Anonymous Referee #2

This study explores the diversity of bacterial communities and their functional potential in several Andean lakes subject to varying levels of anthropogenic pressure. Using amplicon sequencing and functional prediction based on 16S rRNA gene taxonomy, the authors observed greater bacterial diversity in lakes exposed to greater anthropogenic influence than in better preserved environments, as well as differences in the representation of metabolic pathways related to carbohydrate metabolism and xenobiotics degradation. They also observe some spatial heterogeneity and vertical stratifications within the lakes.

This work may be useful as a preliminary study of the condition of the lakes and will be all the more valuable if sampling continues over the years. However, the present study has several significant weaknesses. Firstly, the number of samples is unfortunately too small to allow for statistical analysis: six in total for four lakes, and for Lakes Tota and Fúquene, the two samples taken were from different seasons or years. Secondly, most of the statistical tests presented in the study are not valid given the number of samples, in particular Spearman's correlation. Thirdly, the authors make several claims regarding the results and their significance that the dataset and analyses cannot support.

I therefore recommend that the manuscript be thoroughly revised, that the analyses of the dataset be redone, and that the article be rewritten accordingly. The study would benefit greatly from increasing the size of the dataset, with at least true triplicates per lake, collected simultaneously. If more extensive sampling is not possible, a more descriptive preliminary study is possible, but the analyses of the dataset must be thoroughly revised, with appropriate use of the statistical tests, which will then be limited. Finally, I would have liked to have metadata concerning the samples, such as weather conditions and certain characteristics of the lakes (pH, temperature, etc.). With the exception of the 16S rRNA gene sequencing data, all metadata is missing from the study. These factors could influence bacterial communities and affect the interpretation of the results.

In its current form, the article doesn't meet the journal's publication criteria.

> ***Response:***
> *We thank the reviewer for this careful assessment. We agree that the original sampling design does not support formal statistical testing or strong conclusions about lake identity, preservation status, anthropogenic pressure, sampling depth, or temporal variation. We have therefore substantially revised the manuscript as a descriptive exploratory study.*
>
> *Specifically, we removed inferential analyses that were not appropriate for the dataset, including correlation-based tests, and no longer report p-values or claims of statistically significant differences. Diversity, taxonomic-composition, ordination, and functional-prediction results are now summarized descriptively at the individual-sample level. We also clarified that the two samples from Fúquene and the two samples from Tota were collected in different periods and are not biological or temporal replicates.*
>
> *We further revised the Abstract, Introduction, Materials and Methods, Results, and Discussion to avoid causal or unsupported interpretations. Apparent differences among samples are now described as sample-level patterns only, and we explicitly state that the study was not designed to test the effects of lake identity, preservation status, anthropogenic pressure, sampling depth, spatial position, or temporal variation.*

We also agree that the lack of physicochemical and weather metadata limits interpretation. We have added this limitation to the manuscript and state that environmental drivers such as pH, temperature, conductivity, dissolved oxygen, nutrients, and recent precipitation could not be evaluated directly. We now present these results as a baseline inventory and hypothesis-generating framework for future studies involving replicated, contemporaneous sampling and environmental measurements.

Major comments:

The abstract will need to be revised in accordance with the new analyses and the revision of several statements, such as L25: "vertical stratification driven by …" I have not found any analysis of this kind in the manuscript.

***Response*:**
We agree with the reviewer. We have therefore revised the abstract substantially to align it with the revised descriptive framework of the manuscript.

The new abstract now presents the study as an exploratory survey of bacterial community composition and predicted functional potential in six water samples. We removed unsupported statements about vertical stratification and causal environmental drivers. We also clarified that, because the sampling design was limited, unbalanced, and not fully contemporaneous, the observed patterns cannot be attributed conclusively to lake identity, preservation status, anthropogenic pressure, sampling depth, or spatial position within the lakes.

In addition, we revised the description of PICRUSt2 results to avoid overinterpretation. The abstract now states that PICRUSt2 predictions represent inferred genomic potential based on 16S rRNA gene composition and should not be interpreted as direct evidence of gene presence, gene expression, microbial metabolic activity, or ecosystem process rates. Finally, the abstract now emphasizes that replicated and contemporaneous sampling, together with environmental measurements, shotgun metagenomics, and direct functional assays, will be required to evaluate the environmental processes shaping bacterial community composition and functional potential in these lakes.

