

Predicting climate-change impacts on the global glacier-fed stream microbiome

Corresponding Author: Professor Tom Battin

This file contains all reviewer reports in order by version, followed by all author rebuttals in order by version.

Version 0:

Reviewer comments:

Reviewer #1

(Remarks to the Author)

This manuscript is offering the first prediction of the change in microbial community structure of microbial communities in glacier fed streams under different climate change scenarios. A very timely and most interesting concept.

I consider this manuscript groundbreaking since it is the very first approach to deal with the consequences of climate change and the shift of microbiome structure of glacier fed streams. Our vanishing glaciers are not only changing landscapes, hydrological water cycles but of course also microbial community structures in terms of composition and function.

The authors can successfully proof by the combination of 2.333 procaryotic metagenome-assembled genomes with climatic, glaciological and environmental data from glaciers worldwide that whole bacterial clades - and hence - also functionalities will be lost. As a consequence, this also effects ecosystem functioning.

It has often been speculated that our glaciers are becoming greener with increasing temperature. Here we have a solid proof under various climate change scenarios which I find extremely appealing and noteworthy.

This work is of significance for the field including related fields. I am not aware of any similar studies up to date. To start with, there is an extreme effort in getting these sample sets together which have been collected in relatively short time in intense sampling campaigns (also to avoid any temporal bias [which is, however, not very likely to exclude at all considering different seasons, weather conditions, situations after precipitation, etc.]). The project Vanishing Glaciers provided the ground for this effort which has been successfully done in very remote areas. So far, most studies focus on one area or even one glacier from which it is very tricky and sometimes very speculative to scale up. This manuscript compiles the major mountain ranges except Antarctic glaciers.

To my knowledge this is also the first attempt for a hierarchical, machine-learning based modelling framework to forecast the consequences. The model frameworks GloGEM and CHELSA provide robust base for their projections.

Despite the fact that it is somehow a logical consequence of glaciers "to become greener" with climate change due to increasing phototrophy, it is shown in the most elegant way and backed up by environmental data extremely well. The strain distribution model for 2.333 strain-level MAGs is a suitable tool therefore and also explained very well (despite Supplementary Figure 1 is very difficult to read (text within text fields).

What I also found very intriguing in this study is the prediction of a possible promotion of a parasitic lifestyle, e.g., by Paceibacteria due to a larger size of a host pool.

Sampling also very often offers a massive bias. Here, it has been done in a very consistent way by using the same procedures and probes for environmental data. Also, considering the extreme remote character of the sampling sites, the samples have been processed already in the field to avoid any change. For DOC: have samples been acidified in the field before shipping?

The methodology is very robust and backed up in a solid way. Metagenomic analysis are followed by state-of-the-art protocols and also here they have been conducted consistently with the same procedures to avoid any bias.

Codes are proofed and are reproducible. Generally, there is enough detail provided for the methodology to be reproduced. I would not see any caveats which would require additional work.

In their manuscript, the authors claim the loss of clades under the aspect of climate change and hence, also the change of ecosystem functionalities. They proof it well with solid methodology, sampling and protocols. I do not see any flaws in the concept as well as the presentation of the data. Figures and supplementary ones show the predicted change in the different variables - also on metabolic pathways - very well.

I am very much in favor with this study, the conceptualism thereof, the implementation, interpretation and presentation. Supplementary information is sufficient and provides valuable background data.

This paper is a crucial piece of work showing us the consequences of climate change and the role of microbial communities which literally can change the functionalities in the future. I consider it as crucial to present it to decision makers for actions against climate warming. Chapeau to the authors. I highly recommend it being published.

(Remarks on code availability)

Codes have been able to be installed and run. The authors provided a README file.

Reviewer #2

(Remarks to the Author)

Bourquin et al. created a model to explore the impact that glacial-fed streams will experience under three different SSP climate change scenarios. The model was built using a dataset of samples collected from a wide selection of glacial-fed streams, ensuring it is based on robust data. The results show that shifts in glacial discharges due to glacial retreat will lead to an increase in primary productivity, mainly due to decreased turbidity, and a loss of microbial clades. This study is well-presented and clear, and it is of utmost relevance, making it suitable for publication in Nature Communications. However, I believe it could be improved by emphasizing why certain choices were made, adding some missing methodologies, and potentially expanding on some topics. Please find my detailed comments below:

1. L29-32: You mention both an increase in biodiversity and a loss of entire bacterial clades under various climate change scenarios. These two aspects could be seen as contrasting. Could you add a brief discussion of this in your work?

