



Proteorhodopsin Phototrophy in Antarctic Coastal Waters

 Jerónimo Cifuentes-Anticevic,^a  María E. Alcamán-Arias,^{b,c,d}  Tomás Alarcón-Schumacher,^{a,e}  Javier Tamayo-Leiva,^{a,c}
 Carlos Pedrós-Alió,^f  Laura Farías,^{b,c}  Beatriz Díez^{a,c,g}

^aDepartment of Molecular Genetics and Microbiology, Pontificia Universidad Católica de Chile, Santiago, Chile

^bDepartment of Oceanography, Universidad de Concepción, Concepción, Chile

^cCenter for Climate and Resilience Research (CR)², Santiago, Chile

^dEscuela de Medicina, Universidad Espíritu Santo, Samborondón, Ecuador

^eMax Planck Institute for Marine Microbiology, Bremen, Germany

^fDepartamento de Biología de Sistemas, Centro Nacional de Biotecnología (CSIC), Madrid, Spain

^gCenter for Genome Regulation (CRG), Santiago, Chile

ABSTRACT Microbial proton-pumping rhodopsins are considered the simplest strategy among phototrophs to conserve energy from light. Proteorhodopsins are the most studied rhodopsins thus far because of their ubiquitous presence in the ocean, except in Antarctica, where they remain understudied. We analyzed proteorhodopsin abundance and transcriptional activity in the Western Antarctic coastal seawaters. Combining quantitative PCR (qPCR) and metagenomics, the relative abundance of proteorhodopsin-bearing bacteria accounted on average for 17, 3.5, and 29.7% of the bacterial community in Chile Bay (South Shetland Islands) during 2014, 2016, and 2017 summer-autumn, respectively. The abundance of proteorhodopsin-bearing bacteria changed in relation to environmental conditions such as chlorophyll *a* and temperature. *Alphaproteobacteria*, *Gammaproteobacteria*, and *Flavobacteriia* were the main bacteria that transcribed the proteorhodopsin gene during day and night. Although green light-absorbing proteorhodopsin genes were more abundant than blue-absorbing ones, the latter were transcribed more intensely, resulting in >50% of the proteorhodopsin transcripts during the day and night. *Flavobacteriia* were the most abundant proteorhodopsin-bearing bacteria in the metagenomes; however, *Alphaproteobacteria* and *Gammaproteobacteria* were more represented in the metatranscriptomes, with qPCR quantification suggesting the dominance of the active SAR11 clade. Our results show that proteorhodopsin-bearing bacteria are prevalent in Antarctic coastal waters in late austral summer and early autumn, and their ecological relevance needs to be elucidated to better understand how sunlight energy is used in this marine ecosystem.

IMPORTANCE Proteorhodopsin-bearing microorganisms in the Southern Ocean have been overlooked since their discovery in 2000. The present study identifies taxonomy and quantifies the relative abundance of proteorhodopsin-bearing bacteria and proteorhodopsin gene transcription in the West Antarctic Peninsula's coastal waters. This information is crucial to understand better how sunlight enters this marine environment through alternative ways unrelated to chlorophyll-based strategies. The relative abundance of proteorhodopsin-bearing bacteria seems to be related to environmental parameters (e.g., chlorophyll *a*, temperature) that change yearly at the coastal water of the West Antarctic Peninsula during the austral late summers and early autumns. Proteorhodopsin-bearing bacteria from Antarctic coastal waters are potentially able to exploit both the green and blue spectrum of sunlight and are a prevalent group during the summer in this polar environment.

KEYWORDS Antarctica, photoheterotrophy, proteorhodopsin, marine microbiology, metagenomics, metatranscriptomes, sunlight


Citation Cifuentes-Anticevic J, Alcamán-Arias ME, Alarcón-Schumacher T, Tamayo-Leiva J, Pedrós-Alió C, Farías L, Díez B. 2021. Proteorhodopsin phototrophy in Antarctic coastal waters. *mSphere* 6:e00525-21. <https://doi.org/10.1128/mSphere.00525-21>.

Editor Barbara J. Campbell, Clemson University

Copyright © 2021 Cifuentes-Anticevic et al.

This is an open-access article distributed under the terms of the [Creative Commons Attribution 4.0 International license](https://creativecommons.org/licenses/by/4.0/).

Address correspondence to Beatriz Díez, bdiez@bio.puccl.

 Proteorhodopsin-bearing bacteria are prevalent in Antarctic coastal waters through the austral summer. Proteorhodopsin gene abundance and transcription during day and night in coastal waters @cifuentesjero @liolaemus @TamayoLeiva_J @CarlosPedros7

Received 9 June 2021

Accepted 9 August 2021

Published 18 August 2021

Seasonal light availability during the Antarctic summer plays a critical role in shaping phytoplankton and bacterioplankton communities, which are central players in the biogeochemical cycles and food webs of marine ecosystems such as the Southern Ocean (1, 2). Seasonal variation in sea ice cover and day length modulates light availability, resulting in high productivity in the summer and very low productivity in the winter (3–5). During the spring/summer, photoautotrophs (such as diatoms and haptophytes) use light as a primary energy source via chlorophyll *a* (Chl*a*) (6). However, photoheterotrophic microorganisms possess two additional mechanisms for harvesting sunlight energy: (i) bacteriochlorophyll-based photosystems (7) and (ii) rhodopsin (8, 9). Aerobic anoxygenic phototrophic bacteria use bacteriochlorophyll *a* and several other pigments (i.e., carotenoids) to capture light, as well as sophisticated machinery to transport protons across the membrane (10). In contrast, rhodopsin-based phototrophy, the simplest strategy among phototrophs, consists of a single integral membrane protein with a covalently bonded retinal (11). To date, proteorhodopsin is the most studied rhodopsin because of its presence in different bacterial phyla and its wide distribution in the ocean (12). However, there are few quantitative measurements of proteorhodopsin-bearing bacteria available for marine environments, where it has been suggested that proteorhodopsins are a major energy-conserving strategy to capture sunlight in the surface ocean (13). Particularly, in high latitude environments, which exhibit marked changes in light variability, the presence and abundance of proteorhodopsin-bearing bacteria have been documented in both the Beaufort and Chukchi Seas of the Arctic (14–16). For example, proteorhodopsin-bearing bacteria, mainly affiliated with *Alphaproteobacteria*, accounted for 1% to 45% of the marine bacterial abundance throughout the photic zone during the summer in the Beaufort Sea (16), and proteorhodopsin gene transcripts have been found during both the winter and summer seasons (15).

So far, the presence and expression of proteorhodopsin genes in the Antarctic marine ecosystem remain much more understudied than in the Arctic Ocean (see Table 1 for the limited references of proteorhodopsin genes in the Antarctic marine ecosystem). The first report of a functional blue-absorbing proteorhodopsin (blue-PR) in Antarctica was reported in 2001 at Palmer Station (Anvers Island) on the west coast of the Antarctic Peninsula (8). Subsequently, transcription of proteorhodopsin genes related to *Flavobacteriia* (*Polaribacter*), *Alphaproteobacteria* (SAR11 clade), and *Gammaproteobacteria* (SAR92 clade) in the sea ice microbial community from the Ross Sea region during the austral summer was reported (17). The presence of the gene and protein was demonstrated through metagenomics and metaproteomics in the coastal surface seawater of Palmer Station (18, 19). Finally, concentrations of rhodopsin-based photosystems were determined using retinal as a proxy in the subantarctic waters of the Subtropical Frontal Zone off New Zealand, showing that the abundance of rhodopsin-based photosystems was, on average, 20 times higher than that of Chl*a*-based photosystems (20). However, in Antarctica's coastal waters, the relative contribution of proteorhodopsin-bearing bacteria to the microbial community remains elusive, as the effect of environmental parameters on the transcription of the proteorhodopsin gene. Thus, the present study is the first to identify the taxonomic affiliation and relative abundance of proteorhodopsin-bearing bacteria, as well as proteorhodopsin gene transcription, in coastal marine waters of the West Antarctic Peninsula (WAP) under contrasting environmental conditions (Chl*a* levels) and light availability (during day and night). The information provided here will help to understand how proteorhodopsin-driven phototrophy contributes to enter sunlight energy into this ecosystem. Additionally, we investigated whether proteorhodopsin can act as a light-driven proton pump by analyzing the transcript levels of enzymes involved in the retinal biosynthetic pathway and determining ion pumping and spectral tuning residues of proteorhodopsin protein sequences.

