We thank both the reviewers for their comments, which have greatly improved our manuscript. Note that line numbers correspond to the "TrackedChanges" version of the manuscript.

Reviewer #1

1. The authors have done a great job revising this manuscript. They have addressed all my comments, and I especially appreciate the addition of distance metric comparisons across four real-world data sets, which I find very helpful.

    a. Thank you! We agree the addition of the real-world datasets greatly improved the manuscript.

2. I am happy to recommend the manuscript for publication, and I only have one minor comment: In the abstract, I interpret your formulation as a general claim that your new metric is less sensitive to noise than other metrics such as Bray-Curtis. However, as you show in the new Fig. 4, Bray-Curtis is in fact less sensitive to noise than the new metric at high alpha values. Perhaps this can be clarified before publication.

    a. After reflection, we see how the phrase "comparably insensitive" is confusing – it sounds like GUA is always less sensitive, when it can be more sensitive than BC (at high alpha) or less sensitive (at lower alpha). We've edited the sentence:

    b. Line 31: "…GUA is comparably **sensitive** to realistic noise in load estimates compared to conventional alternatives like Bray-Curtis dissimilarities, particularly at lower α."

Reviewer #2

1. Pendleton and Schmidt present Absolute UniFrac, a modification of Weighted UniFrac that uses absolute abundances from qPCR, dPCR, flow cytometry, or spike-in standards to capture the joint effects of phylogeny and biomass. The revision expands the evidence beyond the original Lake Ontario case and adds practical guidance. You retain the four-taxon toy simulation and add four 16S rRNA datasets spanning a nuclear reactor cooling system, the mouse gut, a stratified freshwater lake, and peanut rhizosphere soil, with absolute load quantified by flow cytometry for the cooling water and lake, droplet digital PCR for the gut, and quantitative PCR (qPCR) for the soil. The datasets span roughly 215 to 24,000 ASVs and loads from approximately 4 × 10^5 cells per milliliter to 2 × 10^12 16S copies per gram. Code, data, and an R environment are provided. The goal is to present guidance on a new, practical, abundance-aware phylogenetic dissimilarity for studies that already measure absolute counts. Overall, the revision resolves the main first-round requests. The remaining items are clarity tweaks and two small robustness checks. Note, at over 2700 words, not including figure legends, this is on the long end for a Brief Communication (expected word count ~1000)

a. We thank the reviewer for the thorough consideration of our manuscript, and feel the paper has greatly improved due to their suggestions. We recognize the manuscript has also grown significantly in size due to these additions, but still feel that it retains the character of a Brief Communication (communicating a new idea), less than an original article presenting a full research project. We will defer these considerations to the editor for their decision.

2. Dataset breadth: Addressed. You added three datasets beyond the original lake analysis and compared UA to BCR, UR, and BCA for each system. See Fig. 2 and the surrounding text in the tracked or clean manuscript.

    a. Thank you for noting that the expanded analyses across systems addressed this point.

3. Alpha guidance: Addressed. You move from coarse choices to α in 0.1 steps, report PERMANOVA R^2 across α, and give practical selection advice. See Fig. 3, its caption, and Box 1.

    a. We appreciate the reviewer's acknowledgment that the revised analyses and guidance adequality addressed this concern.

4. Rarefaction guidance: Addressed in a practical sense. Box 2 provides a concrete workflow and code pointers for incorporating rarefaction while preserving absolute scaling. Please also include a minimal sensitivity check against a no-rarefaction pipeline for one dataset in the supplement, using a Mantel R test to compare the two distance matrices. This will preempt common questions associated with understanding if rarefying your data actually changes the results.

    a. This is a great suggestion, and we're grateful for the encouragement to run these experiments. We tested the impact of rarefaction on GUA values for all four datasets and have included an additional figure in the supplement showing these analyses (Fig. S4). We've also added additional information within the main text describing our findings:

    b. Line 245: "Applying GU^A in practice raises several considerations related to sequencing depth, richness, and computational cost. Both Bray-Curtis and UniFrac can be sensitive to sequencing depth because richness varies with read count [17–19]. To address this, we provide a workflow and accompanying code describing how we incorporated rarefaction into our own analyses (Box 2; available code). This approach minimizes sequencing-depth biases while preserving abundance scaling for downstream β-diversity analysis.

    To explore the sensitivity of GU^A to rarefaction, we calculated GU^A using rarefaction at multiple sequencing depths, and used Mantel tests to assess the correlation between the rarefied GU^A distance matrices the non-rarefied control (Fig. S4; supplemental methods). GU^A was largely insensitive to rarefaction at

high α values, even at depths as low as 250 reads per sample, but sensitivity increased as α decreased (Fig. S4). When rarefying to the minimum sequencing depth, as recommended, the effect of rarefaction was negligible (all Mantel's R > 0.95, Fig. S4)."