2. L85-87: Could you justify why you chose to focus on SSP1, SSP5, and particularly SSP3?

3. L98: Is it phosphorus or phosphate? Check Supplementary Fig. 2.

4. L113-115: While it is true that glacial flour and sediment will decrease with a decrease in glacial meltwater discharge, nutrient input from glaciers to the outflow systems might also decrease (for example, it looks like DIN and phosphate will decrease under SSP3 and SSP5; Supplementary Fig. 2). How will this impact the "GFS greening"? Also, from Supplementary Fig. 2, it appears that the main change in terms of median values will be water temperature. Therefore, it would be interesting to include some discussion on how this could influence algal activity.

5. L135: Define what GAM stands for.

6. L207: As the majority of the strains are expected to increase (L170), why did you only delve into the genomic and functional traits associated with decreased abundance? It might also be interesting to focus on the genomic traits and functional potential of strains that are expected to increase in abundance.

7. L239-241: Wasn't it possible to check this with your model? Maybe by looking at genes involved specifically in chemolithoautotrophy?

8. L252: Should this be "predicted to decrease"?

9. L293: What was the distance among the sampling patches?

10. L322: There is a missing space.

11. L350: I would clarify why samples from the Alaska Range were not sequenced.

12. L375: Which database did you use when running eggNOG-mapper? You worked with the KEGG classification, not the eggNOG classification. Is it correct?

13. L382-383: It would be helpful to include the MAG taxonomic information (with also specifics such as completeness and contamination) in the Supplementary Information (SI).

14. I couldn't find where you reported how chlorophyll-a and bacterial abundance were measured/obtained. Could you please include this information in the Methods section?

15. Some R packages are missing references (e.g., ggtree, ggridges).

16. L442: Correct "as well as and covariates" to "as well as covariates."

17. Fig 2: Do the "Group" colors correspond to the branch colors in the phylogenetic tree? Why are "not significant" and "increase" strains clustered and represented together as just one group? I would also suggest flagging the clades predicted to increase with a different color, as this is also an important result.

18. Supplementary Fig. 1: Please specify what the different colors in panel B represent.

(Remarks on code availability)

Reviewer #3

(Remarks to the Author)

This is one of the first manuscripts that I've encountered that attempted to examine the likely changes to environmental microbial community structure and function under different climate scenarios. They specifically focused on glacier fed streams (GFS)—an important environment as glacier meltwater is a significant contributor to freshwater resources globally and is rapidly changing as a result of melt due to climate change. The work is highly significant, as generalized climate models usually predict the effects of a warming atmosphere on physical parameters (e.g. future climate, ocean circulation, etc.), chemical parameters (e.g. weathering rates and chemical products, carbon and nutrient cycling, etc.), and biological parameters (e.g. plant growth and coverage, species shifts, etc.); however, very little has been done to predict the impact of climate change on specific microbial communities. Since these communities underlie many of the important ecosystem processes and provide significant feedbacks to the climate system, it is crucial to understand how they will respond to environmental change.

The general approach used in this manuscript was to use a space-for-time approach where one sample was collected close to the glacier terminus ("early") and another downstream ("late"). They then examined the differences between the "early" and "late" samples in terms of biogeochemistry, biomass, and alpha diversity metrics. Using Generalized Additive Models (GAMs), they determined the correlation of these parameters with climate change, and predicted how they would change in the future under three climate models (SSPs). These parameters were correlated with a large set of strain-resolved metagenome assembled genomes (MAGs) to examine strain responses to the predicted changes as well as examining the response of functions (based on predicted changes in KEGG orthologs and KEGG categories). They argue that cross validation from several different predictive approaches (combination of various GAMs and random forest models) makes their predictive approach robust.

Altogether, they predict 1) a shift toward more generalist and away from more specialist strains, 2) a shift toward light acquisition pathways, which is consistent with a future "greening" of GFS). This could lead to the loss of entire clades and a major shift in carbon and nutrient cycling in these systems.

Overall, I found this paper and their approach exciting and their results fascinating. I have a few concerns, some significant and others less so.

1. Their space-for-time approach is not well explained. The approach taken might well be reasonable, but it is unclear how the conversion of space for time was accomplished, especially considering that a) there are only two points at each location for space, b) glaciers retreat at dramatically different rates, c) the sediments that they sampled (sandy sediments) are quite mobile in GFS, and may not be the same age as the location (and this would be affected by the elevation changes in the system) and d) the linearity of the space-for-time assumptions is not established. Some further description and justification of the space-for-time approach and its validity would be welcomed.

