

Widespread horizontal transfer and strong selection enhance microbial adaptation in Antarctic soils

Corresponding Author: Professor Chris Greening

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)

In this manuscript, Peng et al. provide a detailed set of analyses characterizing the prevalence of horizontal gene transfer (HGT) for metabolic adaptability as well as the selective pressure for conserving functionality in critical metabolic genes in the microbial communities of Antarctic soils. The study represents an excellent follow-up to Ortiz et al. (2021, PNAS) on the same environmental sample set, which helped establish the current understanding of the importance of high affinity trace gas oxidation ("aerotrophy") in polar desert soils. This study focuses on a dereplicated set of medium/high quality MAGs, enabling a rigorous investigation of gene flow.

The authors determined that genes related to aerotrophy are both widespread across MAGs and widely transferred between MAGs via HGT within and among phyla. Further analysis of these genes showed that there is an apparent "purifying" selective pressure resulting in a conserved ligand-binding site and presumed function that decreases away from catalytically critical and buried moieties. This evidence supports the conclusion that the microbial communities in Antarctic soils are highly adaptable, enabling them to withstand the multi-extreme conditions of their environment and potentially respond to new rapid changes like climate.

The manuscript is incredibly thorough and well-written, including meticulous methodology and excellent visuals for displaying the complex genomic data. The topic of aerotrophy, particularly in extreme environments, continues to grow in relevance and importance as a major metabolism for microbial survival and global biogeochemical cycling. I believe this study significantly advances this field, and I highly recommend it for publication. My review below includes some general questions and suggestions as well as specific, line-by-line comments.

General comments:

I am curious if the authors have any possible explanations or additional thoughts as to why high-affinity methanotrophy does not follow the same trends as H₂ and CO oxidation in this environment. It is common for H₂ and CO oxidation to co-occur in similar arid communities while CH₄ can be more variable. However, previous studies have identified genomic potential for high-affinity methanotrophy in Antarctic soils (Dragone et al. 2022, mSystems; Edwards et al. 2017; Genome Announcements), and atmospheric CH₄ oxidation has been observed in high-latitude Arctic soils (Lau et al. 2015, ISME J; Stackhouse et al. 2017, Geobiology). PmoA is present in some of the metagenomes from this site based on Fig. 2 of Ortiz et al. (2021, PNAS) but seems to be absent from any of the MAGs based on Fig. S4 of this manuscript. I assume that is the primary limitation for this discussion?

Variation in the environmental properties of the soils is mentioned briefly, but not explored further. I recognize the manuscript is already dense and an in-depth analysis of the relationship of the described data with these parameters is likely beyond the scope here, but were geochemical/physicochemical data factored in as a control at any point for statistical purposes? Properties like pH and ion abundances could play a significant role in metabolic preferences and protein structural requirements, which may partially affect some of the observed variation in hydrogenase gene exchange and protein structure.

Specific comments:

Line 114 – Is this mainly due to the fact that Actinos are the most abundant phyla? Could also be interesting to give stats for HGT events per phyla normalized to abundance.

Line 130 – From context, I assume “mobilome” is the collection of MGEs?

Lines 132-133 – It would be useful to provide the numbers for mobilome/MGEs in desert vs. non-dryland soils in the text, just for establishing frame of reference.

Lines 173-178 (and Fig. S4) – Was metabolic pathway completeness within the MAGs also assessed in addition to the primary marker genes (e.g., presence of other gene subunits and intermediates)?

Lines 182-184 – Phyla are not displayed in Fig 2a. Could an additional Fig/Table also be cited here which shows that information for the nine genes of interest? Maybe Fig 2b, which covers eight of the stated nine phyla, focusing on the aerotrophy genes specifically.

Lines 195-197 (and Fig. S6) – Could the degree of incongruence be affected at all by the comparison of MAGs with different quality levels? Specifically, if certain MAGs in the species tree were less complete and, therefore, had fewer SCGs to use for the alignment the tree is based on, potentially resulting in less accurate tree placement.

Lines 243-244 – See general comments. Interesting potential for follow-up analyses beyond this manuscript, regardless.

Line 337 – Just a question of language choice. Does this study prove that the communities are continually metabolizing? They exchange the encoded genetic potential, but active metabolism has not been demonstrated here.

Lines 403-406 – What is the rationale for the different identity thresholds? Some seem lower than I would expect.