We investigated proteorhodopsin-bearing bacterial dynamics across three austral late summers-early autumns (2014, 2016, and 2017) in Chile Bay (South Shetland Islands, WAP). The shallow water (~200 m) of Chile Bay is subjected to both strong intraseasonal and interannual variability, mainly due to tidal and wind-driven processes

TABLE 1 Proteorhodopsin sequences identified in Antarctic marine samples and main findings regarding taxonomy and absorption spectra (July 2021)

Sampling site	Finding(s)	Absorbing spectra	Taxon ^a	Methodology	Reference
Palmer Station	Proteorhodopsin gene present	Blue		Fosmid clone	8
Palmer Station	Proteorhodopsin gene present	Blue	Likely <i>Proteobacteria</i>	Fosmid clone	25
Ross Sea sea ice	Proteorhodopsin gene and transcript in sea ice bacteria	Blue and green	<i>Alphaproteobacteria</i> , <i>Gammaproteobacteria</i> , and <i>Bacteroidetes</i>	PCR amplicons	17
Palmer Station	Proteorhodopsin present in summer and winter		SAR11, OMG	Metaproteome	19
Palmer Station	Proteorhodopsin present in summer and winter			Metagenome	18
King George Island Potter Cove sediments	Proteorhodopsin in viruses and bacteria	Green	Many	Metagenome	26
King George Island	Proteorhodopsin gene present	Blue and green		PCR amplicons	27
Subtropical frontal zone off New Zealand (subantarctic waters)	Rhodopsin photosystem concn			Quantification of retinal as proxy a for rhodopsins	20

^aEntries in the Taxon column correspond to the reported proteorhodopsin-bearing microorganisms in the original reference. OMG, oligotrophic marine *Gammaproteobacteria* group.

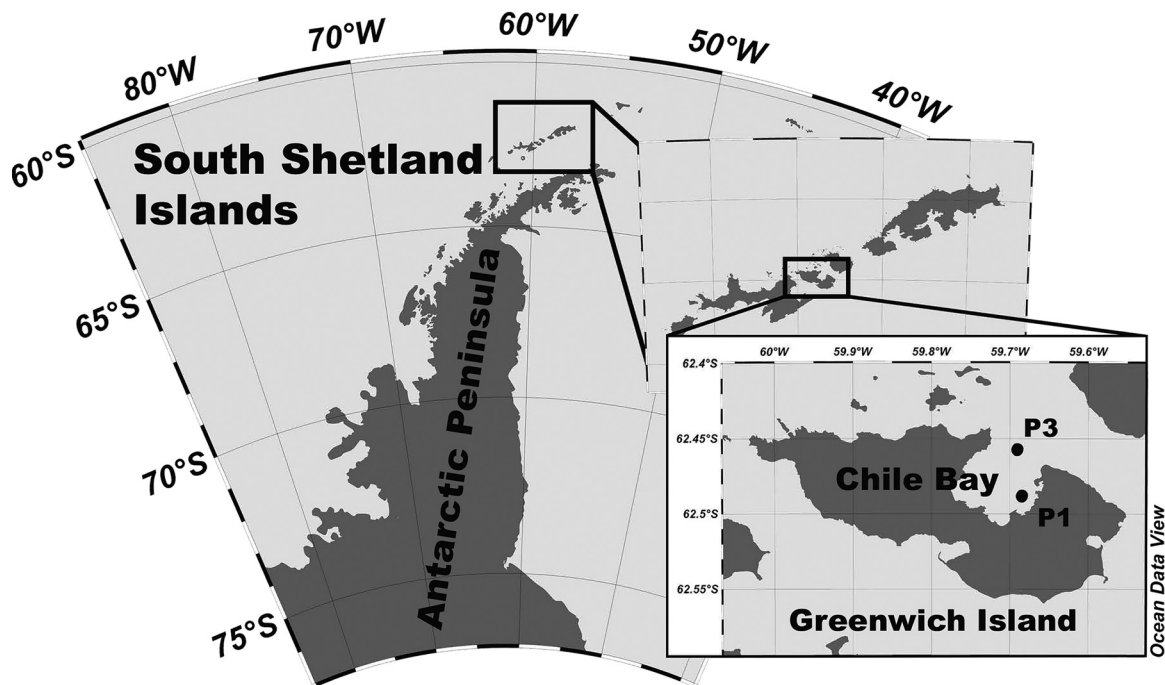


FIG 1 Map of Chile Bay on Greenwich Island, South Shetland Islands, Antarctica. Sampling sites P1 (62.49°S, 59.66°W) and P3 (62.49°S, 59.66°W) are indicated by black circles on the inset map.

(21), with significant modulation by climate events like El Niño–Southern Oscillation and the Southern Annular Mode (22). Finally, Antarctic surface water (23), which is warmer ($>0^{\circ}\text{C}$), fresher (salinity <33.5), and richer in nutrients (nitrate $>15\ \mu\text{M}$), is carried to Chile Bay by the Bransfield Current.

RESULTS

Environmental conditions at the study site. Marine surface (2 m) and subsurface (30 m) waters of the two locations in Chile Bay (Fig. 1) were monitored between February and March during the summers/autumns of 2014 ($n = 4$), 2016 ($n = 10$), and 2017 ($n = 19$). Environmental data from 2014, 2016, and 2017 summers are summarized in Table 2 (metadata for each sample are shown in Table S1 in the supplemental material), and data from 2014 are also published (24, 25). Briefly, in 2014, surface water temperature and Chl a increased from -0.1°C and $0.3\ \text{mg m}^{-3}$ in February to 0.3°C and $2.5\ \text{mg m}^{-3}$ in March (24, 25). In 2016, Chl a concentration varied from 1.0 to $11.4\ \text{mg m}^{-3}$, and seawater temperature varied from -0.2 to 0.8°C . In 2017, the Chl a levels were between 0.2 and $1.29\ \text{mg m}^{-3}$, and the seawater temperature was the highest of the 3 years (1.4 to 2.7°C). Nutrients such as nitrate and phosphate levels were on average $>15.56\ \mu\text{M}$ and $>1.35\ \mu\text{M}$, respectively, for all years (Table 2).

Abundance, phylogenetic analysis, and daily transcriptional activity of proteorhodopsin genes. The proteorhodopsin gene abundance, taxonomic affiliation, and transcriptional expression (day and night) were investigated on two metagenomes and four metatranscriptomes (two daytime, two nighttime) from different sampling periods of the summer of 2014 (24). The metagenomic analysis revealed 66 proteorhodopsin gene sequences. Comparing the abundance of proteorhodopsin reads to single-copy genes, we determined that proteorhodopsin-bearing bacteria represented on average $17\% \pm 4\%$ and $17\% \pm 7\%$, in February and March, respectively, of the total bacteria. The results from meta-omic analyses are shown in Fig. 2, including the relative abundance of each proteorhodopsin operational taxonomic unit (OTU) during the two periods (February and March) and their transcription during the day and night. The taxonomic composition of the proteorhodopsin sequences was very similar for February and March (Fig. 2; see also Fig. S1 in the

TABLE 2 Environmental variables obtained in Chile Bay during the summers of 2014, 2016, and 2017^a

Year	Date (day.month)	Sampling point	Depth (m)	Time	Temp (°C)	Salinity (PSU)	Nitrate (μM)	Nitrite (μM)	Ammonium (μM)	Phosphate (μM)	Silicate (μM)	Chlorophyll (mg m ⁻³)	Analysis	n
2014	11.02	P3	2	N	-0.1 ^b	33.1 ^b	22.4	0.17	nd	1.86	46.8	0.31	MetaT	1
	14.02	P3	2	D	-0.1 ^b	33.2 ^b	19.5	0.16	nd	1.61	43.6	0.36	MetaG-MetaT	1
	03.03	P3	2	D	0.3 ^b	33.5 ^b	20.2	0.16	nd	1.59	44.1	2.38	MetaT	1
2016	04.03	P3	2	N	0.3 ^b	33.9 ^b	17.7	0.10	nd	1.28	37.9	2.53	MetaG-MetaT	1
	18.02–04.03	P1	2	D/N	0.1 ± 0.3	nd	15.56 ± 2.95	0.29 ± 0.03	nd	1.35 ± 0.26	51.57 ± 8.71	5.8 ± 4.3	(RT)-qPCR	4
2017	08.02–25.02	P1	2	D/N	-0.16 ± 1.14	nd	16.34 ± 3.09	0.28 ± 0.04	nd	1.46 ± 0.23	51.28 ± 7.42	1.97 ± 0.98	(RT)-qPCR	6
		P1	2	D/N	2.06 ± 0.41	33.89 ± 0.19	20.09 ± 3.01	0.19 ± 0.02	0.67 ± 0.24	1.43 ± 0.13	49.79 ± 7.98	0.59 ± 0.31	(RT)-qPCR	5
		P3	2	D/N	2.03 ± 0.32	33.95 ± 0.12	22.56 ± 2.11	0.21 ± 0.02	0.56 ± 0.19	1.54 ± 0.12	50.13 ± 8.06	0.5 ± 0.34	(RT)-qPCR	9
		P3	30	D/N	1.53 ± 0.12	34.16 ± 0.01	23.41 ± 4.55	0.20 ± 0.03	0.66 ± 0.14	1.62 ± 0.13	53.24 ± 6.43	0.33 ± 0.17	qPCR	5

^aData are means ± standard deviations. Abbreviations: D, daytime; N, nighttime; PSU, practical salinity unit; nd, no data; Meta-G, metagenome; Meta-T, metatranscriptome; (RT)-qPCR, PR gene and transcript quantification by qPCR; qPCR, PR gene quantification by qPCR.

