nature portfolio

Peer Review File

Unravelling large-scale patterns and drivers of biodiversity in dry rivers

REVIEWER COMMENTS

Reviewer #1 (Remarks to the Author):

Dear Editor, dear Authors,

The manuscript NCOMMS-23-36666-T "Unravelling large-scale patterns and drivers of biodiversity in dry rivers" is an interesting study assessing dry-sediment biodiversity patterns in non-perennial rivers worldwide. The authors did a great job, first being able to collect such a valuable dataset with global data, and then focusing on a highly relevant but still understudied topic. The manuscript is in general well-written and easy to follow, with only a few information missing or sections that need to be adjusted. Below I provide some general and detailed comments. I hope that these comments can help to further improve your manuscript.

Abstract. You provide a (quite impressive) spatial framework, but a temporal context would also be good here. When did the sample collection happen? Over how many months/years?

Introduction. In general, the introduction is well-written and balanced, providing the readers with the right amount of background information.

Lines 174-176. I suggest adding an extra sentence here, stating that the ratio of NPRs will increase in the future. This will highlight even more the importance of your work.

Lines 182-183. "…due to spatial co-occurrence of multiple successional stages after flow resumes". I agree but also suggest including other factors among the causes of beta-diversity increase. For example, you can have a temporal mismatching among dry and wet patches, leading to the presence of different communities even before the rewetting. Also, in case you have local patches in the river networks that are more prone to drying, you could have a more dry-adapted local community.

Line 201. Here you have quite a strong jump from bacterial communities to conservation/restoration actions. I suggest adding a sentence at the beginning of this paragraph to better introduce the topic. Another good option would be to exchange the first and the second sentence of this paragraph.

Lines 210-220. The clarity of this section can be improved. I suggest clearly stating your specific aims and/or questions and then providing matching hypotheses and/or expectations. At the moment, your HPs are not sufficiently elaborated. For example, at line

212 you mention that you want to measure both alpha and beta-diversity, but then you don't have any HP about beta-diversity. Also, in your HPs, you have a statement about abiotic drivers (i.e. resource availability), but biotic factors are not mentioned.

Results. I suggest adding a short paragraph describing the variability of the environmental parameters. You have a huge dataset, so it would be interesting to have an estimation of the environmental variability. As an alternative, also a table (in the main text or in the supplementary) reporting the mean and SD values of the environmental variables for each sampling site would be good. Besides this, results are properly explained, and with enough details.

Discussion. This section is well-framed and supported by the results, with only few parts that need further elaboration (see comments below).

Lines 417-420. I suggest being a bit more cautious with this statement. This weak correlation could be also due to unmeasured variables and/or convergence in niche selection or evolutionary or biogeographic processes. I suggest adding other possible explanations in this paragraph. Also, the fact that all the samples were collected at the same successional stage (i.e. during the dry phase) could affect your findings.

A temporal perspective is missing in the discussion. For example, if possible, it would be interesting to know if the sampling sites are naturally intermittent or if they became intermittent in recent years (and in this case for how long). When this information is not available, this point should be discussed in the limitations of the study. For example, you could mention that communities from naturally intermittent sites (or sites experiencing dry-outs for decades) might be more adapted to these harsh conditions.

A general conclusion, including the future implications of this study and next steps, is also missing, besides a quick mention of this at lines 452-455. I suggest moving this to a separate section and further developing it.

Material and methods. This section is well explained and the methods are suitable for the study design. However, some bits of information missing (see below).

Line 483. Did you try to assess how much precipitation and mean annual temperature changed in the time period 2000-2015/16? This change could be considerable and affect NPRs communities. This bit of information would be interesting, but I also know this task can be really challenging for such an extensive study area, so you can ignore this comment in case it's not feasible.

Lines 489-490. Why did you select a different time period for this? Wouldn't it have been better to be consistent and have again the period 1970-2000?
Line 497. “Land cover was derived using GIS”. Here there are some bits of information missing. For example, you could provide the source of the information (e.g. CORINE data?), and explain at which spatial scale this was calculated (e.g. in the whole basin? In the portion of basin upstream from the sampling site? In a radius of n km around the sampling stations?).
Lines 572-573. I guess here with taxa you mean the previously specified groups Archaea, Bacteria, Fungi, Algae, Protozoa, Nematoda, Arthropoda, and Streptophyta. Please confirm or specify otherwise.
Line 580. It’s not completely clear if you used a custom backward procedure or a built-in function from an R package. From what you wrote looks like a custom procedure, based on the partial dependence plots, but please, state this clearly at the beginning of the sentence. Also, I guess you repeated the whole procedure for each single taxon. Quite impressive, good job.
Lines 605-608. I did not fully understand why you decided to exclude the nestedness component from your analysis. This also provides insights into mechanisms underlying assembly processes, for example, loss of sensitive taxa in some habitat patches.

Reviewer #2 (Remarks to the Author):
This paper presents results of a large-scale (84 non-perennial rivers across 19 countries on five continents) collaborative/community study). The focus on intermittent/non-perennial rivers is timely given recent papers in other Nature journals that identified these types of watercourses form significant parts of river networks worldwide. There is a clear research gap being addressed here in terms of the biodiversity of these systems. Samples were collected from dry river beds, then multiple taxonomic groups identified using molecular methods from bacteria to various invertebrates.

However, the storyline of the manuscript seems to me to be quite inconsistent at present, largely due to unbalanced structure and the lack of a clear set of aims/hypotheses that track through into the results. For example, the paper seems like it should be more strongly

rooted in the metacommunity processes literature with testing of these theories all the way through the remainder of the paper – this is something I expected given the start of the second paragraph of the introduction, but thereafter (from line 180) the major gap presented is not unpacked any further and the text abruptly switches to multiple other ideas (colonisation, biogeochemistry and biological trait effects). The aims and hypotheses from around line 210 further suggest a lack of a clear focus and structure – these important elements somehow should tee up the results/discussion structure, but at present it is difficult to see how they relate to the two sections of the results, or to the three sections of the discussion.

After the introduction, the first part of the results comes across as quite descriptive for the most part. While this cataloguing work is undoubtedly valuable for many reasons, it could be pared back with more reliance given to figures 2/3, opening space for more focus of the results towards the scientific processes questions. Currently the section starting Line 223 reads mainly as too much of a simple 'list' of which groups were found, how many of them are there, and then how they correlate with each other. If the introduction had given some context as to the importance of the various groups of bacteria/fungi etc in rivers, then the reader would be able to make more sense of why it is important that group x is more abundant or richer than group y.

Around line 250 there seems to be a switch in focus of the results from describing the taxonomy towards consideration of biodiversity-environment relationships. This seems to mark the start of some of the metacommunity process results which were discussed earlier, and as such needs some signposting with a new heading. Various correlative statistics are then used to suggest a significant role for environmental variables in controlling the biodiversity. For example Line 254-257 talks about OTU richness mainly controlled by… and bacteria mainly influenced by…. And latitude was a common driver… But all we can infer from this field survey study are correlations/associations, not controls as these variables have not been manipulated to observe responses. Some careful rewording is needed through this section so as not to conflate correlation and causation.

After line 250 around 3 paragraphs are devoted to density-independent (environment)

relationships, then the next main section (line 298) deals with the density-dependent (biotic) metacommunity processes but across only one paragraph of text. The importance of dispersal is considered only minimally (spatial distances) within this section. A more balanced consideration of the three processes would improve the paper's storyline.

Earlier in the manuscript, line 205 onwards introduces the manuscript's focus on dry phase sampling only. I can see the value in sampling during this time, but there needs to be more justification of why only this phase of the 'flow' regime was sampled. Surely a more complete understanding of these river systems would need to also include some assessment of the biodiversity during other phases such as active flow, ponded water, damp sediments, dry sediments? As a minimum, some discussion of how much we can learn about the processes structuring these rivers if we focus only on one flow phase should be included somewhere.

Line 313-314 – given my concerns above about the cataloguing being a little too prominent, perhaps some more direct data comparisons of results against the studies cited here would elevate that part of paper to provide the critical evaluation needed for the reader to truly appreciate the importance or otherwise of these new results from dry rivers.

Line 147 – specifically a lack of biological data? or all types of data? I suggest stating the former given the research gap that is outlined in line 150

Line 150 – a co-ordinated 'survey' is a better description than an 'experiment' as nothing was manipulated.

Methods: 470-478: some details about how this work was undertaken in sterile working environments is missing to rule out contamination, particularly for the microbial groups that were identified.

477-478: Some notes here about whether there may have been some sample degradation from collection to final storage would be useful to know about. I would have expected to see some mention of preservation of these sub-samples much earlier in the work flow,

perhaps in the field using either chemicals or freezing.

Metabarcoding methods: Surprising to see fungi diversity analysed using 18S – my understanding is that ITS is a better region to target for this group (one e.g. https://bsssjournals.onlinelibrary.wiley.com/doi/full/10.1111/ejss.13329)

Reviewer #3 (Remarks to the Author):

Dear Authors,

I have now read your MS entitled "Unravelling large-scale patterns and drivers of biodiversity in dry rivers".

Overall I appreciated reading your MS as it provides a large-scale overview of biodiversity across multiple taxa in NPRs, which is not common. I acknowledge the strength of the survey and of the data, as well as the originality of the work. Nonetheless, I have some major critics that I will detail below, and the descriptive nature of the work make me skeptical regarding some conclusions and the generality of the work. On the one side we need these types of large-scale descriptive studies, but on the other side it is always frustrating that underlying mechanisms (drivers) are poorly revealed. Herebelow I detail some comments that may perhaps help improving conclusions regarding drivers of observed patterns.

1) Abstract and Introduction are well-written and clear, but I think authors should be more transparent regarding previous knowledge on NPRs. Authors argue that data are lacking for NPRs (fist words of the Abstract + many statements like that in the Introduction), which is not totally true. Even though NPRs are less studied than perennial rivers, knowledge have been acquired in these ecosystems, including per several authors of the MS. We actually have some knowledge regarding freshwater organisms (e.g., insects), their biodiversity patterns and underlying drivers, but this is clearly less the case for microbial communities of NPR riverbeds (what is adressing this MS). For freshwater macro-organisms, we know that dispersal process for community assembly in NPRs is important, but we may ask if this is also the case for microbial soil communities in these ecosystems. I think that authors should

be much more transparent in what we know in NPR biodiversity to better frame the actual lack of knowledge and to built coherent specific hypotheses. There is a very general hypothesis in the last sentence of the Introduction whereas more specific hypotheses could be stated based on previous knowledge (that somewhat exist).

2) I also would like authors to be more precise in what they call biodiversity; here authors focus mostly on microbial biodiversity and this should be made more explicit throughout the introduction, as hypotheses and expectation are not the same as when considering macrofauna biodiversity (from insects to birds for instance). You are considering only some facets of biodiversity and I encourage authors to be more transparent by avoiding general wording (e.g., l. 210 "Here, we provide the first global-scale attempt to assess NRP biodiversity" should read ""Here, we provide the first global-scale attempt to assess NRP soil microbiota biodiversity" or something like that).