2. If we accept that their models are correct, they make predictions regarding shifts in mean nearest taxon distance, strain distribution, genomic traits, and functional potential. In the cases where these features can be directly correlated to the environmental parameters, these predictions are reasonable. However, in some cases they go well beyond that direct correlation. For example, they predict that Gamma and Alpha Proteobacteria will increase in abundance, which is reasonable based on the correlation of strains from these clades with environmental parameters. However, they then go on to argue that "Gamma- and Alphaproteobacteria in GFS are also associated with thiamine (vitamin B1) and vitamin B12 biosynthesis, and putatively can alleviate the demands of auxotrophic algae". This argument is based on a broad correlation of these phylum-level clades with those features, and do not specifically relate to the features of the strains characterized in this study. This happens several times in the main text (e.g. lines 181-185, 220-222, 254-255). Only proximal and direct predictions based upon specific changes in their models should be included. Don't go too far beyond the data.

3. While predictions of environmental parameters and biomass based on the models, it is not at all clear to me how alpha diversity is being predicted here based on the environmental data (lines 120-124). The description in the methods is simply "Biomass and diversity metrics for the bacterial communities were modelled analogously." (lines 389-390). Could you explain this more clearly?

4. For the GAMs, which samples were used to train the models and which to test them? This is not clear. Please clearly state.

5. There was no metagenomics performed on the Alaska Range (line 350). Why? Please clarify.

6. I didn't quite follow the co-assembly approach for constructing their MAGs. They state "For each individual assembly, we mapped the reads of the 5 closest samples (Euclidean distances of gps coordinates)" (lines 362-363). Are they co-assembling MAGs from the same location (i.e. both reaches along the stream) as well as the four other closest locations (by Euclidean distance), or is this something else? Why did they approach your co-assembly this way--seems rather random.

7. They don't describe variability very much. While I understand that they are attempting to make predictions about GFS response globally to climate change, I imagine that there must be some fairly significant variation depending upon location as well as other parameters. This will likely require one or several other manuscripts to describe in detail; however, some characterization of how representative their results are and how likely any one particular GFS is likely to see these same changes would help strengthen their argument here.

8. Similar to point 7, the sole or at least primary driver described here for GFS benthic communities is glacial melt input. While this is true near the glacier, as one gets further from the source, the influence of other components (e.g. plant organic matter, sediment and soil erosion, meteorological inputs, anthropogenic inputs) increases. So while their predictions may hold for GFS close to the glaciers, it is unclear to me how well this will translate to changes further downstream. This also adds a bit to my concerns regarding their space-for-time substitution approach. Some statements (caveats or clarifications) regarding this would be helpful.

Once again, I am excited and broadly convinced regarding their approach and their findings. Their findings are consistent with other predicted changes resulting from glacial melt, but makes the shifts quantitative and palpable. It is generally well written and easy to follow. I applaud the authors for making a very complex paper very accessible. I think this paper will be used as a guide for many future studies. Kudos.

(Remarks on code availability)

The code provides a minimal readme file, but seems to have all of the necessary components to reproduce the results.

Version 1:

Reviewer comments:

Reviewer #2

(Remarks to the Author)

The authors have thoroughly addressed my comments and suggestions, and I would like to thank them for their effort and attention to detail. They have made the necessary revisions to clarify key points, improve the structure, and incorporate additional information where needed. In my opinion, the changes have significantly strengthened the overall quality and coherence of the work.

(Remarks on code availability)

Reviewer #3

(Remarks to the Author)

Great job responding to the concerns and critiques of me and the other reviewers. I am happy with the current version and with the responses.

(Remarks on code availability)

Open Access This Peer Review File is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license, and indicate if changes were made.

In cases where reviewers are anonymous, credit should be given to 'Anonymous Referee' and the source.

The images or other third party material in this Peer Review File are included in the article's Creative Commons license, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons license and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder.

To view a copy of this license, visit <https://creativecommons.org/licenses/by/4.0/>

Response to the reviewers for the manuscript “Predicting climate-change impacts on the global glacier-fed stream microbiome”

Reviewer #1

This manuscript is offering the first prediction of the change in microbial community structure of microbial communities in glacier fed streams under different climate change scenarios. A very timely and most interesting concept.

I consider this manuscript groundbreaking since it is the very first approach to deal with the consequences of climate change and the shift of microbiome structure of glacier fed streams. Our vanishing glaciers are not only changing landscapes, hydrological water cycles but of course also microbial community structures in terms of composition and function.

AUTHORS: We are delighted to receive such positive feedback to our work. We would like to express our sincere gratitude to reviewer #1 for these comments.

The authors can successfully proof by the combination of 2.333 procaryotic metagenome-assembled genomes with climatic, glaciological and environmental data from glaciers worldwide that whole bacterial clades - and hence - also functionalities will be lost. As a consequence, this also effects ecosystem functioning.

It has often been speculated that our glaciers are becoming greener with increasing temperature. Here we have a solid proof under various climate change scenarios which I find extremely appealing and noteworthy.

