Phylogenetically and functionally diverse microorganisms reside under the Ross Ice Shelf

Clara Martínez-Pérez1,2,17, Chris Greening3,4, Sean K. Bay3,4, Rachael J. Lappan3, Zihao Zhao1, Daniele De Corte5, Christina Hulbe6, Christian Ohneiser7, Craig Stevens8,9, Blair Thomson10, Ramunas Stepanauskas11, José M. González12, Ramiro Logares13, Gerhard J. Herndl1,14,15, Sergio E. Morales16 & Federico Baltar1,10

Throughout coastal Antarctica, ice shelves separate oceanic waters from sunlight by hundreds of meters of ice. Historical studies have detected activity of nitrifying microorganisms in oceanic cavities below permanent ice shelves. However, little is known about the microbial composition and pathways that mediate these activities. In this study, we profiled the microbial communities beneath the Ross Ice Shelf using a multi-omics approach. Overall, beneath-shelf microorganisms are of comparable abundance and diversity, though distinct composition, relative to those in the open meso- and bathypelagic ocean. Production of new organic carbon is likely driven by aerobic lithoautotrophic archaea and bacteria that can use ammonium, nitrite, and sulfur compounds as electron donors. Also enriched were aerobic organoheterotrophic bacteria capable of degrading complex organic carbon substrates, likely derived from in situ fixed carbon and potentially refractory organic matter laterally advected by the below-shelf waters. Altogether, these findings uncover a taxonomically distinct microbial community potentially adapted to a highly oligotrophic marine environment and suggest that ocean cavity waters are primarily chemosynthetically-driven systems.
ce shelves are permanent floating extensions of grounded sheets of ice that connect to a landmass. The Ross Ice Shelf, by area the largest ice shelf in the world, floats atop an ~54,000 km² ocean cavity that covers about half of the Ross Sea and hugs the coast of Antarctica (Fig. 1a). Generally over 300 m thick, the ice shelf creates a “lid” that isolates the underlying ocean from the atmosphere and from sunlight, and exerts a direct effect on the chemical composition of the water column beneath it (in general ~700 m deep). Waters under the permanent ice shelves are influenced by continental ice-sheet melting and are thus an important intermediary between subglacial outflow from the Antarctic continent and the open Ross Sea, and ultimately the Southern Ocean. Despite their oceanographic significance, sub-ice shelf habitats are among the least-studied ecosystems in the world’s oceans.

Oceanographic and biogeochemical observations of the water cavity beneath the Ross Ice Shelf have been largely concentrated on the shelf margins, in particular at the McMurdo Ice Shelf (northwestern portion of the Ross Ice Shelf). Here, nutrient- and biomass-rich water advected from eastern McMurdo Sound likely plays an important role in sub-ice biogeochemistry of the dark ecosystem beneath the shelf front. Direct observations in the grounding area have also confirmed a stratified and quiescent ocean setting. As a result, water below the Ross Ice Shelf is reported to be exchanged with the Ross Sea with an estimated residence time of 0.9–5.4 years. This allows transport of nutrients and organisms from the sea into the cavity. However, unlike other well-ventilated shelves (e.g., Amery shelf), the proximity to open water is likely a major factor controlling biogeochemical process in the central basin of the Ross Ice Shelf cavity.

Opportunities to directly access the central sub-ice shelf cavity have been greatly limited by logistical constraints and only one expedition to date has sampled the seawater beneath the center of the Ross Ice Shelf. Sampling of the sub-ice water column took place through borehole J9, during the Ross Ice Shelf Project of 1977. The environment beneath the Ross Ice Shelf was described as “similar to the abyssal ocean in being cold and aphotic.” Within these waters, “sparse” populations of bacteria, microbial eukaryotes, and animals were observed. The microbial populations were proven to be heterotrophically active and incorporated radiolabeled organic carbon molecules at very low rates comparable to the abyssal ocean. Autotrophic activity of these microbial communities was subsequently reported and attributed to “nitrifying bacteria.” In these aphotic ecosystems lacking photosynthetic primary production, dark carbon fixation by nitrifying microorganisms may be sufficient to sustain observed microbial and macrofaunal populations. Lateral inputs of organic carbon from the Ross Sea may also support these populations. However, given these studies preceded the advent of molecular techniques, the composition of the microbial communities, their relatedness to open ocean communities, and their possible links to ecosystem function remained unexplored.

In this study we accessed the waters beneath the Ross Ice Shelf to uncover the phylogenetic and functional diversity of the microbial communities under the Antarctic ice shelf. We combined multi-omics techniques (metagenomics, metatranscriptomics, single-cell genomics) with supporting biogeochemical measurements (nutrient measurements and heterotrophic bacterial production). We show that the waters below the shelf harbor a diverse microbial community with a taxonomic composition distinct from other open ocean environments. In addition, we observed the transcription of various genes associated with lithoautotrophic and organoheterotrophic growth, uncovering the basis for previous activities reported in below-shelf waters.

**Results**

The water column under the Ross Ice Shelf is characterized by a steep vertical ammonium gradient. During the Ross Ice Shelf Program in December 2017, an access borehole was created by hot water drilling at site HWD-2 (latitude 80.6577°S, longitude 164.103°E).
Table 1: Biogeochemical data from the water columns below the Ross Ice Shelf at the HWD-2 borehole.

| Depth (m) | Temp (°C) | Practical salinity (psu) | NH₃ (µM) | NOₓ (µM) | PO₄³⁻ (µM) | SiO₂ (µM) | Cell abundance (<18-µm cells ml⁻¹) | LNA (%) | HNA (%) | PHP (µmol C m⁻³ d⁻¹) | Turnover time (d) |
|-----------|-----------|--------------------------|----------|----------|-------------|----------|-------------------------------|--------|--------|-----------------|-----------------|
| 30        | -2.13     | 34.67                    | 0.04 (0.02) | 7.35 (0.23) | 0.72 (0.04) | 165 (2.1) | 0.03 (0.1)                   | 115.7  | 16.0  | 0.720 (0.003) | 461             |
| 180       | -1.96     | 34.69                    | 0.05 (0.02) | 7.32 (0.04) | 0.72 (0.04) | 165 (2.1) | 0.03 (0.1)                   | 115.7  | 16.0  | 0.720 (0.003) | 339             |
| 330       | -1.91     | 34.76                    | 0.04 (0.01) | 7.37 (0.05) | 0.70 (0.004) | 165 (2.1) | 0.30 (0.02)                  | 115.7  | 16.0  | 0.720 (0.003) | 386             |

Mean (standard deviation in parentheses) values are shown (n=3).