Reviewer #2

(Remarks to the Author)

This manuscript by Peng et al. investigates the evolutionary forces, including horizontal gene transfer and natural selection, shaping microbial adaptation in Antarctic soils. The authors employ a comprehensive and integrative framework that combines short- and long-read metagenomic sequencing, population genomic analyses, and functional annotation and prediction. The manuscript is clearly written, and the data visualization is strong and informative. The findings regarding the roles of horizontal gene transfer and purifying selection acting on aerotrophy-related genes are particularly compelling. However, I have several major concerns that warrant careful consideration.

Major comments:

1. The sample size is relatively small, consisting of 16 previously collected samples and one newly generated sample. It is unclear where the sampling sites are located and how representative these sites are of broader Antarctic soil ecosystems. The authors should clarify the geographic coverage and sampling rationale to justify the generalizability of their conclusions. In addition, key findings related to horizontal gene transfer and selection would be strengthened by validation using publicly available Antarctic metagenomic datasets.

2. The use of pN/pS to infer selection across a dataset spanning multiple phyla is not appropriate because pN/pS is a population-level metric designed to quantify standing polymorphism within a relatively coherent evolutionary unit. When applied across deeply divergent taxa, the signal becomes confounded by phylogenetic distance, fixed substitutions accumulated over long evolutionary timescales, differences in mutation spectra, codon usage, recombination rates, and demographic histories. As a result, the ratio no longer reflects recent or ongoing selection but instead mixes polymorphism with divergence, leading to biased or uninterpretable estimates. For cross-phyla comparisons, a phylogenetically informed framework such as dN/dS estimated using codon-based substitution models along reconstructed gene trees would be more appropriate, as it explicitly accounts for evolutionary history and fixed substitutions. Alternatively, pN/pS analyses should be restricted to within-species or within-phylum population clusters where the assumptions of the metric are met.

3. It is not surprising that purifying selection is widespread in microbial communities in Antarctic soils, as strong functional constraint is a pervasive feature of microbial genomes across environments. In most bacteria and archaea, the majority of coding genes are under purifying selection. Therefore, observing widespread purifying selection does not in itself demonstrate environment-specific adaptation; it largely reflects baseline evolutionary constraint common to microbial life. In extreme environments such as Antarctica, where energy availability is limited and growth rates are slow, functional conservation may be more pronounced because organisms depend on tightly optimized metabolic and stress-response systems. However, to substantiate this claim, a comparative analysis of the degree of purifying selection between Antarctic soils and other ecosystems is necessary. Without such a baseline comparison, it is difficult to determine whether the observed patterns reflect environment-specific evolutionary pressure or the general background constraint typical of microbial genomes.

4. To better substantiate claims of adaptive evolution, the authors should also investigate signals of positive selection. This can be done using phylogenetically informed dN/dS analyses with codon-based models, including branch and branch-site models implemented in tools such as PAML or HyPhy. These analyses would provide a more balanced and rigorous

assessment of evolutionary forces beyond the expected background of purifying selection.

5. The code used for data processing and analysis is not available, which limits reproducibility and transparency.

Other comments:

Line 128: It is not clear how MGE count was measured for each phylum (e.g., if it was standardized by the number of MAGs and the size of genome).

Lines 137-138: Need a statistical test to assess the correlation between the average number of detected transfers and co-occurrence coefficients.

Lines 227-230: which statistical test support that the pN/pS for aerotrophy is significantly lower than sample-average pN/pS? Also, p value needs to be provided.

Lines 230-235: same comment for Lines 227-230.

Line 390: how were the 43 conserved single-copy genes selected?

Lines 456-457: MGEs specified here are not consistent with Fig. S3. For example, plasmids and phages are missing.

Version 1:

Reviewer comments:

Reviewer #1

(Remarks to the Author)

All of my questions and comments from the first review have been thoroughly addressed through the additions in the main text and supplemental material. I recommend this manuscript for publication and commend the authors for their work.

Reviewer #2

(Remarks to the Author)

The authors have satisfactorily addressed my comments. The manuscript has been improved accordingly. I have no further substantive comments and recommend acceptance of the manuscript.

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

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

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

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

REVIEWER COMMENTS

Reviewer #1 (Remarks to the Author):

In this manuscript, Peng et al. provide a detailed set of analyses characterizing the prevalence of horizontal gene transfer (HGT) for metabolic adaptability as well as the selective pressure for conserving functionality in critical metabolic genes in the microbial communities of Antarctic soils. The study represents an excellent follow-up to Ortiz et al. (2021, PNAS) on the same environmental sample set, which helped establish the current understanding of the importance of high affinity trace gas oxidation (“aerotrophy”) in polar desert soils. This study focuses on a dereplicated set of medium/high quality MAGs, enabling a rigorous investigation of gene flow.