This work is of significance for the field including related fields. I am not aware of any similar studies up to date. To start with, there is an extreme effort in getting these sample sets together which have been collected in relatively short time in intense sampling campaigns (also to avoid any temporal bias [which is, however, not very likely to exclude at all considering different seasons, weather conditions, situations after precipitation, etc.]). The project Vanishing Glaciers provided the ground for this effort which has been successfully done in very remote areas. So far, most studies focus on one area or even one glacier from which it is very tricky and sometimes very speculative to scale up. This manuscript compiles the major mountain ranges except Antarctic glaciers.

AUTHORS: We are grateful to this reviewer for recognizing the efforts that went into this study — ranging from the field work to sequencing and modeling.

To my knowledge this is also the first attempt for a hierarchical, machine-learning based modelling framework to forecast the consequences. The model frameworks GloGEM and CHELSA provide robust base for their projections.

Despite the fact that it is somehow a logical consequence of glaciers “to become greener” with climate change due to increasing phototrophy, it is shown in the most elegant way and backed up by environmental data extremely well. The strain distribution model for 2.333 strain-level MAGs is a suitable tool therefore and also explained very well (despite Supplementary Figure 1 is very difficult to read (text within text fields)).

AUTHORS: We are very happy to read this positive appraisal of our modeling framework. We do agree that a hierarchical structure, building on robust projections such as those provided by GloGEM and CHELSA, makes sense for predictive modeling of microbial populations. Following this comment, we have revised Supplementary Figure 1 to increase readability and clarity.

What I also found very intriguing in this study is the prediction of a possible promotion of a parasitic lifestyle, e.g., by Paceibacteria due to a larger size of a host pool. Sampling also very often offers a massive bias. Here, it has been done in a very consistent way by using the same procedures and probes for environmental data. Also, considering the extreme remote character of the sampling sites, the samples have been processed already in the field to avoid any change. For DOC: have samples been acidified in the field before shipping?

AUTHORS: We agree with this reviewer on the necessity to properly handle the various samples — starting right in the field. Therefore, we developed SOPs for the Vanishing Glacier Project that were strictly and consistently applied across all study sites as variation in sampling procedures, sequencing and data processing can inflate variance of microbial communities, making predictions very hard.

Regarding DOC samples - they were filtered (ashed GF/F filters) and stored in pre-combusted glass vials at 4°C without acidification. Since shipping of acids is very problematic, we would have had to rely on local supplies of acid of the highest grade, which unfortunately is often not (or not easily) available in the countries that we visited. Therefore, to avoid potential contamination, we preferred not to acidify our DOC samples in the field. However, we have been running assays to determine the biologically available fraction of DOC (BDOC) using Alpine GFS samples. For this, filtered DOC samples are amended with a native bacterial community, and DOC concentration is monitored over 28 days (incubations in the dark and at room temperature). Typically, we find very low BDOC, ranging between 0 and 20% of initial DOC concentrations. We therefore consider the error introduced due to storage of our DOC samples to have minor influence.

The methodology is very robust and backed up in a solid way. Metagenomic analysis are followed by state-of-the-art protocols and also here they have been conducted consistently with the same procedures to avoid any bias.

AUTHORS: Thank you for this comment, which we fully agree with.

Codes are proofed and are reproducible. Generally, there is enough detail provided for the methodology to be reproduced. I would not see any caveats which would require additional work.

AUTHORS: This is indeed an important aspect of our study and we are grateful to the reviewer for highlighting it. We hope that the availability of our code helps other researchers adopt a similar, hierarchical modeling framework.

In their manuscript, the authors claim the loss of clades under the aspect of climate change and hence, also the change of ecosystem functionalities. They proof it well with solid methodology, sampling and protocols. I do not see any flaws in the concept as well as the presentation of the data. Figures and supplementary ones show the predicted change in the different variables - also on metabolic pathways - very well.

AUTHORS: We are grateful for this positive tone here as well.

I am very much in favor with this study, the conceptualism thereof, the implementation, interpretation and presentation. Supplementary information is sufficient and provides valuable background data.

This paper is a crucial piece of work showing us the consequences of climate change and the role of microbial communities which literally can change the functionalities in the future. I consider it as crucial to present it to decision makers for actions against climate warming. Chapeau to the authors. I highly recommend it being published.

AUTHORS: Overall, we would like to reiterate our gratitude for this reviewer for her/his positive feedback and the recognition of our work.

Reviewer #2

Bourquin et al. created a model to explore the impact that glacial-fed streams will experience under three different SSP climate change scenarios. The model was built using a dataset of samples collected from a wide selection of glacial-fed streams, ensuring it is based on robust data. The results show that shifts in glacial discharges due to glacial retreat will lead to an increase in primary productivity, mainly due to decreased turbidity, and a loss of microbial clades. This study is well-presented and clear, and it is of utmost relevance, making it suitable for publication in *Nature Communications*. However, I believe it could be improved by emphasizing why certain choices were made, adding some missing methodologies, and potentially expanding on some topics. Please find my detailed comments below:

AUTHORS: We are pleased to read that reviewer # 2 finds our work to be of great relevance. We thank the reviewer for the constructive feedback, which we address below.