Temperature, Temp

LNA: low nucleic acid content cells, HNA: high nucleic acid content cells, PHP: prokaryotic heterotrophic production.

Below-shelf microbial communities are distinct from open ocean communities. Microbial community composition beneath the Ross Ice Shelf was determined using a combination of 16S
rRNA gene amplicon sequencing and shotgun metagenomic sequencing. The microbial community was dominated by six phyla: Proteobacteria, SAR324, Crenarchaeota (mostly Nitrospina-
sphaerases), Marinisomatota (formerly Marinimicrobia, SAR406
clade), Chloroflexota (mostly SAR202), and Planctomycetota
(Fig. 2b). Consistent with a dark oligotrophic environment, the
eukaryotic community was largely comprised of taxa typically
found in the meso- and bathypelagic open ocean, including
Alveolota, Dinoflagellata, and Rhizaria lineages (Supplementary
Fig. 1a, Supplementary Data 1). With respect to viruses, most
bacteriophages detected in the metagenomic assemblies (~50%)
belonged to uncultured or unclassified taxa (Supplementary
Fig. 1b, Supplementary Data 2), with the most abundant classified
viruses affiliating with the family Myoviridae (~30%).

We used 16S rRNA gene sequences extracted from metage-
nomic reads (miTags) to profile the relatedness of microbial communities beneath the Ross Ice Shelf to those of marine
ecosystems globally (Fig. 2a, c20–23). This approach enabled
comparison of microbial communities from available marine
metagenomic datasets, while circumventing potential biases from
inter-study community composition comparisons based on
amplicon analyses. In agreement with previous global metage-
nomic analyses, beta diversity analysis (Bray-Curtis dissimi-
larly) showed oceanic microbial communities cluster by depth,
though this was less pronounced in polar regions (Fig. 2c, d). In
this global context, the communities beneath the Ross Ice Shelf
form a cluster that is related to, but distinct from, those of
mesopelagic polar open ocean waters (Fig. 2c, d). When
compared to deep (>200 m) open ocean communities worldwide,
compositional differences between open-ocean and below-shelf
microbial communities are evident even at the phylum level
(Fig. 2d). For example, the relative abundances of Chloro-
flexota, Gemmatimonadota, Marinisomatota, Myxococcota, Planctomy-
cetota, and SAR324 were significantly higher under the Ross Ice
Shelf, especially in deeper layers (Kruskal-Wallis test, $p < 0.05$, Supplementary Data 3). The phyla Halobacterota, Anc36, and PAUC34f,
while typically rare in the open dark oceans, showed a tenfold
increase in relative abundance in the cavity beneath the Ross Ice
Shelf. Analyses restricted to polar environments using MGLM-
ANOVA confirmed similar compositional differences between the
ocean cavity and deep (>200 m) open-water polar en-
vironments ($LRT = 17333$, $p = 0.001$, Supplementary Data 3).
In addition, Indicator Species Analysis (Indval) congruently identified 'signature species' of the ocean cavity (with respect to deep
open-water polar communities) belonging to the phyla PAUC34f,
Planctomycetota, and SAR324, as well as the classes Lenti-
sphaeria, and SAR202 ($p = 0.001$–0.002, full $p$ values shown in

Fig. 2 Comparison of bacterial and archaeal communities in the cavity beneath the Ross Ice Shelf with open ocean environments worldwide. a Global map depicting the locations of metagenomic surveys utilized in the analysis and this study. Overlapping of symbols represent locations where multiple depths were sampled. b Phylum-level composition of microbial communities under the Ross Ice Shelf based on 165 rRNA amplicon sequencing (this study). The results for each sequencing triplicate are averaged; results for individual replicates and controls are shown in Supplementary Fig. 2a, b. Comparisons with metagenomic 16S ribosomal RNA genes (miTags) are shown in Supplementary Fig. 2c. c Cluster dendrogram depicting the average linkage hierarchical clustering based on a Bray-Curtis dissimilarity matrix of community compositions, based on the relative abundance of miTags from this study, and community compositions of open-ocean and below-shelf polar communities from previous global metage-
nomic datasets, while circumventing potential biases from
inter-study community composition comparisons based on
amplicon analyses. In agreement with previous global metage-
nomic analyses, beta diversity analysis (Bray-Curtis dissimi-
larly) showed oceanic microbial communities cluster by depth,
These 'signature species' (with IndVal p < 0.05, test statistic >0.5, Supplementary Data 3) represented on average ~10% of the community beneath the Ross Ice Shelf, reaching up to 17% in the mid water column, in comparison to an average abundance of 0.75% in deep polar open waters.

Amplicon sequencing analysis provided additional taxonomic resolution of the communities under the ice shelf and confirmed the depth differentiation anticipated from oceanographic and chemical data. Significant differences in community alpha and beta diversity below the Ross Ice Shelf were observed between the basal boundary layer below the ice (30 m) and the deep water column (330 m) (p = 0.028, Supplementary Data 3, Supplementary Figs. 2 and 3). The species driving these differences are described in the Supplementary Notes.

Nitrifying archaea and bacteria dominate transcription under the shelf. We used a multi-omics approach to uncover the functional capacity of the microbial community beneath the Ross Ice Shelf, integrating genome-resolved metagenomics, single-cell genomics, and metatranscriptomics. We assembled 235 dereplicated partial genomes (Fig. 3, Supplementary Figs. 4 and 5; Supplementary Data 4). These comprised 67 SAGs (single-amplified genomes) and 168 manually curated MAGs (metagenome-assembled genomes), all with completeness >50% and contamination <5% (Fig. 3; Supplementary Data 4). These represent on average 50–60% of each sample’s metagenomic and metatranscriptomic reads, including all phyla with relative abundance above 0.5% (Fig. 2) and the top four most abundant genera (Supplementary Fig. 2b). Their phylogenetic diversity, metabolic traits, and relative abundances are depicted in Fig. 3.