The authors determined that genes related to aerotrophy are both widespread across MAGs and widely transferred between MAGs via HGT within and among phyla. Further analysis of these genes showed that there is an apparent “purifying” selective pressure resulting in a conserved ligand-binding site and presumed function that decreases away from catalytically critical and buried moieties. This evidence supports the conclusion that the microbial communities in Antarctic soils are highly adaptable, enabling them to withstand the multi-extreme conditions of their environment and potentially respond to new rapid changes like climate.

The manuscript is incredibly thorough and well-written, including meticulous methodology and excellent visuals for displaying the complex genomic data. The topic of aerotrophy, particularly in extreme environments, continues to grow in relevance and importance as a major metabolism for microbial survival and global biogeochemical cycling. I believe this study significantly advances this field, and I highly recommend it for publication. My review below includes some general questions and suggestions as well as specific, line-by-line comments.

Response: We thank the reviewer for the time and effort devoted to evaluating our manuscript. We sincerely appreciate the positive feedback on the significance and quality of the study, as well as the support for publication of this manuscript.

General comments:

I am curious if the authors have any possible explanations or additional thoughts as to why high-affinity methanotrophy does not follow the same trends as H₂ and CO oxidation in this environment. It is common for H₂ and CO oxidation to co-occur in similar arid communities while CH₄ can be more variable. However, previous studies have identified genomic potential for high-affinity methanotrophy in Antarctic soils (Dragone et al. 2022, mSystems; Edwards et al. 2017; Genome Announcements), and atmospheric CH₄ oxidation has been observed in high-latitude Arctic soils (Lau et al. 2015, ISME J; Stackhouse et al. 2017, Geobiology). PmoA is present in some of the metagenomes from this site based on Fig. 2 of Ortiz et al. (2021, PNAS) but seems to be absent from any of

the MAGs based on Fig. S4 of this manuscript. I assume that is the primary limitation for this discussion?

Response: The *pmoA* gene is indeed present in the metagenomic assemblies but at relatively low abundance compared to genes responsible for H₂ and CO oxidation. However, it also appears to be restricted to certain Antarctic sites, such as Mount Seuss 4 and Mount Seuss 5 in the study of Ortiz et al., 2021, PNAS (Fig. 2), and is only abundant in two of nine East Antarctic samples in the study of Hutchinson et al., 2026, ISME J (Fig. 2b). Among the 676 MAGs recovered in this study, *pmoA* was detected in only a single MAG affiliated with *Gammaproteobacteria*. This limited recovery constrained our ability to perform HGT and selection analyses for this gene, and therefore it was not further discussed in the manuscript.

Our recent work indicates that methane oxidation occurs in a narrower thermal range (4 to 30 °C), whereas H₂ and CO oxidation show broader temperature ranges (-20 to 50 °C and -20 to 42 °C, respectively; Hutchinson et al., 2026, ISME J). The data in Hutchinson et al.'s recent paper suggests that high-affinity methane oxidation is also more spatially restricted, observable in coastal regions with comparatively milder conditions but not in soils from higher elevation inland mountains, which could help explain the observed differences in distribution and trends relative to H₂ and CO oxidation.

Variation in the environmental properties of the soils is mentioned briefly, but not explored further. I recognize the manuscript is already dense and an in-depth analysis of the relationship of the described data with these parameters is likely beyond the scope here, but were geochemical/physicochemical data factored in as a control at any point for statistical purposes? Properties like pH and ion abundances could play a significant role in metabolic preferences and protein structural requirements, which may partially affect some of the observed variation in hydrogenase gene exchange and protein structure.

Response: In the revised manuscript, we have now included an exploratory principal component analysis (PCA) to examine the relationship between pN/pS^(gene) of hydrogenases and physicochemical variables. This analysis revealed that sample-averaged pN/pS^(gene) of [NiFe] group 1l hydrogenases is significantly associated with multiple key soil properties, including carbon and nitrogen status, electrical conductivity, and pH, whereas [NiFe] group 1h showed a significant association only with phosphorus, suggesting stronger environmental sensitivity of group 1l. Together, these findings suggest that environmental properties may contribute to shaping selection patterns across hydrogenase genes, although a comprehensive mechanistic understanding of these effects is beyond the scope of the present study. These results have now been incorporated into the revised manuscript. Changes made:

Results and Discussion (L262-264): Notably, pN/pS^(gene) of group 1l hydrogenases showed strong associations with multiple environmental properties across samples, including soil carbon and nitrogen status, electrical conductivity, and pH (**Fig. S16 and Table S15**).