3) Also, please avoid in the Introduction paragraphs such as the one l. 201-209. I'm really skeptical that a large-scale correlative study as the one you conducted is "crucial to inform conservation and restoration of freshwater biodiversity facing global change...". If you really think that macro-ecological studies as the one you conducted is crucial for conservation, please explain clearly why and how. My feeling is that you have generated interesting and novel fundamental knowledge, and this already a lot (and enough). I'm not sure that this type of arguments (associated with conservation issues) is required to justify your study. You know that microbial diversity is driven by climate and nutrient, but does this knowledge help resolving conservation issues? If you think this will really help, please explain how in the Discussion. Otherwise I would propose deleting this paragraph (replace it by a better knowledge background in NPRs).

4) Regarding patterns of soil biodiversity, I missed a general and visual approach for describing patterns. here you mostly focus on alpha diversity, which is fine. But you may have plot in Figure 1 the distribution of OTU richness for the most abundant groups, at least for readers to visually see whether this generate clustered patterns of evenly spread patterns of alpha diversity (in other words I suggest you to map biodiversity patterns). Also, I would have like to see a graph and/or analysis testing whether the distribution of the

major groups is similar across the globe or not. This is more or less what you have done in the beta-diversity analysis, but again this is vague and not visual. I would advice you to run some types of multivariate analyses (NMDS, ddRDA, clustering trees) for instance to search for the existence of "clusters" (group of sites that are similar in term of taxa composition) and test whether some environmental variables are associated with this clusters. I would be curious to see (i) if some clusters can be found, (ii) if they are geographically coherent (or not) and (ii) if their are associated with some environmental factors.

5) the randomForest appraoch you used for revealing "drivers" is statistically "fine", but I missed some details and I'm not sure it is necesarilly the best approach. First, there is no mention of collinearity. I'm pretty sure that some of your environmental factors are strongly associated one with each other, which as you know can be a problem for statistical inferences. Please, add a correlation matrix so that readers can judge potential problems associated with collinearity. RandomForest are supposed to be good for dealing with collinearity, but -assuming strong collinearity among some of your variables- I would like to see an complementary analysis in which you would have reduced collinearity, for instance by focusing on PCA axes (for instance, it is highly likely that climate can be reduced by one or two orthogonal PCA axes, the same for environmental variables). I know this is frustrating to reduce dimensionality, but this can also be strongly informative. Second, you choose random Forest to reveal drivers, which is not in my opinion the best option. I know RF haves many advantages, including generating stronger r2 than traditional approaches, but this is a predictive approach, not a mechanistic approach. Given that your goal is to reveal drivers, I would have used a more mechanistic approach to reach this objective. In particular, I think that parametric approaches (GLMs of structural equation models) would have been much more powerful than RF to really reveal likely mechanisms. I would have first reduce dimensionality (or generate latent variables) and then use classical linear models (including quadratic terms if needed). Linear models have the strong advantage to provide parameters, that can be then used to compare more easily drivers among taxa. For instance, parameters (slope coefficients) can be used to test the global effect size (ES) of a factor (e.g., climate) on biodiversity patterns: is the ES for climate significantly different from zero across all taxa? is the global ES for climate higher than the ES for nutrients across all taxa?. Using parametric appraoches allow for testing quantitatively/statistically these hypotheses, which

non-parametric approaches can not do. In addition, you can test statistically whether the effect of a factor (e.g., climate) is homogeneous among taxa or not (using heterogeneity tests on effect sizes). Finally, if you combine linear models to causal approaches, you can test for direct and indirect effects: for instance climate may be related to nutrients availability (the latter being associated with OTU richness) and to OTU richness directly. This can be modeled quite easily as soons as you reduce dimensionality and you use parametric approaches. I think such type of mechanistic approach would be much more informative than a predictive and "black box" approach as the one you used, even if they produce lower r2 (and if for some taxa there is not significant drivers, this is not a problem, it is also an important finding).

6) I may be wrong but you did not included in the models the date at which you sampled the soil since drying. I may be wrong but I guess that soil biodiversity evolve along the drying process: the community composition is likely not the same some days after the drying occur than some moths after the drying occur. Could you please include this parameter in the models?

7)The lasso approach is nice but I'm a bit skeptical regarding the way you included environmental and climatic distances. From what I understood, you estimated the distances based on all factors at a time. If yes, I'm afraid that the resulting distances may be uninformative. For instance, for environmental variables you included all 11 variables, it is likely that you loose all relevant information. For instance the C content is likely more important than the sediment particles mean size to explain beta diversity (as the former is more associated with OTU richness than the latter). So I would rather used distances based on each variable independently, by focusing for instance on the ones that are more associated with OTU richness (Table 1). This lack of precision may obfuscate some patterns (or reinforce your findings if the abiotic factors are still unimportant)

8) The discussion is nice overall but you did not discuss an important finding that finally the variance in richness along the gradient is mostly unexplained by the abiotic factors you selected. This needs to be discussed; either you missed important variables, or community are mostly driven by biotic interactions (as suggested by the lasso analysis)

Reviewer #4 (Remarks to the Author):

This manuscript presents a research work that is based on a powerful data set about a kind of natural systems of relevance in the context of climate change. The knowledge of non-perennial rivers (NPRs) biodiversity and its relationship with environmental drivers is a matter of great significance for understanding the behaviour of these ecosystems and predicting their evolution in this changing context that probably will benefit them in front of the better known perennial rivers.

However, the main pitfall of these data is that the group of samples is clearly dominated by those coming from temperate latitudes and climates, with a low representation of arid/semiarid environments and very few samples from other climate classes. The authors point out this limitation at the end of the discussion (lines 436 to 440), but up to here they always refer to this work as a global-scale assessment. In my opinion, this bias should be taken in account and become more explicit through the whole manuscript. Indeed, from a practical point of view, making more visible this limitation will not discredit the main statements or the relevance of this research, because temperate and semiarid latitudes contain a high amount of NPRs and the expected efects of global warming on rivers will be specially intense in these regions.

Methods and obtained results of biodiversity assessment look appropriated for the purposes of this research, although their significance and the validity of data interpretations are inevitably constrained by the use of OTUs richness as biodiversity indicator and the high level of taxonomic units selected for the analysis and comparison of taxa composition in the community assemblages, as usually happens in this kind of metabarcoding studies. Statistical analyses are mostly out of my scope of expertise, but they seem to be robust and give meaningful information. I am not familiar with methods used for other kind of analyses, such as those for chemical or climatic parameters.

Conclusions emerge as logical inferences after the discussion of results. They must be interpreted using additional informations and hypotheses, which are well-provided by an adequate pool of references. In spite of everything, several generalizations, assumptions or especulations had to be done, but this extent has been taken in account and the language

used in the discussion reflects this fact when it is necessary. Literature is also appropriately used to provide context to the results when it is convenient, as well as to the purposes of the research in the introduction.

Unfortunately, the concluding remarks in the last paragraph (lines 452 to 455) ended up as if they were included in the section named “potential limitations” of the discussion. Although the end of the discussion is the most logical position for these statements, in my opinion the consistence of the story sequence would be improved if they were clearly placed out of this section. Actually, considering the usefulness of knowing these potential limitations before reading the other sections of the discussion (in a similar way that my previous comment about the bias of the samples), probably this section should better be placed at the beginning of the discussion, leaving the concluding remarks at the end.

According to the relevance of this research in the context of climate change, I would expect to find in the discussion some inferences of these results focused on the prediction of the consequences of warming and altered hydrological patterns in these community assemblages and ecological systems. How could they be modified by changes in abiotic factors? Which effects could have the strong biotic interactions found in this work on the biodiversity responses to these changes? Probably more assumptions should be required for this kind of reasonings, but I think that it would be worthy due to the relevance of the results obtained as well as to the expected role of NPRs in temperate regions in a near future. Anyway, it is only a suggestion and the lack of these considerations should not be an objection for its publication.

REVIEWER COMMENTS

Reviewer #1 (Remarks to the Author):

Dear Editor, dear Authors,
The manuscript NCOMMS-23-36666-T "Unravelling large-scale patterns and drivers of biodiversity in dry rivers" is an interesting study assessing dry-sediment biodiversity patterns in non-perennial rivers worldwide. The authors did a great job, first being able to collect such a valuable dataset with global data, and then focusing on a highly relevant but still understudied topic. The manuscript is in general well-written and easy to follow, with only a few information missing or sections that need to be adjusted. Below I provide some general and detailed comments. I hope that these comments can help to further improve your manuscript.

1) Abstract. You provide a (quite impressive) spatial framework, but a temporal context would also be good here. When did the sample collection happen? Over how many months/years?

Samples were collected between April 2015 and March 2016 to allow participants from both hemispheres to collect sediments during the dry season. This is now indicated in the abstract (L154).

"Dry sediments were collected over a one-year period from 84 non-perennial rivers across 19 countries on five continents, and biodiversity was estimated using environmental DNA metabarcoding"

2) Introduction. In general, the introduction is well-written and balanced, providing the readers with the right amount of background information.
Lines 174-176. I suggest adding an extra sentence here, stating that the ratio of NPRs will increase in the future. This will highlight even more the importance of your work.

We have added the suggested sentence (L169-171):

"NPRs, which naturally experience dry phases, constitute more than half of the global river network length[7,8] and are increasing in extent due to climate change and escalating demands for freshwater[9]."

3) Lines 182-183. "…due to spatial co-occurrence of multiple successional stages after flow resumes". I agree but also suggest including other factors among the causes of beta-diversity increase. For example, you can have a temporal mismatching among dry and wet patches, leading to the presence of different communities even before the rewetting. Also, in case you have local patches in the river networks that are more prone to drying, you could have a more dry-adapted local community.

We have modified the sentence to also mention the co-occurrence of different habitats in both space and time (L176-179):

"Extended dry phases generally reduce local diversity (i.e. alpha-diversity) of aquatic taxa (e.g., invertebrates, fishes) but increase compositional changes between communities (i.e. beta-diversity), due to spatial co-occurrence of communities in dry and wet habitats, and species turnover during wet–dry cycles[11]*."*

4) Line 201. Here you have quite a strong jump from bacterial communities to conservation/restoration actions. I suggest adding a sentence at the beginning of this paragraph to better introduce the topic. Another good option would be to exchange the first and the second sentence of this paragraph.

In view of this comment and several others by Reviewer 3, we have removed the reference to restoration and conservation strategies. Instead, we now emphasize the need for a better understanding of community assembly rules to predict responses of biodiversity and ecosystem functioning to global change (L172-174):

"Understanding biodiversity patterns and the underlying assembly rules that govern community establishment during dry phases is crucial to anticipate the response of NPR biodiversity and ecosystem functioning to global change, including changing patterns of river drying."

5) Lines 210-220. The clarity of this section can be improved. I suggest clearly stating your specific aims and/or questions and then providing matching hypotheses and/or expectations. At the moment, your HPs are not sufficiently elaborated. For example, at line 212 you mention that you want to measure both alpha and beta-diversity, but then you don't have any HP about beta-diversity. Also, in your HPs, you have a statement about abiotic drivers (i.e. resource availability), but biotic factors are not mentioned.