The presence and transcription of key metabolic genes in assembled and unassembled reads was used to identify prevailing metabolic pathways in the cavity under the Ross Ice Shelf. By far the most highly transcribed genes involved in autotrophic energy conservation pathways were those for oxidation of ammonium (ammonia monoxygenase, amoA) and nitrite (nitrite oxidoreductase, nxrA) (Fig. 4b). Accordingly, ammonium transporters and amoA
were the most transcribed genes overall (Supplementary Fig. 6). Transcription patterns correlated with ammonium concentrations (Table 1) and relative abundance of the archaeal order Nitrosophaerales (Supplementary Figs. 2b, 4 and 5). Phylogenetic analysis corroborated that the most numerous amoA genes and transcripts were affiliated with *Nitrosopumilus* spp. (Fig. 4c, Supplementary Fig. 7), the most abundant and active archaeal lineage beneath the ice shelf (Supplementary Figs. 2b, 4 and 5), with some gammaproteobacterial amoA reads also detected (Fig. 4a, Supplementary Fig. 7).

The metagenomic and metatranscriptomic reads of the marker gene for nitrite oxidation, *nxrA*, were affiliated with the phyla Nitrospinota and, to a lesser extent Nitrospirota (Supplementary Data 5, Supplementary Fig. 8). In line with an autotrophic lifestyle, we identified the determinants of ammonium- or nitrite-dependent carbon fixation via the archaeal 4-hydroxybutyrate cycle (*hbsC, hbsT* genes) and *Nitrospina* reductive tricarboxylic acid cycle (*aclB* gene) (Fig. 4, Supplementary Figs. 9, 10 and 11; Supplementary Data 3).

Consistent with these results, reconstructed genomes from the genera *Nitrosopumilus* and *Nitrospina* were among those with highest relative transcriptional activity in our dataset (S4, S5). These groups express a small fraction of their genomes (i.e., ~25% of total genes at 30 m) compared to other community members (Supplementary Fig. 4d–f), devoting most of their transcriptional effort to the key processes of carbon fixation and ammonia and nitrite oxidation, respectively. Despite being well-represented in the metatranscriptomic dataset, the relative abundance of the genus *Nitrospina* was low in the metagenomic dataset. For instance, the *Nitrospina* lineage represented by SAG_5 was among the least abundant genomes, but was highly active on the transcriptional level (RNA/DNA ~270; Supplementary Fig. 5) (Supplementary Data 4). These discrepant findings are in line with recent single-cell analyses showing Nitrospinota have high activity despite low abundance; it is proposed that the large cell size or high mortality rates of these nitrite oxidizers are responsible for their low abundance in metagenomes and amplicon datasets compared to ammonium oxidizers.

Various inorganic and organic energy sources likely support below-shelf bacteria. Many members of the microbial community are capable of supporting or surviving beneath the shelf through a

---

**Fig. 4 Energy conservation and carbon fixation strategies of communities beneath the Ross Ice Shelf.** a Dot plot showing the metabolic potential of the 235 metagenome-assembled genomes (MAGs) and single-amplified genomes (SAGs). The size class of each point represents the number of genomes in each class that encode the gene of interest and the shading represents the average genome completeness. b Heatmaps showing the relative abundance of these genes in the three metagenomic and metatranscriptomic unassembled short reads datasets. For metagenome reads, the heatmap shows the abundance of each pathway, expressed as average gene copies per organism (across all genes listed in the pathway) calculated relative to the abundance of 14 universal single-copy ribosomal genes, with scales capped at 1. For metatranscriptome reads, the heatmap shows log10-transformed reads per kilobase million (RPKM). Where genes within the same pathway are collapsed together, the values (community percentage or RPKM) are summed. c Phylogenetic tree of protein sequences of the highly transcribed ammonia monooxygenase subunit A (amoA) gene from archaeal single-amplified genomes and unbinned metagenomic contigs shown in bold compared to reference sequences. See Supplementary Fig. 7 for a detailed version of this tree.
Fig. 5 Relative abundance and transcription of selected carbohydrate active enzyme (CAZyme) classes. Data is displayed for reconstructed genomes (MAGs and SAGs) where CAZyme diversity was highest (top 50 genomes). Bubble plots represent the number of different genes from each CAZyme class per genome (GH, glycosyl hydrolases; GT, glycosyl transferases; CBD, genes containing carbohydrate binding domains). Heatmaps represent the total gene transcription for each CAZyme class, normalized to total transcripts per sample (transcripts per million, TPM). The data used to construct these plots is provided in Supplementary Data 7.

The metatranscriptome also revealed various other processes supporting life beneath the shelf. The heterotrophic majority in the deep ocean environments, as we have recently described for other aerobic heterotrophic bacteria. Genes for formate oxidation were also widespread and highly transcribed (Fig. 4b, Supplementary Fig. 6), with transcriptional effort spreading across a variety of substrate-utilization processes. The metatranscriptome also revealed various other processes supporting life beneath the shelf. The heterotrophic majority in the deep ocean environments, as we have recently described for other aerobic heterotrophic bacteria. Genes for formate oxidation were also widespread and highly transcribed (Fig. 4b, Supplementary Fig. 6), with transcriptional effort spreading across a variety of substrate-utilization processes.
Collectively, our results provide a detailed insight on the ecological strategies adopted by communities living in the world’s most extensive sub-ice shelf system. Oceanic cavities below ice shelf systems are uniquely different from open ocean environments in their dependence on in situ chemosynthesis and on lateral advection of food sources from open-water areas, rather than on vertical fluxes of phytoplankton-derived detrital matter. We estimate that the waters sampled at the borehole location have been in the cavity for as much as four years prior to sampling; this is up to 10-20-fold longer than the time predicted for marine snow from the ocean surface to reach the abyss (≈6000 m)\(^3\). Likewise, the heterotrophic production rates measured in this study and at borehole J9\(^{10}\) were among the lowest measured in marine ecosystems, including environments with similar temperatures\(^3\). It has been suggested that production rates are highly influenced by the supply and concentration of labile dissolved organic material\(^13\), and thus the water column beneath the ice shelf is predicted to be highly oligotrophic with respect to labile organic matter.