Methods (L561-564): Principal component analysis (PCA) was performed using the vegan package in R based on gene-level pN/pS values of [NiFe] group 1h and 1l hydrogenases. Environmental variables were standardized and fitted to the ordination using the ‘envfit’ function with 999 permutations.

Specific comments:

Line 114 – Is this mainly due to the fact that Actinos are the most abundant phyla? Could also be interesting to give stats for HGT events per phyla normalized to abundance.

Response: We agree with the reviewer that this pattern may be driven by the high abundance of *Actinomycetota* MAGs in our dataset. Following the reviewer’s suggestion, we quantified HGT events per phylum normalized by relative abundances. Specifically, HGT counts between phyla were normalized by the product of their relative abundances: $HGT_{norm}(i, j) = HGT(i, j) / (A_i \times A_j)$, where $HGT(i, j)$ is the number of horizontal gene transfer events between phyla i and j , and A_i and A_j represent the relative abundances of the two phyla. This analysis shows that several abundant phyla, including *Pseudomonadota*, *Chloroflexota*, *Acidobacteriota*, *Planctomycetota*, and *Verrucomicrobiota*, exhibit higher-than-expected HGT as donors transferring to *Desulfobacterota_B*. These results have been added to the revised manuscript (Lines 114-119) and are further illustrated in Supplementary Figure 2b.

Methods (L482-485): To control for the effect of MAG abundance on HGT patterns, we normalized HGT counts between phyla by the product of their relative abundances: $HGT_{norm}(i, j) = HGT(i, j) / (A_i \times A_j)$, where $HGT(i, j)$ is the number of HGT events between phyla i and j , and A_i and A_j represent their relative abundances.

Line 130 – From context, I assume “mobilome” is the collection of MGEs?

Response: Yes. In this study, “mobilome” refers to the collection of mobile genetic elements (MGEs). We have now defined the term at its first occurrence in the revised manuscript (Lines 127-128).

Lines 132-133 – It would be useful to provide the numbers for mobilome/MGEs in desert vs. non-dryland soils in the text, just for establishing frame of reference.

Response: We thank the reviewer for this helpful suggestion. We have now added the quantitative comparison of mobilome genes between desert and non-dryland soils in the main text to provide clearer context. Changes made:

Results and Discussion (L135-137): consistent with observations of the high proportion of mobilome genes in desert soils worldwide, compared to non-dryland soils (~1.4-2.5× higher abundance)^{37,38}.

Lines 173-178 (and Fig. S4) – Was metabolic pathway completeness within the MAGs also assessed in addition to the primary marker genes (e.g., presence of other gene subunits and intermediates)?

Response: Our analysis primarily focused on marker genes, which encode essential (typically catalytic) subunits involved in major energy, carbon, and nitrogen acquisition processes and are widely used as indicators of metabolic potential. To provide additional context on metabolic pathway completeness within genomes, we have now assessed pathway/module completeness across all MAGs using DRAM, which estimates completeness based on the presence of pathway-associated genes. Several key metabolic modules were recovered at high completeness (>70%) in substantial fractions of MAGs, including pentose phosphate pathway (72.0%), F-type ATPase (61.7%), Complex I NADH:quinone oxidoreductase (46.9%), glycolysis (44.4%), the TCA cycle (40.8%), and the Calvin cycle (25.1%). The detailed metabolic pathway completeness of each MAG is provided in Supplementary Table 2. Changes made:

Methods (L441-442): To assess metabolic pathway completeness within MAGs, we used DRAM v1.4.6⁷⁸ to estimate pathway/module completeness based on the presence of pathway-associated genes (**Table S2**).

Lines 182-184 – Phyla are not displayed in Fig 2a. Could an additional Fig/Table also be cited here which shows that information for the nine genes of interest? Maybe Fig 2b, which covers eight of the stated nine phyla, focusing on the aerotrophy genes specifically.

Response: Thank you for the suggestion. We have now added a supplementary figure showing the HGT patterns of the nine genes of interest across phyla (Supplementary Figure 6).

