2-6 should be SM2-7.
--Line 413: Figure 3b and SM2-4 should be Figure 3c and SM2-5.
--Line 456: SM1-6 should be SM1-7, the vulnerability values. (SM1-6 is DADA2 script in my document).
We apologise for such errors and made the requested changes – we also checked again the entire document to remove any incorrect reference to figures.

--In the methods I appreciate the inclusion of all R packages and version numbers. In addition, the inclusion of specific functions from each package will help improve the reproducibility of this paper. Along those lines, please make sure the formatting of each package name and function are consistent throughout the methods. It seems they are formatted as package and 'function', but in Line 150 and 163 there is different formatting.
We apologise for this and uniformed all mentions for packages and functions.

--Figure 1 and associated supplementary figures: Figure 1 is generally inconsistent with the figure legend and with the body of the text in the paper associated with the figure. Figure 1a is labeled as "Taxonomic richness (Nb ASVs)". The legend states that this should be phylogenetic richness. Indeed, the associated supplementary figure, SM2-1 has figure a as "Phylogenetic richness (% total Faith's PD)" and a different scale. I think you added the wrong figure here. SM2-3a is supposed to be the figure for taxonomic richness, so maybe that was added to Figure 1a accidentally.
For figure 1b, the numbers are not reported as percentages as indicated in the legend. In SM2-3b, the raw number of ASVs are reported for taxonomic richness, but they are different than the values in Figure 1b, so it is unclear which numbers reported are correct.
For Figure 1 c and d, they are labeled as "Taxonomic richness" and "Taxonomic diversity", respectively. In the legend, they are supposed to be "Phylogenetic richness" and "Phylogenetic diversity", respectively. If they were phylogenetic metrics, they would match the associated Figure SM2-1c,d. It is unclear what the correctly reported figure should be.
We indeed inverted two figures here! Our apologies. We changed the figure, and checked the figure legend. The numbers were correct; confusion was due to the inversion between figures.

--Line 344: Change "clades" to classes. It seems you are reporting bacterial classes in the figure. It might be worthwhile to change "clades" in the body text to classes to be more clear about the taxonomic hierarchy you are considering.
Corrected.

--Line 351-59: Your finding of Arcobacter in the coral microbiome is indeed interesting considering the pristine nature of the sites you sampled. You discuss how this genus is found in thermally stressed and fish effluent stressed corals. It might be important to note that Arcobacter is a common microbe (based on ASVs and OTUs) found associated with coral disease, though your corals were not in any diseased states. Here are some citations that found Arcobacter in coral diseased lesions:
Gignoux-Wolfsohn SA, Vollmer SV (2015) Identification of Candidate Coral Pathogens on White Band Disease-Infected Staghorn Coral. PLoS ONE 10:e0134416.
Sunagawa S, DeSantis TZ, Piceno YM, Brodie EL, DeSalvo MK, Voolstra CR, Weil E, Andersen GL, Medina M (2009) Bacterial diversity and White Plague Disease-associated community changes in the Caribbean coral Montastraea faveolata. ISME J 3:512–521.
Meyer JL, Castellanos-Gell J, Aeby GS, Häse CC, Ushijima B, Paul VJ (2019) Microbial Community Shifts Associated With the Ongoing Stony Coral Tissue Loss Disease Outbreak on the Florida Reef Tract. Front Microbiol 10:2244.
We agree and now mention these relevant papers and the fact that this genus is associated to disease.

--Line 435: "cured" should be "curated"
Corrected.

--Line 463: "such extinction would induce 23% of phylogenetic richness in the ecosystem"
Do you mean that it would induce a "loss of" 23% richness? I found this sentence confusing after the previous sentence discussing loss of ASV-based richness in the community.
Phrasing was indeed unclear here – we rephrased the sentence.

--Line 466-8: "slope of microbial diversity extinction, which increased only slightly with increasing loss of macroscopic species"
I could be mis-interpreting Figure 4b, but it seems that with increasing loss of species, the slope decreased from -0.5 to closer to -0.75. It increased in the negative direction, if that is what is meant by that sentence. I could be interpreting the graph incorrectly, but I found that sentence misleading.

We indeed meant a steeper decrease. We rephrased to avoid such confusion.

--Figure 2 Legend: The legend does not seem to fully fit with the figure and I think you may need to edit it. Specifically, I interpreted the following fixes: “(a)” at the beginning might actually be “(c)”? I think c in the following sentence, “(b and c) Average rank-abundance curve obtained from surface microbiomes of all animal taxa (b) and all 35 planktonic communities (c),” is supposed to be “a”.
We indeed inverted two legends there. We apologise for this and corrected our legend.