1. L29-32: You mention both an increase in biodiversity and a loss of entire bacterial clades under various climate change scenarios. These two aspects could be seen as contrasting. Could you add a brief discussion of this in your work?

AUTHORS: Thank you for pointing this out — this is an important comment, and we apologize for not having been clear enough on this. Indeed, bacterial biodiversity (as Shannon H') is predicted to increase by ~ 6% in future GFSs. We speculate that this is explicable by increased niche availability linked to increased resource availability (due to "greening"). By the way, an accompanying paper, accepted in *Nature Microbiology* (Michoud et al.), suggests indeed that biofilms in GFSs with elevated algal biomass are more complex, hence potentially offering more diverse niches. Furthermore, present-day GFS microbial diversity is dominated by microbes that are well adapted to the selective constraints of the GFS environment. For instance, we previously linked environmental constraints to enhanced microdiversity in GFS bacterial communities (doi.org/10.1038/s41396-021-01106-6). Microdiversity (i.e. the presence of many closely-related taxa) can be assessed using phylogenetic measures, such as nearest taxon distance. Low nearest-taxon-distances reflect phylogenetic clustering - which is high in communities under selection. In future GFSs, increased resource availability may weaken these selective constraints and allow for more biodiversity, however, at the expense of these well-adapted and microdiverse clades. Our model predictions are in line with these expectations - showing an increase in nearest taxon distances (reflecting the loss of microdiversity), the putative loss of entire clades, and at the same time higher overall diversity. We revised the text to better explain the apparent discrepancy between a loss of clades and a concomitant increase in biodiversity.

2. L85-87: Could you justify why you chose to focus on SSP1, SSP5, and particularly SSP3?

AUTHORS: These simply reflect extreme positive (SSP1, also called "Sustainability - Taking the Green Road"), extreme negative (SSP5, also called "Fossil-fueled Development - Taking

the Highway”) and the intermediate (SSP3, also called “Regional Rivalry - A Rocky Road”) climate change scenarios. These scenarios are often used in predictive studies in ecology.

3. L98: Is it phosphorus or phosphate? Check Supplementary Fig. 2.

AUTHORS: Thanks for noticing this - phosphorus concentration in the streamwater was measured as soluble reactive phosphorus (SRP). We corrected this in Supplementary Fig. 2.

4. L113-115: While it is true that glacial flour and sediment will decrease with a decrease in glacial meltwater discharge, nutrient input from glaciers to the outflow systems might also decrease (for example, it looks like DIN and phosphate will decrease under SSP3 and SSP5; Supplementary Fig. 2). How will this impact the “GFS greening”? Also, from Supplementary Fig. 2, it appears that the main change in terms of median values will be water temperature. Therefore, it would be interesting to include some discussion on how this could influence algal activity.

AUTHORS: This is an interesting comment for which we are very grateful. To be honest, and this is the beauty of the science that we all do, we do not know. We discussed this issue to some extent already in a previous paper (doi.org/10.1038/s41561-024-01393-6). Our modeling framework forecasts both a decrease in turbidity and concomitantly, predicts increases in DIN and SRP. We acknowledge that GFS nutrient concentration depends on a multitude of factors, including underlying geology, atmospheric deposition, and catchment vegetation and related soil development. For instance, climate-induced glacier shrinkage reduces the erosive capacity of glaciers, thereby potentially impacting phosphorus supply from bedrock. Smaller glaciers may also receive reduced atmospheric deposition (particularly affecting N in GFS). However, our previous study (doi.org/10.1038/s41561-024-01393-6; Fig. 1b) across a wide range of GFSs showed that benthic algal biomass is stronger linked to turbidity (i.e., hence light availability) than to the availability of inorganic nutrients. Hence, our reflections on a future “greening” of GFSs and streamwater turbidity seem supported.

5. L135: Define what GAM stands for.

AUTHORS: Thank you for pointing this out. GAM stands for Generalized Additive Model. We corrected this in the revised manuscript.

6. L207: As the majority of the strains are expected to increase (L170), why did you only delve into the genomic and functional traits associated with decreased abundance? It might also be interesting to focus on the genomic traits and functional potential of strains that are expected to increase in abundance.

AUTHORS: This is a valid point for which we are grateful. Our choice was motivated by the fact that we have genomic information available for microbes abundant in present-day GFSs, including those that are predicted to decrease in abundance. Focusing on these strains therefore provides robust interpretation.

