# Response to reviewers

## Reviewer #1 (Remarks to the Author):

Nature Ecology and Evolution review: Global Drivers and vulnerabilities of coral reef fish functions

My overall opinion of this manuscript is positive. The global distribution of these fish functions on coral reefs has not been previously assessed, nor has their overlap or vulnerability. There’s clear novelty, broad interest, compelling results, the manuscript is generally well written and well-structured, and the authors use multiple extensive datasets and a wide array of sound laboratory and statistical methods. Therefore, my overall recommendation is that this article will ultimately be publishable in Nature Ecology and Evolution. However, I am recommending a major revision because this manuscript needs to be substantially improved. While there is little to criticize regarding most of the data, methods, and results, the paper is severely lacking in sufficient detail and information. In many instances, not nearly enough information is provided for the reader to fully comprehend what is being presented, and in some instances key data or methods elements are not provided. There are also too many bold and unnecessary statements that are not properly backed up by the given citations. The authors should also check for grammar errors, as I encountered many throughout the paper. Lastly, there are two instances where I believe the methods are not optimal and should be improved (calculating correlations between functions and estimating species’ length infinity). I will go through the paper line by line and provide my comments and suggestions:

*Thank you for your constructive and detailed review. We have now provide more details in methods and supplementary methods. Further, we updated the approach to estimating correlations by applying multivariate models as suggested. We also thoroughly checked for grammar errors and modified the main text substantially. Overall, we incorporated all comments with exception of the suggested approach to estimating species length infinity. We provide more details of the modifications below.*

Line 3: why not just say “Preserving ecosystem functioning is a critical challenge…”   
*We now state “Preserving coral reef functioning is a critical challenge of the 21st century.”*

Line 10: what is meant here by “diverging,” perhaps a better word choice is needed.   
*Changed to “varying”*

Line 20: these first two statements are overly bold and not well-focused. The authors would do better to precisely outline the issues they fill with this research article, and to do succinctly, with concise, information-driven, factual sentences. Have people really been managing coral reef fisheries with a surplus-production model for thousands of years? Even if true, does this information help place your reader in the context of your study? The citation given referencing millennia of production-based management schemes is not an article about historical ecosystem impacts, but a modern, long-term plant biodiversity experiment. The authors should improve these aspects throughout the paper and ensure proper citations are given.   
*We have now substantially changed the first paragraph of the introduction to quickly place the reader in the context of coral reef functioning.*

Line 32: Coral reef ecosystem services are crucial for people throughout the tropics, but not necessarily all of humanity. You might want to replace humanity with something more focused. The sentence that follows about cycling elements to thrive in nutrient-poor waters is out of place and breaks the flow. What you want is to connect the importance of reef ecosystem services to their threats, so this sentence should be removed.   
*We removed “to humanity” and the sentence about efficient cycling of elements.*

Line 54: this should say coral reefs, not just reefs, since you do not include other types of reefs, e.g., rocky reefs from temperate locations.   
*Done.*

A general comment for the entire results section: very few of the actual results are presented - there is little to no information on the analyses used, there are no numbers/data/statistics provided, and the main results are very quickly glossed over. While the format of NEE dictates that much information be provided in the supplementary material, sufficient information must still be given for the reader to understand how you achieved your conclusions. A good example comes on line 126, where the authors state that by “combining vulnerability scores and species contributions,” they demonstrate how fishing and coral loss will impact fish functions, then a figure reference is provided. The reader has no idea how these things were combined or what analyses were done to reach these conclusions. In the paragraph beginning on line 54, when the authors state that fish functions vary substantially, as a reader I am expecting to find out how they vary, which ones covary, which areas of the world have high or low functions or high or low multifunctionality, yet all that is stated is that herbivory is negatively correlated with phosphorus excretion. So, while the authors state that they demonstrate clear tradeoffs in services and that biomass alone is insufficient, I must take them for the word and rely only on a figure to provide any detail.

*We now wrote a new, separate results section. Further, following the recommendation of another reviewer, we now removed this part of analysis entirely. Functional vulnerability is now discussed by combining our results with the known effects of human stressors on fish community structure.*

Paragraph beginning on line 86: no information is given here as to how these contributions were calculated. The reader will have to constantly reference the supplementary methods to understand the results of the paper.   
*We now provide more information in the results section.*

Coming back to line 126: no information is provided as to how this analysis was done or what was quantified, and no actual quantitative results are given. As a reader it’s hard to believe that anything has been truly demonstrated when the only information given is a figure.   
*This paragraph is now removed entirely because as we removed this part of the analysis as mentioned above.*