Given comment and similar comments by Reviewers 2 and 3, we have revised this section to clarify our aims and hypotheses. We now also outline current knowledge on the relative strength of dispersal, environmental filtering and biotic interactions on dry riverbed communities in previous sections (L185-210), to better introduce our specific aim and hypotheses. Furthermore, we have included a second hypothesis on the potential mechanisms (including biotic interactions) controlling beta-diversity in dry riverbeds (L226-228).

"*Our aim was to characterize the alpha- and beta-diversity of dry-riverbed communities and to identify their major potential drivers. We analyzed the diversity of eight major taxa (Algae, Archaea, Bacteria, Fungi, Protozoa, Arthropoda, Nematoda and Streptophyta), and estimated the relative abundances of copiotrophic and oligotrophic microorganisms. We then used a random forest approach and correlation networks inferred from graphical models to assess the relative contribution of potential abiotic and biotic drivers of biodiversity patterns. Our first hypothesis was that the availability of key resources such as organic carbon, inorganic nutrients and water is positively related to the taxonomic richness (i.e. alpha-diversity) of dry-riverbed communities and that constraints imposed by limited resources promote microorganisms adapted to coping with the harsh conditions in dry sediments. Our second hypothesis was that spatial community turnover (i.e. beta-diversity) in dry riverbeds is primarily controlled by environmental filtering and to a lesser extent by biotic interactions.*"

6) Results. I suggest adding a short paragraph describing the variability of the environmental parameters. You have a huge dataset, so it would be interesting to have an estimation of the environmental variability. As an alternative, also a table (in the main text or in the supplementary) reporting the mean and SD values of the environmental variables for each sampling site would be good. Besides this, results are properly explained, and with enough details.

We have added summary statistics of climatic, physicochemical and land-use variables (Supplementary Table S1) and a brief description of variability in key environmental parameters, at the beginning of the Results section (L232-238).

"*The 84 sediment samples were collected in regions spanning the main five Köppen climate classes, with most collected in temperate and arid regions (A: tropical n=2, B: arid: n=14, C: temperate n=66, D: continental n=1, E: polar n=1) (Fig. 1). Our sampling design covered a wide gradient of physicochemical conditions and land use (Table S1) and included river channels up to 11 m wide (mean ± sd = 3.5 ± 2.4 m), in many cases in riparian forest (64 ± 33%). Sediment organic carbon content ranged from 0.1–8% (1.3 ± 1.7%) (Table S1). The duration of the dry phase preceding sample collection ranged from 7–800 days."*

Discussion. This section is well-framed and supported by the results, with only few parts that need further elaboration (see comments below).

7) Lines 417-420. I suggest being a bit more cautious with this statement. This weak correlation could be also due to unmeasured variables and/or convergence in niche selection or evolutionary or biogeographic processes. I suggest adding other possible explanations in this paragraph. Also, the fact that all the samples were collected at the same successional stage (i.e. during the dry phase) could affect your findings.

We mentioned in the original manuscript that when interpreting the results, we assumed that the important environmental variables were included in our analyses. ("Strong linkages occurred among the beta-diversity of Bacteria, Fungi, Algae, Protozoa and Nematoda, which, assuming adequate representation of the major environmental variables, suggest that biotic interactions might be more important to the assembly of dry riverbed communities than dispersal or environmental filters.") We agree that other possible explanations should be considered and now discuss them in greater detail (L429-444).

"The observed limited influence of dispersal on the assembly of dry-riverbed communities should be interpreted with caution, because our sites spanned many river networks and our Euclidean measure of spatial distances did not account for hydrological connectivity among sites. Hydrological connectivity is likely to constrain the movement of aquatic taxa and could thus have led to underestimation of the role of dispersal[22]. Given the low explanatory power of the variables included in the random forest models of OTU richness for some taxa, the weak correlations between environmental variables and the beta-diversity of most taxa also suggest a limited influence of environmental filtering on community assembly in dry riverbeds. This result is surprising, because the analyzed environmental variables (e.g. organic carbon content, pH, climate) are known drivers of soil and freshwater microbial diversity[48,62,63]. Dormant or dead aquatic organisms detected in dry sediments by eDNA

metabarcoding may have partially obscured biodiversity–environment relationships. In addition, samples were collected at sites with different drying histories. At sites that have experienced seasonal dry phases over evolutionary timescales, some aquatic species have evolved specific adaptations to tolerate dry phases. In contrast, at newly intermittent sites communities may be prone to higher biodiversity loss in response to unpredictable drying events[6]."

8) A temporal perspective is missing in the discussion. For example, if possible, it would be interesting to know if the sampling sites are naturally intermittent or if they became intermittent in recent years (and in this case for how long). When this information is not available, this point should be discussed in the limitations of the study. For example, you could mention that communities from naturally intermittent sites (or sites experiencing dry-outs for decades) might be more adapted to these harsh conditions.

Unfortunately, we do not have information about the history of intermittency at all study sites, although most were naturally intermittent. We now acknowledge this issue when discussing the unexpectedly small influence of environmental variables on diversity patterns (see our response to the comment above and L440-444).

9) A general conclusion, including the future implications of this study and next steps, is also missing, besides a quick mention of this at lines 452-455. I suggest moving this to a separate section and further developing it.

The conclusion has been further developed with a new, titled *Conclusions* section (L466-476) to include the implications of our study, as suggested:

*"**Conclusions***
As temperatures and aridity increase due to ongoing climate change in many global regions, NPRs are becoming increasingly prevalent in both space and time. Our results suggest that biodiversity of most major taxa in dry riverbeds is likely to be increasingly affected by increasing climatic extremity, especially under conditions of low resource availability. Management strategies are thus needed to support the biodiversity and thus ecosystem functions and services of NPRs affected by diminishing flows[69]. In addition, the tight linkages we observed between the beta-diversity of multiple major taxa in dry riverbeds suggest that recognizing biotic interactions in such strategies could maximize the effectiveness of actions taken to protect biodiverse NPR communities as they adapt to longer, more frequent and higher-intensity dry phases."

10) Material and methods. This section is well explained and the methods are suitable for the study design. However, some bits of information missing (see below).
Line 483. Did you try to assess how much precipitation and mean annual temperature changed in the time period 2000-2015/16? This change could be considerable and affect NPRs communities. This bit of information would be interesting, but I also know this task can be really challenging for such an extensive study area, so you can ignore this comment in case it's not feasible.

Assessing temporal changes in precipitation and mean annual temperature is beyond the scope of our study. As acknowledged by the reviewer, this would indeed be a huge undertaking for

such a large study area. Therefore, although such an analysis could no doubt produce interesting results, we have not revised the manuscript as suggested.

11) Lines 489-490. Why did you select a different time period for this? Wouldn't it have been better to be consistent and have again the period 1970-2000?

Mean annual temperature and precipitation were accessed from the WorldClim database, which provides data for the period 1970-2000. Potential evapotranspiration and the aridity index were accessed from the Global Potential Evapotranspiration (Global-PET) and Global Aridity Index (Global-Aridity) (version 1) datasets provided by the CGIAR Consortium for Spatial Information (CGIAR-CSI). The difference in time period is due to the fact that Global-PET and Global-Aridity were each provided in the form of a single grid layer representing the annual average over the period 1950-2000, rather than for the period 1970-2000.

Following a suggestion by Reviewer 3 (see comment #5), we removed PET and Aridity from our statistical analyses. Indeed, although random forests are robust to collinearity, we reran the analyses after excluding the most highly correlated variables: as Aridity and PET are, virtually by definition, highly correlated with precipitation and temperature, respectively, both have been excluded in the statistical analyses from the revised version of the manuscript. This does not alter any of the main results or interpretations.

12) Line 497. “Land cover was derived using GIS”. Here there are some bits of information missing. For example, you could provide the source of the information (e.g. CORINE data?), and explain at which spatial scale this was calculated (e.g. in the whole basin? In the portion of basin upstream from the sampling site? In a radius of n km around the sampling stations?).

It is now clarified in the manuscript (L513-516) that: “*The proportion of the catchment area upstream of the sampling site covered by forest (%) was determined by each participant using local knowledge or GIS-based information derived from the most up-to-date national land cover maps for each country*.”

13) Lines 572-573. I guess here with taxa you mean the previously specified groups Archaea, Bacteria, Fungi, Algae, Protozoa, Nematoda, Arthropoda, and Streptophyta. Please confirm or specify otherwise.

Yes, we confirm that we mean the previously specified groups when using the expression« selected taxa ». To avoid ambiguity, we replaced « selected taxa » by the « eight taxa » wherever necessary throughout the manuscript.

14) Line 580. It’s not completely clear if you used a custom backward procedure or a built-in function from an R package. From what you wrote looks like a custom procedure, based on the partial dependence plots, but please, state this clearly at the beginning of the sentence. Also, I guess you repeated the whole procedure for each single taxon. Quite impressive, good job.

Thanks. Yes, indeed, we repeated the whole procedure for each single taxon. This was done iteratively. The procedure is based on the Recursive Feature Elimination algorithm

(Gregorutti et al. 2017) consisting in iteratively training the model for each taxon, inspecting the variable importance plot and then removing the lowest rank variable until reaching the best model (based on the % variance explained). We now indicate this in L604-607.

"*...we selected variables with the recursive feature elimination algorithm[87] by iteratively training the model for each taxon, inspecting the variable importance plot and then removing the lowest ranked variable until reaching the best model, as indicated by the % variance explained.*"

15) Lines 605-608. I did not fully understand why you decided to exclude the nestedness component from your analysis. This also provides insights into mechanisms underlying assembly processes, for example, loss of sensitive taxa in some habitat patches.

We decided to present only the turnover component because of the consistently low (i.e. 1-14%) contribution of the nestedness component to beta-diversity of the eight selected taxa. We now indicate this in L629-631.

"Analyses were only conducted on the true turnover component of the Bray-Curtis index, because the contribution of the nestedness component to beta-diversity of the eight taxa was consistently low (1–14%)."

Reviewer #2 (Remarks to the Author):

This paper presents results of a large-scale (84 non-perennial rivers across 19 countries on five continents) collaborative/community study). The focus on intermittent/non-perennial rivers is timely given recent papers in other Nature journals that identified these types of watercourses form significant parts of river networks worldwide. There is a clear research gap being addressed here in terms of the biodiversity of these systems. Samples were collected from dry river beds, then multiple taxonomic groups identified using molecular methods from bacteria to various invertebrates.

1) However, the storyline of the manuscript seems to me to be quite inconsistent at present, largely due to unbalanced structure and the lack of a clear set of aims/hypotheses that track through into the results. For example, the paper seems like it should be more strongly rooted in the metacommunity processes literature with testing of these theories all the way through the remainder of the paper – this is something I expected given the start of the second paragraph of the introduction, but thereafter (from line 180) the major gap presented is not unpacked any further and the text abruptly switches to multiple other ideas (colonisation, biogeochemistry and biological trait effects). The aims and hypotheses from around line 210 further suggest a lack of a clear focus and structure – these important elements somehow should tee up the results/discussion structure, but at present it is difficult to see how they relate to the two sections of the results, or to the three sections of the discussion.

