nature portfolio

Peer Review File

New SAR11 isolate genomes and global marine metagenomes resolve ecologically relevant units within the Pelagibacterales

Corresponding Author: Professor Michael Rappe

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)
General Comments:

The manuscript describes a new phylogeny for the abundant marine bacterium SAR11. This phylogeny is based only on high quality isolate and single cell-amplified genomes, and is enabled by the addition of 81 new isolates from the ocean near Hawaii, an impressive contribution as only 25 were previously available. The authors additionally map over a thousand ocean metagenomes from multiple datasets to explore the geographic distribution of their SAR11 clades, and based on a combination of the phylogenetic and ecological groupings suggest new family, genus, and species-level names and type strains. A more clearly delineated SAR11 phylogeny fills a clear need in ocean microbial ecology, and this approach that combines ecology and phylogeny into a manually curated taxonomy calls to mind a widely adopted freshwater taxonomy that has been incredibly valuable to the freshwater community for over a decade (https://journals.asm.org/doi/10.1128/mmbr.00028-10). I think this will serve as a similarly valuable, and long overdue, resource.

My main concern is around the construction of the phylogenetic tree, and although phylogenetics is not my primary expertise I did share the public preprint with a colleague whose main focus is phylogenies and they shared my concerns. To create the final tree, the authors excluded genomes that did not have multiple closely related references, as well as excluding groups that are ambiguously part of the SAR11 order. However, all of the clade's diversity should be included to obtain the most reliable branch structure. Instead of showing both versions in Fig 1 to highlight unstable branches, why not show the single large tree along with confidence values for the branches? And instead of excluding subgroups V and VI why not include them and clarify their debated relationship to SAR11? Limiting the downstream mapping analysis to 95ANI clusters is required methodologically, but the authors can just clarify which leaves each ANI cluster combines. (In other words, re-making the tree, which should not take too long, should not require repeating all of the downstream mapping.) Additionally, the authors throughout the paper and in the methods refer to the tree as being constructed from genes, but that is highly unusual. Typically trees are constructed from protein sequences.

Overall the paper is well written and the figures are beautiful. I have a few additional comments and suggestions regarding clarity of figures and discussion of results which I detail below.

Specific Comments:

59- I suggest you add the number of proposed genera specifically here

62- this phrasing makes it sound like you recommend someone do this in the future, not that you do it here in this paper

141- in presenting your results in table 1, you could consider adding a brief statement as to what your results suggest is the best approach for future culturing attempts

234- can you add any statistical analysis to support cohesion within groups beyond eyeballing the heat map?

237- again, can support with any statistics?

244 and Fig2- clarify how group 1 and 3 are distinguished. How is “low-latitude, high diversity” different than “low-latitude”? Confusing because “high-latitude, low-diversity” is explicitly labelled but “low-latitude” does not have any extra description.

253- Can’t two phylogenetically distinct units form the same ecological unit? I think “ecological unit” is perhaps not the best word choice here- if I understand correctly you are trying to say that both ecology and phylogeny, but neither alone, distinguish all the SAR11 “units”, so it’s confusing to call the unit “ecological” when ecology can’t fully resolve it.

255- typo/grammar

Fig 2- why not make the text at the bottom a little bigger, because you could extend the table to the left edge of the teal Pelagibacteriaceae label. It’s just pretty hard to read and you have white space there you could fill.

Fig 2- You should clarify how many metagenomes are in each group category, because it’s easy to assume each value is 1 metagenome but you mapped almost 1500 (!!). A legend to the heatmap shading could also help clarify how you combined coverage/abundance values from different metagenomes, which isn’t clear to me.

280- suggest defining RED (rel evol divergence I assume, but this is not super common term and it does not show up on google easily)

Fig 3. It seems like some panels are lacking labels

316- Is there some information on the extent, or significance, of the phylogenomic structure in Fig 3a?

320- what are these “detection values”? I don’t see units in the Fig 3 barplots either.

Fig 3b- I’m confused about this panel, although it looks very cool I’m not sure what’s on it. At first glance I thought it was the contigs of the two sags along the “x-axis”, and their coverage in the heatmap. But then there are also multiple metagenomes from each location, so maybe it is the sample on the “x-axis” and the abundance in the metagenome in the heat map? Adding a legend with units would help.

Fig 3b- this looks cool but if you stretched it into a straight line along the whole figure width it would be much easier to read. It’s confusing because for example on the left BATS is the bottom row but on the right BATS becomes the top row. In the end, interpretability should trump coolness I think, unfortunately. Also, then you could add an x-axis label which would clear up my comment above.

Fig 3x- not sure what panel to call this, but can you put station ALOHA on the map as well? Since it is in the panel labelled b. How do these stations relate to the metagenome "groups" in Fig 2, that could be clarified.

336- typo/grammar

339- I’m finding it confusing that in figures the labels are Kane’ohe Bay but in the text the abbreviation KByT is used.

379- confidence in the tree could help support these group breakdowns

Fig 4. I would swap order of Type material and Subgroup, so Type is on the right

Fig 4- your superscripts are very hard to see/read/distinguish. Consider bold and underline instead? Also, why put the superscript in the Species column- wouldn’t putting it in the Type column make more sense?

409- instead of deep/shallow I think fine/coarse is more clear and unambiguous.