L59-61: If all these described processes have already been extensively documented, what novelties your manuscript is bringing? This sentence or the objective of the study needs to be reworded.

***Response**:*
We agree with the reviewer that the previous wording did not clearly define the novelty of the manuscript. We revised this section to distinguish between ecological processes that are already well documented in Colombian lakes and the specific contribution of the present study. The revised text now clarifies that, although previous work has begun to characterize microbial communities in Colombian highland lakes, comparable ASV-based metataxonomic surveys across multiple Colombian Andean lentic systems remain scarce. We therefore frame the manuscript as one of the first exploratory inventories of bacterial community composition and predicted functional potential across contrasting highland lakes in this region, rather than as a study demonstrating new anthropogenic-pressure effects.

L120: In what way is the sample taken from Lake Tota in 2018 considered an internal control? How was it taken into account in the analyses? Further details are required.

***Response**:*
We agree with the reviewer that the previous description of the December 2018 Tota sample as an "internal control" was unclear and inappropriate. We have removed this terminology from the manuscript. The 2018 Tota sample was not treated as an internal control, biological replicate, temporal replicate, or evidence of temporal change.

We clarified that this sample was retrieved from Forero-Pineda et al. (2021) and reprocessed in the present study using the same bioinformatic workflow applied to the 2019 samples. It was included only to provide a historical reference for Tota under a comparable ASV-based analytical framework. Because it was non-contemporaneous with the 2019 samples, it was analyzed separately and excluded from ordination analyses and cross-lake summaries of the five-sample 2019 dataset.

L116-118: Unfortunately, the dataset is too limited to reveal any variation between the conditions (at least three samples per site, taken on the same date, would be required). Furthermore, the experimental protocol and the selection of samples need to be reconsidered. We have a mix of conditions: a single sample for two lakes, and two samples for two other lakes, but two of these are from the same year but a different season, and two others are from a different year. No trends can be observed with such experimental protocol.

***Response**:*
We agree with the reviewer that the dataset is too limited and unbalanced to reveal trends among conditions or to test lake-level, seasonal, temporal, or preservation-status effects. We have revised the Materials and Methods to clarify the sampling structure and to avoid presenting the samples as replicated comparisons.

The revised text now states that the primary exploratory dataset comprised five samples collected in 2019: two from Fúquene in April and September, one each from Calderona and Colorado in October, and one from Tota in November. We further clarify that, because these samples were collected at different times and were not obtained under a balanced, replicated design, each sample was treated as an individual observational unit. No sample was considered a biological or temporal replicate of another, and no condition-level trends were inferred from this sampling design.

Accordingly, the analyses and interpretation have been revised throughout the manuscript to present sample-level descriptive patterns only, without statistical testing or claims of trends driven by lake identity, preservation status, season, depth, or temporal variation.

L117: I found the term "semester" rather vague and would recommend specifying the date or month instead (or the season if there are several samples). This makes it even more difficult to identify a seasonal trend, or is this aspect relevant given the local climate?

***Response:** We agree with the reviewer that "semester" was vague. We replaced this term with the specific sampling months. We also clarified that, in the Colombian Andean region, seasonal variation is more closely related to rainfall regime than to temperate climatic seasons. However, because the samples were not replicated across months or lakes, rainfall-related or seasonal patterns were not evaluated. The manuscript now states that each sample was treated as an individual observational unit, that no sample was considered a biological, seasonal, or temporal replicate of another, and that no seasonal trends were inferred from this sampling design.*

L139: I would have liked a brief explanation as to why this particular protocol was chosen. Why was a home-made protocol chosen rather than a commercial kit, which would have allowed for a better comparison of the data with that from other publications? Is this due to technical reasons and the type of samples?

***Response**:*
We thank the reviewer for this suggestion. We have added a brief explanation for the use of the CTAB/SDS-based extraction protocol. The protocol was chosen because it had been used for all filters included in the study, including the previously published Tota 2018 sample. We also clarified that CTAB/SDS-based extraction is suitable for environmental filter material containing mixed microbial biomass and potential organic inhibitors. We agree that commercial kits can facilitate comparisons among studies, but in this case consistency across the samples included in the present dataset was prioritized.