^bDuring 2014, temperature and salinity were measured using a multiparameter sensor (Oakton PCD650).

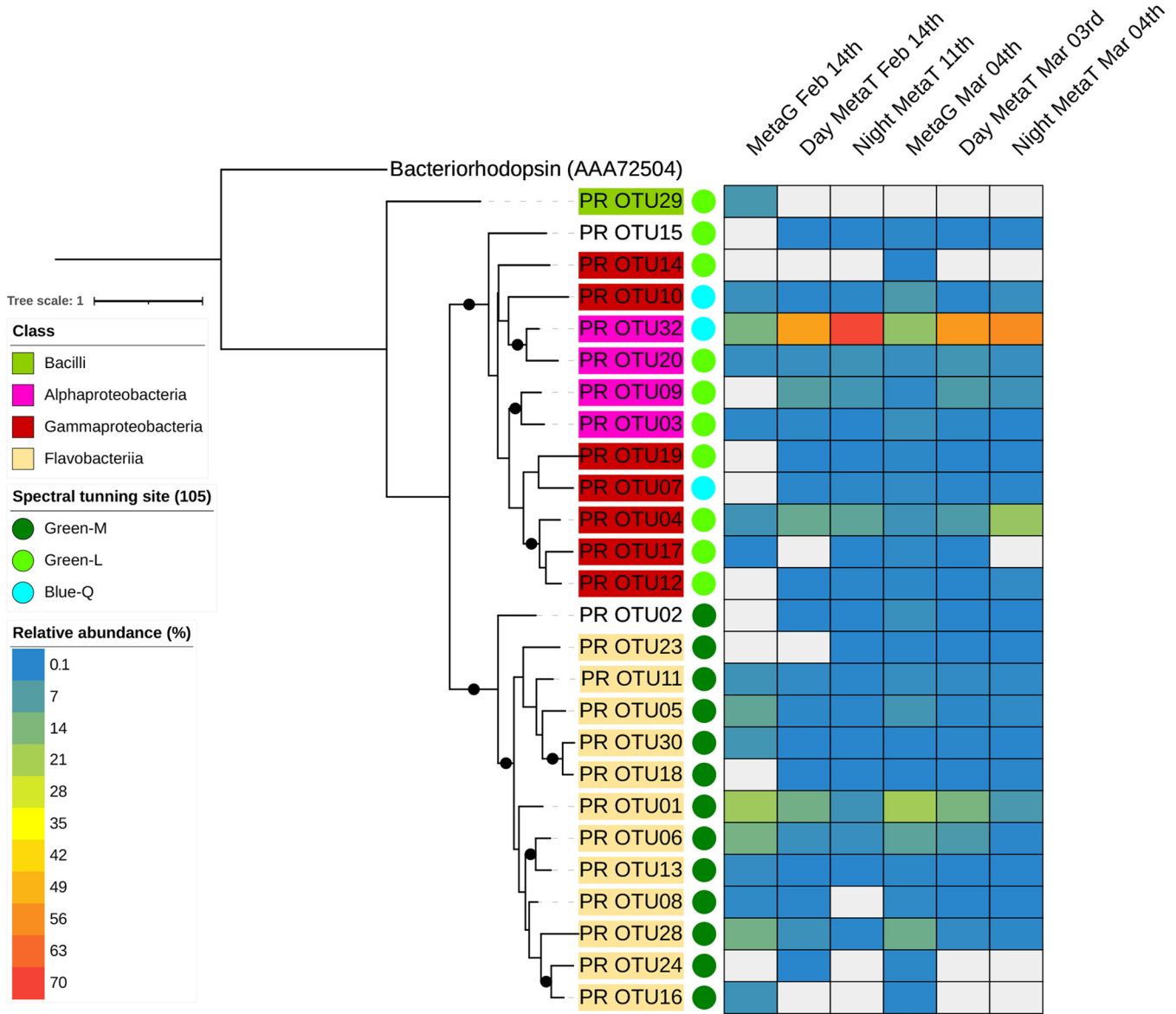


FIG 2 Phylogenetic reconstruction of proteorhodopsin proteins (PR OTUXX) and their relative abundance in metagenomes (MetaG) and metatranscriptomes (MetaT) from austral summer 2014 in Chile Bay. The color scale (blue-yellow-red) represents those sequences with a relative abundance from >0.1% to 70%, while light gray represents those sequences with a relative abundance from 0 to 0.1%. Taxonomic classification was inferred from the phylogenetic placement of proteorhodopsin sequences from Chile Bay in the phylogenetic reconstruction with reference sequences (see Fig. S1 in the supplemental material). Taxonomy of the reference sequences was obtained from metadata available in the MicRhoDE database (56). The tree is rooted to *Halobacterium salinarum* bacteriorhodopsin (NCBI accession no. AAA72504). Black circles above the nodes indicate >95% Ultra Fast bootstrap support and >80% SH-*alrt* branch support.

supplemental material). Proteorhodopsin reads in both metagenomes were mainly affiliated with *Bacteroidetes* (70 to 55%; primarily *Flavobacteriia*), *Alphaproteobacteria* (17 to 25%; SAR11 and SAR116), and *Gammaproteobacteria* (8 to 16%; SAR86 and SAR92) (Fig. S1 and Table S2). Conversely, the metatranscriptomic reads for all samples showed that *Alphaproteobacteria* accounted for a larger fraction than *Flavobacteriia*, for which all members were underrepresented in the metatranscriptomes. A blue-PR from SAR11 clade was the most transcribed proteorhodopsin gene, while a green-absorbing proteorhodopsin (green-PR) from this same clade was transcribed to a lesser extent. *Gammaproteobacteria* were also proportionally more abundant in the metatranscriptomes than in the metagenomes. SAR92 and SAR86 were the main *Gammaproteobacteria* that transcribed the proteorhodopsin gene, while PR OTU01 and PR OTU06 were the main proteorhodopsin transcribed genes among the *Flavobacteriia*. The differences found between night and day may be

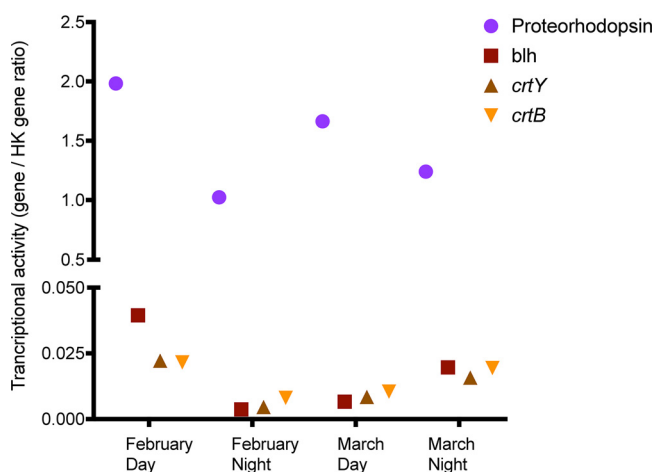


FIG 3 Relative transcriptional levels of proteorhodopsin gene and retinal biosynthetic genes during the austral summer of 2014 under low (February) and high (March) chlorophyll *a* concentrations in the seawater. Relative transcript abundances were calculated using the ratio of transcripts for each gene: proteorhodopsin, β -carotene dioxygenase (*blh*), lycopene cyclase (*crtY*), and phytoene synthase (*crtB*) to the mean transcript abundance of the housekeeping (HK) genes (*rplB*, *rpoB*, and EF-Tu).

attributed to the fact that transcription was generally higher in daytime, depicting that light availability may regulate proteorhodopsin gene transcription differently for every taxa (Fig. 3). It appears that *Flavobacteriia* transcription was considerably reduced at night, resulting in a larger proportion of the expression being attributed to “*Candidatus Pelagibacter*.” Furthermore, in searching for retinal biosynthetic enzymes, we found gene sequences for *blh*, *crtY*, and *crtB* in both metagenomes, and they were transcribed during both the day and night (Fig. 3). Thus, proteorhodopsin-bearing bacteria should be able to synthesize retinal or acquire it from the environment and use it as a chromophore. However, retinal biosynthetic enzymes were considerably less transcribed than the proteorhodopsin gene (>10-fold), with no observed difference between day and nighttime samples.