Based on these heterotrophic rates and assuming a heterotrophic prokaryotic growth efficiency of ≈5% (typical of deep oceanic waters, e.g.,\(^46\)), we estimate a total organic carbon demand (i.e., the combined carbon incorporation into biomass and respiration) of ≈6–12 µmol C m\(^{-2}\) d\(^{-1}\). This total carbon demand is in the same range as the carbon fixation rates reported from the environment beneath the J9 borehole (8.3 µmol C m\(^{-2}\) d\(^{-1}\)). While the contribution of exogenous organic matter remains to be quantified, the close coupling between in situ dark carbon fixation and organic carbon demand suggests that the ecosystem beneath the Ross Ice Shelf is largely sustained by dark carbon fixation. This would differ from deep open ocean environments, where heterotrophic carbon demand significantly relies on the vertical fluxes of particulate organic carbon generated in the euphotic layer\(^41,42\).

Our multi-omic results support this hypothesis, while uncovering the mediators and pathways responsible for the autotrophic and heterotrophic activities under the Ross Ice Shelf (Fig. 6). Among the lineages represented by MAGs and SAGs with the highest transcriptional activity are those originating from the chemolithoautotrophic genera *Nitrosopumilus* and *Nitrospina*. Overall, this agrees with previous reports that aerobic ammonium-oxidizing microorganisms are widespread in Antarctic marine environments (e.g.,\(^43\)) and that ammonium oxidation occurs beneath Antarctic shelves and sea ice\(^12,44\). These and other inferred facultative chemolithoautotrophs (such as facultative sulfur-oxidizing bacteria) are likely to be responsible for dark carbon fixation rates previously observed beneath borehole J9\(^{12}\) and thus provide a supply of organic carbon to an ecosystem shielded from sunlight.

The importance of dark carbon fixation has been recognized in various oceanic regions during the polar winter. Microbial lineages (e.g., *Nitrospina*, *Nitrosopumilus*, SAR324, and Marinisomatota\(^{45-47}\) and enzymes (such as those mediating ammonium, nitrite, and sulfur oxidation\(^48\)) that mediate chemolithoautotrophy have been observed to increase in Antarctic waters during the transition to the winter season. Likewise, comparable lineages and genes capable of sulfur compound oxidation have been detected in winter open waters and the central basin under the Ross Ice Shelf. Together with mounting evidence that sulfur compound oxidizers sustain carbon fixation in the wide dark open ocean (e.g.,\(^49\)) and the diverse sources of reduced sulfur compounds in marine oxic environments (e.g.,\(^50\)), it is plausible that these clades can also contribute to chemolithoautotrophy in the oceanic cavity beneath the Ross Ice Shelf.

It is likely that ammonium is a primary energy source sustaining primary production in aphotic Antarctic waters. Consistent with this idea, ammonium oxidation rates have been reported to be higher in Antarctic coastal waters during the austral winter and to significantly support the heterotrophic demand\(^13\). In the absence of direct rate measurements in this study, we estimated the ammonium oxidation rates potentially supported by the standing ammonium concentrations in the water column. Our estimates for the basal layer (≈90 nM NH\(_4\)\(^+\) d\(^{-1}\)) are in accordance to rates measured in the Southern Ocean (62 nM NH\(_4\)\(^+\) d\(^{-1}\)) with...
comparable ammonium concentrations (0.7 μM NH₄⁺), and could support the heterotrophic demand in the oceanic cavity under the shelf (Supplementary Notes). These estimates suggests that the microbial communities beneath the Ross Ice Shelf can sustain ammonium oxidation at similar rates to those in the winter Antarctic Ocean and have the potential to be significant primary producers.

The ammonium profile beneath the Ross Ice Shelf is intriguing. Contrary to other nutrient concentrations measured (which do not vary significantly through the water column), ammonium concentrations are significantly higher in the ice basal boundary layer compared to the deeper water samples, but comparable to those in the periphery of the shelf. This profile (exclusive for ammonium with respect to other nitrogen species) is consistent with the reports beneath borehole J9. The proposed circulation model beneath the shelf, by which the cavity is filled southward by dense water masses that reach its interior via deep cavity circulation, renders it unlikely that the high ammonium concentrations detected in the fresh, northward flowing waters beneath borehole HW2D or J9 originate from the open Ross Sea. If externally sourced, nutrient concentrations would be expected to be highest in deeper waters, or else be homogenized in the water column as water masses evolve and mix in the cavity. The latter appears to be the case for the other nutrients measured in this and the J9 expedition. The exception observed in the ammonium profile suggests that this compound could be sourced beneath the ice shelf. In particular, terrestrial-origin sediments in the basal ice layer may be a significant source of ammonium to the seawater circulating beneath. Deployment of cameras at HW2D revealed sedimentary englacial debris in the lower 20 meters of the ice shelf. While ice melting and freezing can plausibly result in the rainout of the pellets in a sub-ice-shelf cavity, we did not witness this effect; no sediments were retrieved from the pumping samples and the microbial communities sequenced from the englacial debris and the water column were unrelated (Supplementary Fig. 2). However, temperature and salinity data from our study site (Fig. 1b) clearly showed ice-shelf basal melting and a supply of freshwater to the upper region of the water column, a phenomenon that could result in the observed replenishment of ammonium concentrations in this system. In free-floating sea ice, as well as in subglacial lakes, ammonium enrichments have been traditionally attributed to wet and dry atmospheric deposition, as well as in situ organic matter regeneration in brine channels, especially within older and thicker ice. The latter may be also a mechanism for ammonium accumulation in deep layers of the ice shelf, subject to solubilization and transport by fresh melt water. If such is the case, the ammonium transported by the ice basal boundary layer could be sourced locally (at borehole HW2D) or elsewhere upstream. Dissolved nutrients in the ice sheet or englacial debris are eventually diluted as they circulate the interior of the shelf, which could explain the observed higher concentrations in the water column from borehole J9, 330 km upstream from our study site. While the driving factors of the nutrient profile in the water column remain unclear, the tenfold decrease in ammonium concentrations correlate with changes in relative transcriptional activity of the ammonium-oxidizing genus Nitrosopumilus (Supplementary Fig. 4). As described in Supplementary Notes, we observed depth-related differences in microbial community composition, metabolic capabilities, and gene expression, though additional depth profiles would be required to confirm this.