L138-149: Please elaborate a little further on the description of the methodology here. Was a pre-amplification PCR carried out prior to sequencing? Were the PCR and MiSeq sequencing steps carried out in-house or performed by a sequencing platform? If in-house, which kits were used?

***Response**:*
We thank the reviewer for requesting additional methodological detail. We have revised this section to clarify which steps were performed locally and which were performed by the sequencing provider. Specifically, we now state that amplification of the 16S rRNA gene V3–V4 region was performed locally using primers 341F and 805R, whereas library preparation, indexing, quality control, and Illumina MiSeq sequencing were performed by Macrogen Inc.

Because the downstream library-preparation and sequencing workflow was performed by the external sequencing provider, we did not add kit-specific details unless explicitly documented in the sequencing report. We also clarified that the historical Tota sample was generated using the same DNA extraction protocol, primer pair, target region, sequencing platform, and library-preparation procedure as the 2019 samples.

L160: I think there's a sentence missing here to explain how the ASVs were trained, and what parameters were used for the clustering?

***Response**:*
Response: We thank the reviewer for pointing out that the ASV-inference procedure required further clarification. We have revised the Methods to explain that sequence variants were inferred using the DADA2 denoising framework and were not clustered into operational taxonomic units at a fixed sequence-similarity threshold. After quality filtering, forward and reverse reads were dereplicated, error models were learned separately for each read direction, and exact amplicon sequence variants were inferred using the DADA2 core algorithm. Because the retained forward and reverse reads did not provide sufficient overlap, the denoised reads were concatenated using mergePairs with justConcatenate = TRUE. A sequence table was then constructed, and chimeric sequences were assessed using the consensus method implemented in removeBimeraDenovo. The corresponding filtering and denoising parameters have now been reported explicitly in the revised Methods.

L177: Spearman's correlation test, like correlation tests in general, requires a considerable number of samples to produce reliable results. Please review your analysis.

***Response**:*
We agree with the reviewer that Spearman's correlation-based co-occurrence analysis is not appropriate for this dataset because the number of samples is too small to support reliable correlation inference. We have therefore removed the co-occurrence analysis from the manuscript, including the associated Spearman correlation test, network results, figure/table references, and interpretation.

The revised manuscript no longer presents co-occurrence patterns or correlation-based associations among taxa. Instead, microbial community structure is summarized descriptively at the individual-sample level using taxonomic composition, alpha-diversity summaries, and descriptive ordination. We also state explicitly in the Methods that no correlation analyses were performed and that no claims of statistically significant correlations or co-occurrence relationships are reported.

L210: I am not entirely sure how the differential analysis carried out here with 2 samples per treatment group (or less?). Please revise your analysis to ensure you only use statistical tests that are appropriate for your dataset.

***Response**:*
We agree with the reviewer that the previous ALDEx2 differential-abundance analysis was not appropriate for the sampling design, because the dataset lacked sufficient replication within lakes or preservation-status categories. We therefore removed the ALDEx2 analysis from the manuscript, including the associated statistical results, p-values, and interpretations.

The revised Methods now state that PICRUSt2 outputs were used only for descriptive visualization of sample-level predicted functional profiles. Predicted pathway abundances were converted to within-sample relative abundances, and no pooling by lake or preservation status and no inferential differential-abundance tests were conducted. We also revised the Results and Discussion to avoid claims of statistically significant differences in predicted functional pathways.

L233: Please could you explain and clarify in the 'Methodology' section how you filtered out ASVs 'with no name available'? Did this filtering apply only to unclassified ASVs, or also to those with multiple affiliations at certain taxonomic levels?

***Response**:*
We thank the reviewer for identifying this ambiguity. We have revised the Methods to clarify that the phrase "ASVs with no name available" was imprecise. ASVs lacking an assignment at the phylum level were excluded from the bacterial-community summaries, as were sequences classified as chloroplasts or other non-target eukaryotic lineages. In contrast, ASVs assigned at the phylum level but unresolved at the family, genus, or species level were retained and represented as unclassified categories at the corresponding rank. Taxonomic assignment was performed hierarchically using the RDP training set with assignTaxonomy; unresolved lower ranks were recorded as NA or unclassified rather than treated as multiple taxonomic affiliations. We have revised the Methods and Results accordingly.