Proteorhodopsin functional domain. All retrieved proteorhodopsin sequences harbored the most conserved proton pumping domain (aspartic acid, threonine, and glutamic acid at positions 97, 101, and 108, respectively), except for one sequence from *Exiguobacterium*, which had a lysine at position 108. Thus, all of the proteorhodopsin genes encoded a fully functional protein that, in turn, could act as a light-driven proton pump (Fig. 4). In the light tuning position (position 105), proteorhodopsin sequences mostly harbored leucine or methionine (common in green-PR), while some had glutamine (typical in blue-PRs). During both months, green-PR genes (83 to 73% in February and March, respectively) were more abundant than blue-PR genes (Fig. 2 and Table S2). However, the transcription of green-PR was comparable to that of blue-PR during the day and night, were blue-PR accounted for more than 50% of the proteorhodopsin transcripts during day and night (Table S2). Throughout the summer of 2014, SAR11 was the main taxon transcribing blue-PR, while other proteorhodopsin-bearing bacteria related to the *Flavobacteriia* class SAR92 and SAR86 clade transcribed green-PR (Fig. 2 and Table S2).

Daily temporal variability of the proteorhodopsin gene abundance and transcriptional activity. An estimation of the relative abundance of proteorhodopsin-bearing bacteria in Chile Bay was determined by quantitative PCR (qPCR) for samples taken at 2 m and 30 m during 2016 ($n = 10$) and 2017 ($n = 19$) at the two different sampling sites (Fig. 1). We did not find any significant differences (Kruskal-Wallis [KW], $P > 0.5$) in the relative abundances of proteorhodopsin-bearing bacteria from the targeted taxa between the two sampling sites, therefore both points were considered replicates (Tables S3 and S4). Through qPCR, we estimated that the relative abundance of proteorhodopsin-bearing bacteria from the SAR11 clade, SAR92 clade, and *Flavobacteria* from the NASB clade (Flavo-NASB-like) ranged from 1.1 to 5.7% of the bacterial community

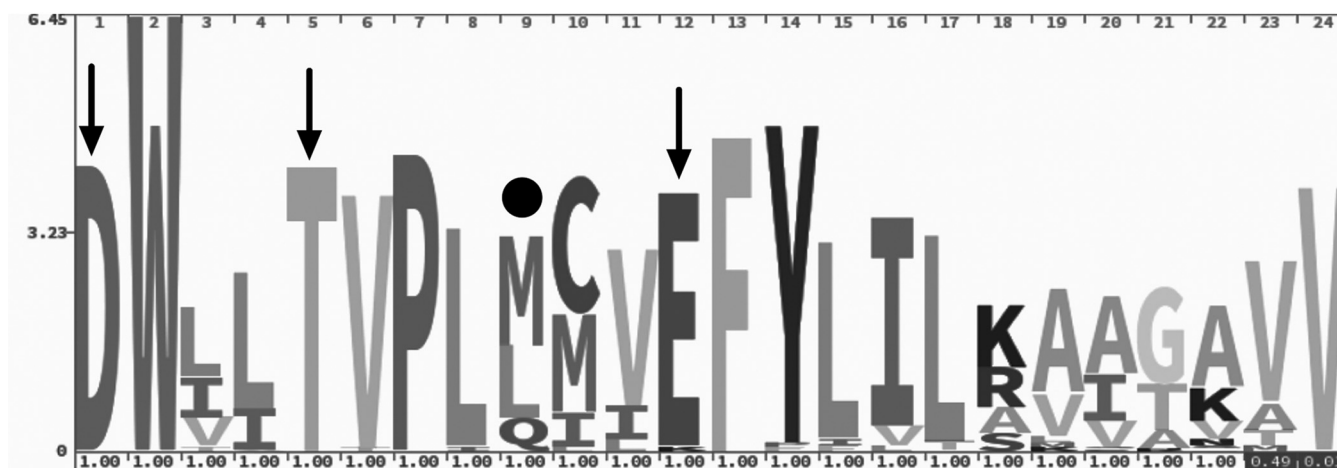


FIG 4 HMM logo of the proteorhodopsin proton pumping domain and spectral tuning amino acid retrieved from Chile Bay. The arrows indicate amino acids involved in proton pumping at positions 97, 101, and 108. The solid black circle indicates the amino acid of the spectral tuning position at position 105.

in 2016, whereas a higher relative abundance was found in 2017 (12.6 to 63.3%). In both years, proteorhodopsin sequences of the SAR11 clade were more abundant than those of the SAR92 clade and Flavo-NASB-like (Table S4 and Fig. S2). Additionally, SAR11 proteorhodopsin sequences were more abundant in 2017 (at both depths) than in 2016 (KW, $P = 6e-5$ [Fig. 5A]), while no significant differences between years and depths were found for the SAR92 clade or the Flavo-NASB-like clade (Table S4 and Fig. 5A). When comparing the relative abundances of SAR11 proteorhodopsin genes between the 2 years, it was higher during 2017, inversely to the *Chla* levels from Chile Bay, which were higher during 2016 (KW, $P = 8.5e-5$; Fig. 5B). Spearman's correlation (80) was used to estimate correlations between physicochemical and biological data. Environmental variables, such as temperature, *Chla*, nitrite, nitrate, and phosphate, correlated with the relative abundances of proteorhodopsin genes from the SAR11 clades (Fig. S3). Temperature ($r = 0.64$, $P < 0.05$), ammonia ($r = 0.73$, $P < 0.05$), and nitrate ($r = 0.48$, $P < 0.05$) correlated positively with the SAR11 proteorhodopsin gene abundance, while *Chla* levels ($r = -0.51$, $P < 0.05$) and nitrite ($r = -0.71$, $P < 0.05$) correlated negatively with this taxon. The relative abundance of the Flavo-NASB-like proteorhodopsin gene positively correlated with ammonia levels ($r = 0.57$, $P < 0.05$). Other environmental variables, including salinity, oxygen (both available only for 2017), silicate, and phosphate, did not significantly correlate with any proteorhodopsin gene abundances (Fig. S3). A permutational multivariate analysis of variance (PERMANOVA) showed that there were significant differences in the relative abundances of the proteorhodopsin gene from SAR11 and SAR92 clades in the marine microbial community with temperature (PERMANOVA, SAR11 $R^2 = 0.418$, $P < 0.01$) and depth (PERMANOVA, SAR92 $R^2 = 0.237$, $P < 0.05$), respectively (Table S5).

Reverse transcription-qPCR (RT-qPCR) was used to quantify proteorhodopsin gene transcripts from SAR11 and Flavo-NASB-like for daytime and nighttime surface (2 m) samples from 2016 and 2017 to confirm the daily variability of the proteorhodopsin gene transcriptional activity observed in 2014. Results show that proteorhodopsin gene transcription was active during both the daytime and nighttime (Fig. 6). Furthermore, no significant difference in transcription was found between daytime and nighttime for the SAR11 proteorhodopsin gene during both summer periods (Fig. 6). Transcription of the Flavo-NASB-like proteorhodopsin gene in 2016 was below the assay detection limit, while in 2017, no significant difference was found between the day and night (Fig. 6).

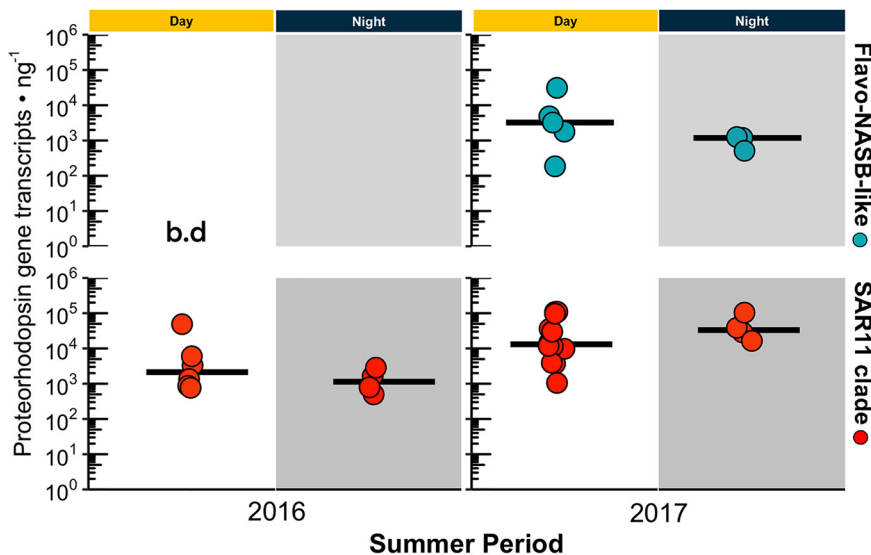


FIG 6 Transcript levels of proteorhodopsin genes of the SAR11 clade and Flavo-NASB-like clade in RNA samples obtained during the day and night in the austral summers of 2016 and 2017. The horizontal line in the graph represents the mean. Flavo-NASB-like proteorhodopsin transcripts were below the detection level (b.d) during the summer of 2016. For 2016, day, $n = 7$; night, $n = 3$. For 2017, day, $n = 10$; night, $n = 3$.

different in polar regions, proteorhodopsin sequences described in the Arctic (14–16) are phylogenetically placed in the same clades as the proteorhodopsin proteins identified in Antarctica (8, 17). Further analyses should be performed to evaluate whether the proteorhodopsin proteins are adapted to cold or whether the shared sequences of the poles respond to the adaptation of the marine bacteria that carry these proteorhodopsin genes.