Lines 195-197 (and Fig. S6) – Could the degree of incongruence be affected at all by the comparison of MAGs with different quality levels? Specifically, if certain MAGs in the species tree were less complete and, therefore, had fewer SCGs to use for the alignment the tree is based on, potentially resulting in less accurate tree placement.

Response: We agree that variation in MAG quality and completeness could, in principle, influence species tree reconstruction and thereby affect the inferred degree of incongruence between gene tree and genome phylogeny. To minimize this potential bias, we applied quality filtering prior to tree construction, retaining only MAGs above our completeness and contamination thresholds (completeness > 50%, contamination < 10%). In addition, for horizontal gene transfer inference, the species tree was inferred from concatenated single-copy core genes, and alignment columns that were sparsely represented across genomes were removed, reducing the influence of incomplete MAGs on tree topology. We have now added a clearer description of these criteria in the revised Methods section.

To further address this concern, we examined the distribution of 43 universal SCGs across the 676 MAGs used in this study and found that 458 out of 676 MAGs contained ≥ 30 SCGs (a median of 34 SCGs per MAG). This indicates that most MAGs contribute substantial phylogenetic signal to the species tree. Changes made:

Methods (L469-472): A corresponding species tree was generated using 43 universal single-copy genes (SCGs) as defined by CheckM⁶⁵, after removing columns in the concatenated alignment that were present in <50% of genomes and had an amino acid consensus <25%.

Lines 243-244 – See general comments. Interesting potential for follow-up analyses beyond this manuscript, regardless.

Response: As suggested in the general comments, we have added an analysis of the relationship between $pN/pS^{(gene)}$ and physicochemical variables. This revealed that selection pressure varies with environmental properties across samples, and these results have now been incorporated into the revised manuscript. Changes made:

Results and Discussion (L262-264): Notably, $pN/pS^{(gene)}$ of group 11 hydrogenases showed strong associations with multiple environmental properties across samples, including soil carbon and nitrogen status, electrical conductivity, and pH (**Fig. S16 and Table S15**).

Line 337 – Just a question of language choice. Does this study prove that the communities are continually metabolizing? They exchange the encoded genetic potential, but active metabolism has not been demonstrated here.

Response: We thank the reviewer for pointing this out and have revised the sentence to improve clarity. Changes made:

Conclusion (L356-358): This study provides critical evidence that Antarctic microbial communities are not static or inactive, as once thought, but rather metabolically capable and genetically adapting to persist and grow under polyextreme conditions.

Lines 403-406 – What is the rationale for the different identity thresholds? Some seem lower than I would expect.

Response: In our analysis, a default identity threshold of 50% was applied to most gene families to balance sensitivity and specificity. For certain genes or protein families, different thresholds were applied based on their sequence conservation. These thresholds were selected based on previous literatures (Ortiz et al., 2021, PNAS; Greening, et al., 2019, ISME J; Cordero, et al., 2019, ISME J) and preliminary benchmarking of our dataset to maximize detection of true homologs while minimizing negative hits. In addition, to ensure the reliability of our results, the short sequences (less than 20% of the average length for a given gene family) were further inspected using two approaches: (i) structural prediction to verify the presence of conserved functional domains, and (ii) phylogenetic analysis to confirm that they cluster appropriately within their respective gene families.

Reviewer #2 (Remarks to the Author):

This manuscript by Peng et al. investigates the evolutionary forces, including horizontal gene transfer and natural selection, shaping microbial adaptation in Antarctic soils. The authors employ a comprehensive and integrative framework that combines short- and long-read metagenomic sequencing, population genomic analyses, and functional annotation and prediction. The manuscript is clearly written, and the data visualization is strong and informative. The findings regarding the roles of horizontal gene transfer and purifying selection acting on aerotrophy-related genes are particularly compelling. However, I have several major concerns that warrant careful consideration.

Response: We thank the reviewer for their careful reading of our manuscript and for the constructive comments. We have carefully considered all the concerns raised and have addressed them in the revised manuscript.

Major comments:

1. The sample size is relatively small, consisting of 16 previously collected samples and one newly generated sample. It is unclear where the sampling sites are located and how representative these sites are of broader Antarctic soil ecosystems. The authors should clarify the geographic coverage and sampling rationale to justify the generalizability of their conclusions. In addition, key findings related to horizontal gene transfer and selection would be strengthened by validation using publicly available Antarctic metagenomic datasets.