410- a similar combination of phylogeny and biogeography was used to refine taxonomic groups in abundant freshwater lakes heterotrophs, albeit for full length 16S, in 2012. I see a lot of parallels here, you could consider pointing to that as an example of the utility of this approach as it has greatly expanded our understanding of freshwater ecology. (https://journals.asm.org/doi/10.1128/mmbr.00028-10 , has about a million citations now)

420- this citation (Zhao et al., 2024) must be unlinked to your reference manager because it is in a different format

424- I don’t understand how ecotypic differentiation is in conflict with recombination. Couldn’t recombination be the mechanism leading to ecological differences? Why does ecotypic differentiation indicate speciation instead?

432- I don’t see how time series further this. In fact, I would think more spatial resolution would be more clarifying. After all, you have “global” metagenomes but they are just a few sites in total, and wouldn’t spatial span more environmental variables?

458 – how is your P and N story distinct from the companion paper? May be worth a sentence to mention in the discussion.

530- primers 515F-926R is V4-V5 not V4. Also, did you use Parada's 515F-Y or 515F-C, Parada has two 515's in the paper. This is why you should include the actual primer sequence in parentheses as well, because it gets super confusing with citations alone. Also it's likely that you used the Walters version of the Parada primers. I would just double check this with whoever performed the amplicon sequencing.

530- If you went back and did full length 16S of your isolates that could extend the impact of your phylogeny to 16S amplicon surveys as well.

598- But you are making a comprehensive tree to clarify the SAR11 phylogeny; why would you leave them out? You should include them, and then tell us with confidence are they part of SAR11 or not.

602- shouldn't you be building trees using proteins not genes?

621- what was your outgroup? This is important, and should further be chosen appropriately to decide if those other groups are SAR11 or not (line 598)

626- Why are you using a pruned phylogeny, why not just include confidence values in the branches and then limit downstream analysis to the major groups only. You shouldn't exclude genomes from your tree just because you don't have a lot of them, the tree structure should be shaped by the full diversity. Did any 95% ANI clusters disagree with the tree structure? Why not make the tree, then do the whole-genome ANI clusters and show mapping for the branch-level that each 95ANI maps back to?

(Remarks on code availability)
The NCBI BioProject is a placeholder (line 1108- "NCBI associated with BioProject XXXXXXX")
It looks like it does exist on Figshare though, but I suggest they use the figshare DOI to cite instead of the in-text link.

There is some code and data available on github. It's certainly not all the code used in the analyses but it looks like figure code is there. I guess most stuff was part of the ANVI'O workflow though, so maybe there's not actually a lot of free-standing code with the paper analysis. I did not look in detail, just glanced through.

Reviewer #2

(Remarks to the Author)
The manuscript submitted by Freel et al present an original and impactful study investigating the ecology and evolution of SAR11 marine bacteria. The researchers have significantly expanded the cultured representation of SAR11 diversity, investigated biogeography through genomics and metagenomic fragment recruitment, identified ecologically distinct clades of SAR11 and then proposed a taxonomic framework and naming convention for SAR11… but I guess we don't call it that anymore!

The study is well executed and the paper is well written.

My only main concern is the decision to remove genomes from their original phylogenetic tree. The researchers generated a phylogeny of all 481 genomes and then removed "SAGs that did not fall into a 90% gANI cluster of at least three genomes…" Second, they dereplicated remaining genomes using a 95% gANI cutoff. This left 268 SAR11 genomes, which are the genomes that are also included in the final taxonomic framework and naming convention. I am okay with this sort of pruning and dereplicating in phylogenetic analysis and it will certainly give you a "cleaner" and better resolved tree with nice well supported clades. However, removing this "messiness" might have implications later. So, I would be curious to see what happens to the phylogeny and the taxonomic structure if these genomes are added back in. Does it impact the monophyly of the named genera for example? Also, is it possible to include some bootstrap values on the phylogenies. It would be nice to know how well supported (or not well supported) many of these clades (as well as they deeper relationships) are.

(Remarks on code availability)

Reviewer #3

(Remarks to the Author)
In their study, Freel et al. generated an extensive collection of SAR11 genomes from cultures and combined these with publicly available genomic data to reconstruct the evolutionary and ecological units within this lineage. The study is highly relevant given the global distribution of SAR11 and its distinctive evolutionary history, which challenges traditional taxonomic classification methods based n cutoffs. Below, I provide several suggestions on the phylogenomics part of the study, that I hope will help the authors refine and strengthen their study:

Lines 162–163: The presence/absence of genes is generally not a robust criterion for selecting marker genes within a lineage. Since the study is grounded in the vertical evolutionary history of the clade, relying on markers that may have experienced complex horizontal gene transfer events could obscure or distort the phylogenetic relationships. This might explain why subclades Ia and Ib form a monophyletic group in the analysis.

Lines 169–170: The assignment of families appears somewhat arbitrary and is largely based on previous classifications that define SAR11 as an order. However, the authors have compelling results that could support the development of a more consistent and evolutionarily grounded framework for defining taxonomic levels within SAR11. I particularly appreciate the approach of integrating ecological and evolutionary evidence. The same comment applies to the proposed genera shown in Figure 4.

Figure 1: I recommend testing outgroup-independent rooting methods, such as Minimal Ancestor Deviation (MAD), to evaluate the influence of the chosen outgroup on the topology and stability of the resulting groups.