The spectral tuning ability of proteorhodopsin has drawn attention since 2001, when Béjà et al. (8) reported blue-PR in the Southern Ocean (Palmer Station) bacterioplankton community. For the Antarctic marine environment, the presence of blue-PR and green-PR has been reported only in sea ice microbial communities from the Ross Sea (17). We retrieved green- and blue-absorbing proteorhodopsins, and even during the high-*Chla* period (March) reported in Chile Bay 2014 (24), the relative abundance of green- and blue-PR genes did not change from February to March. As blue-PR using organisms might have been outcompeting for the sunlight resource with phototrophic eukaryotes at that period of the year in Antarctica, the green-PR was generally more abundant as they can exploit a different wavelength than *Chla*-based microorganisms. This result is similar to data from the summer Arctic coastal marine environment (16), thereby confirming the higher relative abundance of green-PR found at high latitude coastal sites (9). Considering that new ice-free coastal areas could appear in the WAP (32), is important to study green-PR photo(hetero)trophy to determine the significance of these microorganisms in these (new) polar coastal areas.

The abundance of proteorhodopsin-bearing bacteria recovered in Chile Bay is at the lower end of the previously reported abundance at diverse environments such as the North Sea (35%) (29), the Chesapeake Bay (40%) (33), the Baltic Sea (46%) (30), the Sargasso Sea (65%) (34), the ultraoligotrophic Eastern Mediterranean Sea, and the San Pedro Channel where the proteorhodopsin-bearing bacteria represented more than 70% of the marine microbial community (35, 36). This low abundance of proteorhodopsin-bearing bacteria in Chile Bay, compared to other nonpolar oceanic regions, may be related to the particular marine community structure or (a)biotic interactions between the proteorhodopsin-bearing bacterial community and the environment. However, the proportion of proteorhodopsin-bearing bacteria has been estimated only in a few studies of polar waters. In the Chukchi and Beaufort Seas, less than 0.4% of the

bacterioplankton harbored a proteorhodopsin gene (2×10^{-3} proteorhodopsin/16S rRNA) (14). Boeuf et al. (16) estimated that in the Southern Beaufort Sea, 32% of the SAR11 bacteria carried the proteorhodopsin gene and that these accounted 63% of all proteorhodopsin-bearing prokaryotes. Moreover, the number of proteorhodopsin-bearing cells represented up to 45% of the total cells (16). In our 3-year study, proteorhodopsin-bearing bacteria ranged widely from 1.2 to 63.3% of the Antarctic coastal water bacterial community. Although meta-omics showed that *Bacteroidetes* was the main proteorhodopsin-bearing phylum in 2014, the qPCR results suggest a low relative abundance of these bacteria in 2016 and 2017. This may be due to a bias in the primer coverage for *Flavobacteriia* proteorhodopsin gene sequences, designed using North Atlantic and Arctic sequences. However, due to the inability to design universally conserved qPCR primers that cover most currently described proteorhodopsins (16), it is difficult to accurately determine the proteorhodopsin-bearing bacterial composition and abundance through qPCR. Comparison of our results with others from previous reports should be made with caution, due to the different methodologies used in each study, and estimating only the relative abundance of the proteorhodopsin-bearing bacteria that can be retrieved with the qPCR primers used.

Proteorhodopsin gene expression in polar waters. Although studies showing the presence of proteorhodopsin gene sequences in polar waters are few, those showing proteorhodopsin gene transcription are substantially less frequent, with the expression of this gene determined in an extremely few cases (17, 19). A study off Palmer Station in Antarctica, which analyzed metaproteomes from six summers and three winters, found only three proteorhodopsin sequences out of 1,061 proteins: two associated with SAR11 and one with the oligotrophic marine *Gammaproteobacteria* group (19). In a seasonal study from early winter to spring in the Amundsen Gulf (Southeastern Beaufort Sea, Arctic Ocean), proteorhodopsin gene transcription was observed at winter darkness in January, after ice breakup in May, and associated with phytoplankton blooms in late June (15). *Gammaproteobacteria* were always the most abundant proteorhodopsin-bearing bacteria, whereas the abundance of proteorhodopsin-bearing *Alphaproteobacteria* and *Bacteroidetes* varied, the former being more abundant in late winter. Koh et al. (17) retrieved only 17 proteorhodopsin sequences from cDNA clone libraries of Ross Sea ice core samples from five locations and three depths. Finally, Boeuf et al. (16) obtained proteorhodopsin sequences from DNA and cDNA along an Arctic coastal transect in the summer, but no quantitative data for proteorhodopsin gene transcription was shown. Thus, our results from Antarctica, which combine metatranscriptomics and RT-qPCR, significantly increase the amount of polar proteorhodopsin gene transcriptional data. Furthermore, the 2014 metatranscriptomic analysis is the first to compare blue- and green-PR transcription levels in polar marine environments, describing that these bacteria use two different sunlight spectra to conserve light energy. Our results show that proteorhodopsin sequences from the WAP exhibited color adaptation, with the blue-PR much more expressed than green-PR, despite its lower abundance. Additionally, SAR11 clade presented mostly blue-PR (but also green-PR at the lower level), whereas *Flavobacteriia* presented (and transcribed) only green-PR.

During the summer of 2014, an interesting pattern emerged regarding the transcriptional activity and the blue/green proteorhodopsin ratio. Green-PR sequences were more abundant in the DNA samples (~90%), as has been described in the Arctic (16), while blue-PR accounted for about 50% of the transcription. Although most of the green-PR belonged to *Bacteroidetes*, their expression was lower than that of the SAR11 blue-PR. This pattern has been reported only for temperate open ocean waters (ALOHA station) (37), but not for polar marine environments. Although RT-qPCR did not reveal significant differences in proteorhodopsin gene transcription during the day and night between the summers of 2016 and 2017, an overall higher transcription was found during the daytime of summer 2014. Previously, circadian transcriptional activity of the proteorhodopsin gene has been reported only from tropical and subtropical marine environments (38–40). It is also worth noting that the relative abundance of *Flavobacteriia*

proteorhodopsin transcripts was lower at night. This suggests a reduction in proteorhodopsin gene transcription at night/dark by *Flavobacteriia* as previously reported (41, 42), resulting in an apparent higher proportion of *Alphaproteobacteria* proteorhodopsin reads.

The (RT-)qPCR results from this study should be taken with caution because the primers used likely did not recover a substantial fraction of the proteorhodopsin diversity found by meta-omics in the 2014 samples. However, some conclusions can be derived. First, the analysis shows large variability in both relative abundance and community composition of proteorhodopsin-bearing bacteria across the three summers. Second, there is a negative correlation between the contribution of SAR11 proteorhodopsin with *Chla*, which were very different during 2016 and 2017 in Chile Bay at the sampling time. This pattern has been previously described in other polar and temperate oceans (13, 14, 16, 35) and may be related to the oligotrophic lifestyle of proteorhodopsin-bearing SAR11 that benefit from low-molecular-weight dissolved organic matter concentrations (43).

Ubiquity of proteorhodopsin-bearing bacteria during summer in the WAP. In the WAP region, aerobic anoxygenic phototrophic bacteria account for up to 8% of the community (7), while photosynthetic cyanobacteria are a minor component of the marine community (19, 44–46). Thus, photoautotrophy in this system mostly relies on eukaryotic microorganisms, whose activity increases during blooms, like those described during the austral summer in Chile Bay (24, 25). Our results demonstrate that every late summer and early autumn in Chile Bay, proteorhodopsin-bearing bacteria are a ubiquitous and dominant group of phototrophic microorganisms, and this may be extrapolated to the WAP coastal waters. Here, we demonstrate not only the presence but also the transcriptional activity of proteorhodopsin-bearing bacteria across three different austral late summers, and we identified that differences in their abundance may be related to environmental parameters. The relative abundance of SAR11 proteorhodopsin negatively correlates with *Chla* levels, while SAR92 proteorhodopsin positively correlates with phosphate levels, as previously reported for the Arctic (16). The negative correlation between the SAR11 clade proteorhodopsin abundance and *Chla* concentrations has also been previously reported for the North Atlantic Ocean (14, 47).

It becomes necessary to determine how these environmental variables drive proteorhodopsin-bearing bacterial composition, abundance, and activity in Antarctic waters to model proteorhodopsin-bearing bacterial dynamics over the incoming years. Further ecological and biochemical studies will also be required to fully understand how proteorhodopsin contributes to microbial energetic metabolism and how light availability might influence or affect rhodopsin photoheterotrophy, particularly in rapidly changing environments such as Antarctica. This is especially important under the present climate change scenario because as some environmental variables change in the ocean (48), it will be relevant to predict how sunlight energy will be used by marine bacteria in the Southern Ocean.