7. L239-241: Wasn't it possible to check this with your model? Maybe by looking at genes involved specifically in chemolithoautotrophy?

AUTHORS: No, it is not possible to obtain this information directly from our models. The models predict shifts in the abundance of MAGs and from these we then inferred changes in the functional potential of the microbiome by looking at the genes that the respective MAGs encode. Directly resolving individual genes that may be more or less prevalent in the future,

would be valid and interesting but would (potentially) require a different modeling framework (e.g considering functional redundancy, orthologs, etc).

8. L252: Should this be "predicted to decrease"?

AUTHORS: Yes. We corrected this mistake.

9. L293: What was the distance among the sampling patches?

AUTHORS: Sediment patches were sampled within less than 10m apart. We added this information to the revised manuscript.

10. L322: There is a missing space.

AUTHORS: Thank you. We corrected this mistake.

11. L350: I would clarify why samples from the Alaska Range were not sequenced.

AUTHORS: Alaska was our last expedition and samples simply came in too late for sequencing. However, we opted to maintain the Alaska GFSs for modeling the environment.

12. L375: Which database did you use when running eggNOG-mapper? You worked with the KEGG classification, not the eggNOG classification. Is it correct?

AUTHORS: We used eggNOG-mapper (v2.1.9) against the eggNOG v5.0 database, which also contains KEGG orthologues. We added the database to the revised manuscript.

13. L382-383: It would be helpful to include the MAG taxonomic information (with also specifics such as completeness and contamination) in the Supplementary Information (SI).

AUTHORS: We refer the reviewer to the accompanying paper in *Nature Microbiology* for a detailed breakdown of MAG-related information.

14. I couldn't find where you reported how chlorophyll-a and bacterial abundance were measured/obtained. Could you please include this information in the Methods section?

AUTHORS: Thank you for pointing this out. We apologize for this lapsus and have now added the respective information in the Methods section.

15. Some R packages are missing references (e.g., ggtree, ggridges).

AUTHORS: Thank you for spotting this, we have included the missing references.

16. L442: Correct "as well as and covariates" to "as well as covariates."

AUTHORS: Corrected.

17. Fig 2: Do the "Group" colors correspond to the branch colors in the phylogenetic tree? Why are "not significant" and "increase" strains clustered and represented together as just one group? I would also suggest flagging the clades predicted to increase with a different color, as this is also an important result.

AUTHORS: Thank you for pointing this out. Given that we focused on MAGs predicted to decrease in abundance, we initially represented only the decreasing clades on this phylogenetic tree. We have now added clades predicted to increase in the tree.

18. Supplementary Fig. 1: Please specify what the different colors in panel B represent.

AUTHORS: We revised Supplementary Figure 1 to increase readability (in response to a comment by Reviewer # 1) and clarified the meaning of the different colors (i.e. dataset sources) in the legend.

Reviewer #3

This is one of the first manuscripts that I've encountered that attempted to examine the likely changes to environmental microbial community structure and function under different climate scenarios. They specifically focused on glacier fed streams (GFS)--an important environment as glacier meltwater is a significant contributor to freshwater resources globally and is rapidly changing as a result of melt due to climate change. The work is highly significant, as generalized climate models usually predict the effects of a warming atmosphere on physical parameters (e.g. future climate, ocean circulation, etc.), chemical parameters (e.g. weathering rates and chemical products, carbon and nutrient cycling, etc.), and biological parameters (e.g. plant growth and coverage, species shifts, etc.); however, very little has been done to predict the impact of climate change on specific microbial communities. Since these communities underlie many of the important ecosystem processes and provide significant feedbacks to the climate system, it is crucial to understand how they will respond to environmental change.

AUTHORS: We would like to express our gratitude also to this reviewer for recognizing the significance of our work. Quite obviously, we strongly agree with the reviewer regarding the importance of these unique microbial communities threatened by climate change.

The general approach used in this manuscript was to use a space-for-time approach where one sample was collected close to the glacier terminus ("early") and another downstream ("late"). They then examined the differences between the "early" and "late" samples in terms of biogeochemistry, biomass, and alpha diversity metrics. Using Generalized Additive Models (GAMs), they determined the correlation of these parameters with climate change, and predicted how they would change in the future under three climate models (SSPs). These parameters were correlated with a large set of strain-resolved metagenome assembled genomes (MAGs) to examine strain responses to the predicted changes as well as examining the response of functions (based on predicted changes in KEGG orthologs and KEGG categories). They argue that cross validation from several different predictive approaches (combination of various GAMs and random forest models) makes their predictive approach robust.

AUTHORS: This summary of our modeling framework is correct - and we appreciate the more detailed comments below which, in our opinion, further strengthened the manuscript.