L236-242: I don't think it's possible to observe a statistically significant difference between just two samples. This echoes my general comments on the proper use of statistical tests.

***Response**:*

We agree that a statistically significant difference cannot be reliably inferred from only two individual samples without biological replication. We therefore removed the corresponding statistical comparisons, p-values, significance symbols, and post hoc tests from the manuscript and figures. The Results were revised to describe differences among samples only descriptively. We also now state explicitly that inferential testing was not performed because the study comprised only six samples and temporal sampling was uneven among lakes.

L271: I believe the authors are referring to 'alpha diversity' in this section? Please use standard terminology for diversity analysis. Once again, the authors must be very cautious when interpreting the results of statistical tests carried out on such a small sample.

***Response**:*
We agree and have revised the terminology throughout this subsection. Coverage-standardized Hill numbers are now explicitly described as measures of alpha diversity, whereas Bray–Curtis dissimilarities and principal coordinates analysis are described as measures of beta diversity. We also removed inferential statistical tests and restricted the interpretation to descriptive patterns among individual samples because of the small and unevenly replicated dataset.

L281-283: This is an overstatement that is not sufficiently supported by your results and statistical tests. It is not possible to demonstrate temporal or spatial heterogeneity by comparing just two samples. The same applies to several of the claims made in the discussion.

***Response**:*
We agree and have revised the terminology throughout this subsection. Coverage-standardized Hill numbers are now explicitly described as measures of alpha diversity, whereas Bray–Curtis dissimilarities and principal coordinates analysis are described as measures of beta diversity. We also removed inferential statistical tests and restricted the interpretation to descriptive patterns among individual samples because of the small and unevenly replicated dataset.

L215-295: As mentioned above, I recommend either omitting this part of the analysis or re-running it using a larger sample size. The Spearman's correlation test has been misapplied here.

***Response**:*
We agree with the reviewer that the limited and unevenly replicated dataset does not support reliable inferential statistical comparisons. We therefore comprehensively revised the analytical framework and removed all hypothesis-testing procedures that treated the six samples as sufficient biological replication. Specifically, we removed the Kruskal–Wallis and post hoc comparisons of ASV abundances, the statistical comparisons between sampling dates, the ALDEx2 differential-abundance analysis and associated FDR thresholds and effect sizes, the Spearman correlation-based co-occurrence network, and any related p-values or significance statements. We also did not apply PERMANOVA to the Bray–Curtis dissimilarities. The revised Results now present taxonomic composition, coverage-standardized alpha-diversity estimates, Bray–Curtis ordination, and predicted functional profiles descriptively at the individual-sample level. The corresponding Methods, figures, Results, and Discussion were revised accordingly, and the limitations imposed by the small and uneven sampling design are now stated explicitly.

L305: once again, it is not possible to draw this conclusion from a sample of this size.

***Response**:*
We agree with the reviewer that the limited sampling design does not support conclusions regarding temporal stability or temporal variation. We therefore removed the former subsection entitled "Temporal bacteriome composition in Tota and Fúquene aquatic systems" and deleted the statement suggesting temporal stability in dominant taxa. Differences between the two available sampling dates for Fúquene and Tota are now mentioned only descriptively within the taxonomic-composition section, with an explicit statement that each date was represented by a single sample and that no temporal inference was made.

L319: Why this interest? Please provide a more detailed explanation of the interest in these taxa and mention it in the introduction.

***Response**:*
We thank the reviewer for requesting a clearer justification. We expanded the Introduction to explain that the interest in these taxa extends beyond human health and is framed within a One Health context. The genus Mycobacterium includes environmental nontuberculous mycobacteria associated with humans, cultured and wild fish, reptiles, and other aquatic organisms, while other detected genera, such as Leptospira and Legionella, include lineages of potential human or animal-health relevance. We nevertheless emphasize that genus-level detection based on the V3–V4 region cannot demonstrate the occurrence of pathogenic species, viable organisms, host infection, virulence, or health risk. Accordingly, the analysis is presented only as exploratory phylogenetic placement and taxonomic observation.