Concluding paragraph: I suggest the authors spend some time reflecting about the management implications of this work. Are there certain high or low functioning areas that should be prioritized? If we have limited resources but need to prioritize some functions, or combinations of functions, to preserve, which should they be?   
*This paper aims to demonstrate existing trade-offs among functions and what drives them. In order to have applicable management recommendations, we first need to find out how much of a certain function is ‘good’. We are currently still lacking understanding of what a healthy reef really looks like in terms of process-based functions. We added the following sentences: “It is necessary to gauge the state of reef ecosystems based on multiple, complementary, process-based functions, yet our comprehension of process-based functioning of coral reefs or what a “functional” coral reef really means is still in its infancy. To appropriately manage coral reef functioning with respect to local needs and stressors, future studies should seek to establish functional baselines for global coral reefs. Until then, our study suggests that coral reef fish functions can be managed by enhancing standing stock biomass, protecting local key species, and promoting a balanced community structure."*

Figure 3: I am a bit confused on how to interpret the x axis. Negative coefficients indicate higher contribution to function while positive indicate higher contribution to biomass? So, Lethrinids contribute highly to biomass but not piscivory? Please try to make this clearer for the reader.

*The x-axis was showing the difference between the contribution to function and the contribution to biomass. This means that higher values indicate that a certain family tends to contribute more than a certain function compared to its biomass. We now removed this figure, because it seemed to be more confusing than informative.*

Methods: something critical has been missed here despite the extensive use of Bayesian mixed effects models in this paper. These data have an important, intrinsic hierarchical structure that has not been accounted for nor mentioned. In these data, transects are nested in sites, which are nested in localities. This needs to be stated upfront for the reader to understand the structure of the data and its implications. Furthermore, some models may need to be adjusted to account for this hierarchy. Observations from the same spatial units are non-independent and likely exhibit some degree of autocorrelation, so the statistical methods must account for this non-independence. Hence, the models used should be adjusted to include varying intercept random effects for site and locality.

*We adapted all models to include random effects for site and locality and we modeled all five functions simultaneously in a multivariate model structure. Further, we now explicitly discuss the hierarchical structure of the uvc dataset at the start of the method section.*

Line 8: when you selected families for which you have data, was there a minimum in terms of the number of species in this family for which you had data? For instance, did you select an entire family even if you only had data on one species within the family?

*To add in supp methods.*

Line 26/27: here I am getting a bit lost. You say you ran “the model” yet no model has been presented. You then say you ran this model for every combination of body sizes and sea surface temperatures, yet as a reader I don’t know anything about body sizes or sea surface temperatures at this point in the methods/data, so it’s impossible for me to gauge what exactly you are doing. Furthermore, there has been no mention of what environmental data were used or where they came from. You mention sea surface temperatures, but not whether these were measured in-situ during fish transects or came from satellite sources or a climate model. For all instances of environmental data in the paper, the source and resolution of these data must be provided.

*The model mentioned here is talking about the bioenergetic model provided in the R package ‘fishflux’ that estimates fluxes of carbon, nitrogen, and phosphorus. We now reworded this section to clarify. We also added mentions of body size and sea surface temperature in the first section of the methods.*

Line 45: a reference should be given for temperature.

*We added the reference for sea surface temperature in the first section of the methods.*

Line 47-50: you’ve stated that a mixed effect model was performed but the equation given clearly shows a simple linear model with no random effects, so this isn’t actually a mixed effects model. However, it should be. If I am clearly understanding this section the authors have performed, for each function, a simple liner regression of function ~ log(biomass) + SST, and they have extracted the transect-level residuals. Then, they’ve tested the raw Pearson correlations between the residuals for the different pairs of functions. There are multiple issues here. 1) this approach does not account for the spatial hierarchy of the data, so it ignores spatial non-independence as well as uneven sampling effort. Locations with the highest sampling intensity will thus drive the resulting patterns. This needs to be a mixed effects model with random effects for site and locality. Curiously, the authors state just afterwards that they did run a mixed effects model of this nature, but no further information is given. 2) Running a correlation on the residuals of a linear model instead of running a multiple regression is an inappropriate analysis. Please see Freckleton, R.P. (2002), On the misuse of residuals in ecology: regression of residuals vs. multiple regression. Journal of Animal Ecology, 71: 542-545 for a thorough review of this issue. Please also see McElreath, Richard. Statistical rethinking: A Bayesian course with examples in R and Stan. CRC press, 2020, particularly Chapter 5 page 137, which states:

“Residuals are parameters, not data. There is a tradition, especially in parts of biology, of using residuals from one model as data in another model. For example, a biologist might regress brain size on body size and then use the brain size residuals as data in another model. This procedure is always a mistake. Residuals are not known. They are parameters, variables with unobserved values. Treating them as known values throws away uncertainty. The right way to adjust for body size is to include it in the same model, preferably a model designed in light of an explicit causal model.”

If all parameters are standardized, regression coefficients are equivalent to correlation coefficients. Thus, in the simplest sense, if you wanted to know the correlation strength between, for example, herbivory and phosphate excretion while accounting for the spatial hierarchy of your data and controlling for biomass and SST, you could run a Bayesian mixed effects model of herbivory ~ phosphate + biomass + SST, or phosphate ~ herbivory + biomass + SST (both of which would provide the same correlation strength). However, the most appropriate way to do this analysis is probably to perform a multivariate mixed effects model (like others performed in this paper), with [herbivory, piscivory, phosphate excretion, nitrate excretion, production] ~ biomass + SST, and to extract the correlation coefficients between the functions from this model. This will provide equivalent results to what several sets of mixed effects regressions between each pair of standardized functions would provide, but in one holistic analysis.