Response: We acknowledge that the sample set is somewhat limited, but note that the McMurdo Dry Valleys are the largest ice-free region in Antarctica and therefore of considerable ecological significance. To strengthen our findings, however, we have incorporated nine additional metagenomic datasets from across East Antarctica (Hutchinson et al., ISME J, 2026). We have clarified the geographic coverage of the existing and new datasets by adding a supplementary figure showing the broad sampling locations (Supplementary Figure 21). Selection analyses on the additional datasets generated results consistent with those reported from the Mackay Glacier region. Changes made:

Results and Discussion (L234-238): Consistent patterns were observed in an independent set of nine Antarctic metagenomes from geographically distinct regions across East Antarctica (Dronning Maud Land, Robinson Ridge, and Bunger Hills), where aerotrophy genes also exhibited low pN/pS values indicative of strong purifying selection (0.09 on average; **Fig. S9 and Table S11**).

Methods (L526-530): To strengthen the generalizability of our conclusions, we incorporated nine additional Antarctic metagenomes from three geographically disjunct regions: Dronning Maud Land (n = 3), Robinson Ridge (n = 3), and Bunger Hills (n = 3) (**Fig. S21 and Table S1**)³⁴. The same population-level polymorphism analysis workflow was applied to assess the robustness and reproducibility of the observed patterns.

2. The use of pN/pS to infer selection across a dataset spanning multiple phyla is not appropriate because pN/pS is a population-level metric designed to quantify standing polymorphism within a relatively coherent evolutionary unit. When applied across deeply divergent taxa, the signal becomes confounded by phylogenetic distance, fixed substitutions accumulated over long evolutionary timescales, differences in mutation spectra, codon usage, recombination rates, and demographic histories. As a result, the ratio no longer reflects recent or ongoing selection but instead mixes polymorphism with divergence, leading to biased or uninterpretable estimates. For cross-phyla comparisons, a phylogenetically informed framework such as dN/dS estimated using codon-based substitution models along reconstructed gene trees would be more appropriate, as it explicitly accounts for evolutionary history and fixed substitutions. Alternatively, pN/pS analyses should be restricted to within-species or within-phylum population clusters where the assumptions of the metric are met.

Response: In our study, pN/pS was estimated independently within each metagenome-assembled genome (MAG) using read-mapping-derived variants, as implemented in *anvi'o*. The MAGs were dereplicated at 99% ANI and thus approximate strain-level populations, allowing pN/pS to reflect within-population polymorphism.

We acknowledge that comparing pN/pS across MAGs from different phylogenetic groups has inherent limitations, as differences in genomic background and evolutionary history may influence the estimates. Accordingly, we interpret pN/pS primarily as a relative measure of selective constraint within a consistent analytical framework, rather than as a direct comparison of selection strength across deeply divergent taxa.

To address this limitation, we have incorporated complementary phylogeny-aware analyses on phylogenetically coherent groups (within the same phylum) using codon-based substitution models (dN/dS) implemented in HyPhy. Consistent with the pN/pS results, these analyses indicate that most sites in aerotrophy genes are under strong purifying selection ($dN/dS < 1$), with only a small proportion of sites evolving neutrally or under diversifying selection (Supplementary Figures 10-12). Detailed results are provided in the responses to the comments below. Changes made:

Results and Discussion (L238-243): Codon-based substitution model (dN/dS) analyses further support strong purifying selection, with most sites in aerotrophy genes exhibiting $dN/dS < 1$, while only a small proportion show evidence of neutral or diversifying evolution (**Figs. S10-12 and Table S12**). As a complement to pN/pS, dN/dS captures selection acting on fixed substitutions and thus reflects historical evolutionary constraints, enabling inference across both population-level and divergence-level timescales.

Methods (L531-555):

Phylogenetic codon-based selection analysis

To conduct selection analyses on Antarctic aerotrophy genes and compare them with homologs from other soil ecosystems, we retrieved additional reference sequences from 40,351 species representatives in the genomic catalog of soil microbiomes (SMAG)⁹⁷. Gene and protein

sequences from each MAG were predicted using Prodigal v2.6.3⁷⁰. Homologs of group 1h (n = 2787) and 1l [NiFe]-hydrogenases (n = 421), CO dehydrogenases (*coxL*, n = 9102), and the Type IE RuBisCO large subunit (*rbcL*, n = 160) were identified using the same custom database and pipeline as for the Mackay Glacier samples. For each phylum, aerotrophy genes identified from SMAG were combined with those from Antarctic samples, and duplicate sequences were removed. Protein sequences were aligned using MAFFT v7.520⁸⁶, and poorly aligned regions were trimmed with trimAl v1.5 (-automated1)⁷⁶. Sequences with more than 50% gap characters across the alignment were subsequently removed using a custom script to ensure sufficient alignment coverage. The filtered protein sequences were then realigned with MAFFT v7.520⁸⁶ and re-trimmed with trimAl v1.5 (-automated1)⁷⁶. Maximum likelihood gene trees were inferred from the resulting alignments using IQ-TREE2 v2.3.6⁶⁸ with model selection (-MFP) and 1,000 ultrafast bootstrap replicates. Corresponding codon alignments were generated by mapping nucleotide sequences to the protein alignments using PAL2NAL v14⁹⁸.