Lines 640–652: I find it surprising that Chloroflexota genomes were used as the outgroup, considering the considerable phylogenetic distance between this phylum and SAR11. This disparity increases the risk of long-branch attraction artifacts, which could distort tree topology and affect the RED estimates. In addition, different marker genes (120) are used to get this estimates. Why aren't the authors using the marker genes used to build the SAR11 tree?

(Remarks on code availability)

Version 1:

Reviewer comments:

Reviewer #1

(Remarks to the Author)
I have no additional concerns and appreciate the author's detailed responses. They pushed back on a few of my suggestions/criticisms, but I think in all those cases they supported their decision well.

(Remarks on code availability)
The authors addressed my concerns here in their response letter.

Reviewer #2

(Remarks to the Author)
The authors have addressed my comments sufficiently. Great work and a great paper. Thank you.

(Remarks on code availability)

Reviewer #3

(Remarks to the Author)
Thank you for addressing the suggestions provided.

(Remarks on code availability)

Response to Reviewers

Nature Communications manuscript **NCOMMS-25-18926**

Freel, K. C. et al. *New isolate genomes and global marine metagenomes resolve ecologically relevant units of SAR11*.

Page, line, figure, and table numbers included here refer to the manuscript document that includes tracked changes.

REVIEWER COMMENTS

Reviewer #1 (Remarks to the Author):

General Comments:

The manuscript describes a new phylogeny for the abundant marine bacterium SAR11. This phylogeny is based only on high quality isolate and single cell-amplified genomes, and is enabled by the addition of 81 new isolates from the ocean near Hawaii, an impressive contribution as only 25 were previously available. The authors additionally map over a thousand ocean metagenomes from multiple datasets to explore the geographic distribution of their SAR11 clades, and based on a combination of the phylogenetic and ecological groupings suggest new family, genus, and species-level names and type strains. A more clearly delineated SAR11 phylogeny fills a clear need in ocean microbial ecology, and this approach that combines ecology and phylogeny into a manually curated taxonomy calls to mind a widely adopted freshwater taxonomy that has been incredibly valuable to the freshwater community for over a decade (https://journals.asm.org/doi/10.1128/mmbr.00028-10). I think this will serve as a similarly valuable, and long overdue, resource.

We greatly appreciate the positive feedback.

My main concern is around the construction of the phylogenetic tree, and although phylogenetics is not my primary expertise I did share the public preprint with a colleague whose main focus is phylogenies and they shared my concerns. To create the final tree, the authors excluded genomes that did not have multiple closely related references, as well as excluding groups that are ambiguously part of the SAR11 order. However, all of the clade's diversity should be included to obtain the most reliable branch structure. Instead of showing both versions in Fig 1 to highlight unstable branches, why not show the single large tree along with confidence values for the branches? And instead of excluding subgroups V and VI why not include them and clarify their debated relationship to SAR11? Limiting the downstream mapping analysis to 95ANI clusters is required methodologically, but the authors can just clarify which leaves each ANI cluster combines. (In other words, re-making the tree, which should not take too long, should

not require repeating all of the downstream mapping.) Additionally, the authors throughout the paper and in the methods refer to the tree as being constructed from genes, but that is highly unusual. Typically trees are constructed from protein sequences.

We thank the reviewer for their careful consideration of our work, and for their suggested improvements. Other reviewers also raised some concerns with our phylogenies, which we have addressed with new analyses and clarifications in our revised manuscript. As suggested, we reconstructed a broad, comprehensive phylogeny that now includes subgroups IV and V (which were previously classified as SAR11) as well as orphan SAGs that were not contained within larger clades. We now include additional text in the manuscript regarding the evolutionary relationships between subgroups IV and V and other subgroups within the order Pelagibacterales. *We highlight that a comprehensive understanding of the relationships between subgroups I-III and IV requires additional, focused analyses that are beyond the scope of this work that is largely focused within the family* Pelagibacteraceae *(formerly subgroup I).*

We have corrected the text throughout the manuscript in order to clarify that protein sequences were used to construct the phylogenies. Thank you for catching that.

Overall the paper is well written and the figures are beautiful. I have a few additional comments and suggestions regarding clarity of figures and discussion of results which I detail below.

Thank you for the kind comments.

Specific Comments:

59- I suggest you add the number of proposed genera specifically here

Done (page 2, line 41).

62- this phrasing makes it sound like you recommend someone do this in the future, not that you do it here in this paper

We agree with the reviewer and have changed the phrasing to now more clearly articulate that this is what was accomplished in this study (page 2, lines 44-45).

141- in presenting your results in table 1, you could consider adding a brief statement as to what your results suggest is the best approach for future culturing attempts

While we understand and appreciate the motivation behind the reviewer's comment, offering appropriate and effective suggestions is complex and requires addressing considerations that may detract from the focus of the current study. The determination of favorable culture conditions is notoriously difficult and varies a great deal between different subclades, and final decisions for best approaches will ultimately depend on the motivation behind the future work (e.g. total strains vs.

diversity of strains, etc). After considering this suggestion at length, we finally concluded that a separate study that is dedicated to describing the best practices for the cultivation of SAR11 with lessons from prior efforts and the inclusion of additional investigators would be a more effective way to realize this important suggestion. Given the space considerations in our already dense manuscript, such a discussion would have unlikely reached its intended utility and impact. We hope that the reviewer agrees with our resolution.