*We agree with the reviewer on both arguments made. We now included random effects of localities and sites. Further, to estimate correlation coefficients we opted for the most elegant option - multivariate mixed-effects models. This enabled us to estimate correlations on the site, locality, and residual level, as is now shown in figure 2. The only obstacle for using multivariate models was that some transects have missing values (false zeros) for herbivory or piscivory. We tackled this issue by including these observations as missing values, essentially handling them as parameters in the model. This is now explained in the methods section as followed:* *“In the underwater visual transect database, 291 transects (3%) did not contained herbivores and 4467 transects (49%) did not contain piscivores yielding false zeros for herbivory and piscivory, respectively. We considered these zeros as missing values because the absence of the observation of a herbivore or piscivore does not prove the actual absence of herbivores and piscivores. In fact, it is highly unlikely that a coral reef fish community contains no herbivores and piscivores at all. To avoid removing all transects with missing values for herbivory or piscivory (n = 4620) from our database when running multivariate analyses, we imputed the missing values by adding these observations as parameters in the multivariate models.”*

Line 48: does community-level refer to transect? If so, state that.   
*Indeed, we now stated this.*

Line 58: because the structure of the data has not yet been presented, the reader is not aware of what a “locality” is, so make sure to sufficiently introduce the data in the beginning of the methods.   
*We now introduce the structure of the data in the beginning of the methods.*

Line 72: You should state why each function was log-transformed.   
*Functions were log-transformed to ensure a normal distribution of residual errors and to incorporate the allometric relationship with biomass of most functions due to metabolic theory. We now state is in the methods.*

Line 76: again, this appears to be a simple linear model with no random effects, and therefore not a mixed effects model. Perhaps I am missing something? In either case, this model should also be a mixed effects model that accounts for the data structure: transects nested in sites nested in localities, otherwise your data could be highly biased by uneven sampling effort between localities and by autocorrelation within localities.   
*As stated earlier, we now included random effects.*

Line 82: What were these weakly informative priors? Standard practice is to provide this as part of the model equation/formula to be completely transparent with the model. Secondly, no information has been given as to how the models were evaluated? There must have been some assessment of model quality? Minimally R hat, chain convergence, posterior predictive checks, etc.   
*We now provide the priors we used, and specified how we assessed the model quality. Specifically, we checked, the R hat, trace plots, and posterior prediction plots.*

Line 104: I’d encourage the authors to read carefully over their manuscript. This sentence is an example where the phrase is ambiguous and not clear. “To know how often species are contributing more than average” to me means: in how many transects is there a species that dominates the function. However, what the authors mean is for each species, how often does it dominate a function. So, it is a species-by-species analysis and not a transect-level analysis, and this needs to be clear.   
*We now modified the wording to: “Finally, we quantified the frequency of dominance per species, i.e. the number of transects in which a species is dominant for a given function divided by the total number of transects in which that species is observed.”*

Line 117: It’s unclear to me whether the authors developed their own model based on the scores from Graham et al, or if they used a model already produced by Graham et al.   
*We used the scores from Graham et al. to create a new model and used the model to predict the scores for species that are not provided in that paper. Graham et al. only used three families, so we are aware that this is a big extrapolation. Following reviewer’s recommendation, we now removed this analysis from the paper.*

Line 120: why are model formulas/equations provided for other Bayesian models but not here? *This part was removed.*

Line 121: Reference the supplementary figure otherwise the reader has no basis on which to believe there was a good model fit. *Same*

Line 127-130: this information is passed over very quickly without enough detail. Please expand this section to better explain exactly how this analysis was done and make it clearer for the reader as to how it should be interpreted. *Same.*

Line 10: Here the authors have extracted max lengths from FishBase and substituted them for length infinity. Max length and length infinity, while similar values, are not interchangeable. Froese & Binohlan (2005) specifically published methods on how to estimate length infinity from FishBase max length and provided the formula to do so on the FishBase downloads page. The authors could easily convert these raw max length values to estimated length infinity values using this approach. The authors should either make this adjustment or specify why this was not done.

Froese, R. and Binohlan, C. (2000), Empirical relationships to estimate asymptotic length, length at first maturity and length at maximum yield per recruit in fishes, with a simple method to evaluate length frequency data. Journal of Fish Biology, 56: 758-773.