MATERIALS AND METHODS

Sampling site. Seawater samples were collected from Chile Bay on Greenwich Island, South Shetland Island, Antarctica (Fig. 1). Samples for meta-omics were collected during 2014 (February and March) (24). Samples for quantitative PCR (qPCR) were collected as a time series during the late summer of 2016 (18 February to 4 March 2016) and 2017 (8 February to 25 February 2017) (see Table S1 in the supplemental material). Surface (2 m) and subsurface (30 m) seawater was sampled during the day and night at two locations in Chile Bay: (i) P1, which was close to the “Fuerza Aérea” glacier (62°29'2'' S – 59°40'6'' W); and (ii) P3, which was more exposed to the open ocean (62°27'6'' S – 59°40'6'' W). In the summer of 2016, on February 18, sunrise at 05:30 h and sunset at 20:53 h, while for March 4, the sunrise and sunset at 6:14 h and 20:04 h, respectively. During the summer of 2017, sunrise/sunset was at 05:01/21:22 h and 05:53/20:28 h, at the beginning and end of this period, respectively. From a small inflatable boat, seawater was first collected using a hand-operated membrane pump, deposited into clear 20-liter acid-washed (HCl [10%]) bottles, and then transported to the INACH (Instituto Antártico Chileno) laboratory at Chile Bay for processing.

Environmental variables. Seawater temperature (°C), salinity, and oxygen (milliliter liter⁻¹) at the P1 and P3 sampling locations were obtained using a CTD (conductivity, temperature, depth) profiling sensor (Seabird 19; Sea-Bird Electronics, Bellevue, WA, USA). Salinity, oxygen, and ammonium (NH₄⁺)

were measured only for the 2017 samples. To determine nutrients, namely, nitrite, nitrate, phosphate, and silicic acid, triplicate prefiltered (0.7- μm GF/F glass fiber filter) seawater samples were collected in 15-ml polyethylene flasks at P1 (2-m depth) and P3 (2-m and 30-m depth) and stored at -20°C until further analysis. Nutrient concentrations were determined using standard colorimetric techniques with a segmented flow Seal AutoAnalyzer3 (SEAL Analytical GmbH, Norderstedt, Germany) at Universidad de Concepcion as described previously (24). NH_4^+ was measured as previously described (49). For each location and depth, Chl *a* was determined in triplicate by filtering 1 liter of seawater through 0.7- μm GF/F glass fiber filters, which were frozen at -20°C until laboratory analyses by acetone extraction and fluorometric measurements (50).

DNA and RNA extraction. For DNA and RNA analysis of the 2016 and 2017 samples, microbial biomass from 3 to 4 liters of seawater was prefiltered through 200- μm nylon mesh followed by a 20- μm polycarbonate filter using a peristaltic pump (6 to 600 rpm) (model no. 7553-70; Cole Parmer, Vernon Hills, IL, USA) at 50 to 100 ml min^{-1} . The subsequent filtrate was concentrated onto 0.22- μm -pore-size Sterivex units (Millipore, Burlington, MA, USA). The filters for RNA analysis were preserved in RNAlater (Invitrogen, Carlsbad, CA, USA). Both RNA and DNA filters were maintained at -80°C until laboratory processing at the Pontifical Catholic University of Chile, Santiago, Chile. DNA was extracted according to Tillett and Neilan (51) with modifications. Briefly, filters were resuspended in lysis buffer (1% potassium ethyl xanthogenate [Sigma-Aldrich, St. Louis, MO, USA], 100 mM Tris-HCl [pH 7.4], 20 mM EDTA [pH 8], 800 mM ammonium acetate). Sterile glass beads were added and then shaken in a BeadBeater for 30 s. Next, the mixture was incubated with sodium dodecyl sulfate (SDS) (1% final concentration) at 65°C for 2 h and then placed on ice for 30 min. DNA was extracted with phenol-chloroform-isoamyl alcohol (25:24:1), and the residual phenol was eliminated with chloroform-isoamyl alcohol (24:1). The extract was cleaned by overnight precipitation with cold isopropanol and then washed with 70% ethanol. DNA was quantified using the Qubit 2.0 fluorometer (Thermo Fisher Scientific, MA, USA), the quality (A_{260}/A_{280}) was assessed spectrophotometrically, and the integrity was checked by agarose gel electrophoresis. RNA was extracted from the filters using TRIzol (Invitrogen) and the RNA Clean & Concentrator kit (Zymo Research, USA). To eliminate any remaining DNA, 1 μg of RNA from each sample was treated with DNase (Turbo DNase; Invitrogen), and the absence of DNA contamination was assessed analyzing the 16S rRNA gene by PCR using the primers and conditions described below.

Identification of proteorhodopsin gene sequences from 2014 meta-omes. Identification of proteorhodopsin protein-encoding genes and proteorhodopsin gene transcription analysis were performed using previously obtained sequence data from the 0.22- to 20- μm bacterial fraction of the 2014 samples (24). Briefly, surface seawater samples were collected in 2014 on February 11 (nighttime, 21:00 h [local time UTC-3]) and 14 (daytime, 11:00 h) and March 3 (daytime, 13:30 h) and 4 (nighttime, 21:00 h) (Table 2). DNA and RNA were extracted and processed as previously described (24). The metagenomic and metatranscriptomic sequencing data from the 0.22- to 20- μm bacterial fractions used in this study are available at NCBI under BioProject accession no. [PRJNA421008](https://www.ncbi.nlm.nih.gov/bioproject/PRJNA421008). Quality trimming of metagenomic reads was performed using Prinseq (83): a hard clipping of the first 7 leftmost bases and 9 leftmost bases for February and March metagenomes, respectively, mean read quality of 30, and 3' trimming for bases with quality below 30. Similarly, for the metatranscriptomes, a minimum quality of 30, a minimum length of 30 bp, and a hard clipping of the first 11 bases was performed. Low complexity sequences and undetermined bases were filtered (-ns_max_p 0 -lc_method dust -lc_threshold 7) (83) as described previously (52). Trimmed reads were assembled with SPAdes software v3.10.1 (meta option) (53). Contigs larger than 500 bp were used for protein and gene prediction via Prodigal v2.6.3 with meta mode and bypassing the Shine-Dalgarno sequence (54). Identification of proteorhodopsin candidates was made with predicted proteins from both metagenomes using DIAMOND (BLASTP; E value $\leq 10^{-7}$) (55) against the curated database MicRhoDE (56), particularly with those annotated as proteorhodopsin. Next, the potential proteorhodopsins were evaluated using the Pfam database (57) (HMMER 3.0; <http://hmmer.org/>), identifying the HMM profile Bac_rhodopsin (PF01036; hmmscan; trusted cutoff) and selecting those belonging to the InterPro family proteorhodopsin (IPR017402) when classified with InterProScan (58). The resulting proteins were manually curated, identifying the presence of the proton pumping functional domain at the C-helix and those larger than 100 amino acids.

Abundance and expression of the proteorhodopsin gene in the 2014 meta-omes. To determine the relative abundance of proteorhodopsin in the 2014 metagenomes, Bowtie2 v2.2.6 (59) (sensitive; default) was used to recruit reads from the metagenome to the corresponding proteorhodopsin gene. Next, single-copy housekeeping genes *recA*, *rplB*, *rpoB*, and EF-Tu were identified in the predicted proteins of the metagenomes using HMMsearch with the HMM profiles PF00154, PF00181, PF04563, and PF00009, respectively, and then they were verified as belonging to the InterPro families IPR013765, IPR002171, IPR015712, and IPR004541, respectively (33, 34, 36). The average normalized abundance of the proteorhodopsin gene in the metagenomic sequence data was calculated as the average of each ratio of the proteorhodopsin gene to single-copy housekeeping gene, where the number of reads recruited to each gene was normalized by the recruited gene length. To estimate proteorhodopsin gene transcription, we analyzed the RNA sequencing data. Quality trimming of metatranscriptomic reads was performed as described previously (52), and the remaining rRNA sequences were removed using SortMeRNA (60) (default parameters). Nonaligning reads were mapped with Bowtie2 (59) (sensitive; default) to the set of proteorhodopsin genes. Enzymes of the retinal biosynthetic pathway, namely, 15,15'- β -carotene dioxygenase (*blh*), phytoene synthase (*crtB*), and lycopene cyclase (*crtY*), were identified and quantified in the same manner as the single-copy housekeeping genes, but using the HMM profiles PF15461, PF00494, and PF05834, respectively (31). To assess whether there was a difference in the transcription levels between the four metatranscriptomes, we analyzed the *rplB*, *rpoB*, and EF-Tu

genes. Single-copy housekeeping gene reads from the four metatranscriptomes were obtained by the same procedure described above. The relative abundance in each metatranscriptome of the proteorhodopsin gene and retinal biosynthetic pathway enzymes was expressed as the transcript abundance of each gene times the average abundance of the housekeeping genes. Although the *recA* gene has been used as a housekeeping gene to normalize the expression of functional genes in metatranscriptomes (61), it apparently does not exhibit constitutive levels of expression in Antarctic marine waters (62, 63). Therefore, we used *recA* only to normalize proteorhodopsin gene abundance in metagenomes but not to normalize transcriptional activity.