234- can you add any statistical analysis to support cohesion within groups beyond eyeballing the heat map?

We thank the reviewer for this suggestion. Our revision now includes appropriate statistical tests that we describe in Methods (page 14, lines 655-660), Results (page 5, lines 198-201), and Supplementary Fig. 4.

237- again, can support with any statistics?

We have also addressed this, and revised our Methods (page 14, lines 655-660), Results (pages 5, lines 216-217), and Supplementary Fig. 5 to describe the analyses.

244 and Fig2- clarify how group 1 and 3 are distinguished. How is “low-latitude, high diversity” different than “low-latitude”? Confusing because “high-latitude, low-diversity” is explicitly labelled but “low-latitude” does not have any extra description.

We thank the reviewer for catching that. We have now changed the grouping to “low-latitude, low-diversity” in order to distinguish between the two low-latitude groups (Figure 2; pages 5, lines 212-217).

253- Can’t two phylogenetically distinct units form the same ecological unit? I think “ecological unit” is perhaps not the best word choice here- if I understand correctly you are trying to say that both ecology and phylogeny, but neither alone, distinguish all the SAR11 “units”, so it’s confusing to call the unit “ecological” when ecology can’t fully resolve it.

We agree, and thank the reviewer for pointing this out. We have rephrased this sentence in hopes of articulating a more accurate view of our sentiment on this topic while avoiding circular language (page 5, lines 225-228).

255- typo/grammar

No longer appears in the rephrased sentence (page 5, lines 225-228).

Fig 2- why not make the text at the bottom a little bigger, because you could extend the table to the left edge of the teal Pelagibacteriaceae label. It’s just pretty hard to read and you have white space there you could fill.

We appreciate the perspective, and have modified Figure 2 as suggested.

Fig 2- You should clarify how many metagenomes are in each group category, because it's easy to assume each value is 1 metagenome but you mapped almost 1500 (!!). A legend to the heatmap shading could also help clarify how you combined coverage/abundance values from different metagenomes, which isn't clear to me.

We thank the reviewer for pointing this out and agree that the original presentation could be misinterpreted as representing a single metagenome per column. To clarify the number of metagenomes included in each category, we revised the top section of the figure to display a heatmap that functions as a scale bar, indicating the number of metagenomes per column. We also added Supplementary Table 11, which lists the specific metagenomes included in each numbered column, along with the total number of metagenomes per column. Additionally, we updated the figure legend to clarify the data shown. The values in the figure represent detection values, defined as the proportion of nucleotides in a given genome covered at least 1x. Each column shows the average detection value across all metagenomes grouped into that category (as defined in Supplementary Tables 8 and 11).

280- suggest defining RED (rel evol divergence I assume, but this is not super common term and it does not show up on google easily)

We now define RED (relative evolutionary distance) in the text (page 6, line 252).

Fig 3. It seems like some panels are lacking labels

We agree with the reviewer. We have added labels to additional sections of Figure 3, as well as some other minor modifications to improve its clarity.

316- Is there some information on the extent, or significance, of the phylogenomic structure in Fig 3a?

We have updated the figure caption to indicate that the subtree of Ia.3.VI genomes is extracted from the full phylogeny shown in Figure 2 (page 28, lines 969-970) and have now included ultrafast bootstrap values at all nodes.

320- what are these "detection values"? I don't see units in the Fig 3 barplots either.

We have added the definition of detection values at its first occurrence in the text (page 5, lines 198-199), as well as the figure legend. These values refer to the proportion of nucleotides with at least 1x coverage from read mapping.

Fig 3b- I'm confused about this panel, although it looks very cool I'm not sure what's on it. At first glance I thought it was the contigs of the two sags along the "x-axis", and their coverage in the heatmap. But then there are also multiple metagenomes from each location, so maybe it is the sample on the "x-axis" and the abundance in the metagenome in the heat map? Adding a legend with units would help.

Fig 3b- this looks cool but if you stretched it into a straight line along the whole figure width it would be much easier to read. It's confusing because for example on the left

BATS is the bottom row but on the right BATS becomes the top row. In the end, interpretability should trump coolness I think, unfortunately. Also, then you could add an x-axis label which would clear up my comment above.

While we are fond of the aesthetics of the initial figure, we recognize that, at this point, we are probably over-familiar with this work and thus greatly appreciate the fresh input on this. We went ahead and arranged the genomes as two horizontal lines as suggested, and added a legend for the coverage values. This is now panel e, as we tried to be more thorough in segregating and explaining each of the individual elements of the concatenated figure. We have also included a more descriptive legend for Figure 3 in order to provide additional context and clarity (page 28, lines 968-982).

Fig 3x- not sure what panel to call this, but can you put station ALOHA on the map as well? Since it is in the panel labelled b. How do these stations relate to the metagenome "groups" in Fig 2, that could be clarified.

We have revised the maps and added Station ALOHA to clarify its location with respect to Kāneʻohe Bay. This is now specifically identified as panel d. We have also added text to indicate how these stations relate to the metagenome groups in Figure 2 (page 7, lines 283-291).

336- typo/grammar

Corrected.

339- I'm finding it confusing that in figures the labels are Kane'ohe Bay but in the text the abbreviation KByT is used.