*Thank you for the suggestion. This is the only comment that we decided not to include. The paper by Froese and Binohlan is over twenty years old and the data they used does not include reef fishes. Moreover, asymptotic length can vary significantly across locations depending on temperature and other environmental variables. We therefore opted for a method proposed and applied by Renato Morais and colleagues. They standardize asymptotic lenth to maximum length to avoid underestimating growth. In this method, the parameter k, is adapted to kmax in order to represent a realistic growth curve. Doing this may underestimate growth of juvenile fishes, while the potential overestimation of adult fishes would be very minimal. As visual census data often does not accurately record juveniles and we removed small fishes from this specific analysis, we argue that our method minimizes potential bias considering the data available. The risk of using a general conversion factor would be too big as we would estimate a growth equal to zero for a lot of individuals and potentially dramatically underestimate biomass production.*

Line 11: what is this standardized coefficient? Is there a reference? *We added the definition and reference. “  
For k, we used a standardized coefficient that describes the potential growth trajectory of an individual if linf were to be equal to its maximum length (i.e., kmax; Morais & Bellwood, 2018).”*

Line 14: Be careful to reword this sentence, you’ve filtered out all other species than the species in your species list, not the other way around. *We changed it to “We selected the species of our species list”*

Line 15: I am a bit lost here, what is the “length-frequency method” is this also from Morais? *The length-frequency method is another, arguably unreliable way to estimate growth trajectories, that is used for some of the data provided on fishbase. We opted to only use growth curve parameters that were estimated from otolith data. We clarified this by saying: “We selected the species of our species list and only included the k estimates coming from otolith studies.”*

Line 17: again, what temperatures are you referring to? No information has been given about temperature to this point in the manuscript.   
*As mentioned above, we now added text on temperature at the start of the methods section.*

Line 18: “additional” otolith data implies that there were otolith date previously referenced in the manuscript. If this is not the case be sure not to confuse the reader.   
*We now clarified this by mentioning otolith data in the previous paragraph.*

Line 19: “All fishes” - there is no information given on how many species were collected or how many families these species represented. You should also mention how these fishes were collected – what method and were they euthanized, etc.?   
*Since initial submission, this data was published under the form of a data paper. We now added this reference.*

Line 41: Here you state that you fitted Bayesian Hierarchical models, but no information is given about these models? How are they hierarchical? Why are they hierarchical? What likelihood functions were used? What priors were used? This information should be presented for each of the Bayesian models you conduct in the paper.   
*We added more details on the model.*

Line 43: It is not clear how many species and individuals actually had otolith data? 496 estimates yes, but how many species?   
*“After combining the two data sources, we obtained 496 estimates of kmax for 181 species across varying temperatures.”*

Line 52: natural log transformed.   
*Done.*

Line 53: what were these priors? Put them with the model formula.   
*To do!*

Line 59: Is this value considered high? Low? Medium? It would be good to give some context for the reader.   
*It means that most of the variation (independent of size and sst) is explained by phylogeny. This value is thus considered quite high. We now specify this.*

Line 69: Here you say, “for all 100 trees.” Again, you really have to consider the information you have given to the reader at this point. The reader hasn’t heard anything about any phylogenetic trees, but now you are saying you’ve got 100 of them. Why 100? Where did they come from? This information is really the basic necessities that need to presented before you rush into describing the analyses.   
*We added a paragraph on the phylogenetic trees.*

Comment for phylogenetic extrapolation: overall I’m quite impressed with the work you’ve done, and I do believe your phylogenetic analyses are properly executed. However, is it possible to perform some form of cross validation to ensure the accuracy of these extrapolations given that you’ve extrapolated a huge number of species? For instance, image that all the hard data you collected came from closely related species and families, this might produce unreliable “out-of-sample” predictions.   
*To do.*

Line 110: It could be good to say how many families and how many species per family since you use these data for extrapolation. *To do.* Line 123: combined how?   
*Essentially by taking a weighted average of nutrient contents across trophic groups where the weights represent the probability of being in a diet group.*

Line 135: The level of detail given here is really insufficient. Where were these individuals caught? What habitat type? As a reader I don’t know what part of the world they came from or whether they were caught in a lagoon or on a reef crest.   
*To add.*

Line 153: how did you define what an outlier is? *To add.*

Line 161: If you used an informative prior you must give the reasoning as to why and to how you chose the prior. *To add.*

Extended Data Figure 2: it might be easier to read if you put the names of the functions as figure titles. *We added the function names.*

Extended Data Figure 5: just a small thing, but some figures have capital letters for the panel labels, and some have lower case.   
*Fixed.*

Final comment: I want to reiterate that although my review is detailed and critical, I tried to be constructive because this paper has clear potential to be published in NEE, and you need to tie up any loose ends in this manuscript before it can be published in a top tier journal. Overall, I remain positive about the manuscript. The main substance in terms of data and results is undeniable and I will ultimately support the publication of this manuscript.

I am signing my review for transparency and I invite the authors to contact me if they have any questions.