Phylogenetic analysis of proteorhodopsin sequences. A reference tree was built with the retrieved proteorhodopsin proteins from the metagenomes and the “strain_only=strain” sequences from the MicRhoDE database (56). The protein sequences from Chile Bay and the MicRhoDE database were first clustered separately using cd-hit (64, 65) to 82% identity (16) and then aligned with MAFFT (G-INS-i) (66). The multiple sequence alignment was trimmed using trimAl (-gt 0.2) (67), and then the tree was reconstructed with IQ-TREE (automatic model detection, -bb 10000 -alrt 10000) (68, 69). Short amplicon protein sequences from polar marine environments (8, 14–17) were phylogenetically placed using EPA-ng algorithm (70) to the reference proteorhodopsin tree. The resulting tree was processed with GAPP (71) and then visualized in iTOL (72). To determine the proton pumping amino acids at positions 97, 101, and 108 and the spectral tuning switch amino acid at position 105, the protein sequences were screened using the multiple sequence alignment with the positions previously identified in the references sequences and visualized in Skyline (73).

Quantification of proteorhodopsin gene abundances and transcription. qPCR was used to determine the abundances of three proteorhodopsin gene types and the bacterial 16S rRNA gene in the 2016 and 2017 DNA samples using the primers listed in Table S3. In this study, specific primers were used to target the proteorhodopsin gene of *Flavobacteriia* from the NASB clade (47), *Alphaproteobacteria* (clade SAR11) (42), and *Gammaproteobacteria* (clade SAR92) (16), and universal bacterial 16S rRNA gene primers (74). To obtain the standard curve (10^8 to 10^2 copies) for proteorhodopsin gene quantification, the three proteorhodopsin gene types were cloned into the pGEM-T Easy Vector (Promega, Madison, WI, USA). Plasmid DNA was linearized with SacI (Thermo Fisher) and quantified using NanoDrop (Thermo Fisher Scientific). Standard curves for the 16S rRNA gene were generated by amplifying *Escherichia coli* DNA and then purifying the amplicon with GeneJET Gel Extraction kit (Thermo Fisher Scientific). qPCRs were performed in triplicate using $1 \mu\text{l}$ of DNA ($1 \text{ ng } \mu\text{l}^{-1}$) in a final volume of $15 \mu\text{l}$ with the SensiMix kit (Bioline GmbH, Luckenwalde, Germany) and a LightCycler 480 (Roche Holding AG, Basel, Switzerland) real-time qPCR device. The program was as follows: (i) 95° for 10 min and (ii) 40 cycles, with 1 cycle consisting of 10 s at 95°C , annealing for 20 s at the primer-specific temperature (Table S6), and 30 s at 72°C . Proteorhodopsin gene copy numbers were normalized to the 16S rRNA gene copy numbers, assuming 1.9 copies of 16S rRNA and 1 copy of proteorhodopsin gene per genome (16, 34). cDNA was synthesized from 750 ng of DNA-free RNA using the iScript Select cDNA Synthesis kit (Bio-Rad, Hercules, CA, USA) with the same SAR11 and *Flavobacteriia* primers used for the qPCR. Quantification was performed as described above for the DNA samples.

Statistics. The conducted statistical analyses were performed in R with the stats (75), vegan (76), hmisc (77), and corrplot (78) packages. Oceanographic variables were standardized using the z-score method (mean 0, variance 1). Missing values in the temperature factor were computed by linear interpolation with the imputets (79) package. Spearman's rank correlation (ρ) (80) was performed to estimate simple correlations ($P < 0.05$) between standardized oceanographic factors (i.e., physicochemical and biological) ($N = [2016] \text{ day} = 7, \text{ night} = 3; [2017] \text{ day} = 15, \text{ night} = 4$). For multiple comparisons, adjusted P values were obtained using the Stats package and Benjamini-Hochberg (BH) *post hoc* tests (81). To estimate the variance of taxonomic group proteorhodopsin percentage abundances by sampling dates, depths, and between taxonomic groups, the Kruskal-Wallis rank sum analysis was applied (81) ($df = 2, n = N$). A permutational analysis of variance (PERMANOVA) (82) with marginal effect was applied to taxonomic group as the explanatory factor ($df = 1, n = N$) using the function `adonis2` of the R vegan package (9,999 permutations) to estimate the impact of the oceanographic variables on the proteorhodopsin percentage abundances.

Data availability. Raw metagenomic data for each metagenome and metatranscriptome were deposited in the Sequence Read Archive database under BioProject accession number [PRJNA421008](https://www.ncbi.nlm.nih.gov/bioproject/PRJNA421008).

SUPPLEMENTAL MATERIAL

Supplemental material is available online only.

FIG S1, PDF file, 1.5 MB.

FIG S2, PDF file, 0.3 MB.

FIG S3, PDF file, 0.6 MB.

TABLE S1, XLSX file, 0.01 MB.

TABLE S2, XLSX file, 0.01 MB.

TABLE S3, XLSX file, 0.01 MB.

TABLE S4, XLSX file, 0.01 MB.

TABLE S5, XLSX file, 0.01 MB.

TABLE S6, XLSX file, 0.01 MB.

ACKNOWLEDGMENTS

We are grateful to the personnel of Armada de Chile at the Arturo Prat Station and Instituto Antártico Chileno (INACH) for their support with logistical operations. We are also thankful to Beatriz Fernández, Leslie Abarzúa, Josefa Verdugo, and Giovanni Testa for their technical assistance in sample collection, nucleic acid extraction, and nutrient analysis. We thank Christina Ridley for her help in the editing the language of this paper and for her valuable advice on it.

This work was financially supported by INACH RT_15-10 and RT_04-19, ANID for international cooperation DPI20140044, ANID/FONDAP 15110009 and 15200002, ANID MSc. scholarship 22172327, ANID National Doctoral Scholarship 21171048, and grant CTM2016-80095-C2-1-R from the Spanish Ministry of Economy and Competitiveness. Carlos Pedrós-Alió is a member of the PTI POLARCSIC.

J.C.-A. and B.D. designed the study. J.C.-A., M.E.A.-A., T.A.-S., and J.T.-L. processed and analyzed sequence data. C.P.-A. and L.F. helped with the Discussion and critically reviewed the manuscript for journal submission. J.C.-A. and B.D. wrote the manuscript. All authors contributed to the article and approved the submitted version.

We declare that this research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