The community members with highest relative abundance and transcriptional activity throughout the water column included nitrifying autotrophic taxa and organoheterotrophic bacteria (Supplementary Figs. 4, 5 and 6). It is likely that the genomes with highest relative transcriptional activity represent two opposite adaptive strategies to the conditions beneath the Ross Ice Shelf. Based on the proportion of their genome expressed, nitrifiers are likely to effectively exploit the surrounding environment by expressing a reduced set of genes encoding a few metabolic pathways. The opposite is observed in the highly expressed heterotrophic clades (Supplementary Fig. 4). By expressing up to 95% of their genome (e.g., in members of Latecibacterota and Verrucomicrobiota), the transcriptional effort of the latter is spread across a variety of process and in particular, to the exploitation of multiple substrates. These observations are consistent with previous studies combining expression and genomic datasets, which suggest that activity levels, substrate utilization and transcriptome diversity may be linked in defining ecological niches of microbial communities.

In particular, our results suggest that the most active heterotrophic organisms are adapted to degrade complex organic compounds, including most of the enriched phyla in this environment, such as Myxococccota and Planctomycetota. Their capacity to degrade complex organic material from a range of sources, including potentially of both autochthonous and allochthonous origin, likely confers a major selective advantage in this highly oligotrophic ecosystem. Heterotrophy based on the consumption of recalcitrant dissolved organic carbon has been considered as one possibility for sustaining the oceanic Antarctic winter food web, and could also be an additional support for life under the Ross Ice Shelf. Unlike organic carbon in Antarctic winter waters, which may have accumulated during the highly productive summer season, organic substrates beneath the Ross Ice Shelf potentially consist of vertically transported exudates and necromass derived from lithoautotrophic primary producers, but also recalcitrant complex organic compounds laterally transported from the Ross Sea into the shelf cavity. Decomposition of phytoplankton entering the shelf cavity is estimated at a scale of ~ 10 years. Together with previous reports of diatoms in below-shelf waters, this indicates that some photoautotrophically-derived organic matter can reach the center of the oceanic cavity. However, the metagenomes suggest that photosynthetic eukaryotes (i.e., class Bacillariophyceae) make a small fraction of the eukaryotic community (0.05%); this finding is also consistent with undetectable concentrations of chlorophyll beneath borehole J9. Despite potentially serving as a substrate for organoheterotrophs beneath the ice shelf, phytoplankton are therefore unlikely contributors to the dissolved organic matter pool, whereas detrital sources of bacterial substrates may be more important. Further work is now needed to discriminate organic matter sources and nutrient exchange processes within the shelf.

Overall, microorganisms under Antarctica’s ice shelves can thrive in some of the coldest and possibly carbon-limited marine waters, while playing a crucial role in the remineralization of nutrients to the Southern Ocean. Our results not only suggest that the waters below the Ross Ice Shelf are driven by chemolithoautotrophic processes, but also uncover the mechanisms responsible for sustaining that activity. Alongside other recent reports of oceanic dark carbon fixation, this study also emphasizes the importance of inorganic energy sources in driving marine communities in the absence of photosynthesis. Finally, our results suggest that ammonium associated with fresh melt waters at the base of the ice is an important supply of inorganic electron donors supporting chemolithoautotrophs, and thus has a significant influence in the composition and activity of the microbial community. Ocean-driven basal melting, a source of freshwater and thus potentially of ammonium in the sub-ice cavity, may increase in a warming climate scenario. Assuming that our observations are representative of the central region of the cavity under the Ross Ice Shelf, increased basal ice melting could result in an increased vulnerability of communities supported by sub-ice shelf processes, potentially leading to shifts in
the relative biogeochemical importance of chemolithoautotrophic processes in this extensive ecosystem. These insights emphasize the importance of baseline data from existing sub-ice shelf ecosystems, such as the Ross Ice Shelf, to inform the prediction of biogeochemical impacts of climate change in the Southern Ocean.

**Methods**

**Site selection and description.** Sampling took place in December 2017 and was conducted by members of the Aotearoa New Zealand Ross Ice Shelf Program. Samples were collected from the sub-shelf water column at a site in the central region of the ice shelf, borehole HWD-2 (Latitude –80.6577 N, Longitude 174.4626 W), ~300 km from the shelf front and 330 km northwest of borehole JH (Fig. 1a). The sampling site is near the glaciological boundary between ice originating from the West Antarctic Ice Sheet and ice flowing from East Antarctica through Transantarctic and Antarctic Mountain glaciers (Fig. 1a). Sediment of terrestrial origin was observed in the lowest ~60 m of the ice.

**Hot water drilling and sampling.** A hot water drilling system built and operated by the Victoria University of Wellington Drilling Office was used to bore through the ice shelf, creating an access borehole with a maximum diameter of 30 cm. The borehole was used for direct sampling of water and sea floor sediments, and to conduct in situ measurements in the water column. These activities were conducted in a tent that facilitated both operations in all weather conditions. Seawater samples were obtained from three depths (400 m, 550 m, and 700 m from the top of the shelf, which correspond to 30 m, 180 m, and 330 m deep from the bottom of the ice shelf, respectively). These were chosen to characterize the water column under the Ross Ice Shelf while keeping the sampler ca. 40–50 m away from the seafloor and from ice crystals and sediment in the ice shelf basal layer. The drilling water was fresh (<15 psu) and relatively warm (between 27 °C and 50 °C) so it remained stable floating in the borehole and did not sink into deeper layers. This, together with the advection of seawater below the ice shelf, precluded any contamination of collected seawater with the drilling water (Supplementary Fig. 2a, b). The lack of intrusion of the freshwater used for the drilling was routinely checked by salinity and temperature-depth profiles.