Sincerely, Matthew McLean

**Reviewer #2 (Remarks to the Author):**

This manuscript offers a summary of how coral reef fish communities affect ecosystem functions across the globe by analyzing a database of reef fish species abundance in reefs around the world. The authors then calculated five ecosystem functions using multiple, complex methodologies and analyze the role of fish community structure (i.e., richness, body size, trophic level, and immaturity) in individual functions. As presented, no one individual aspect of community structure is revealed to explain ecosystem functioning across global reefs, but the authors suggest that the variation among communities and between regions means that there is no clear objective for protecting coral reef function and that there are no global keystone species to target.

Overall, this manuscript needs to be rewritten, line by line, in order to improve its clarity. Sentences are imprecise throughout, which muddies the take home messages of the discussion and confuses the main objectives of the study. This is an impressive piece of work that has used multiple data sources and synthesized them in unique, groundbreaking ways. When a manuscript involves big data like this, the authors absolutely have to be fully precise with their explanations of why things were done in the way they were. This manuscript unfortunately is not there, from both an ecological and methodological standpoint. Clearer gaps in knowledge need to be addressed in the introduction, more precise language on how these gaps will be addressed needs to be added/changed, and the big picture and small scale implications of what we have learned from this global analyses needs to be outlined better.

*Thank you. We have now rewritten the manuscript substantially. We focused on improving clarity, we added a results section, and provided much more details in the methods.*

One major point of confusion for me was the lack of any attempt to combine the ecosystem functions mentioned into a multifunctionality framework. The authors even specifically state that it is necessary to examine multiple functions at once in order to understand how reefs may be changing with different fish communities (line 154-6). Well, the data exist in this paper for this to happen! Even a simple average geometric mean of multifunctionality could be informative, especially for identifying hotspots in fish-driven reef function.

*We now added estimates of multifunctionality per locality and and biomass based on a geometric mean as suggested. We agree that suh a measure can be infomative, but we argue that multifunctionality should be interpreted with care as this summarizing measure may obscure the trade-offs between individual functions.*

Also, based on the title and abstract of this paper, I expected a discussion of how fish community structure as a whole, and not just species diversity, need to be explicitly considered when estimating how fish affect functioning in coral reefs. There are a couple of places where this topic is briefly touched upon, but this should be directly addressed in both the introduction and discussion. This would also allow the authors to discuss a key finding, that biomass is the key predictor for ecosystem functioning across reefs. I know that, and agree with, the biomass-corrected variables used in the paper. But I don’t see why the authors couldn’t use and discuss both biomass effects, and then biomass independent effects. That seems like a huge part of this story, and that there actually IS a global pattern in how fish communities affect ecosystem functions in reefs. The authors mention that their work reveals that there is no clear objective for protecting coral reef functioning, and there are no global keystone species to target. It is unclear how exactly the authors evaluated those statements, and is surprising that the authors have no recommendations that would help address those clear conservation goals from their global analyses. *We now mention the role of community structure in the introduction. In the discussion, we both discuss the effect of biomass and community structure as a whole. We also modified the discussion on conservation.*

*We now say: “It is necessary to gauge the state of reef ecosystems based on multiple, complementary, process-based functions, yet our comprehension of process-based functioning of coral reefs or what a “functional” coral reef really means is still in its infancy. To appropriately manage coral reef functioning with respect to local needs and stressors, future studies should seek to establish functional baselines for global coral reefs. Until then, our study suggests that coral reef fish functions can be managed by enhancing standing stock biomass, protecting local key species, and promoting a balanced community structure."*

Section/Line comments Abstract: This section is very strong. The authors should use the focus of each of these sentences to guide the topic sentences of the manuscript sections. Clear and concise! *Thank you. We aimed to clarify all topical sentences.*

Introduction: The introduction could use a rewrite with clarity and concision in mind. Between lines 25-42, I counted way too many uses of words and phrases that contradict the statement made before (e.g., “However, ….” , “…but simultaneously quantifying”, “Yet, the integrity of reefs” etc.). This is indicative of an unfocused introduction. The central message is simple enough that the authors should be more specific and precise with their wording and sentences. *We rewrote the introduction and avoided using many contradicting words. We believe the introduction is now more focused.*

3- “integrity” of ecosystems is vague, especially when the authors actually mean “functioning.” Could actually just remove this opener and start with the next sentence, which is pretty good. In that sentence, and throughout, be careful with informal language. “The world’s coral reefs” can be replaced with “Coral reefs” (I don’t think any readers are going to be thinking of any coral reefs on other worlds!) and “today’s coral reef conservation efforts” to “current coral reef conservation efforts”.   
*We removed the word ‘integrity’ throughout the manuscript. We also avoided using informal language and removed the given examples.*

21- “integrity” repeated here again. Still unsure what the integrity of an ecosystem means. *Agreed, it is a rather confusing and vague term. We removed it.*

22- careful with switching tenses here. This sentence could probably be removed. *Done.*

23-25 why does sustaining multiple functions require both high richness and a variety of species assemblages? Also, what does “a variety of species assemblages” mean? Are the authors trying to say that community structure as a whole, and not just biodiversity, is key to a multifunctional ecosystem? If so, that is a fantastic point to make and represents a huge gap in ecology today! *This section has been changed entirely.*