5. Computational scaling: Addressed. You benchmark GUA against Fast UniFrac and Bray-Curtis, showing quadratic growth with richness, and note caching opportunities. Please add the CPU model, RAM, R version, and package versions directly in the Fig. 4 caption or Methods so the timing is contextualized without following external links.

    a. We are grateful for the suggestion. We've added:

    b. Line 277: "All benchmarks were run on an AMD EPYC 64-core processor with 1014 GB system memory (R v4.3.3, vegan v2.7-1, phyloseq v1.52.0, GUnifrac v1.8.1)."

6. Copy-number assumptions: Partially addressed. Because readers will wonder, add a sentence justifying your choice not to apply per-taxon 16S copy-number correction, and, if possible, include a brief sensitivity check for one dataset in the supplement. This will preempt another common question set associated with multicopy 16S genes, which can be: 1) divergent within a single genome and 2) number >15 in some lineages (e.g., Proteobacteria).

    a. We also enjoyed following up on this suggestion. We tested the effect of copy-number normalization (using PICRUSt2) on GUA, and have incorporated these analyses into a new supporting figure (Fig. S6). We also included results text discussing these findings:

    b. Line 310: "GU^{A} was also insensitive to normalization approaches that adjust ASV abundances based on the predicted 16S rRNA gene copy number (Fig. S6; supplemental methods). We used PICRUSt2 (v2.6) to normalize ASV abundances within each dataset based on predicted 16S copy number per genome, and then applied our standard rarefaction pipeline prior to GU^{A} calculation (Box 2) [20, 21]. Correlations between GU^{A} distance matrices calculated from 16S copy number-normalized datasets and those from the original, non-normalized datasets were consistently near unity across all datasets (Mantel's R > 0.98, Fig. S6A), demonstrating that 16S copy number-normalization had a negligible effect on GU^{A}. Sensitivity of GU^{A} to 16S copy number-normalization generally decreased with increasing values of α in the cooling reactor, freshwater, and soil datasets, but not in the mouse gut dataset (Fig. S6A). In the mouse gut dataset, several highly abundant ASVs belonging to the genus *Faecalibaculum* had predicted 16S copy numbers of seven copies per genome, explaining the relatively greater (though still weak) sensitivity of this dataset to copy number-normalization (Fig. S6B). Finally, we note that 16S copy number-normalization does not account for variation in genome copies per cell (ploidy), which can vary across several orders of magnitude between species and growth phase [22–24]."

Minor Comments

7. Abstract: Early in the abstract, it might be useful to indicate clearly that "load" and "abundance" are synonymous. This recommendation is based on the fact that "load" tends to be more of a medical-focused term, whereas the new metric is useful for other microbial environments (e.g., soils, oceans) where the term "load" is less frequently used compared to "abundance".

    a. The first sentence now reads:

    b. Line 16: "β-diversity is central to microbial ecology, yet commonly used metrics overlook changes in microbial load (or "absolute abundance")…"

8. Abstract line 31: if using alpha in the abstract, need to define as "tunable ecological dimension" or similar.

    a. We've edited this sentence to make this clearer:

    b. Line 28: "To balance these effects, we also incorporate absolute abundance into the generalized extension (GU^A) that has a tunable, continuous ecological parameter (α) that modulates the relative contribution of rare versus abundant lineages to β-diversity calculations."

9. Novelty statement is clearer. Add cautious "to our knowledge" phrasing on line 152.

    a. This is a fair point. We've reworded:

    b. Line 109: "Although substituting absolute for relative abundances is mathematically straightforward, to our knowledge there is no prior work that examines UniFrac in the context of absolute abundance…"

10. Box 1 is useful. Add a one-line note in the text that high α may produce biomass-dominated ordinations and that NMDS or partial ordinations can help when Axis 1 tracks load. Aligns Box 1 advice with lines 198-200.

    a. We are glad that Box 1 is useful! And the recommendation for NMDS or partial ordination is an excellent idea. We've added a sentence (after discussing the ordinations) to provide this option:

    b. Line 214: "Beyond tuning α to modify the influence of abundances, approaches such as NMDS or partial ordination can also be used to modulate the sensitivity of ordinations to microbial load [16]."

11. Line 268: Possible word error: "cell number of 16S copy number" to "cell number or 16S copy number"

    a. This typo has been corrected.

12. Figure 3 Legend Line 331: Is the Freshwater dataset needing to be mentioned in this sentence?

    a. It was originally unmentioned as no samples were subsetted from this dataset; to clarify this, we've added:

    b. Line 224: "…and only mature rhizosphere samples for the soil dataset (no samples were excluded from the freshwater dataset)."

13. Figure S3: Define CR and SBR in the figure legend.

    a. This suggestion has been incorporated, by defining conventional and sod-based rotation strategies directly in the legend of Fig. S3.