Samples were collected by in situ filtration using a McLane WTS-LV-Bore Hole filter pump fitted with a 142 mm diameter, 0.22 μm pore-size filter (Super membrane filters, Pall Corporation). Before and after deployment, the filter holder was thoroughly cleaned to avoid sample cross-contamination. The filter holder was also flushed after every deployment with fresh water to prevent salt crystal formation and sample contamination. This sampling approach was aimed at obtaining the most realistic representation of the microbial community’s composition and activity with the minimum possible sampling biases.

**Metagenomic and metatranscriptomic approaches.** Approximately 200 L of water were filtered at each depth within ca. 2 h. After filtration, the filtered water was filtered through Whatman GF/F filters, collected in acid-cleaned HDPE bottles, and stored frozen until analysis in the home laboratory, following procedures recommended by the Joint Global Ocean Flux Study (JGOFS)68. The liquid samples for the determination of microbial cell abundance, prokaryotic heterotrophic production, and the generation of single-cell amplified genomes (SGAs) were collected in acid-cleaned Nalgene™ opaque amber HDPE bottles, stored at 2 °C, and transported within 48 h to Scott Base to perform further laboratory analyses. The samples were imported to New Zealand under Ministry for Primary Industry permit number 2017063583 (Permit to import Restricted Biological Products of Animal Origin) issued to the University of Otago Department of Marine Science.

To check for potential contamination, samples were also collected from the following sites: freshly melted snow near the camp area, drilling water from a reservoir tank, and sediments dislodged from the ice shelf (identified as englacial debris) and collected with the reaming tool. Water samples were filtered onto 0.22 μm polycarbonate filters (47 mm filter diameter, Millipore), and all samples were stored in cryovials and frozen.

**Physicochemical measurements.** A SBE 19plusV2 SeaCat Profiler CTD (Seabird Electronics, Inc.) was used to measure temperature, salinity and depth within the borehole, and in the water under the Ross Ice Shelf for a detailed characterization of the water column. Furthermore, a self-contained single channel logger (RBR Solo) was attached to the frame of the WTS-LV-Bore Hole pump (at the opposite side of the water intake) for an accurate determination of the temperature and depth of the sampling casts. Samples for determining the concentrations of nitrate, dissolved reactive phosphorus, dissolved inorganic phosphorus and Si(OH)₄ were collected and processed using methods described elsewhere69. Measurements of nutrient concentrations were routinely corrected with reference blank solutions in each sample run. No anomalies were detected in the blanks, indicating no source of detectable contamination during the measurements.

**Prokaryotic abundances and heterotrophic production.** Prokaryotic abundance was determined by flow cytometry. Samples (1.6 mL) were preserved with glutaraldehyde (2% final concentration), left at 4 °C in the dark for 15 min, flash-frozen in liquid nitrogen, and stored at ~80 °C until analysis. Prior to analysis, the fixed samples were thawed, stained in the dark with a DMS-diluted SYTO-13 dye ( Molecular Probes Inc., 2.5 µM final concentration) for 5 min, and run on a BD AccuriTM flow cytometer with a laser emitting at 488 nm wavelength. Samples were run at low or medium speed until 10,000 events were captured. A suspension of yellow-green–1 µm latex beads (10¹⁰–10¹⁵ beads mL⁻¹) was added as an internal standard (Poly sciences, Inc.).