We thoroughly revised the abstract and introduction in the light of these comments and now also state the objective earlier on. Specifically, we first elaborate on existing knowledge regarding the relative strength of dispersal, environmental filtering and biotic interactions on dry riverbed communities (L185-210) before presenting our refined aims and hypotheses (L216-228). The Results/Discussion sections were also re-organized to better reflect this structure.

2) After the introduction, the first part of the results comes across as quite descriptive for the most part. While this cataloguing work is undoubtedly valuable for many reasons, it could be pared back with more reliance given to figures 2/3, opening space for more focus of the results towards the scientific processes questions. Currently the section starting Line 223 reads mainly as too much of a simple 'list' of which groups were found, how many of them are there, and then how they correlate with each other. If the introduction had given some context as to the importance of the various groups of bacteria/fungi etc in rivers, then the reader would be able to make more sense of why it is important that group x is more abundant or richer than group y.

As our study is the first large-scale report of dry riverbed diversity covering a wide range of taxonomic groups, the presentation of alpha- and beta-diversity patterns along with the identification of the potential underlying drivers is arguably an important first step and major strength of the paper. That said, we have significantly shortened this description, which is mainly based on Figures 2 and 3, as suggested (L239-251). This has allowed us in turn to expand on the relative influence of the potential mechanisms (i.e. dispersal, environmental filtering and biotic interactions) controlling community assembly in dry riverbeds.

"eDNA metabarcoding analysis produced 2629 eukaryotic, 2035 bacterial and 82 archaeal operational taxonomic units (OTUs) (Table S2). Actinobacteria were the most common phylum detected with the 16S marker, followed by Proteobacteria and Bacteroidetes (Fig. 2). Within the Proteobacteria, the most abundant classes were Alphaproteobacteria (mean ± sd = 11.5 ± 6.4% of the total number of reads, n=84), Gammaproteobacteria (8.6 ± 9.1%) and Betaproteobacteria (4.6 ± 4.5%). Within Eukaryota (18S marker), the most abundant phyla corresponded to Fungi (49 ± 22%, n=76), with Ascomycota (26 ± 19%) and Basidiomycota (11 ± 13%) being the most common classes (Fig. 3). OTUs for Algae and Protozoa contributed on average 5% and 14% of the reads, respectively. Twelve metazoan phyla were detected, with the highest relative abundances corresponding to Arthropoda (8 ± 12%) followed by Nematoda (3 ± 4%). Sequences assigned to Arthropoda included Collembola (3 ± 7% of the total number of reads), Ostracoda (2 ± 6%), Arachnida (2 ± 3%, corresponding mainly to mites). Around 4% of the reads corresponded to Streptophyta."

3) Around line 250 there seems to be a switch in focus of the results from describing the taxonomy towards consideration of biodiversity-environment relationships. This seems to mark the start of some of the metacommunity process results which were discussed earlier, and as such needs some signposting with a new heading.

We modified the subheadings to clarify the structure of the Results section and better reflect our aims and hypotheses. The subheading "The composition of dry-riverbed communities" was added in L231 to delineate the paragraph presenting the taxonomy and composition of the

communities of dry riverbeds. The paragraph about biodiversity-environment relationships is now integrated under the subheading "Patterns and drivers of alpha-diversity in dry riverbeds" (L253).

4) Various correlative statistics are then used to suggest a significant role for environmental variables in controlling the biodiversity. For example Line 254-257 talks about OTU richness mainly controlled by… and bacteria mainly influenced by…. And latitude was a common driver… But all we can infer from this field survey study are correlations/associations, not controls as these variables have not been manipulated to observe responses. Some careful rewording is needed through this section so as not to conflate correlation and causation.

We screened the manuscript to remove allusions to causal inference and to clarify our aim to identify « potential » drivers of dry riverbed biodiversity, acknowledging that our analyses can only inform about correlations/associations between variables.

L270-272: Dry-phase durations were negatively correlated with the OTU richness of Archaea and Algae, which declined sharply within 100 days of dry-phase onset.

L275-279: Most variables related to resource availability in sediments (e.g. organic carbon content and soluble reactive phosphorus [SRP], but not dissolved inorganic nitrogen [DIN]) were important predictors of the variation in OTU richness for most taxa. Carbon content was positively related to the OTU richness of heterotrophic organisms (e.g. Protozoa and Arthropoda) but was a poor predictor of Algae and Streptophyta richness

L288-290: The OTU richness of most taxa was weakly correlated with variables related to river size (i.e. river width and catchment area).

5) After line 250 around 3 paragraphs are devoted to density-independent (environment) relationships, then the next main section (line 298) deals with the density-dependent (biotic) metacommunity processes but across only one paragraph of text. The importance of dispersal is considered only minimally (spatial distances) within this section. A more balanced consideration of the three processes would improve the paper's storyline.

We have developed the 'biotic' paragraph (L320-325) to include a more detailed description of the relationships between beta-diversity of the eight selected taxa and specific environmental variables as well as spatial distance, as suggested by Reviewer 3 (see comment #7). The revised text reads:

"The spatial distance between sampling sites was positively correlated with the beta-diversity of Bacteria and Streptophyta. Environmental (i.e. climatic, physiochemical and land-use) variables were not generally associated with the beta-diversity of any of the taxa (Fig. 5). However, the beta-diversity of Bacteria was correlated with temperature and to a lesser extent with SRP and river width, and that of Fungi and Archaea was correlated with catchment area and SRP, respectively (Fig. 5)."

6) Earlier in the manuscript, line 205 onwards introduces the manuscript's focus on dry phase sampling only. I can see the value in sampling during this time, but there needs to be more

justification of why only this phase of the 'flow' regime was sampled. Surely a more complete understanding of these river systems would need to also include some assessment of the biodiversity during other phases such as active flow, ponded water, damp sediments, dry sediments? As a minimum, some discussion of how much we can learn about the processes structuring these rivers if we focus only on one flow phase should be included somewhere.

We focused our large-scale analysis on the dry phase of non-perennial rivers, as this represents almost complete *terra incognita*, whereas large-scale studies on biodiversity during the aquatic phase already exist (e.g. Crabot et al. 2021, Ecography, Soria et al. 2017, Oikos). We have, however, revised the Introduction to draw attention to the current state of knowledge on the aquatic phase and to stress the large research gap on communities in dry riverbeds. In addition, in the discussion we added some perspectives on potential further work to address the biodiversity of NPR during dry and wet phases within a standardized framework. (L460-465):

"Methodological differences between morphology-based and molecular approaches, or among molecular analysis protocols, complicate comparison of biodiversity patterns documented by different studies. Harmonizing the molecular approaches and workflows used to characterize communities during dry and wet phases could thus provide a more complete picture of the aquatic–terrestrial biodiversity of NPRs."

7) Line 313-314 – given my concerns above about the cataloguing being a little too prominent, perhaps some more direct data comparisons of results against the studies cited here would elevate that part of paper to provide the critical evaluation needed for the reader to truly appreciate the importance or otherwise of these new results from dry rivers.

As suggested, we have now expanded on the comparison of our results with data from perennial rivers and, to some extent, also from soils:

L382-386: "Freshwater fungal communities are generally dominated by Ascomycota and Chytridiomycota, whereas Basidiomycota are much more common in terrestrial habitats[20,46]. However, although Basidiomycota were particularly abundant in our dry riverbed samples, we found no indication that dry-phase durations affected the ratio of Ascomycota to Basidiomycota."

L391-394: "Our results provide new insight into the biodiversity patterns of a broad range of taxa that have been largely disregarded in NPRs. In particular, Ciliophora and Cercozoa dominated the protozoan community, as reported in soils[48], perennial river sediments[49], temporary ponds and other NPR sediments[50]."

L398-405: "Most identified arthropods were Collembola, Ostracoda and Arachnida (mainly mites), indicating that we captured the diversity of benthic or interstitial meiofauna, whereas macroinvertebrate eDNA was poorly detected by our metabarcoding approach. Sediments can provide a refuge for some aquatic invertebrates during dry phases[53], promoting their detection by eDNA analyses. In contrast, our detected arthropod community was dominated by largely terrestrial taxa such as Collembola and Arachnida, supporting previous reports that colonization of dry riverbeds by terrestrial arthropods from adjacent riparian areas contributes to dry-phase community assembly[29,54]."

8) Line 147 – specifically a lack of biological data? or all types of data? I suggest stating the former given the research gap that is outlined in line 150

We now clarified that we specifically mean a lack of data on biological communities (L146-148):

"More than half of the world's rivers periodically dry up, and the extent of river drying is increasing due to climate change and water resource use. However, we still lack knowledge about the biological communities in dry riverbeds."

9) Line 150 – a co-ordinated 'survey' is a better description than an 'experiment' as nothing was manipulated.

Yes, we now refer to our study as a "large-scale coordinated survey" (L150, L212, L332).

10) Methods: 470-478: some details about how this work was undertaken in sterile working environments is missing to rule out contamination, particularly for the microbial groups that were identified.

We now indicate (L492-494) that precautions were taken, including sterilization of sieves between sites, to avoid cross-contamination between samples. We also added a reference to a large-scale study on soil fungal communities (Tedersoo et al. 2014), which also included air-drying at room temperature before sample storage in zip-lock bags.

"In the laboratory, the sediments were sieved (2 mm) and air-dried for one week (Tedersoo et al. 2014). To minimize the risk of cross-contamination, sieves were sterilized using 10% bleach between processing of samples from different sites."

Tedersoo, L. et al. 2014 Global diversity and geography of soil fungi. Science 346, 1256688.

11) 477-478: Some notes here about whether there may have been some sample degradation from collection to final storage would be useful to know about. I would have expected to see some mention of preservation of these sub-samples much earlier in the work flow, perhaps in the field using either chemicals or freezing.

The simple preservation and storage protocol we adopted was an unavoidable tribute to our goal of covering a number of locations around the world as large as possible, including in remote areas and where effective chemicals for DNA preservation (e.g. Gray et al. 2013) are not readily available. Note also that sample origin has been found to have a greater influence on soil community composition than sample storage conditions prior to DNA extraction (Lauber et al. 2010).

We have added the rationale justifying our choice(L499-503): "*These simple preservation and storages measures were designed to maximize the number of surveyed sites, including sites in remote areas and where effective chemicals for DNA preservation*[72] *are not readily available. Moreover, the potential effects of sample storage conditions on large-scale comparisons of soil communities are small*[73]"

Gray, M. A., Z. A. Pratte, and C. A. Kellog. 2013. Comparison of DNA preservation methods for environmental bacterial community samples. FEMS Microbiology Ecology 83:468–477.