We have revised the manuscript text and figures to use "Kāneʻohe Bay" in place of 'KByT' in all instances except a few cases where referencing the time-series study specifically was more appropriate.

379- confidence in the tree could help support these group breakdowns

We have added bootstrap support values to the main phylogeny in Figure 1 as well as the condensed phylogeny of Figure 4.

Fig 4. I would swap order of Type material and Subgroup, so Type is on the right

Done.

Fig 4- your superscripts are very hard to see/read/distinguish. Consider bold and underline instead? Also, why put the superscript in the Species column- wouldn't putting it in the Type column make more sense?

We replaced the superscripts with bold and underlined font to more clearly distinguish the relevant species, and have made these annotations in the "Type" column as suggested.

409- instead of deep/shallow I think fine/coarse is more clear and unambiguous.

Done (page 8, line 358-359).

410- a similar combination of phylogeny and biogeography was used to refine taxonomic groups in abundant freshwater lakes heterotrophs, albeit for full length 16S, in 2012. I see a lot of parallels here, you could consider pointing to that as an example of the utility of this approach as it has greatly expanded our understanding of freshwater ecology. (https://journals.asm.org/doi/10.1128/mmbr.00028-10 , has about a million citations now)

We thank the reviewer for highlighting this citation, which we have now included in the Discussion (page 8, lines 364-366). There are notable parallels between our work and that of Newton et al. (2011), particularly in the shared emphasis on integrating ecological information into the framework. We also concur with their view that such frameworks should evolve dynamically as new data become available. In that spirit, we hope this publication, like that of Newton et al. (2011), will serve as a useful guide for future efforts in this continually developing field.

420- this citation (Zhao et al., 2024) must be unlinked to your reference manager because it is in a different format

Thank you for catching that - it is now fixed (page 8, line 372).

424- I don't understand how ecotypic differentiation is in conflict with recombination. Couldn't recombination be the mechanism leading to ecological differences? Why does ecotypic differentiation indicate speciation instead?

Zhao et al. argue that homologous recombination occurs frequently within SAR11, even between DNA with relatively low sequence similarity. They suggest that this process plays a major role in maintaining genetic cohesion, and is responsible for the lower average nucleotide identity (ANI) gap observed between SAR11 species (86–91%), compared to the typical 95% threshold used for other bacteria. In contrast, they found that non-homologous recombination, which can lead to ecological diversification, was relatively rare.

We offer a different perspective via our study. While Zhao et al. propose that homologous recombination can preserve species-level cohesion even at low sequence identity (<91%), our findings suggest that ecological diversification can still be resolved across a continuum of genomic similarity, including above the ANI gap of 86-91%. This allowed us to define genus-level groupings as well as detect instances of ecological speciation among closely related SAR11 genomes. In other words, the cohesive effect of homologous recombination does not appear strong enough to counteract the forces of environmental selection that drive ecological

differentiation. We also propose that non-homologous recombination, although less frequent, may introduce novel genetic material into populations. If these genetic changes confer a selective advantage, they may spread and contribute to ecological diversification.

432- I don't see how time series further this. In fact, I would think more spatial resolution would be more clarifying. After all, you have "global" metagenomes but they are just a few sites in total, and wouldn't spatial span more environmental variables?

We fully agree that increasing the spatial resolution of global ocean sampling for marine microbial metagenomes offers many important benefits. However, the specific point we aim to emphasize here relates to understanding the genetic determinants that drive either competitive exclusion or co-existence among SAR11 populations. In our view, geographically constrained time series studies offer a powerful approach to uncover these genetic determinants. Our perspective is informed by our group's experience working in such a system (Ramfelt et al. 2024; Tucker et al. 2025; Tucker et al. 2024; Tucker et al. 2021; Ramfelt et al. 2025). In this example, the Kāne'ohe Bay Time-series focuses on the well-defined and environmentally constrained system of Kāne'ohe Bay, where samples have been collected regularly for over eight years. Within this limited geographic area, we have observed that SAR11 populations separated by only a few nautical miles can exhibit distinct genomic profiles coinciding with sharp environmental gradients (Tucker et al., 2024). The limited spatial scale helps constrain environmental variability, allowing for more precise investigation of the metabolic and ecological factors that differentiate co-existing populations of Pelagibacterales.

458 – how is your P and N story distinct from the companion paper? Maybe a sentence to mention in the discussion.

This study highlights a unique case in which an isolate genome from Kāne'ohe Bay and a single-amplified genome (SAG) from the Bermuda Atlantic Time-series Station within one genus (Ia.3.VI) each contain genes that appear to be specific to the environments in which they were collected. By examining differences in ecologically relevant gene content, we demonstrate potential fine-scale ecological speciation between these closely related genomes. We interpret regional differences in nitrogen and phosphorus availability between the Pacific and Atlantic Oceans as a likely driver of these genetic distinctions, especially in genes involved in phosphorus and nitrogen acquisition. Our focus in this analysis is on two specific pathways: the C–P lyase pathway and type IV pilus assembly.

In parallel work (Tucker et al., 2024), we identified key metabolic traits and selective pressures that drive the distinct coastal and offshore distributions of co-occurring Pelagibacteraceae genera in the tropical Pacific. That study presents a broader analysis of gene content related to nutrient acquisition, organic carbon utilization, and cellular stress responses. The results reveal consistent differences between nearshore and offshore genera, suggesting that biogeochemical variation,

particularly in nitrogen and organic carbon availability, has played a role in shaping ecological divergence in gene content at the genus level.