Altogether, they predict 1) a shift toward more generalist and away from more specialist strains, 2) a shift toward light acquisition pathways, which is consistent with a future "greening" of GFS). This could lead to the loss of entire clades and a major shift in carbon and nutrient cycling in these systems. Overall, I found this paper and their approach exciting and their results fascinating. I have a few concerns, some significant and others less so.

AUTHORS: Thank you again for the positive and constructive feedback.

1. Their space-for-time approach is not well explained. The approach taken might well be reasonable, but it is unclear how the conversion of space for time was accomplished,

especially considering that a) there are only two points at each location for space, b) glaciers retreat at dramatically different rates, c) the sediments that they sampled (sandy sediments) are quite mobile in GFS, and may not be the same age as the location (and this would be affected by the elevation changes in the system) and d) the linearity of the space-for-time assumptions is not established. Some further description and justification of the space-for-time approach and its validity would be welcomed.

AUTHORS: We thank the reviewer for these thoughtful comments. We extended the manuscript to better address these points. Space-for-time substitution is a powerful approach to study ecological processes over longer time-scales and has been repeatedly used in ecological studies (e.g., doi.org/10.1038/s41559-017-0426-x). However, we agree that certain criteria require careful consideration for this approach. Importantly, work on vegetation succession in proglacial forefields ([doi/abs/10.2307/2937039](https://doi.org/10.2307/2937039)) has shown that space-for-time substitution is particularly useful when deterministic processes drive the assembly of rather low-diversity communities where locations along the chronosequence follow similar ecological trajectories (doi.org/10.1111/j.1365-2745.2010.01664.x).

In this context, our previous work (doi.org/10.1038/s41396-021-01106-6) showed that benthic bacterial communities in GFS are indeed characterized by a rather low diversity and shaped by deterministic assembly (i.e. homogeneous selection). Our selection of sampling points was motivated by the necessity to maximize the space-for-time substitution gradient - but at the same time constrained by being part of the same ecological trajectories. We therefore sampled sediments from an upstream reach as close as possible to the glacier snout (typically a few meters from the snout; limited only by field safety aspects) and a reach close to the terminal moraine demarking the last glacial maximum (i.e., the “Little Ice Age”; LIA) approximately 150 to 170 years ago. Below LIA, soils are typically better developed and stream channels more stable. This induces a rather non-linear change in environmental conditions. However, it is reasonable to assume at least a quasi-linearity upstream from LIA. We acknowledge that this is complex, depending on glacier length recession dynamics (also including periods of positive mass balances). It is unfortunate that this cannot simply not be resolved for each glacier individually, at least not within the scope of this study. We also want to draw the attention to the fact that we were simply limited logistically to two reaches per GFS only. On average, sampling one single reach took up to four hours, hence the triplicate time per GFS would simply not have been feasible. However, we used penalized smooth splines (i.e. non-linear relationships smoothed with a penalty on the number of knots) to model relationships between covariates.

Regarding the sandy sediment fraction - it is correct that it is constantly transported downstreams in GFS therefore potentially reducing the differences between up- and downstream benthic microbial communities. Bacterial dispersing along their sedimentary substrate may indeed add random noise to these communities - nevertheless we find that deterministic assembly processes predominantly shape benthic bacterial communities in GFS. Moreover, given the importance of algal biomass highlighted in previous work (doi.org/10.1038/s41561-024-01393-6) and our models, we argue that the time-scales of sediment reworking (in downstream reaches) and growth of benthic algae appear sufficiently long to sustain microbial niches. We have implemented these considerations in the revised manuscript.

2. If we accept that their models are correct, they make predictions regarding shifts in mean nearest taxon distance, strain distribution, genomic traits, and functional potential. In the cases where these features can be directly correlated to the environmental parameters, these predictions are reasonable. However, in some cases they go well beyond that direct correlation. For example, they predict that Gamma and Alpha Proteobacteria will increase in abundance, which is reasonable based on the correlation of strains from these clades with environmental parameters. However, they then go on to argue that "Gamma- and Alphaproteobacteria in GFS are also associated with thiamine (vitamin B1) and vitamin B12 biosynthesis, and putatively can alleviate the demands of auxotrophic algae". This argument is based on a broad correlation of these phylum-level clades with those features, and do not specifically relate to the features of the strains characterized in this study. This happens several times in the main text (e.g. lines 181-185, 220-222, 254-255). Only proximal and direct predictions based upon specific changes in their models should be included. Don't go too far beyond the data.

AUTHORS: We are grateful for this critical comment, based on which we have removed the speculation regarding interactions between bacteria and algae. However, we also believe that hypotheses that arise from our models in combination with independent lines of evidence are interesting for other readers - but agree that they needed to be more clearly phrased as such. We therefore rephrased these instances to more clearly highlight them as hypotheses that, for instance, may be addressed in future work.