Lauber, C. L., N. Zhou, J. I. Gordon, R. Knight, and N. Fierer. 2010. Effect of storage conditions on the assessment of bacterial community structure in soil and human-associated samples. FEMS Microbiology Letters 307:80–86.

12) Metabarcoding methods: Surprising to see fungi diversity analysed using 18S – my understanding is that ITS is a better region to target for this group (one e.g. https://bsssjournals.onlinelibrary.wiley.com/doi/full/10.1111/ejss.13329)

We have now added the reasons for our choice in the Methods section (L527-531):

"*We opted for 18S primers to meet our main aim of capturing a broad suite of eukaryotic taxa, including Algae, Protozoa, Metazoa and Fungi (Li et al. 2020). Although the internal transcribed spacer (ITS) is the preferred target for Fungi, 18S is also informative (Banos et al. 2018), and both ITS and 18S markers have limitations and introduce biases, particularly in fungal community analyses (Tedersoo et al. 2022).*"

Banos, S., G. Lentendu, A. Kopf, T. Wubet, F. O. Glöckner, and M. Reich. 2018. A comprehensive fungi-specific 18S rRNA gene sequence primer toolkit suited for diverse research issues and sequencing platforms. BMC Microbiology 18:190.

Li, F., F. Altermatt, J. Yang, S. An, A. Li, and X. Zhang. 2020. Human activities' fingerprint on multitrophic biodiversity and ecosystem functions across a major river catchment in China. Global Change Biology 26:6867–6879.

Tedersoo, L., M. Bahram, L. Zinger, R. H. Nilsson, P. G. Kennedy, T. Yang, S. Anslan, and V. Mikryukov. 2022. Best practices in metabarcoding of fungi: From experimental design to results. Molecular Ecology 31:2769–2795.

Reviewer #3 (Remarks to the Author):

Dear Authors,

I have now read your MS entitled "Unravelling large-scale patterns and drivers of biodiversity in dry rivers".
Overall I appreciated reading your MS as it provides a large-scale overview of biodiversity across multiple taxa in NPRs, which is not common. I acknowledge the strength of the survey and of the data, as well as the originality of the work. Nonetheless, I have some major critics that I will detail below, and the descriptive nature of the work make me skeptical regarding some conclusions and the generality of the work. On the one side we need these types of large-scale descriptive studies, but on the other side it is always frustrating that underlying mechanisms (drivers) are poorly revealed. Herebelow I detail some comments that may perhaps help improving conclusions regarding drivers of observed patterns.

1) Abstract and Introduction are well-written and clear, but I think authors should be more

transparent regarding previous knowledge on NPRs. Authors argue that data are lacking for NPRs (fist words of the Abstract + many statements like that in the Introduction), which is not totally true. Even though NPRs are less studied than perennial rivers, knowledge have been acquired in these ecosystems, including per several authors of the MS. We actually have some knowledge regarding freshwater organisms (e.g., insects), their biodiversity patterns and underlying drivers, but this is clearly less the case for microbial communities of NPR riverbeds (what is adressing this MS). For freshwater macro-organisms, we know that dispersal process for community assembly in NPRs is important, but we may ask if this is also the case for microbial soil communities in these ecosystems. I think that authors should be much more transparent in what we know in NPR biodiversity to better frame the actual lack of knowledge and to built coherent specific hypotheses. There is a very general hypothesis in the last sentence of the Introduction whereas more specific hypotheses could be stated based on previous knowledge (that somewhat exist).

We have revised the Abstract, Introduction, and Results/Discussion sections to align with our stated objectives. We specifically address the relative influences of dispersal, environmental filtering, and biotic interactions, incorporating existing knowledge on communities in dry riverbeds (L185-210). Furthermore, we clarified and rearranged the presentation of our aims and hypotheses (L216-228). Finally, in line with comments also made by Reviewers 1 and 2, we introduced a second hypothesis exploring potential mechanisms, including biotic interactions, that influence beta-diversity in dry riverbeds (L226-228).

Abstract (L146-151) "*More than half of the world's rivers periodically dry up, and the extent of river drying is increasing due to climate change and water resource use. However, we still lack knowledge about the biological communities in dry riverbeds. In particular, it remains poorly understood how dispersal, environmental filtering and biotic interactions drive biodiversity in dry rivers. We conducted a large-scale coordinated survey of patterns and potential drivers of biodiversity in dry riverbeds.*"

2) I also would like authors to be more precise in what they call biodiversity; here authors focus mostly on microbial biodiversity and this should be made more explicit throughout the introduction, as hypotheses and expectation are not the same as when considering macrofauna biodiversity (from insects to birds for instance). You are considering only some facets of biodiversity and I encourage authors to be more transparent by avoiding general wording (e.g., l. 210 "Here, we provide the first global-scale attempt to assess NRP biodiversity" should read ""Here, we provide the first global-scale attempt to assess NRP soil microbiota biodiversity" or something like that).

We have clarified in the Abstract and Introduction that we present data on microbes, invertebrates and plants. Although different microbes receive most attention, their biodiversity in NRPs being least known, both plants and animals are also considered:

L150-153: *"We conducted a large-scale coordinated survey of patterns and potential drivers of biodiversity in dry riverbeds. We focused on eight major taxa, including microorganisms, invertebrates and plants: Algae, Archaea, Bacteria, Fungi, Protozoa, Arthropoda, Nematoda, and Streptophyta."*

L216-211: *"Our aim was to characterize the alpha- and beta-diversity of dry-riverbed communities and to identify their major potential drivers. We analyzed the diversity of eight major taxa (Algae, Archaea, Bacteria, Fungi, Protozoa, Arthropoda, Nematoda and Streptophyta), and estimated the relative abundances of copiotrophic and oligotrophic microorganisms. We then used a random forest approach and correlation networks inferred from graphical models to assess the relative contribution of potential abiotic and biotic drivers of biodiversity patterns."*

3) Also, please avoid in the Introduction paragraphs such as the one l. 201-209. I'm really skeptical that a large-scale correlative study as the one you conducted is "crucial to inform conservation and restoration of freshwater biodiversity facing global change...". If you really think that macro-ecological studies as the one you conducted is crucial for conservation, please explain clearly why and how. My feeling is that you have generated interesting and novel fundamental knowledge, and this already a lot (and enough). I'm not sure that this type of arguments (associated with conservation issues) is required to justify your study. You know that microbial diversity is driven by climate and nutrient, but does this knowledge help resolving conservation issues? If you think this will really help, please explain how in the Discussion. Otherwise I would propose deleting this paragraph (replace it by a better knowledge background in NPRs).

We deleted this paragraph, as also suggested by other reviewers.

4) Regarding patterns of soil biodiversity, I missed a general and visual approach for describing patterns. here you mostly focus on alpha diversity, which is fine. But you may have plot in Figure 1 the distribution of OTU richness for the most abundant groups, at least for readers to visually see whether this generate clustered patterns of evenly spread patterns of alpha diversity (in other words I suggest you to map biodiversity patterns). Also, I would have like to see a graph and/or analysis testing whether the distribution of the major groups is similar across the globe or not. This is more or less what you have done in the beta-diversity analysis, but again this is vague and not visual. I would advice you to run some types of multivariate analyses (NMDS, ddRDA, clustering trees) for instance to search for the existence of "clusters" (group of sites that are similar in term of taxa composition) and test whether some environmental variables are associated with this clusters. I would be curious to see (i) if some clusters can be found, (ii) if they are geographically coherent (or not) and (ii) if their are associated with some environmental factors.

We have now included maps of OTU richness for the eight taxa as supplementary material (Supplementary Fig. S1-S8) to visualize alpha-diversity patterns. We have also included principal coordinate analyses to visualize beta-diversity patterns (Supplementary Fig. S.18-25) and have used permutational analysis of variance (PERMANOVA) to assess variation in community composition between continents and climate classes. To better identify associations between beta-diversity of the eight investigated taxa and environmental factors, we also re-ran the Graphical Lasso approach with distances computed on individual environmental variables. The association between all environmental variables and beta-diversity patterns for each taxon is visualized in the revised Figure 5. This analysis supports the idea that biotic interactions are more important than environmental conditions. In addition,

the new presentation of the data allowed us to visualize the dependencies between environmental variables, as suggested in comment #5.

5) the randomForest appraoch you used for revealing "drivers" is statistically "fine", but I missed some details and I'm not sure it is necesarilly the best approach. First, there is no mention of collinearity. I'm pretty sure that some of your environmental factors are strongly associated one with each other, which as you know can be a problem for statistical inferences. Please, add a correlation matrix so that readers can judge potential problems associated with collinearity. RandomForest are supposed to be good for dealing with collinearity, but - assuming strong collinearity among some of your variables- I would like to see an complementary analysis in which you would have reduced collinearity, for instance by focusing on PCA axes (for instance, it is highly likely that climate can be reduced by one or two orthogonal PCA axes, the same for environmental variables). I know this is frustrating to reduce dimensionality, but this can also be strongly informative. Second, you choose random Forest to reveal drivers, which is not in my opinion the best option. I know RF haves many advantages, including generating stronger r2 than traditional approaches, but this is a predictive approach, not a mechanistic approach. Given that your goal is to reveal drivers, I would have used a more mechanistic approach to reach this objective. In particular, I think that parametric approaches (GLMs of structural equation models) would have been much more powerful than RF to really reveal likely mechanisms. I would have first reduce dimensionality (or generate latent variables) and then use classical linear models (including quadratic terms if needed). Linear models have the strong advantage to provide parameters, that can be then used to compare more easily drivers among taxa. For instance, parameters (slope coefficients) can be used to test the global effect size (ES) of a factor (e.g., climate) on biodiversity patterns: is the ES for climate significantly different from zero across all taxa? is the global ES for climate higher than the ES for nutrients across all taxa?. Using parametric appraoches allow for testing quantitatively/statistically these hypotheses, which non-parametric approaches can not do. In addition, you can test statistically whether the effect of a factor (e.g., climate) is homogeneous among taxa or not (using heterogeneity tests on effect sizes). Finally, if you combine linear models to causal approaches, you can test for direct and indirect effects: for instance climate may be related to nutrients availability (the latter being associated with OTU richness) and to OTU richness directly. This can be modeled quite easily as soons as you reduce dimensionality and you use parametric approaches. I think such type of mechanistic approach would be much more informative than a predictive and "black box" approach as the one you used, even if they produce lower r2 (and if for some taxa there is not significant drivers, this is not a problem, it is also an important finding).

In this comment, the reviewer presents a thought-provoking perspective by delving into the ongoing debate about the benefits of statistical inferences versus machine learning inferences. While we understand the arguments, we remain convinced that Random Forests (RF) are preferable for our purposes. A key reason is that RFs are much less sensitive to collinearity problems, as the reviewer also acknowledges, since only a subset of the predictors is used to construct trees (i.e. one-third of the predictors are randomly sampled as candidates at each split). Additionally, consistent bootstrapping of the data further mitigates collinearity issues, thus reducing the likelihood of correlated variables in smaller data subsets and preventing overfitting.