31-42 I was surprised that this paragraph doesn’t fully focus on the role of fishes in nutrient cycling/ecosystem functioning in coral reefs.   
*The second paragraph of the introduction now focuses on fishes.*

33- “integrity” repeated again. Do you mean function?   
*No more integrity in this manuscript.*

34- An example of an edit to be more concise: the authors wrote “a plethora of stressors” and then added “such as….”. Pick one of the two of these. I would suggest getting rid of “a plethora of”. *Done as suggested.*

37 Generally it is better to write out exactly why a topic is “at the forefront of scientific discourse” as opposed to just stating that it is. Because the reader is left to ask, why?   
*We removed this sentence and now say: “Drastic declines in habitat quality and fish biomass have evoked serious concerns about the persistence of coral reef functioning.”*

40-2 These two sentences are a good example of where this introduction goes astray. First, the sentence starts with “Conversely, …” but the sentence does not offer information that is counter to the previous sentence (not to mention that the previous sentence began with an “However” further obscuring the main point. I suppose “additionally” would be more appropriate, but it’s still unclear. I’m not sure that it is true that we know little about elemental fluxes and their drivers, as claimed. Are you talking about the role of reef fish in elemental fluxes? Be precise. Next, the Allgeier paper cited here as a “but see” example examines exactly what the authors claim that we know comparatively little about. That citation seems like a very good jumping off point for this paper, considering the differences in scale in this paper (global) vs that one (multi-site). Finally, the ending sentence offers no reasoning. Why does “this” (again even the “this” is unclear, is it that we have only proxies of functioning, or that we know little about elemental fluxes? Or both?) constitute a severe limitation.   
*As mentioned above, we rewrote the introduction entirely, and hope you will find it has improved in clarity.*

Paragraph on line 86 This paragraph talks a lot about “species” roles, but the figure shows taxa groups in families. This gets confusing when claims like “functions consistently hinge [upon] a few dominant species” (line 92-3) are made. Do you mean, “a few dominant families”? Or are there key species that provide high functioning?   
*We understand the confusion, and now removed the part of the figure showing the contribution of fish families as it was not a main result and distracts the reader from the main point which is indeed that on a local scale a few key species provide high functioning while the identity of these key species vary across sites.*

97-100 These results seem like a very interesting implication of this research, but I don’t know where in the figures the reader can see this. Without a visual, this is difficult to interpret. Determine how important this finding is, and decide how to get this across in a more clear way. Also, please make sure that this isnt directly repetitive from the introduction.   
*We now have a separate results paragraph accompanying figure 4. We hope this clarifies these findings.*

102-105 See below comment on line 122. This reads like a topic sentence. Readers are going to be very interested in the details behind the claim that both species richness and functionally dominant key species are important to maximize the functions provided by reef fish. Regardless, this sentence doesn’t quite mesh with the opening sentence of this paragraph, so please edit for clarity here.   
*The topical sentence of this paragraph is now: “Our results also reveal that functions consistently hinge on a few dominant species, but the identities of local, dominant species vary strongly across sites.”*

108 I’m not sure what a “severe change” is, especially in a broad term like community structure. *We modified the sentence as followed: “Fishing and climate-induced coral loss have caused declines in reef fish biomass and shifts in community structure.”*

112 “effect” used twice in one sentence. *Changed to:   
“When accounting for the effect of biomass, these community shifts can…”*

114 is this what Fig 2 shows? Do you mean Fig 4? Unclear   
*We removed the mention of the figure in this discussion paragraph as we now have a seperate results section.*

114-116 “On the other hand” and “However” open back to back sentences. This is a clear indication that whatever points are trying to be made, are unfocused.   
*We edited the section and replaced ‘on the other hand’ to ‘further’, and ‘However’ to ‘Simultaneously’.*

122 This closing sentence for this paragraph is a great example of something that happens a lot in this paper. The paragraph above it is wavering, with an unclear main theme. This last sentence is actually quite good and brings up a fantasticthough point. It reads much more like a topic sentence. Considering the topic sentence of this paragraph is pretty vague (i.e., “shines new light on” doesn’t tell us anything specific), the authors should consider a reordering of the logic flow here, in above paragraphs, and in the discussion section as a whole.   
*The topical sentence of this paragraph now reads: “Our global analysis of multiple functions suggests pathways in which human-induced shifts in reef fish community structure may impact coral reef ecosystems.”*

122-4 I’m not sure that something can be both nuanced and holistic. Rephrase.   
*We removed the sentence.*

125 “Similarly…” It is unclear what is similar here. Rephrase   
*We removed this paragraph as suggested below.*

134 What does this sentence mean? How do herbivores fishes have a high variability in species roles? Should you have grouped the species roles differently then, into categories like “grazers of macroalgae” “bioeroders” “grazers of microalgae”? You defined what an herbivore is, so it’s strange to see that high variability is explained by your too-broad definition of herbivore. 135,137 Back to back sentences beginning with “While”.