Together, these investigations support our proposed taxonomic framework by showing that ecologically meaningful traits and selective pressures align with genus-level distinctions within the Pelagibacteraceae.

We thank the reviewer for their suggestions. To reflect this connection, we have added some text to our Discussion (page 9, lines 410–413) emphasizing how these complementary studies contribute to a broader understanding of SAR11 niche partitioning through a metabolism-focused lens.

530- primers 515F-926R is V4-V5 not V4. Also, did you use Parada's 515F-Y or 515F-C, Parada has two 515's in the paper. This is why you should include the actual primer sequence in parentheses as well, because it gets super confusing with citations alone. Also it's likely that you used the Walters version of the Parada primers. I would just double check this with whoever performed the amplicon sequencing.

We are thankful for the attention of our reviewer. We have added the primer sequences to the manuscript (page 11, lines 483-487) and clarified that we used the 515F-Y primer targeting the V4-V5 region. We can confirm that these are the primers from Parada et al., 2016, which is also listed in Table 1 as the reference for these primers in the Walters publication.

530- If you went back and did full length 16S of your isolates that could extend the impact of your phylogeny to 16S amplicon surveys as well.

We added a folder on FigShare with all available full-length 16S rRNA gene sequences from these genomes (80 available; https://doi.org/10.6084/m9.figshare.29983285; page 15, lines 678-683). We previously determined that the amplicon surveys fail to distinguish all of the genera we describe here [e.g. Litoralipelagibacter (Ia.3.III) and Undatipelagibacter (Ia.3.VI)]. While certainly of value, our conclusion is that the 16S rRNA gene cannot be relied upon for definitive genus-level resolution in SAR11 bacteria.

598- But you are making a comprehensive tree to clarify the SAR11 phylogeny; why would you leave them out? You should include them, and then tell us with confidence are they part of SAR11 or not.

We appreciate the reviewer's comment and acknowledge that other reviewers also raised concerns regarding our phylogenomic analyses. There is a long history of studies attempting to resolve the evolutionary relationship between subgroup V and SAR11 subgroups I-III. Based on some of the most rigorous and sophisticated analyses of alphaproteobacterial evolution available (e.g., Muñoz-Gómez et al., 2019), we are confident that subgroup V is not specifically related to subgroups I-III. As such, we do not believe that repeating those analyses here would contribute meaningfully to the specific focus of this manuscript.

In contrast, subgroup IV has not been subjected to the same level of scrutiny, largely due to a historical lack of representative genomes. While we successfully isolated and sequenced the genome of a representative of subgroup V (HIMB59) from Kāneʻohe Bay nearly two decades ago (Brandon 2006; Thrash et al., 2011; Grote et al., 2012), environmental genomes from subgroup IV have only recently become available (Haro-Moreno et al., 2020). Although existing genomic and phylogenetic data suggest that subgroup IV is also not part of the SAR11 lineage (Haro-Moreno et al., 2020), the evidence is less definitive than for subgroup V.

While resolving the relationship between subgroup IV and SAR11 subgroups I-III is certainly of scientific interest, the precedent set by subgroup V shows that doing so in a definitive manner would require extensive additional analyses beyond the scope and resources of our current study. Importantly, we want to emphasize that this question, though relevant to broader phylogenetic discussions, is tangential to the central focus of our work and does not alter or diminish the core findings presented in our manuscript.

602- shouldn't you be building trees using proteins not genes?

Our apologies; all of these trees were built with proteins and not genes and we have modified the text to clarify this point.

621- what was your outgroup? This is important, and should further be chosen appropriately to decide if those other groups are SAR11 or not (line 598)

Lines 592-593 of the original manuscript describe the outgroup. To investigate the robustness of our family- and genus-level groupings, we first reconstructed four different phylogenies that included subgroup IV and V genomes. The phylogenies constructed were: (1) with an outgroup that included isolate genomes from the order Pseudomonadales*; (2) with an outgroup that included ten broadly dispersed isolate genomes from multiple orders within the* Gammaproteobacteria*; (3) with the original outgroup of our submitted manuscript (ten genomes from the Rhodobacteraceae); and (4) with no outgroup.*

To further assess the influence of outgroup selection on tree topology, we applied the Minimal Ancestor Deviation (MAD) method to evaluate rooting stability based on a suggestion from Reviewer 3. When we ran MAD on the dataset without an outgroup, and with prior knowledge that subgroups IV and V are likely basal, we found that these groups were placed on long internal branches, raising concerns about the reliability of this rooting (see additional figures on figshare: 10.6084/m9.figshare.28127633). To cross-validate, we also applied midpoint rooting, which placed subgroup V as the outgroup. Importantly, across all phylogenies, the genus- and family-level groupings identified in this study remained consistent regardless of the outgroup strategy.

We hope that this additional analysis provides confidence in the robustness of our taxonomic framework for SAR11. We have updated the Methods section to indicate that we used the midpoint rooting approach (pages 12-13, lines 578–580).

626- Why are you using a pruned phylogeny, why not just include confidence values in the branches and then limit downstream analysis to the major groups only. You shouldn't exclude genomes from your tree just because you don't have a lot of them, the tree structure should be shaped by the full diversity. Did any 95% ANI clusters disagree with the tree structure? Why not make the tree, then do the whole-genome ANI clusters and show mapping for the branch-level that each 95ANI maps back to?