The blank-corrected leucine incorporation rates were converted into prokaryotic heterotrophic production (PHP) using the theoretical conversion of 1.55 kg mol⁻¹ leucine incorporated71. The rates of leucine incorporation obtained at the incubation temperature (4 °C) were converted to the in situ temperature of -2 °C using an activation energy of 72 kJ mol⁻¹72.
paired-end reads were quality-assessed with FastQC v0.11.7 and MultiQC v1.08.0 for downstream analysis. Observed richness (Chao1) was calculated using the estimate richness function in Phyloseq. Normality of the distribution of alpha-diversity estimates was confirmed using a Shapiro-Wilk test and a one-way analysis of variance (ANOVA) to test for significant differences in richness across depth profiles. As a post-hoc test, the Tukey multiple comparison of means was used to confirm which pairs of sites showed significant differences. For beta-diversity analysis on amplicon and miTag data, Bray-Curtis distance matrices were calculated in Vegan and visualized using a principal coordinate analysis (PCoA). Independent permutation-based analysis of variance (PERMANOVA) based on the Bray-Curtis dissimilarities values were calculated with the adonis function in Vegan (999 random permutations), to test for significant differences in community structure between depth profiles. Finally, a beta-dispersion test (PERMDISP) was applied to confirm that observed differences were not influenced due to dispersion. As a post-hoc evaluation of taxa responsible for differences in microbial community structure, we performed indicator species analysis. We used the indval function in Vegan (999 random permutations), to test for significant differences in community structure between depth profiles. Alpha- and beta-diversity analyses of 16S rRNA amplicons and extracted miTAGs. All statistical analyses were carried out in R v3.5.3. Data manipulation was performed using the R package tidyr and all visualizations were made using ggplot2. Community richness and beta-diversity was calculated using the R packages Phyloseq26 and Vegan v2.5-26. In total, nine samples representing a transect of depth profiles were used for downstream analysis of the ASVs' (Supplementary Fig. 3, Supplementary Data 3). Rarefaction curves were constructed to confirm that sequencing depth adequately captured richness in each sample and rarefied using the Phyloseq rarely_even_depth function with a sample size of 15,400, which represented the minimum sequencing depth to retain 100% of sequence reads. Observed richness (Chao1) was calculated using the estimate richness function in Phyloseq. Normality of the distribution of alpha-diversity estimates was confirmed using a Shapiro-Wilk test and a one-way analysis of variance (ANOVA) to test for significant differences in richness across depth profiles. As a post-hoc test, the Tukey multiple comparison of means was used to confirm which pairs of sites showed significant differences. For beta-diversity analysis on amplicon and miTag data, Bray-Curtis distance matrices were calculated in Vegan and visualized using a principal coordinate analysis (PCoA). Independent permutation-based analysis of variance (PERMANOVA) based on the Bray-Curtis dissimilarities values were calculated with the adonis function in Vegan (999 random permutations), to test for significant differences in community structure between depth profiles. Finally, a beta-dispersion test (PERMDISP) was applied to confirm that observed differences were not influenced due to dispersion. As a post-hoc evaluation of taxa responsible for differences in microbial community structure, we performed indicator species analysis. We used the indval function in Vegan (999 random permutations), to test for significant differences in community structure between depth profiles. Alpha- and beta-diversity analyses of 16S rRNA amplicons and extracted miTAGs. All statistical analyses were carried out in R v3.5.3. Data manipulation was performed using the R package tidyr and all visualizations were made using ggplot2. Community richness and beta-diversity was calculated using the R packages Phyloseq26 and Vegan v2.5-26. In total, nine samples representing a transect of depth profiles were used for downstream analysis. Observed richness (Chao1) was calculated using the estimate richness function in Phyloseq. Normality of the distribution of alpha-diversity estimates was confirmed using a Shapiro-Wilk test and a one-way analysis of variance (ANOVA) to test for significant differences in richness across depth profiles. As a post-hoc test, the Tukey multiple comparison of means was used to confirm which pairs of sites showed significant differences. For beta-diversity analysis on amplicon and miTag data, Bray-Curtis distance matrices were calculated in Vegan and visualized using a principal coordinate analysis (PCoA). Independent permutation-based analysis of variance (PERMANOVA) based on the Bray-Curtis dissimilarities values were calculated with the adonis function in Vegan (999 random permutations), to test for significant differences in community structure between depth profiles. Finally, a beta-dispersion test (PERMDISP) was applied to confirm that observed differences were not influenced due to dispersion. As a post-hoc evaluation of taxa responsible for differences in microbial community structure, we performed indicator species analysis. We used the indval function in Vegan (999 random permutations), to test for significant differences in community structure between depth profiles. Alpha- and beta-diversity analyses of 16S rRNA amplicons and extracted miTAGs. All statistical analyses were carried out in R v3.5.3. Data manipulation was performed using the R package tidyr and all visualizations were made using ggplot2. Community richness and beta-diversity was calculated using the R packages Phyloseq26 and Vegan v2.5-26. In total, nine samples representing a transect of depth profiles were used for downstream analysis. Observed richness (Chao1) was calculated using the estimate richness function in Phyloseq. Normality of the distribution of alpha-diversity estimates was confirmed using a Shapiro-Wilk test and a one-way analysis of variance (ANOVA) to test for significant differences in richness across depth profiles. As a post-hoc test, the Tukey multiple comparison of means was used to confirm which pairs of sites showed significant differences. For beta-diversity analysis on amplicon and miTag data, Bray-Curtis distance matrices were calculated in Vegan and visualized using a principal coordinate analysis (PCoA). Independent permutation-based analysis of variance (PERMANOVA) based on the Bray-Curtis dissimilarities values were calculated with the adonis function in Vegan (999 random permutations), to test for significant differences in community structure between depth profiles. Finally, a beta-dispersion test (PERMDISP) was applied to confirm that observed differences were not influenced due to dispersion. As a post-hoc evaluation of taxa responsible for differences in microbial community structure, we performed indicator species analysis. We used the indval function in Vegan (999 random permutations), to test for significant differences in community structure between depth profiles. Alpha- and beta-diversity analyses of 16S rRNA amplicons and extracted miTAGs. All statistical analyses were carried out in R v3.5.3. Data manipulation was performed using the R package tidyr and all visualizations were made using ggplot2. Community richness and beta-diversity was calculated using the R packages Phyloseq26 and Vegan v2.5-26. In total, nine samples representing a transect of depth profiles were used for downstream analysis. Observed richness (Chao1) was calculated using the estimate richness function in Phyloseq. Normality of the distribution of alpha-diversity estimates was confirmed using a Shapiro-Wilk test and a one-way analysis of variance (ANOVA) to test for significant differences in richness across depth profiles. As a post-hoc test, the Tukey multiple comparison of means was used to confirm which pairs of sites showed significant differences. For beta-diversity analysis on amplicon and miTag data, Bray-Curtis distance matrices were calculated in Vegan and visualized using a principal coordinate analysis (PCoA). Independent permutation-based analysis of variance (PERMANOVA) based on the Bray-Curtis dissimilarities values were calculated with the adonis function in Vegan (999 random permutations), to test for significant differences in community structure between depth profiles. Finally, a beta-dispersion test (PERMDISP) was applied to confirm that observed differences were not influenced due to dispersion. As a post-hoc evaluation of taxa responsible for differences in microbial community structure, we performed indicator species analysis. We used the indval function in Vegan (999 random permutations), to test for significant differences in community structure between depth profiles.
improves assembly (i.e., number of scaffolds reduced) and consequently bin metrics such as contig length and purity of bins. Completeness and quality of final assemblies were assessed by CheckM v1.0.7.50, with bins with >50% completeness and <5% contamination (i.e., high and medium quality bins) retained for further analysis.53

**Genome de-replication, classification, and phylogenetic analysis.** Metagenomic bins and single-cell-assembled genomes with >50% completeness were defined as MAGs and SAGs, respectively, and collectively as ‘genomes’ for simplicity. Comparison and de-replication of genomes were performed with dRep pipeline106. In brief, genomes were grouped at an average nucleotide identity (ANI) of 99%. Representative genomes from each cluster were selected based on the highest ‘genome score’106. This analysis provided a de-replicated genomic database of 235 MAGs/SAGs was performed using ribosomal protein sequences retrieved from CheckM v1.0.7.50 (Fig. 3). The concatenated marker sequence for each genome was aligned using MAFFT109 and an approximate maximum-likelihood phylogenetic tree was generated using FastTree 210 with default parameters. The tree was then visualized and annotated using the web-based tool iTOL v6 (https://itol.embl.de).