Furthermore, we respectfully disagree with the characterization of RFs as a black-box approach. Each tree within the ensemble can be identified and mapped, and the final output results from the weighted aggregation of these trees. Although it is true that effect sizes cannot be determined with RFs, similar to other regression-tree methods, it is possible to derive relative variable importance values through permutations, which have in fact been shown to be more robust than effect size estimates obtained by linear modeling. Overall, RF approaches are less sensitive to violations of statistical assumptions. Importantly, moreover, RFs enable the implicit consideration of interactions between variables at different levels and they also accommodate non-linear relationships between predictors, which is an important aspect that cannot be achieved with linear modeling.

Conversely, we caution against the reviewer's suggestion to apply unsupervised dimension reduction. The reason is that the suggested approaches focus on the predictor variables without considering the response variable. This can complicate interpretations, because latent or composite variables may lack meaningful and readily interpretable representations. Furthermore, unsupervised dimension reductions lead to information loss, especially when the number of selected latent variables in SEMs is large.

To summarize, although we acknowledge the reviewer's view, we remain unconvinced that the suggested alternatives are better suited to meet our objectives and to analyze our data than the machine-learning approach we have chosen to explore the importance of predictors driving biodiversity in NPRs.

Although random forests are robust to collinearity, we reran the analyses after excluding the most highly correlated variables: because aridity and PET are, almost by definition, highly correlated with precipitation and temperature, respectively, and latitude with temperature, aridity, PET and latitude have been excluded in the statistical analyses in the revised manuscript. These exclusions did not alter any of our main results or their interpretation. A correlation matrix is now included as a supplementary material (Fig. S28) to assess correlation between climatic, physicochemical, land-use and spatial variables.

6) I may be wrong but you did not included in the models the date at which you sampled the soil since drying. I may be wrong but I guess that soil biodiversity evolve along the drying process: the community composition is likely not the same some days after the drying occur than some moths after the drying occur. Could you please include this parameter in the models?

Yes, this is an important consideration. It is captured in our models by the duration of the dry period experienced by the riverbed communities before sample collection, corresponding to the variable « Dry phase duration », which was indeed identified as a variable that was significantly correlated with the OTU richness of several taxa (Archaea, Fungi, Algae, Protozoa and Arthropoda). We clarified that this variable refers to "the period between the onset of the dry phase and sample collection" (L509).

7)The lasso approach is nice but I'm a bit skeptical regarding the way you included environmental and climatic distances. From what I understood, you estimated the distances based on all factors at a time. If yes, I'm afraid that the resulting distances may be uninformative. For instance, for environmental variables you included all 11 variables, it is likely that you loose all relevant information. For instance the C content is likely more

important than the sediment particles mean size to explain beta diversity (as the former is more associated with OTU richness than the latter). So I would rather used distances based on each variable independently, by focusing for instance on the ones that are more associated with OTU richness (Table 1). This lack of precision may obfuscate some patterns (or reinforce your findings if the abiotic factors are still unimportant)

This is an excellent point, which also echoes our response to the comment above on Random Forests. Calculating an overall distance matrix is indeed of limited interest, because some potential key drivers can be blurred (as in any unsupervised dimension reduction). Consequently, we have re-run the Graphical Lasso approach with distances now computed on the individual environmental variables, as suggested (L638-640). Figure 5 was adapted accordingly, which is also an efficient way to visualize dependencies among the environmental variables, as suggested in comment #5. Our main conclusion that biotic interactions are more important than environmental variables was not affected by these changes.

8) The discussion is nice overall but you did not discuss an important finding that finally the variance in richness along the gradient is mostly unexplained by the abiotic factors you selected. This needs to be discussed; either you missed important variables, or community are mostly driven by biotic interactions (as suggested by the lasso analysis)

We have expanded our original discussion on the weak correlations between environmental variables and beta-diversity of the eight selected taxa, which we had indeed interpreted – and still interpret – as a greater contribution of biotic interactions to beta-diversity patterns in dry riverbeds. In particular, we have now included a section discussing in more detail the limited explanatory power of the environmental factors we considered (L429-444).

"The observed limited influence of dispersal on the assembly of dry-riverbed communities should be interpreted with caution, because our sites spanned many river networks and our Euclidean measure of spatial distances did not account for hydrological connectivity among sites. Hydrological connectivity is likely to constrain the movement of aquatic taxa and could thus have led to underestimation of the role of dispersal[22]. Given the low explanatory power of the variables included in the random forest models of OTU richness for some taxa, the weak correlations between environmental variables and the beta-diversity of most taxa also suggest a limited influence of environmental filtering on community assembly in dry riverbeds. This result is surprising, because the analyzed environmental variables (e.g. organic carbon content, pH, climate) are known drivers of soil and freshwater microbial diversity[48,62,63]. Dormant or dead aquatic organisms detected in dry sediments by eDNA metabarcoding may have partially obscured biodiversity–environment relationships. In addition, samples were collected at sites with different drying histories. At sites that have experienced seasonal dry phases over evolutionary timescales, some aquatic species have evolved specific adaptations to tolerate dry phases. In contrast, at newly intermittent sites communities may be prone to higher biodiversity loss in response to unpredictable drying events[6]."

Reviewer #4 (Remarks to the Author):

This manuscript presents a research work that is based on a powerful data set about a kind of natural systems of relevance in the context of climate change. The knowledge of non-perennial rivers (NPRs) biodiversity and its relationship with environmental drivers is a matter of great significance for understanding the behaviour of these ecosystems and predicting their evolution in this changing context that probably will benefit them in front of the better known perennial rivers.

1) However, the main pitfall of these data is that the group of samples is clearly dominated by those coming from temperate latitudes and climates, with a low representation of arid/semiarid environments and very few samples from other climate classes. The authors point out this limitation at the end of the discussion (lines 436 to 440), but up to here they always refer to this work as a global-scale assessment. In my opinion, this bias should be taken in account and become more explicit through the whole manuscript. Indeed, from a practical point of view, making more visible this limitation will not discredit the main statements or the relevance of this research, because temperate and semiarid latitudes contain a high amount of NPRs and the expected efects of global warming on rivers will be specially intense in these regions.

We now refer to our study as a "large-scale coordinated survey" and also indicate at the beginning of the Results section the prevalence of temperate and arid locations.

L232-234: The 84 sediment samples were collected in regions spanning the main five Köppen climate classes, with most collected in temperate and arid regions (A: tropical n=2, B: arid: n=14, C: temperate n=66, D: continental n=1, E: polar n=1) (Fig. 1).

L150: We conducted a large-scale coordinated survey of patterns and potential drivers of biodiversity in dry riverbeds.

L211-213: Previous efforts to quantify the dry-phase biodiversity of NPRs have been restricted, both geographically and taxonomically[29–31], with large-scale coordinated surveys in needed to fill this gap.

L214-215: Here, we present the first large-scale attempt to assess NPR biodiversity during dry phases.

L448-449: … our large-scale assessment was biased towards temperate and, to a lesser extent, arid climates, which is a frequent limitation of global biodiversity studies[33,34].

Methods and obtained results of biodiversity assessment look appropriated for the purposes of this research, although their significance and the validity of data interpretations are inevitably constrained by the use of OTUs richness as biodiversity indicator and the high level of taxonomic units selected for the analysis and comparison of taxa composition in the community assemblages, as usually happens in this kind of metabarcoding studies. Statistical analyses are mostly out of my scope of expertise, but they seem to be robust and give meaningful information. I am not familiar with methods used for other kind of analyses, such as those for chemical or climatic parameters.

Conclusions emerge as logical inferences after the discussion of results. They must be interpreted using additional informations and hypotheses, which are well-provided by an adequate pool of references. In spite of everything, several generalizations, assumptions or especulations had to be done, but this extent has been taken in account and the language used in the discussion reflects this fact when it is necessary. Literature is also appropriately used to provide context to the results when it is convenient, as well as to the purposes of the research in the introduction.

2) Unfortunately, the concluding remarks in the last paragraph (lines 452 to 455) ended up as if they were included in the section named "potential limitations" of the discussion. Although the end of the discussion is the most logical position for these statements, in my opinion the consistence of the story sequence would be improved if they were clearly placed out of this section. Actually, considering the usefulness of knowing these potential limitations before reading the other sections of the discussion (in a similar way that my previous comment about the bias of the samples), probably this section should better be placed at the beginning of the discussion, leaving the concluding remarks at the end.

As suggested, we have re-structured this part of the Discussion, and the concluding remarks have been developed under a separate heading entitled "Conclusions" (L466-476), which is distinct from the discussion of "Potential limitations". Despite the importance of recognizing potential limitations, we refrained from moving the "Potential limitations" section to the beginning of the Discussion because our key findings need to be presented first. Information on the predominance of samples from temperate and arid climates is now also provided already at the beginning of the Results section, along with information on the variability of environmental conditions (as also suggested by Reviewer 1) (L232-238). The limitations associated with eDNA metabarcoding are also raised when discussing the limited influence of environmental variables on diversity patterns in dry riverbeds (L438-444).

3) According to the relevance of this research in the context of climate change, I would expect to find in the discussion some inferences of these results focused on the prediction of the consequences of warming and altered hydrological patterns in these community assemblages and ecological systems. How could they be modified by changes in abiotic factors? Which effects could have the strong biotic interactions found in this work on the biodiversity responses to these changes? Probably more assumptions should be required for this kind of reasonings, but I think that it would be worthy due to the relevance of the results obtained as well as to the expected role of NPRs in temperate regions in a near future. Anyway, it is only a suggestion and the lack of these considerations should not be an objection for its publication.

The Conclusions section has been further developed and now addresses potential implications of our results (L467-476):

"As temperatures and aridity increase due to ongoing climate change in many global regions, NPRs are becoming increasingly prevalent in both space and time. Our results suggest that biodiversity of most major taxa in dry riverbeds is likely to be increasingly affected by increasing climatic extremity, especially under conditions of low resource availability. Management strategies are thus needed to support the biodiversity and thus ecosystem

functions and services of NPRs affected by diminishing flows[69]. In addition, the tight linkages we observed between the beta-diversity of multiple major taxa in dry riverbeds suggest that recognizing biotic interactions in such strategies could maximize the effectiveness of actions taken to protect biodiverse NPR communities as they adapt to longer, more frequent and higher-intensity dry phases."

REVIEWER COMMENTS

Reviewer #1 (Remarks to the Author):

This manuscript presents the results of a huge collaborative worldwide study about dry riverbed communities, which represents a valuable step forward on a still poorly studied topic. The manuscript greatly improved compared to the previous version, but there are still some parts that need to be reworked, to better focus the storyline and clarify some sections.

ABSTRACT

Lines 159-161: “Covariation among taxa, particularly Bacteria, Fungi, Algae and Protozoa, explained more spatial variation in community composition (i.e. beta-diversity) than dispersal or environmental gradients. This suggests that biotic interactions may strongly influence communities during dry phases in non-perennial rivers, altering biodiversity responses to global changes.” I’m a bit confused with this interpretation of the covariation among taxa exclusively as biotic interactions. More than biotic interactions, I suggest that this finding may result from evolutionary history, especially considering that your study was performed on a global scale.