3. While predictions of environmental parameters and biomass based on the models, it is not at all clear to me how alpha diversity is being predicted here based on the environmental data (lines 120-124). The description in the methods is simply "Biomass and diversity metrics for the bacterial communities were modelled analogously." (lines 389-390). Could you explain this more clearly?

AUTHORS: Thank you for pointing out this lack of clarity. What we meant is that these models were created the same way as the environmental parameters ones. We revised this section to be clearer.

4. For the GAMs, which samples were used to train the models and which to test them? This is not clear. Please clearly state.

AUTHORS: This work is based on cross-validation. Models were created and validated using 10-fold cross-validation (each time training on 90% of the glacier-fed streams, and testing on the remaining ones). We rephrased the respective section to better clarify this important point.

5. There was no metagenomics performed on the Alaska Range (line 350). Why? Please clarify.

AUTHORS: This point was also raised by reviewer #2. Alaska was our last expedition and samples simply came in too late for sequencing. However, we opted to maintain the Alaska GFSs for modeling the environment.

6. I didn't quite follow the co-assembly approach for constructing their MAGs. They state "For each individual assembly, we mapped the reads of the 5 closest samples (Euclidean distances of gps coordinates)" (lines 362-363). Are they co-assembling MAGs from the same location (i.e. both reaches along the stream) as well as the four other closest locations (by

Euclidean distance), or is this something else? Why did they approach your co-assembly this way--seems rather random.

AUTHORS: We apologize for this unclarity. In fact, we assembled reads into contigs from each GFS reach individually (i.e. technically spoken they were not co-assembled). Then, to bin contigs into MAGs, we mapped reads from five samples to the contigs. This is called multi-coverage binning, which is now commonly done and increases the completeness and decreases contamination of MAGs (doi.org/10.1038/s41592-023-01934-8). Unfortunately, multi-coverage binning is computationally expensive and mapping reads from all metagenomes was infeasible. Hence, we mapped reads of five geographically near sites, including up- and downstream reaches of the same GFS (where available). We revised this section to be clear about this important step.

7. They don't describe variability very much. While I understand that they are attempting to make predictions about GFS response globally to climate change, I imagine that there must be some fairly significant variation depending upon location as well as other parameters. This will likely require one or several other manuscripts to describe in detail; however, some characterization of how representative their results are and how likely any one particular GFS is likely to see these same changes would help strengthen their argument here.

AUTHORS: We are grateful for this very valid point. Please, note that we report median values of change across all GFSs, and where applicable interquartile ranges as well. Hence this provides a good first appreciation of variability related to the global data set. As correctly recognized by this reviewer, exploring forecasts further for regionality would be great but certainly beyond the scope of this manuscript and not trivial either (e.g., for reasons of statistical power). We have provided forecasts for three climate change scenarios. These forecasts do not differ qualitatively, but rather in the magnitude of individual responses. While we agree that better understanding regional (or even local) variability will be interesting and relevant, this putatively requires temporal datasets (of specific locations) and is beyond the scope of this work.

8. Similar to point 7, the sole or at least primary driver described here for GFS benthic communities is glacial melt input. While this is true near the glacier, as one gets further from the source, the influence of other components (e.g. plant organic matter, sediment and soil erosion, meteorological inputs, anthropogenic inputs) increases. So while their predictions may hold for GFS close to the glaciers, it is unclear to me how well this will translate to changes further downstream. This also adds a bit to my concerns regarding their space-for-time substitution approach. Some statements (caveats or clarifications) regarding this would be helpful.

AUTHORS: Thank you for raising this important point. Indeed, our predictions are restricted to the GFS reaches that are relatively close to the glaciers (within the LIA chronosequence); we have made this point clearer in the revised manuscript. However, we also argue that because of the absence of major confounding variation within the upper reaches under direct influence of glaciers, GFSs represent a suitable environment for such predictions, allowing us to model proximal climate change-related shifts in microbial communities. We consider this as a first step before downstream confounding variation (e.g. land use) can be considered in even more complex models. The reviewer is also right in the sense that our space-for-time substitution approach was tailored to resolve the most recent deglaciation history - and hence allowing for a "mirrored" forecasting horizon of a few decades.

Once again, I am excited and broadly convinced regarding their approach and their findings. Their findings are consistent with other predicted changes resulting from glacial melt, but makes the shifts quantitative and palpable. It is generally well written and easy to follow. I applaud the authors for making a very complex paper very accessible. I think this paper will be used as a guide for many future studies. Kudos.

AUTHORS: We are very pleased with this positive assessment. Also we would like to reiterate our gratitude to this reviewer for her/his thoughtful comments. They have further improved the manuscript.