Here again, aesthetics may have got in the way of clearly articulating the message of this analysis. In our initial phylogeny, we did include all relevant genomes in order to create a comprehensive and unpruned phylogeny. Genomes were only excluded to create a pruned phylogeny in an attempt to provide a "clean" view of the genus-level groupings, and to highlight the genomes used in the read recruitment analyses. Confidence values were included on all of the supplemental phylogenies, but we agree with the reviewer that they should be more apparent. We now include all of the genomes in a revised Figure 1, as well as confidence values. We did not have any 95% ANI clusters that disagreed with the tree structure.

Reviewer #1 (Remarks on code availability):

The NCBI BioProject is a placeholder (line 1108- "NCBI associated with BioProject XXXXXXX")

The NCBI BioProject has been updated (page 15, line 680). All of the final genome sequences and underlying raw data (via the SRA) are now publicly available and the relevant accession numbers have been added to Supplementary Table 3.

It looks like it does exist on Figshare though, but I suggest they use the figshare DOI to cite instead of the in-text link.

We have updated the BioProject number as noted above, and also include the FigShare DOI as reference.

There is some code and data available on github. It's certainly not all the code used in the analyses but it looks like figure code is there. I guess most stuff was part of the ANVI'O workflow though, so maybe there's not actually a lot of free-standing code with the paper analysis. I did not look in detail, just glanced through.

We can confirm that an extensive amount of the code was from anvi'o. We have reviewed the code available on GitHub to ensure that all code used in the analyses are included.

We thank the reviewer again for the input and constructive criticism.

Reviewer #2 (Remarks to the Author):

The manuscript submitted by Freel et al present an original and impactful study investigating the ecology and evolution of SAR11 marine bacteria. The researchers have significantly expanded the cultured representation of SAR11 diversity, investigated biogeography through genomics and metagenomic fragment recruitment, identified ecologically distinct clades of SAR11 and then proposed a taxonomic framework and naming convention for SAR11… but I guess we don't call it that anymore!

The study is well executed and the paper is well written.

We thank the reviewer for their positive feedback.

My only main concern is the decision to remove genomes from their original phylogenetic tree. The researchers generated a phylogeny of all 481 genomes and then removed "SAGs that did not fall into a 90% gANI cluster of at least three genomes…" Second, they dereplicated remaining genomes using a 95% gANI cutoff. This left 268 SAR11 genomes, which are the genomes that are also included in the final taxonomic framework and naming convention. I am okay with this sort of pruning and dereplicating in phylogenetic analysis and it will certainly give you a "cleaner" and better resolved tree with nice well supported clades. However, removing this "messiness" might have implications later. So, I would be curious to see what happens to the phylogeny and the taxonomic structure if these genomes are added back in. Does it impact the monophyly of the named genera for example? Also, is it possible to include some bootstrap values on the phylogenies. It would be nice to know how well supported (or not well supported) many of these clades (as well as they deeper relationships) are.

We appreciate the reviewer's comments. Similar perspectives regarding the main phylogeny were shared by the other reviewers. The phylogeny shown in the revised Figure 1 now includes all 484 genomes and no dereplication. It also includes bootstrap support values, which we had previously included only in the supplemental figures for aesthetic reasons. Our original decision to dereplicate the tree was also driven primarily by aesthetic considerations, but informed by analyses that showed that it did not affect the monophyly of the genera we name in this study. It also did not affect the bootstrap support found at the deeper, family-level nodes of the tree.

Reviewer #3 (Remarks to the Author):

In their study, Freel et al. generated an extensive collection of SAR11 genomes from cultures and combined these with publicly available genomic data to reconstruct the evolutionary and ecological units within this lineage. The study is highly relevant given the global distribution of SAR11 and its distinctive evolutionary history, which challenges traditional taxonomic classification methods based n cutoffs. Below, I provide several

suggestions on the phylogenomics part of the study, that I hope will help the authors refine and strengthen their study:

Lines 162–163: The presence/absence of genes is generally not a robust criterion for selecting marker genes within a lineage. Since the study is grounded in the vertical evolutionary history of the clade, relying on markers that may have experienced complex horizontal gene transfer events could obscure or distort the phylogenetic relationships. This might explain why subclades Ia and Ib form a monophyletic group in the analysis.

We appreciate the sentiment of this comment. However, it is important to clarify that we first started with a robust set of genes defined specifically for the phylogenetic analysis of alphaproteobacterial genomes (Wang and Wu 2013; Muñoz-Gómez et al., 2019). We then refined this gene set to focus specifically on genes that are informative for high-quality SAR11 genomes. This gene set is predominately single-copy and rarely laterally transferred (Wang and Wu 2013).

Regarding the monophyly of subclades Ia and Ib. We appreciate that the reviewer is bringing this up, since it represents one of the strengths of our work. The expectation of monophyly within Ia and Ib is an example of historical labels impeding the rational and systematic categorization of sub-entities within the SAR11 lineage. When 16S rRNA gene databases for environmental marine bacteria began expanding in the early 1990s, it became clear that "strategic" sequence selection was required to maintain Ib as a separate entity from Ia. Even so, it has always been difficult to resolve Ib to a monophyletic clade, as opposed to several separate lineages that branch near the base of Ia. The outcomes of genome-based phylogenies of SAR11 now have sufficient coverage to mirror those early rRNA-based studies (e.g. Haro-Moreno et al. 2020). After extensive analysis of the Ia-Ib relationship, we are confident that (i) the genomes considered to be part of the Ib subgroup do not form a monophyletic lineage on their own, and (ii) the most parsimonious approach, based on the largest available genome dataset, combines Ia and Ib into a single monophyletic lineage.