L380: Please reformulate the sentence by removing the word "significant". This applies to statistical tests, and your results do not support such a claim.

***Response**:*
We agree with the reviewer. The term "significant" was inappropriate because the limited and unbalanced sampling design did not support reliable inferential comparisons of predicted functional profiles. We therefore removed the ALDEx2 differential-abundance analysis, the associated p-values and false-discovery-rate results, and all statements referring to statistically significant functional differences. Predicted pathways are now described only as sample-level patterns inferred by PICRUSt2.

L381: I do not see how these results could be linked to the size of the lake and its spatial heterogeneity. Please rephrase your argument and be careful not to overstate your point.

***Response**:*
We agree with the reviewer. Lake size and within-lake spatial heterogeneity were not evaluated by the sampling design and therefore cannot be invoked as explanatory variables. The sentence was removed and the opening of the Discussion was rewritten to present the study as an exploratory description of bacterial community composition in six water samples from four Andean lakes. References to lake size and spatial heterogeneity as factors explaining the observed patterns were removed throughout the Discussion and Conclusions.

The revised text now interprets differences strictly at the sample level and explicitly states that the observed variation cannot be attributed causally to lake size, spatial heterogeneity, anthropogenic pressure, preservation status, sampling period, water-column depth, or any other unmeasured environmental driver.

L400: Please modify terms such as “stable”, “heterogeneity”, “significant”, etc. The results don’t support these claims.

***Response**:*
We agree. Terms implying inferential support or ecological properties not evaluated by the study, including “stable,” “community stability,” “heterogeneity,” “significant,” “resilience,” and “dynamic,” were removed or replaced with explicitly descriptive language.

The revised Discussion now states that community composition and coverage-standardized diversity “differed descriptively among the six samples.” Calderona and Colorado are described only as showing lower observed richness and diversity and high relative abundances of a limited number of taxa. No claims are made regarding community stability or statistically significant differences.

We also added a specific clarification that assessment of community stability, resistance, or resilience requires a defined disturbance and an adequately replicated temporal design, which was not available in this study.

L401-403: overstatement

***Response**:*
We agree and removed this interpretation entirely. The revised manuscript does not state that preservation status or anthropogenic pressure caused the observed diversity patterns. Although Fúquene showed comparatively higher richness and diversity than Calderona and Colorado, this observation is now reported only as a descriptive sample-level pattern.

The Discussion explicitly states that no concurrent physicochemical, hydrological, or land-use variables were measured and that the limited and unbalanced sampling design prevents causal evaluation of preservation status, anthropogenic pressure, or other environmental drivers.

L421-423: Based on what? Your dataset can’t support such statement.

***Response**:*
We agree. The statement attributing differences between the Tota samples to seasonal change was removed. The revised Discussion clarifies that the two Tota samples represent only two sampling dates and that one corresponds to a historical December 2018 sample reprocessed alongside the primary 2019 dataset. These samples therefore do not constitute replicated temporal or seasonal sampling.

The observed differences are now interpreted only as variation between individual samples and not as evidence of seasonal dynamics, temporal instability, or responses to nutrient, temperature, oxygen, hydrological, or biological gradients.

L429-441: See my previous comments about co-occurrence and Spearman test. I suggest strongly to remove this part.

***Response**:*

We agree and removed the co-occurrence analysis entirely from the revised manuscript. The network analysis, Spearman correlations, associated figure, Methods description, Results, Discussion, and Conclusions were deleted.

Accordingly, statements regarding shared niches, indirect microbial interactions, ecosystem resilience, conservation predictions, and a specialized relationship between Mycobacterium and Polynucleobacter were also removed. The revised manuscript does not infer ecological interactions from taxonomic co-occurrence.

L474-476: overstatement

***Response**:*
We agree. Statements linking predicted functional profiles to preservation status, anthropogenic pressure, or nutrient availability were removed. The revised functional analysis is entirely descriptive and presents variation among individual samples without assigning samples to preservation-status groups for inferential comparison.