INTRODUCTION

Paragraphs 1 and 2 could be merged.

Line 181: Please, add “the” before understanding.

Line 187: Please, clarify what you mean by “macroorganisms”, it is a quite general and vague term.

Lines 198-192: “The dendritic structure of river networks and the unidirectional passive transport of organisms in flowing water could generate patterns of decreasing alpha-diversity from headwaters to downstream reaches and high beta-diversity among headwaters”. I suggest mentioning here that this statement refers to microbial diversity because this is not true for all the groups you consider in your study.

The introduction improved a lot compared to the previous version, but it is now too focused on microorganisms, which are not the only ones considered in the study. I suggest extending the examples reported in the different paragraphs, also including other groups of

organisms. Even in the hypotheses, only microorganisms are mentioned.
Line 223: It's not clear what you mean here with water. Is it the water content in the samples? In the methods, you state that the soil samples were sieved and air-dried, but you do not mention a weighing step before and after drying. Please, clarify this.
In the hypotheses it would be useful to repeat some of the references reported earlier in the introduction, to better support them.

RESULTS
Lines 239-240: Instead of grouping them, it would be better to present the number of OTUs for all the eight taxa you considered in this study.
Line 263: Later in the methods you include also dry-phase duration in the climatic variables, whereas here it is presented separately. Please be consistent. For example, here you could delete "climatic variables" and start the sentence with "Mean annual temperature and precipitation".
Line 312, 315: Which other taxa? All of them?

DISCUSSION
Lines 334-337: Why do you interpret the linkage between beta-diversities of the different taxa as assemblies being driven by biotic interactions instead of dispersal limitations and/or evolutionary history? The only explanation of this is the reference to the original paper (ref. 94) in the methods but without any further explanation here. This would need to be better described either here or in the introduction (see a similar comment for the abstract section). Another possible explanation, which is a bit overlooked in the manuscript, is the timing of the colonization process from nearby terrestrial habitats.
Line 339: I suggest changing "Resource availability and climate as drivers of OTU richness in dry riverbeds" to "Resource availability and climate as drivers of OTU alpha-diversity in dry riverbeds".
Lines 362-364: This statement is only partially supported by the results. In fact, while it is true that Algae were negatively affected by dry period duration, partial dependence plots of Fungi, Protozoa, and Arthropoda show peaks at approx. 100 days (as also specified in the results).
The arthropods paragraph at lines 398-405 is a bit confusing, as at the beginning you state

that you “captured the diversity of benthic or interstitial meiofauna”, which suggests aquatic invertebrates, but a few lines after you say that “our detected arthropod community was dominated by largely terrestrial taxa”. Please, clarify this part.
Another part that is missing a bit in the discussion is the shift from aquatic to terrestrial communities. The taxa considered in the study are seldom identified as prevalently aquatic or terrestrial (besides the Algae). Partially refocusing from this perspective could help to better interpret the results. For example, your results about the effects of the duration of the dry phase could be better framed when considering the aquatic-terrestrial shift in community composition and the colonization from riparian habitats.

MATERIAL & METHODS
Samples and data collection: it would be good to have a bit more information about the sampling sites. For example, were they headwaters, mid or lowlands? Had them a similar Strahler order? What is the elevation range? I also wonder where the samples were collected exactly. I mean, in a dry riverbed you have the portion previously wetted but, depending on the river morphology, you can also have portions of the active riverbeds that have been dry for a longer time (e.g. since the last flooding event). This is happening for example in braided rivers, where the active riverbed is way wider than the wetted riverbed. In such situations, it is important to clarify if samples were collected from areas previously inundated or not, as this might affect the colonization stage from terrestrial organisms and consequently the results of your study.
Line 506: Please, better clarify the precipitation parameter. Annual cumulated precipitation? Or?
Lines 514-515: Why did you only quantify the forested area? Considering that drying is also related to agricultural land use, this would have been an interesting variable to incorporate in the study.
Lines 614-619: This part about bacterial and fungi classification is not really fitting in the statistical analyses section. I suggest moving it somewhere else in the methods.
Lines 627-631: I don’t get the rationale behind partitioning beta-diversity into its two components and then only using one. I mean, I agree that nestedness values are very low, but then why do the partition at all?
Lines 633-637: This part should be aligned with the aims and hypotheses, as the comparison

between continents and climate classes is never mentioned before.

SUPPLEMENTARY

Table S1 is appreciated (as was missing before) but not really informative. You obviously have huge ranges of variability for all the variables. I suggest splitting the values for continent or climatic area.

Reviewer #3 (Remarks to the Author):

Dear Authors,

I have now read the revised version of your MS as well as the replies you've made to the reviewers. This is a very serious work, and the MS has been now greatly improved. I've read your arguments regarding the use of RF models versus more "traditional" approaches (linear models). I'm perfectly fine with your choice, and I don't think using an approach rather than another one would have changed the main conclusions. Nonetheless (and as a side note), my main point is that RF is a non-parametric approach, and one can not build a mechanistic model of biodiversity without any parameters. This is fact; all biological laws are driven by some parameters at some points. I think this is important (as biologists) to try (as far as possible) to estimate these parameters, as they can be compared among situations and can be used to generalise (or not) patterns across contexts. But again, I respect the choice you've made, and I'm sure the models have been carefully built.

That being said I have no additional comments. The MS reads very well, and the discussion is honest and well structured. Congratulations.

Reviewer #4 (Remarks to the Author):

I have accurately read the answers of the authors to my suggestions and also made a faster reading to the new version of the manuscript. The authors have satisfactorily addressed my critical remarks, changing the text and/or properly arguing their statements. I also think that the manuscript has been successfully improved with the modifications made by the authors

in response to other reviewers' comments. Therefore, in my opinion the manuscript could be published in its present form.

REVIEWER COMMENTS

Reviewer #1 (Remarks to the Author):

This manuscript presents the results of a huge collaborative worldwide study about dry riverbed communities, which represents a valuable step forward on a still poorly studied topic. The manuscript greatly improved compared to the previous version, but there are still some parts that need to be reworked, to better focus the storyline and clarify some sections.

ABSTRACT

1. Lines 159-161: "Covariation among taxa, particularly Bacteria, Fungi, Algae and Protozoa, explained more spatial variation in community composition (i.e. beta-diversity) than dispersal or environmental gradients. This suggests that biotic interactions may strongly influence communities during dry phases in non-perennial rivers, altering biodiversity responses to global changes." I'm a bit confused with this interpretation of the covariation among taxa exclusively as biotic interactions. More than biotic interactions, I suggest that this finding may result from evolutionary history, especially considering that your study was performed on a global scale.

We have refined our interpretation of the covariation among taxa as influenced by biotic interactions, to include the potential influence of unmeasured environmental variables and co-evolutionary processes in the *Abstract* and *Discussion* sections.

In the Abstract, we now state at L177: "This suggests that biotic interactions or unmeasured environmental and evolutionary factors may markedly influence communities during dry phases in non-perennial rivers, altering biodiversity responses to global climate change."

In the Discussion, we now state at L477–484: "Another non-exclusive explanation for the weak effect of environmental variables on beta-diversity could be that the observed strong partial relationships between taxa also reflect unmeasured environmental variables. This would imply the limited relevance of our measured variables, which is unlikely given the body of evidence of their biotic effects. Finally, the relationships inferred between changes in the community composition of different taxa may also

reflect long-term co-evolution among these taxa. Such co-evolutionary processes involve biotic interactions and may also hide subtler biogeographic histories underlying community assembly."

These also address comment 12.

INTRODUCTION

2. Paragraphs 1 and 2 could be merged.

We would prefer to keep this text as two paragraphs, given that the "Understanding..." sentence that ends paragraph 1 establishes a key knowledge gap that we address, then the "Spatial..." sentence starts a more specific description of conditions, and biodiversity, in dry channels. We would be happy to reconsider this response in light of editorial feedback.

3. Line 181: Please, add "the" before understanding.

The manuscript has been carefully proofread by a native English speaker to rectify any language errors, and have modified this sentence accordingly.

4. Line 187: Please, clarify what you mean by "macroorganisms", it is a quite general and vague term.

We now refer more specifically to "macroinvertebrates" at L202.

5. Lines 198-192: "The dendritic structure of river networks and the unidirectional passive transport of organisms in flowing water could generate patterns of decreasing alpha-diversity from headwaters to downstream reaches and high beta-diversity among headwaters". I suggest mentioning here that this statement refers to microbial diversity because this is not true for all the groups you consider in your study.

We have added a sentence in L207-209 to tone down our statement, while recognizing that the pattern is not unique to microbes: "This pattern particularly applies to microbial diversity, but also holds for active dispersers such as fish and aquatic insects with flying adult stages."

6. The introduction improved a lot compared to the previous version, but it is now too focused on microorganisms, which are not the only ones considered in the study. I suggest extending the examples reported in the different paragraphs, also including other groups of organisms. Even in the hypotheses, only microorganisms are mentioned.

We now also refer to taxa other than microbes at L226–232: "Such increases in the relative abundance of oligotrophic Bacteria (e.g. Actinobacteria, Alphaproteobacteria) or Fungi (e.g. Basidiomycota) also occur in response to experimental sediment desiccation[32]. In other taxa, such as Algae, resistance to dry conditions is facilitated by dormant spores and protective carotenoids[33], whereas Nematoda often undergo anhydrobiosis to survive dry phases[34]. In addition, Arthropoda have a wide range of adaptations conferring resistance or resilience to dry conditions[35]. Dry phases should thus support organisms with traits reflecting the ability to cope with these harsh environmental conditions."

We have also reformulated our first hypothesis so that it is no longer restricted to microorganisms (L245–249): "Our first hypothesis was that the availability of key resources, i.e. organic carbon and

inorganic nutrients, is positively related to the taxonomic richness (i.e. alpha-diversity) of dry-riverbed communities, whereas limited water availability, caused by low precipitation and long dry-phase durations, decreases richness[20].”

7. Line 223: It’s not clear what you mean here with water. Is it the water content in the samples? In the methods, you state that the soil samples were sieved and air-dried, but you do not mention a weighing step before and after drying. Please, clarify this.

This point has now been clarified by our rephrased first hypothesis, as described in response to comment 6 (L244–248).

8. In the hypotheses it would be useful to repeat some of the references reported earlier in the introduction, to better support them.
We have now added several references to better support our hypotheses.

RESULTS
9. Lines 239-240: Instead of grouping them, it would be better to present the number of OTUs for all the eight taxa you considered in this study.

The number of OTUs is now stated for each of the eight taxa (L265–267): “eDNA metabarcoding identified 131 operational taxonomic units (OTUs) of Algae, 82 of Archaea, 2035 of Bacteria, 744 of Fungi, 520 of Protozoa, 158 of Arthropoda, 93 of Nematoda and 51 of Streptophyta.”