Lines 169–170: The assignment of families appears somewhat arbitrary and is largely based on previous classifications that define SAR11 as an order. However, the authors have compelling results that could support the development of a more consistent and evolutionarily grounded framework for defining taxonomic levels within SAR11. I particularly appreciate the approach of integrating ecological and evolutionary evidence. The same comment applies to the proposed genera shown in Figure 4.

While we thank the reviewer for their perspective here, we do respectfully disagree with the sentiment that the taxonomic assignments we propose are arbitrary and largely based on previous classifications.

The primary impetus of this study is very much focused on resolving the intersection between ecology and genome evolution at the fine tips of the SAR11 phylogeny. While a systematics for SAR11 was not our original intention, we have now spent

considerable effort in resolving what we have come to propose as SAR11 genera, as noted in Figure 4. Much of our manuscript illustrates how these genera are both ecologically consistent and firmly grounded in a robust evolutionary framework. Prior to our study, there has been broad and pervasive inconsistency in the assignment of taxonomic ranks to different levels of diversity within the SAR11 lineage. Whereas the GTDB has attempted to remove the arbitrary aspect of bacterial taxonomy by applying a set of rules based on genome relatedness to assign taxonomic levels, we show that these rules do not fit when one considers ecology as a relevant consideration of the taxonomic outcome. Thus, for a taxonomic hierarchy to be meaningful and useful for SAR11, some human intervention to meld the ecological and evolutionary aspects is certainly required, and we are not of the opinion that this is an arbitrary endeavor.

It is entirely consistent with other orders within the Alphaproteobacteria to consider the SAR11 lineage as the order-level Pelagibacterales. This has been well-established in robust phylogenies (e.g. Muñoz-Gómez et al., 2019). Ours is the first study to propose four families within this order, which are wholly consistent with both evolutionary history and ecological information.

Figure 1: I recommend testing outgroup-independent rooting methods, such as Minimal Ancestor Deviation (MAD), to evaluate the influence of the chosen outgroup on the topology and stability of the resulting groups.

We used the Minimal Ancestor Deviation (MAD) method to evaluate the influence of the outgroup on the stability of the overall phylogeny, as suggested. In this figshare folder (https://doi.org/10.6084/m9.figshare.28127633), we include the phylogeny generated using the MAD method when no outgroup was included; however, the clades previously referred to as SAR11 subgroups V and IV were included. We know in this case that SAR11 subgroup V genomes should be identified as the outgroup, the relationship between this group and the Pelagibacterales has been extensively studied*. However, this was not the phylogeny that the MAD method yielded. The genomes in the expected outgroup (purple box) were distributed as long internal branches.*

To further explore outgroup-independent rooting methods, we then implemented a mid-point rooting method to root this phylogeny (figshare: https://doi.org/10.6084/m9.figshare.28127633). While this method can also be biased if an evolutionary rate is miscalculated, it is a simple algorithm and in common use. In this case, "SAR11 subgroup V" was established as an outgroup as expected from previous literature.

While we are only now becoming familiar with the MAD method, we note that the root identified using the midpoint method aligns with what we know regarding SAR11 evolution. Most crucially, the midpoint rooted tree demonstrates that the genera we identified in our study, using genomes of the Rhodobacteraceae *as an outgroup for the phylogenies, remain stable. (trees available in figshare: https://doi.org/10.6084/m9.figshare.28127633).*

Lines 640–652: I find it surprising that Chloroflexota genomes were used as the outgroup, considering the considerable phylogenetic distance between this phylum and SAR11. This disparity increases the risk of long-branch attraction artifacts, which could distort tree topology and affect the RED estimates. In addition, different marker genes (120) are used to get this estimates. Why aren't the authors using the marker genes used to build the SAR11 tree?

Our study used the default workflow for determining RED scores that is provided by the GTDB-Tk ("de_novo_wf"). As defined in Parks et al., 2018, the aim of this tool is to lend objective taxonomic assignments of genomes by placing the genomes in domain-specific protein reference trees, which relies on the bac120 marker gene set.

*We followed previous guidance provided by the authors of this workflow and chose "*p__Chloroflexota*" as the outgroup (see GitHub issue: https://github.com/Ecogenomics/GTDBTk/issues/120), this issue was also referenced in a more recent question regarding the election of outgroup (see GitHub issue: https://github.com/Ecogenomics/GTDBTk/issues/390). These standardized methods are used for classification through GTDB and our intent was to replicate this method in order to assess how standardized RED scores align with our ecoevolutionary framework.*

In phylogenies constructed with a variety of different outgroups, we can confirm that the genera identified here remain stable within the Pelagibacterales. No other distortion was evident. While we fully appreciate that the choice of outgroup can distort the phylogeny and impact order of branching, we note that the RED scores themselves are generated independent of the outgroup.

We thank the reviewer for their careful consideration of our study and their valuable suggestions.