**Metabolic profiling of MAGs, SAGs, and assembled unbinned reads.** ORFs in binned and unbinned contigs were predicted using Prodigal v.2.6.3.96, with default noise-cut-off.s filtered by manual filtering using HMM cut-off scores previously described.118 The remaining ORFs were annotated with homology searches against curated protein databases described above and to the predicted ORFs that matched the additional 10 HMMs described above (Supplementary Data 6). DIAMOND mapping was performed with a query coverage threshold of >80% and a gene specific threshold of 40% (RHO), 60% (AtpA, AmoA, MmoX, Coxl, NraA, NuoF and Rbcl), 75% (HbsT), 70% (PsaA, YgrK, ARO, IsoA), (80%) PsaA, or 50% (all other databases), with data further parsed to retain only group 1 and 2 [NiFe]-hydrogenase hits. For the metagenomic data, forward reads with at least 12 bp in length were used. For the metatranscriptomic data, paired-end reads were merged with BBMerge v38.51 and merged reads of at least 124 bp in length were used. From sample replicates combined for this analysis. The abundance of each gene was converted to reads per kilobase million (RPKM).

\[ \text{RPKM} = \frac{X}{x} \times \text{total sample reads} \times 10^6 \] (3)

where \( X \) = reads aligned to a gene/ gene length (kbp).

The gene abundances in RPKM from the metagenomic data were further used to estimate the proportion of the community encoding these functions. The processed metagenomic reads were aligned to each of the 14 universal single-copy ribosomal marker genes available in SingleM (https://github.com/wvwood/singlem) with DIAMOND using a query coverage threshold of 80%. Alignments with a bitscore below 40 were removed; the alignment counts were converted to RPKM as described above and averaged across the 14 genes to represent the abundance of a universal single-copy gene. Metabolic gene RPKM values were divided by this value to obtain the average gene copies per organism in each sample (abundance relative to a single-copy gene). Heatmaps representing the community percentage (metagenomic data) and RPKM abundance (metatranscriptomic data) were generated in R with ggplot2 (Fig. 4b). Where genes within the same pathway are collapsed together, the values (community percentage or RPKM) are summed.

**Data availability**

The data and code underlying Fig. 2a, c are provided in the github repository https://github.com/ClaMtnez/Ocean_tags. The data underlying Figs. 3, 4 & 5 are provided as a Supplementary Data Files. The sequence data generated in this study have been deposited in the EMBL Nucleotide Sequence Database (ENA) database under Bioproject PRJEB35712 (metagenomic and metatranscriptomic raw reads, metagenomic and metatranscriptomic assemblies, metagenomic assembled genomes, and single-cell amplified genomes) and in the NCBI Sequence Read Archive (SRA) under Bioproject PRJNA593284 (16S rRNA gene amplicon reads). The following public databases were used in this study: Swiss-Prot database, https://www.uniprot.org/, release_2018_10; Genome Taxonomy Database, https://gtdb.ecogenomic.org/, release 80; SILVA non-redundant SSU Ref database, https://www.arb-silva.de/, v.138; UniRef 100 VIROME database, http://www.ncbi.nlm.nih.gov/genome/annotation_prok/tigrfams/, release 150.

**Reporting summary**

Further information on research design is available in the Nature Research Reporting Summary linked to this article.

**References**

1. Fretwell, P. et al. Bedmap2: improved ice bed, surface and thickness datasets for Antarctica. Cryosphere 7, 375–393 (2013).
2. Holland, D. M., Jacobs, S. S. & Jenkins, A. Modelling the ocean circulation beneath the Ross Ice Shelf. Antarct. Sci. 15, 13–23 (2003).
3. Robinson, N. J., Williams, M. J. M., Barrett, P. J. & Pyne, A. R. Observations of flow and ice-ocean interaction beneath the McMurdo Ice Shelf, Antarctica. J. Geophys. Res. Ocean. 115, 1–10 (2010).
4. Vick-Majors, T. J. et al. Biogeochemistry and microbial diversity in the marine cavity beneath the McMurdo Ice Shelf, Antarctica. Limnol. Oceanogr. 61, 572–586 (2016).
5. Beaman, R. R. et al. Ocean Stratification and Low Melt Rates at the Ross Ice Shelf Grounding Zone. J. Geophys. Res. Ocean. 123, 7438–7452 (2018).
6. Smette, H. W. M. Jr & Jacobs, S. S. Circulation and melting under the Ross Ice Shelf: estimates from evolving CFC, salinity and temperature fields in the Ross Sea. Deep Sea Res. Part I Oceanogr. Res. Pap. 52, 959–978 (2005).
7. Michel, R. L., Linick, T. W. & Williams, P. M. Tritium and carbon-14 distributions in seawater from under the Ross Ice Shelf Project ice hole. Science 80(–), 203, 445–446 (1979).
8. Post, A. L. et al. Modern sedimentation, circulation and life beneath the Amery Ice Shelf, East Antarctica. Cont. Shelf Res. 74, 77–87 (2014).
S.K.B., D.D.C., B.T., J.M.G., F.B., and C.G. analyzed the data. C.M.P., C.G., and F.B. wrote the manuscript with assistance from all coauthors.

Competing interests
The authors declare no competing interests.

Additional information
Supplementary information The online version contains supplementary material available at https://doi.org/10.1038/s41467-021-27769-5.

Correspondence and requests for materials should be addressed to Sergio E. Morales or Federico Baltar.

Peer review information Nature Communications thanks Jeff Bowman and the other, anonymous, reviewer(s) for their contribution to the peer review of this work. Peer reviewer reports are available.

Reprints and permission information is available at http://www.nature.com/reprints

Publisher’s note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Open Access This article 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. The images or other third party material in this article 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 http://creativecommons.org/licenses/by/4.0/.

© The Author(s) 2022