10. Line 263: Later in the methods you include also dry-phase duration in the climatic variables, whereas here it is presented separately. Please be consistent. For example, here you could delete “climatic variables” and start the sentence with “Mean annual temperature and precipitation”.

We have removed the reference to dry-phase duration as a climatic variable in the *Methods* section.

11. Line 312, 315: Which other taxa? All of them?

We now indicate the relationships with other taxa at L338–344: “The beta-diversity of Fungi and Bacteria was strongly associated with the beta-diversity of other taxa (e.g. Protozoa, Algae, Nematoda) in terms of both strength (weighted degree values of 0.82 and 0.77, respectively) and the number of links with other taxa (degree value of 5). The highest partial correlations were observed between Fungi and Bacteria. Protozoa (weighted degree value of 0.69, degree value of 5) and Algae (0.52, 5) were also strongly associated with other taxa, in particular Fungi and Bacteria.”

DISCUSSION
12. Lines 334-337: Why do you interpret the linkage between beta-diversities of the different taxa as assemblies being driven by biotic interactions instead of dispersal limitations and/or evolutionary history? The only explanation of this is the reference to the original paper (ref. 94) in the methods but without any further explanation here. This would need to be better described either here or in the introduction (see a similar comment for the abstract section). Another possible explanation, which is a bit overlooked in the manuscript, is the timing of the colonization process from nearby terrestrial habitats.

We addressed the potential influence of dispersal limitation on beta-diversity patterns by including spatial distance between sites. Apart from a modest partial correlation coefficient observed between spatial distance and the beta-diversity of Bacteria and Streptophyta, no association was found with the beta-diversity of other taxa. This suggests that dispersal limitation did not have a dominant influence on the observed beta-diversity patterns. Spatial distance was also used to address potential spatial autocorrelation of unmeasured environmental variables, which might otherwise have contributed to co-variation in beta-diversity patterns among taxa. Again, spatial distance did not emerge as a significant driver of beta diversity for the majority of taxa.

This is stated in L348-350: "The Euclidean spatial distance between sampling sites, which represents an indirect measure of dispersal limitation, was weakly positively correlated with the beta-diversity of Bacteria and Streptophyta."

Please see also our response to comment 1 above.

13. Line 339: I suggest changing "Resource availability and climate as drivers of OTU richness in dry riverbeds" to "Resource availability and climate as drivers of OTU alpha-diversity in dry riverbeds".

We would prefer to keep reference to "OTU richness" here to align with our terminology throughout the manuscript and thus promote clarity.

14. Lines 362-364: This statement is only partially supported by the results. In fact, while it is true that Algae were negatively affected by dry period duration, partial dependence plots of Fungi, Protozoa, and Arthropoda show peaks at approx. 100 days (as also specified in the results).

Thank you for highlighting the opportunity to better describe our results. We have added a sentence to clarify that the OTU richness of Fungi, Protozoa and Arthropoda was promoted under conditions of moderate-duration dry phases, and we have modified the last sentence of this paragraph as follows (L391–394): "In contrast, the OTU richness of Arthropoda, Fungi and Protozoa increased during dry phases of moderate duration (i.e. up to 100 days), which may reflect the colonization of terrestrial and semi-aquatic taxa from riparian habitats and their contribution to dry-riverbed communities"

15. The arthropods paragraph at lines 398-405 is a bit confusing, as at the beginning you state that you "captured the diversity of benthic or interstitial meiofauna", which suggests aquatic invertebrates, but a few lines after you say that "our detected arthropod community was dominated by largely terrestrial taxa". Please, clarify this part.

We now refrain from mentioning "benthic or interstitial," stating only that we captured the diversity of meiofauna (L429–431): "Most Metazoa identified by our metabarcoding approach were Arthropoda such as Collembola, Ostracoda, Arachnida (mainly mites) and Nematoda, representing some important meiofauna groups, whereas macroinvertebrate eDNA was poorly detected."

16. Another part that is missing a bit in the discussion is the shift from aquatic to terrestrial communities. The taxa considered in the study are seldom identified as prevalently aquatic or terrestrial (besides the Algae). Partially refocusing from this perspective could help to better interpret the results. For example, your results about the effects of the duration of the dry phase

could be better framed when considering the aquatic-terrestrial shift in community composition and the colonization from riparian habitats.

We have added a sentence to the *Discussion* to consider the effects of the dry-phase duration on OTU richness (L391–394): "In contrast, the OTU richness of Arthropoda, Fungi and Protozoa increased during dry phases of moderate duration (i.e. up to 100 days), which may reflect the colonization of terrestrial and semi-aquatic taxa from riparian habitats and their contribution to dry-riverbed communities."

MATERIAL & METHODS

17. Samples and data collection: it would be good to have a bit more information about the sampling sites. For example, were they headwaters, mid or lowlands? Had them a similar Strahler order? What is the elevation range? I also wonder where the samples were collected exactly. I mean, in a dry riverbed you have the portion previously wetted but, depending on the river morphology, you can also have portions of the active riverbeds that have been dry for a longer time (e.g. since the last flooding event). This is happening for example in braided rivers, where the active riverbed is way wider than the wetted riverbed. In such situations, it is important to clarify if samples were collected from areas previously inundated or not, as this might affect the colonization stage from terrestrial organisms and consequently the results of your study.

We have added information on stream order and elevation to the *Results* (L261–263): "Sampling sites were predominantly in headwater streams (Strahler order 1 = 38%, 2 = 25%, 3 = 24%, 4 = 8%, 5 = 4% and 6 = 1%), and at a mean ± sd elevation of 436 ± 469 m (range 24–2852 m)."

We have also expanded the description of our sampling protocol at L530–532 to clarify that our pooled samples captured variability in habitat conditions, including drying history, in the riverbeds:

"Riverbed sediment samples were collected from each quadrat (sediment depth: 0–10 cm) and pooled into a single composite sample per site. This sample encompassed within-reach variability in dry riverbed habitats, including areas with different drying patterns, and their associated biodiversity."

18. Line 506: Please, better clarify the precipitation parameter. Annual cumulated precipitation? Or?

We have clarified at L546 that this is "mean annual precipitation (mm)".

19. Lines 514-515: Why did you only quantify the forested area? Considering that drying is also related to agricultural land use, this would have been an interesting variable to incorporate in the study.

We also quantified agricultural and urban land use, as is now stated at L554: "The proportion of the catchment area upstream of the sampling site covered by forest, agricultural land or urban areas (%) was determined by each participant using local knowledge or GIS-based information derived from the most up-to-date national land cover maps for each country."

However, because of the strong negative correlation (r = -0.95) between the percentage of forest and agricultural land cover, we included only forest cover to avoid collinearity in our models (L644–646): "The percentage of forest and agricultural land cover in the catchment were strongly negatively correlated ($r=-0.95$, $p<0.05$, Fig. S28), so only the percentage of forest was included in the analyses."

We have now clarified this point at L398–401: "Increased resource availability in the form of SRP favoured copiotrophic bacteria, whereas prolonged dry phases, organic matter recalcitrance (indicated

by high C:N ratios) and increased catchment-scale forest cover (corresponding to a decrease in agricultural land cover) caused a shift towards taxa with an oligotrophic strategy."

The proportion of urban land was generally low (4.8%). Nevertheless, in response to this feedback, we reran all analyses (random forest models and lasso approach) with this variable included, but it was never selected as a potential explanatory variable of OTU richness or beta-diversity for any of the taxa. We have revised the text accordingly at L318–320): "Land use variables (percentage of forest or urban land cover in the catchment) were never selected as predictors of OTU richness (Table 1)."

20. Lines 614-619: This part about bacterial and fungi classification is not really fitting in the statistical analyses section. I suggest moving it somewhere else in the methods.

We have moved this section to the end of the eDNA metabarcoding section (L618–624).

21. Lines 627-631: I don't get the rationale behind partitioning beta-diversity into its two components and then only using one. I mean, I agree that nestedness values are very low, but then why do the partition at all?

We did not use the calculated nestedness component because differences in richness could arise from differences in sequencing depth between samples in metabarcoding analyses. We minimized such differences by rarefying our metabarcoding datasets so that comparisons between samples were not affected by differences in sequencing depth. In response to this comment, we reran the analyses without partitioning beta-diversity and obtained similar results, and we have therefore removed all reference to partitioning of the Bray-Curtis index (L666–667): "For each of the eight taxa, a community dissimilarity matrix was computed based on Bray-Curtis distances between all pairs of samples"

23. Lines 633-637: This part should be aligned with the aims and hypotheses, as the comparison between continents and climate classes is never mentioned before.

The comparison between continents and climate classes is now mentioned in the last paragraph of the *Introduction* (L241–242): "We analyzed the diversity of eight major taxa (Algae, Archaea, Bacteria, Fungi, Protozoa, Arthropoda, Nematoda and Streptophyta), and estimated the relative abundances of copiotrophic and oligotrophic microorganisms. We investigated differences in community composition between continents and climate classes, then used random forest modelling and correlation networks inferred from graphical models to assess the relative contribution of potential abiotic and biotic drivers of biodiversity patterns."

SUPPLEMENTARY

24. Table S1 is appreciated (as was missing before) but not really informative. You obviously have huge ranges of variability for all the variables. I suggest splitting the values for continent or climatic area.

Table S1 has been updated to now display the summary statistics for each continent. Note that variability is high also within continents.

Reviewer #3 (Remarks to the Author):

Dear Authors,

I have now read the revised version of your MS as well as the replies you've made to the reviewers. This is a very serious work, and the MS has been now greatly improved. I've read your arguments regarding the use of RF models versus more "traditional" approaches (linear models). I'm perfectly fine with your choice, and I don't think using an approach rather than another one would have changed the main conclusions. Nonetheless (and as a side note), my main point is that RF is a non-parametric approach, and one can not build a mechanistic model of biodiversity without any parameters. This is fact; all biological laws are driven by some parameters at some points. I think this is important (as biologists) to try (as far as possible) to estimate these parameters, as they can be compared among situations and can be used to generalise (or not) patterns across contexts. But again, I respect the choice you've made, and I'm sure the models have been carefully built.
That being said I have no additional comments. The MS reads very well, and the discussion is honest and well structured. Congratulations.

Thank you for your thorough review, for the time taken to consider our revision, and for your positive feedback. We particularly appreciate your understanding of our use of random forest models rather than more traditional linear approaches.

Reviewer #4 (Remarks to the Author):

I have accurately read the answers of the authors to my suggestions and also made a faster reading to the new version of the manuscript. The authors have satisfactorily addressed my critical remarks, changing the text and/or properly arguing their statements. I also think that the manuscript has been successfully improved with the modifications made by the authors in response to other reviewers' comments. Therefore, in my opinion the manuscript could be published in its present form.

Thank you for your careful reassessment of our manuscript and your positive feedback.

REVIEWERS' COMMENTS

Reviewer #1 (Remarks to the Author):

Dear Authors,

I now went through the last round of comments and answers. The manuscript greatly improved, being now perfectly readable and understandable in every section. Thanks a lot for addressing all my comments. In my opinion, the manuscript can be published in its present form.