137 …But can you not do this, with the data you have collected? The whole end of this paragraph is confusing and needs to be revisited. *We removed this paragraph altogether.*

Figure Comments:

Fig 1- This figure encapsulates the scale and scope of these analyses in a really nice way. There isnt a lot of specific information that can be gathered from this figure, besides that there were a lot of sites all throughout the world (which is a great thing!). Do the authors want to highlight that there are hotspots of function in different places, for different functions? If so, probably just include the colored points. Also, would there be any value in a figure that shows the highest multifunctional sites, i.e., where multiple processes are occurring at high rates? Identifying the global multifunctional hotspots of coral reef fish communities could be of value to conservation and management.   
*We now edited this first figure to show the actual function estimates per location, as well as a measure of multifunctionality.*

Fig 2. In the caption (181), do you mean “effect size” or “average effect size estimate” as opposed to “fixed effect values”? Regardless, the community variables besides richness mapped out in this figure are unclear. What is the difference between the 97.5% size line and the 2.5% size line? Is 97.5% size mean that the fish community is mostly made up of large fish? Maybe come up with more descriptive variable names than this? I definitely do not know what the immaturity variables are supposed to represent, and how they are different to size. This is further unclear when looking back at the ms text in line 78 that implies that small/immature fish are the same. \*This is now figure 3. We modified the x-axis label to “standardized effect size” and the percentages to lower and upper. We provided extra information in the legend. An individual fish can both be small and mature, so both variables are quite different. Immaturity is a measure that represents the position in the growth curve. We now also added a definition in the legend.

Fig 3. The axis on 3a needs to be explained better. Dots to the right of the 0 line indicate that this family contributes more to this function than as predicted by biomass? Actually, all axes need to be explained better   
*We removed figure 3a.*

Fig 4. I understand what the authors are trying to do here, but I’m not sure that this figure is necessary. Which functions are vulnerable to climate change and overfishing can be discussed in words without too much confusion. This figure, and the paragraph in the main text surrounding it, seem out of place.   
*As suggested, we removed this figure and the paragraph surrounding it. Instead, we now provide a conceptual figure that supports this part of the discussion.*

Methods: Section 3 Did I miss a mention of the priors used here? I see uninformative priors are used in section 4 80 “fitted” should be “fit” ?

*We added more information on the priors everywhere.*

Supp Methods A strength of this manuscript is the synthesis of multiple data sources and analytical techniques. However, this creates a lot of methods and thus, requires that even the supplementary methods need to be extra clear. It would be helpful to add a topic sentence to each relevant section of the Supp Methods so that the reader knows why each method is being used and explained. For example, line 3 should say exactly what parameters are going to be explained in this Supp Methods section, instead of “a number of parameters are required.” Carry this clarity throughout this Supp Methods section. For example, start section 1 (or 1.1) with something along the lines of, “To estimate the growth parameters of fish species (or families?) in our database, we used maximum length from fishbase and collected otolith data…”. Without this, it is difficult to tell why techniques like the Bayesian phylogenetic regression model are necessary.   
*We now added much more information and introductory sentences in the supplementary methods.*

14 is kmax defined? k is defined as a growth rate parameter in the Methods line 68, but each Methods section should be able to stand alone.

*We added the definition.*

**Reviewer #3 (Remarks to the Author):**

Review of Global drivers and vulnerabilities of coral reef fish functions

I enjoyed reading this paper. Using coral reef fishes as a model system, it contains a simple but important message: a given ecosystem function can exist on a coral reef due to a relatively small number of species being present (and the identify of these species contributing to this particular function varies across geographies). The paper also contains an important applied message: that maintaining high levels of species richness across coral reef regions is essential to sustain global ecosystem function on coral reefs. The paper is well written, although I think some statements could be clarified in places (see minor comments below).

*Thank you, we are delighted that you enjoyed reading our paper. We clarified the text as specified below.*

My only major comment is the need to better justify the five key ecosystem functions that you focus on for the reader. I understand these are taken from the Brandl et al. paper, but many readers may not realise. Also, why 5 and not the 8 presented by Brandl et al.? This needs to come early on so the reader feels comfortable with your focused choice – your paper hinges on this to some extent. I suggest building this out in paragraph three (lines 43-53). You can do this in 2-3 strong sentences.   
*We substantially modified the introduction and added a paragraph on fish-mediated functions specifically. We agree that quantifying all functions proposed by Brandl et al. would have been even better. However, we did not have access to data to do so. For example, CaCO3 production by corals requires on detailed data on coral cover including the community composition and the sizes of corals. We did not have any data on benthic cover for this specific project. We thus only analyzed the functions that are related to fishes based on the data available and the confidence in methodology.*

Minor comments

Lines13-16. Link between these two sentences is not clear for the reader: “Furthermore, functions are locally dominated by few species, but worldwide, 70% of the 1110 species in our dataset contribute disproportionally to functioning in at least one local community. This leads to disparate vulnerabilities of functions to anthropogenic stressors” *We removed the second sentence.*