Selection analyses were conducted in HyPhy v2.5.29⁹⁹ using the inferred gene trees. We applied RELAX to test for changes in the intensity of selection¹⁰⁰, MEME (Mixed Effects Model of Evolution) to detect episodic positive selection¹⁰¹, and FEL (Fixed Effects Likelihood) to identify pervasive positive and negative selection across sites¹⁰². To account for potential imbalance in the number of sequences between test (Antarctic) and reference (non-Antarctic) groups in RELAX analyses, we performed 50 independent subsampling replicates, in which reference sequences were randomly subsampled to match the number of test sequences. The results were then aggregated across replicates to ensure robustness of the inference.

3. It is not surprising that purifying selection is widespread in microbial communities in Antarctic soils, as strong functional constraint is a pervasive feature of microbial genomes across environments. In most bacteria and archaea, the majority of coding genes are under purifying selection. Therefore, observing widespread purifying selection does not in itself demonstrate environment-specific adaptation; it largely reflects baseline evolutionary constraint common to microbial life. In extreme environments such as Antarctica, where energy availability is limited and growth rates are slow, functional conservation may be more pronounced because organisms depend on tightly optimized metabolic and stress-response systems. However, to substantiate this claim, a comparative analysis of the degree of purifying selection between Antarctic soils and other ecosystems is necessary. Without such a baseline comparison, it is difficult to determine whether the observed patterns reflect environment-specific evolutionary pressure or the general background constraint typical of microbial genomes.

Response: We agree that purifying selection is widespread in microbial genomes and that strong functional constraint may be more pronounced in extreme environments such as Antarctic soils. To further address this point, we performed a comparative analysis of selection intensity on aerotrophy genes between Antarctic soil microbiomes and other soil ecosystems using the RELAX framework implemented in HyPhy. Specifically, we leveraged the genomic catalog of soil microbiomes reported by Ma et al., Nat Comm, 2023, comprising 40,351 species representatives. From this dataset, we identified homologs of group 1h and 1l [NiFe]-

hydrogenases, CO dehydrogenases (*coxL*), and Type IE RuBisCO large subunit genes (*rbcL*), and conducted codon-based selection analyses.

These analyses indicate that group 11 [NiFe]-hydrogenases (from *Acidobacteriota* and *Chloroflexota*) and CO dehydrogenases (*coxL*, from *Actinomyces*) show significant evidence of selection intensification relative to homologs from other soil ecosystems, based on comparisons across 50 replicate subsampled datasets. These results suggest that, while baseline purifying selection is pervasive, some of the observed patterns suggest environment-associated variation in selection intensity (Supplementary Figure 13). We have added these results and a brief discussion to the manuscript. Changes made:

Results and Discussion (L248-253): Codon-based analyses also show that the group 11 hydrogenase genes exhibit significant evidence of selection intensification relative to homologs from other soil ecosystems (average $K = 1.93$, $P < 0.05$), and no significant pervasive positive selection across sites was detected (**Figs. S13-14 and Tables S13-14**). These results suggest that, while baseline purifying selection is pervasive, some of the observed patterns suggest environment-associated variation in selection intensity.

Methods (L531-555): **Phylogenetic codon-based selection analysis**

4. To better substantiate claims of adaptive evolution, the authors should also investigate signals of positive selection. This can be done using phylogenetically informed dN/dS analyses with codon-based models, including branch and branch-site models implemented in tools such as PAML or HyPhy. These analyses would provide a more balanced and rigorous assessment of evolutionary forces beyond the expected background of purifying selection.

Response: We performed additional phylogenetically informed analyses to test for signals of positive selection using codon-based models implemented in HyPhy, including FEL (Fixed Effects Likelihood) and MEME (Mixed Effects Model of Evolution).