The Discussion now explains that PICRUSt2 predictions are derived from ASV phylogenetic placement and available reference genomes and therefore may partly reflect differences in taxonomic composition. However, they are not interpreted as evidence of gene presence, expression, metabolic activity, or statistically supported associations with environmental conditions.

Because nutrient concentrations and other physicochemical variables were not measured concurrently, the predicted functional patterns are not attributed to eutrophication, nutrient availability, or other environmental gradients.

L492-498: I suggest removing this part too. I have serious doubt about the robustness of such differential abundance analyses. The size of the dataset doesn't support it.

***Response**:*
We agree and removed the differential-abundance analysis from the revised manuscript. The ALDEx2 analysis, adjusted p-values, group comparisons, and all claims of significant differences among preservation-status categories were deleted from the Methods, Results, figures, Discussion, and Conclusions.

The revised Figure 2 now presents only:

- *sample-level weighted NSTI values.*
- *the 30 pathways with the highest total predicted abundances;*
- *a descriptive Bray–Curtis PCoA of the complete predicted functional profiles;*

No inferential differential-abundance tests or PERMANOVA were conducted. Functional variation is interpreted descriptively at the individual-sample level, and the two sampling events available for Fúquene and Tota are not treated as biological replicates or evidence of temporal patterns.

Figures:

Fig1. Panel a does not really provide any information and could be incorporated into a table, for example Table1. The captions for panels e and f are difficult to follow. I would recommend using a more common visualisation, such as the one in panel d.

Fig2. Panel a does not currently contribute anything to the analysis or discussion. I would recommend removing this figure or replacing it with a box plot containing a larger number of samples.

Fig4. The captions are not readable.

Response:
We thank the reviewer for these recommendations. The figures were comprehensively reorganized in the revised manuscript. The former Figure 1 was replaced by a new composite figure containing only the principal taxonomic and diversity results. The uninformative panel from the previous version was removed from the main figure, and sequencing-depth and rarefaction information is now presented separately in Supplementary Figure S1. The less intuitive panels were replaced by more conventional visualizations, including stacked relative-abundance profiles, coverage-standardized Hill diversity, and Bray–Curtis principal coordinates analysis. The Figure 1 caption was rewritten to describe each panel explicitly.

The former Figure 2 was removed. The co-occurrence network was deleted because the six-sample dataset was insufficient for a reliable correlation-based network analysis, and the comparisons between sampling dates were removed because the study design did not support conclusions regarding temporal variation or stability. The remaining informative diversity results were incorporated into the revised Figure 1.

The former Figure 3, showing the exploratory maximum-likelihood placement of Mycobacterium ASVs, was moved to Supplementary Figure S2. This change reduces the emphasis placed on an analysis based on the limited species-level resolution of the V3–V4 region.

Consequently, the revised manuscript now contains two main figures: Figure 1, summarizing taxonomic composition and alpha and beta diversity, and Figure 2, presenting the descriptive PICRUSt2 functional predictions. Supplementary Figure S1 presents sequencing-depth and rarefaction assessments, whereas Supplementary Figure S2 presents the exploratory Mycobacterium phylogeny.

Minor comments:

L93-94: "the sea level, and is the largest remaining…"? This comment applies to several places in the manuscript. The text will need to be checked to ensure it does not contain errors of this kind.

Response:
We thank the reviewer for identifying this grammatical error. The sentence was removed during the broader revision of the study-area description. The revised text now reads: "The study included six water samples from four Andean lentic systems in the Eastern Cordillera of Colombia: Fúquene Lagoon, Tota Lake, Calderona Lagoon, and Colorado Lagoon (Table 1)."

L126: Please avoid giving the reference twice and write instead "Paver and coworkers (2020).

Response:

We agree and revised the citation accordingly. Similar instances of duplicated narrative and parenthetical citations were also checked and corrected throughout the manuscript to ensure consistent reference formatting.

L138: Please correct the title by removing "Assembly". I don't think any assembly step was done here.

Response:
We agree. No sequence-assembly step was performed in this analysis. The subsection title was corrected to: 2.2. DNA extraction and 16S rRNA gene amplicon sequencing