Lines 18-19 – are you suggesting here that conservation strategies to maximize function will always need to be context dependent – I think you are. If so, I would state this more clearly, as this is an important message from the data. *We now say: “Our results underline the need for a nuanced, local, and context-dependent approach to coral reef conservation that considers multiple functions beyond the effect of standing stock biomass.”*

Line 32-33. Not just efficient recycling of nutrients that’s key to their persistence in oligotrophic waters, it’s also about the feedbacks that raise primary production levels around coral reefs (Gove et al. 2016 Nature Comms 7: 10581) and the flexible trophic ecology of corals to exploit these resources (Fox et al. 2018 Current Biology 28: 3355-3363). *Agreed. However, we now removed this sentence following the suggestion of reviewer 1. MAYBE I SHOULD INCLUDE THIS??*

Line 37. Suggest adding Williams & Graham (2019) Funct Ecol 33: 942-947 to references here also as this editorial captures the messages from a suite of studies talking about functional changes to coral reefs in the Anthropocene. *Done.*

Line 41 – “but see 18”. Please tell us what Allgeier et al. found – annoying for reader to have to go and read another paper to understand this statement.   
*Done. We now state this explicitly.*

Line 41-42. How exactly is this hindering effective management of reefs? You leave the reader guessing here – poor style.   
*We now say: “This sparsity of quantification of functions constitutes a severe limitation to efficient management of coral reef ecosystem functioning.” I AM NOT SURE THAT THIS SUFFICES. ANY SUGGESTIONS?*

Line 69. Suggest avoid the term “holistic”, nothing can ever truly be holistic, without mentioning the connotations of unscientific holistic medicines.   
*We removed the word ‘holistic’.*

Line 78. “and/or” is poor style. You can just write “and”. If a mechanic was advertising that their garage fixes exhausts “and” steering systems, you could take your car there to have both fixed, or just one….they are not implying it has to be both at the same time!   
*Good point, we changed it to ‘and’.*

Line 93. I think this is meant to read “consistently hinge on”. *Modified as suggested.*

Line 107. Contributions to what? *Contributions to functions. We removed this specific sentence though.*

Line 108-109. The Hughes et al. reference is not appropriate here as it does not address “reef fish”. *Agreed. We removed the reference here.*

Line 119-120. Why does a shift in algal symbiont dominance matter? I think you need to make it clearer to reader here why this (and algal growth) are considered bad (if that is what you are implying).   
*We simplified this section and now say: … so a shift to herbivore dominance and the subsequent decline of community-level P excretion may change the balance of nutrient cycling on coral reefs, potentially favoring algal growth in comparison with corals.*

Line 123. Same comment about use of “holistic”. *We removed the word holistic.*

Line 125. Poor style to start a new paragraph with “similarly”. Too much connection/reliance on previous sentence that sits in a different paragraph. *We removed this entire paragraph.*

Line 132-133. Can impact smaller browsers also, see Heenan et al. (2016) Proc B 283: 1843 (Fig. 3). *Same.*

Line 164. Suggest replace “through” with “using”. *Done as suggested.*

Figure 1. Why is there no key for the bubble values? The size differences between bubbles is also not that clear – can this be emphasized? Also not clear what the “categorical assignments” are from the figure legend alone – please expand the explanation. *We added extra details in the legend.*

Figures 2-4. Don’t take this the wrong way, but I find the choice of primary colours in your figures rather off-putting (in particular Fig 4). Also, the choice of some color combinations could be more inclusive (e.g. your use of red/green to denote different groups). There are plenty of color-blind friendly palettes out there now.   
*The color palette was created based on the natural colors of a coral reef fish species: Callanthias australis. We regret that the reviewer found the colors off-putting. We tested the color combination by simulating color-blindness using the website* [*https://coolors.co/*](https://coolors.co/) *and the red green combination seem to be distinguishable due to differences in clarity. Moreover, as functions are kept in the same order throughout the paper, the results could be interpreted even if there were no colors at all. Therefore, we decided to keep the colors for now, but if this is a big issue for the reviewer or the editor agrees with the reviewer, we will change the color palette.*

Figure 2 seems unnecessarily large – this could be a single column figure. *Agreed. We decreased the size.*

Figure 4 – Suggest proportion values radiating out run to 1 so piscivory falls within the axis bounds.   
*This is a great point, thanks for noticing. However, following a recommendation of reviewer 2, we now actually removed this figure.*

All Figures – there are some inconsistencies in font use/size across your figures. This might seem picky, but important for consistency. *We standardized sizes and fonts as much as possible.*

Did you synthesize this data set yourself or did you use an existing database that has been previously published from? Please make this clear in the main manuscript.   
*We used an existing underwater visual census database that has been previously published from. We collected extra data ourselves to parametrize bioenergetic models. We stated this more clearly now.*