For group 11 [NiFe]-hydrogenases, MEME identified a small number of sites with evidence of episodic diversifying selection, indicating that a subset of codon sites may have experienced positive selection along specific lineages. In contrast, FEL did not detect significant pervasive positive selection across sites, instead indicating that the majority of sites are under strong purifying selection (Supplementary Figure 14). Together, these results suggest that while purifying selection dominates the evolution of group 11 [NiFe]-hydrogenases, there is limited but detectable evidence of episodic positive selection at specific sites and branches. We have added these results to the revised manuscript. Changes made:

Results and Discussion (L248-253): Codon-based analyses also show that the group 11 hydrogenase genes exhibit significant evidence of selection intensification relative to homologs from other soil ecosystems (average $K = 1.93$, $P < 0.05$), and no significant pervasive positive selection across sites was detected (**Figs. S13-14 and Tables S13-14**). These results suggest that, while baseline purifying selection is pervasive, some of the observed patterns suggest environment-associated variation in selection intensity.

Methods (L531-555): **Phylogenetic codon-based selection analysis**

5. The code used for data processing and analysis is not available, which limits reproducibility and transparency.

Response: We thank the reviewer for pointing this out. To improve reproducibility and transparency, we have now made all scripts used for data processing and analysis publicly available at GitHub. Changes made:

Methods (L571-574):

Code Availability

The scripts used for HGT prediction, population-level polymorphism, and phylogenetic codon-based selection analysis are available in the GitHub repository: <https://github.com/GreeningLab/AntarcticevolutionManuscript>.

Other comments:

Line 128: It is not clear how MGE count was measured for each phylum (e.g., if it was standardized by the number of MAGs and the size of genome).

Response: In Fig. 1a, the MGE counts for each phylum were standardized by the number of MAGs assigned to that phylum. We have clarified this normalization step in both the figure legend and the corresponding section of the manuscript to make the calculation of MGE counts clearer. Changes made:

Results and Discussion (L126-128): Supporting this pattern, *Desulfobacterota_B*, *Planctomycetota*, *Armatimonadota*, and *Chlamydiota* exhibited high average numbers of mobile genetic elements (MGEs) standardized by the number of MAGs in each phylum (**Fig. 1a**).

Figure legends (L831-833): The bar plot on the right shows the average number of MGEs per MAG (i.e., standardized by the number of MAGs in each phylum), with each MGE type represented by a different color.

Lines 137-138: Need a statistical test to assess the correlation between the average number of detected transfers and co-occurrence coefficients.

Response: We have now included a statistical test to support the observed trend. Changes made:

Results and Discussion (L140-143): As expected, we observed a positive correlation between the average number of detected transfers and co-occurrence (Pearson correlation $r = 0.80$, $P = 0.01$ for binned co-occurrence coefficients and the corresponding mean transfer counts; **Fig. 1b and Table S5**).

Lines 227-230: which statistical test support that the pN/pS for aerotrophy is significantly lower than sample-average pN/pS? Also, p value needs to be provided.

Response: We have now conducted the statistical test and reported the results in the revised manuscript. Changes made:

Results and Discussion (L232-233): confirming the elevated strength of natural selection acting on these genes in Antarctic soils (one-sided Wilcoxon signed-rank test, $P = 2.2 \times 10^{-16}$; **Fig. 3**).

Lines 230-235: same comment for Lines 227-230.

Response: We have now included the statistical test in the revised manuscript. Changes made:

Results and Discussion (L245-246): Furthermore, group 1l hydrogenases exhibited significantly lower pN/pS^(gene) values than group 1h (Wilcoxon rank-sum test, $P = 7.8 \times 10^{-4}$).

Line 390: how were the 43 conserved single-copy genes selected?

Response: The 43 conserved single-copy genes (primarily including ribosomal proteins and universally conserved markers RNA polymerase domains) were selected as universal single-copy marker genes as defined in the widely used tool CheckM (Parks et al., Genome Res, 2015). These genes are highly conserved across prokaryotic genomes and are typically found as single copies, making them reliable for assessing genome completeness and for robust reconstruction of species phylogenies.

Lines 456-457: MGEs specified here are not consistent with Fig. S3. For example, plasmids and phages are missing.

Response: We have revised the Methods section to ensure consistency. Changes made:

Methods (L477-481): Genes associated with the mobilome, including integrons, conjugative elements (CEs), insertion sequence (IS) elements and transposons, and phage recombinases were identified by performing HMM searches against the 68 marker HMM profiles provided by proMGE (<http://promge.embl.de/>)⁹¹.