REFERENCES

1. Azam F, Smith DC, Hollibaugh JT. 1991. The role of the microbial loop in Antarctic pelagic ecosystems. *Polar Res* 10:239–244. <https://doi.org/10.1111/j.1751-8369.1991.tb00649.x>.
2. Deppeler SL, Davidson AT. 2017. Southern Ocean phytoplankton in a changing climate. *Front Mar Sci* 4:40. <https://doi.org/10.3389/fmars.2017.00040>.
3. Vernet M, Martinson D, Iannuzzi R, Stammerjohn S, Kozłowski W, Sines K, Smith R, Garibotti I. 2008. Primary production within the sea-ice zone west of the Antarctic Peninsula. I. Sea ice, summer mixed layer, and irradiance. *Deep Sea Res Part II Top Stud Oceanogr* 55:2068–2085. <https://doi.org/10.1016/j.dsr2.2008.05.021>.
4. Giovannoni SJ, Vergin KL. 2012. Seasonality in ocean microbial communities. *Science* 335:671–676. <https://doi.org/10.1126/science.1198078>.
5. Wilkins D, Yau S, Williams TJ, Allen MA, Brown MV, Demaere MZ, Lauro FM, Cavicchioli R. 2013. Key microbial drivers in Antarctic aquatic environments. *FEMS Microbiol Rev* 37:303–335. <https://doi.org/10.1111/1574-6976.12007>.
6. Arrigo KR, Robinson DH, Worthen DL, Dunbar RB, DiTullio GR, VanWoert M, Lizotte MP. 1999. Phytoplankton community structure and the draw-down of nutrients and CO₂ in the Southern Ocean. *Science* 283:365–367. <https://doi.org/10.1126/science.283.5400.365>.
7. Kirchman D, Stegman M, Nikrad M, Cottrell M. 2014. Abundance, size, and activity of aerobic anoxygenic phototrophic bacteria in coastal waters of the West Antarctic Peninsula. *Aquat Microb Ecol* 73:41–49. <https://doi.org/10.3354/ame01709>.
8. Bèjà O, Spudich EN, Spudich JL, Leclerc M, DeLong EF. 2001. Proteorhodopsin phototrophy in the ocean. *Nature* 411:786–789. <https://doi.org/10.1038/35081051>.
9. Pinhassi J, DeLong EF, Bèjà O, González JM, Pedrós-Alió C. 2016. Marine bacterial and archaeal ion-pumping rhodopsins: genetic diversity, physiology, and ecology. *Microbiol Mol Biol Rev* 80:929–954. <https://doi.org/10.1128/MMBR.00003-16>.
10. Kolber ZS, Plumley FG, Lang AS, Beatty JT, Blankenship RE, VanDover CL, Vetriciani C, Koblizek M, Rathgeber C, Falkowski PG. 2001. Contribution of aerobic photoheterotrophic bacteria to the carbon cycle in the ocean. *Science* 292:2492–2495. <https://doi.org/10.1126/science.1059707>.
11. Bèjà O, Aravind L, Koonin EV, Suzuki MT, Hadd A, Nguyen LP, Jovanovich SB, Gates CM, Feldman RA, Spudich JL, Spudich EN, DeLong EF. 2000. Bacterial rhodopsin: evidence for a new type of phototrophy in the sea. *Science* 289:1902–1906. <https://doi.org/10.1126/science.289.5486.1902>.
12. Finkel OM, Bèjà O, Belkin S. 2013. Global abundance of microbial rhodopsins. *ISME J* 7:448–451. <https://doi.org/10.1038/ismej.2012.112>.
13. Gómez-Consarnau L, Raven JA, Levine NM, Cutter LS, Wang D, Seegers B, Arístegui J, Fuhrman JA, Gasol JM, Sañudo-Wilhelmy SA. 2019. Microbial rhodopsins are major contributors to the solar energy captured in the sea. *Sci Adv* 5:eaaw8855. <https://doi.org/10.1126/sciadv.aaw8855>.
14. Cottrell MT, Kirchman DL. 2009. Photoheterotrophic microbes in the Arctic Ocean in summer and winter. *Appl Environ Microbiol* 75:4958–4966. <https://doi.org/10.1128/AEM.00117-09>.
15. Nguyen D, Maranger R, Balagué V, Coll-Lladó M, Lovejoy C, Pedrós-Alió C. 2015. Winter diversity and expression of proteorhodopsin genes in a polar ocean. *ISME J* 9:1835–1845. <https://doi.org/10.1038/ismej.2015.1>.
16. Boeuf D, Lami R, Cunnington E, Jeanthon C. 2016. Summer abundance and distribution of proteorhodopsin genes in the western Arctic Ocean. *Front Microbiol* 7:1584. <https://doi.org/10.3389/fmicb.2016.01584>.
17. Koh EY, Atamna-Ismaeel N, Martin A, Cowie ROM, Beja O, Davy SK, Maas EW, Ryan KG. 2010. Proteorhodopsin-bearing bacteria in Antarctic sea ice. *Appl Environ Microbiol* 76:5918–5925. <https://doi.org/10.1128/AEM.00562-10>.
18. Grzymalski JJ, Riesenfeld CS, Williams TJ, Dussaq AM, Ducklow H, Erickson M, Cavicchioli R, Murray AE. 2012. A metagenomic assessment of winter and summer bacterioplankton from Antarctica Peninsula coastal surface waters. *ISME J* 6:1901–1915. <https://doi.org/10.1038/ismej.2012.31>.
19. Williams TJ, Long E, Evans F, Demaere MZ, Lauro FM, Raftery MJ, Ducklow H, Grzymalski JJ, Murray AE, Cavicchioli R. 2012. A metaproteomic assessment of winter and summer bacterioplankton from Antarctic Peninsula coastal surface waters. *ISME J* 6:1883–1900. <https://doi.org/10.1038/ismej.2012.28>.
20. Hassanzadeh B, Thomson B, Deans F, Wenley J, Lockwood S, Currie K, Morales SE, Steindler L, Sañudo-Wilhelmy SA, Baltar F, Gómez-Consarnau L. 2021. Microbial rhodopsins are increasingly favoured over chlorophyll in high nutrient low chlorophyll waters. *Environ Microbiol Rep* 13:401–406. <https://doi.org/10.1111/1758-2229.12948>.
21. Llanillo PJ, Aiken CM, Cordero RR, Damiani A, Sepúlveda E, Fernández-Gómez B. 2019. Oceanographic variability induced by tides, the intraseasonal cycle and warm subsurface water intrusions in Maxwell Bay, King George Island (West-Antarctica). *Sci Rep* 9:18571. <https://doi.org/10.1038/s41598-019-54875-8>.
22. Moffat C, Meredith M. 2018. Shelf-ocean exchange and hydrography west of the Antarctic Peninsula: a review. *Philos Trans A Math Phys Eng Sci* 376:20170164. <https://doi.org/10.1098/rsta.2017.0164>.
23. Capella JE, Ross RM, Quetin LB, Hofmann EE. 1992. A note on the thermal structure of the upper ocean in the Bransfield Strait-South Shetland Islands region. *Deep Sea Res Part A Oceanogr Res Papers* 39:1221–1229. [https://doi.org/10.1016/0198-0149\(92\)90065-2](https://doi.org/10.1016/0198-0149(92)90065-2).
24. Alcaman-Arias M, Fariás L, Verdugo J, Alarcon-Schumacher T, Diez B. 2018. Microbial activity during a coastal phytoplankton bloom on the Western Antarctic Peninsula in late summer. *FEMS Microbiol Lett* 365:fny090. <https://doi.org/10.1093/femsle/fny090>.
25. Fuentes S, Arroyo JI, Rodríguez-Marconi S, Masotti I, Alarcón-Schumacher T, Polz MF, Trefault N, De la Iglesia R, Diez B. 2019. Summer phyto- and bacterioplankton communities during low and high productivity

- scenarios in the Western Antarctic Peninsula. *Polar Biol* 42:159–169. <https://doi.org/10.1007/s00300-018-2411-5>.
26. de la Torre JR, Christianson LM, Béjà O, Suzuki MT, Karl DM, Heidelberg JF, DeLong EF. 2003. Proteorhodopsin genes are distributed among divergent marine bacterial taxa. *Proc Natl Acad Sci U S A* 100:12830–12835. <https://doi.org/10.1073/pnas.2133554100>.
 27. López JL, Golemba M, Hernández E, Lozada M, Dionisi H, Jansson JK, Carroll J, Lundgren L, Sjöling S, Cormack WPM. 2017. Microbial and viral-like rhodopsins present in coastal marine sediments from four polar and subpolar regions. *FEMS Microbiol Ecol* 93:fiw216. <https://doi.org/10.1093/femsec/fiw216>.
 28. Kim SH, Jung BH, Hong SG, Jung KH. 2020. Temperature dependency of proton pumping activity for marine microbial rhodopsin from Antarctic Ocean. *Sci Rep* 10:1356. <https://doi.org/10.1038/s41598-020-58023-5>.
 29. Riedel T, Tomasch J, Buchholz I, Jacobs J, Kollenberg M, Gerdt G, Wichels A, Brinkhoff T, Cypionka H, Wagner-Döbler I. 2010. Constitutive expression of the proteorhodopsin gene by a Flavobacterium strain representative of the proteorhodopsin-producing microbial community in the North Sea. *Appl Environ Microbiol* 76:3187–3197. <https://doi.org/10.1128/AEM.02971-09>.
 30. Brindefalk B, Ekman M, Ininbergs K, Dupont CL, Yooseph S, Pinhassi J, Bergman B. 2016. Distribution and expression of microbial rhodopsins in the Baltic Sea and adjacent waters. *Environ Microbiol* 18:4442–4455. <https://doi.org/10.1111/1462-2920.13407>.
 31. Olson DK, Yoshizawa S, Boeuf D, Iwasaki W, DeLong EF. 2018. Proteorhodopsin variability and distribution in the North Pacific Subtropical Gyre. *ISME J* 12:1047–1060. <https://doi.org/10.1038/s41396-018-0074-4>.
 32. Peck LS, Barnes DKA, Cook AJ, Fleming AH, Clarke A. 2010. Negative feedback in the cold: ice retreat produces new carbon sinks in Antarctica. *Global Change Biol* 16:2614–2623. <https://doi.org/10.1111/j.1365-2486.2009.02071.x>.
 33. Maresca JA, Miller KJ, Keffer JL, Sabanayagam CR, Campbell BJ. 2018. Distribution and diversity of rhodopsin-producing microbes in the Chesapeake Bay. *Appl Environ Microbiol* 84:e00137-18. <https://doi.org/10.1128/AEM.00137-18>.
 34. Venter JC, Remington K, Heidelberg JF, Halpern AL, Rusch D, Eisen JA, Wu D, Paulsen I, Nelson KE, Nelson W, Fouts DE, Levy S, Knap AH, Lomas MW, Nealson K, White O, Peterson J, Hoffman J, Parsons R, Baden-Tillson H, Pfannkoch C, Rogers YH, Smith HO. 2004. Environmental genome shotgun sequencing of the Sargasso Sea. *Science* 304:66–74. <https://doi.org/10.1126/science.1093857>.
 35. Sieradzki ET, Fuhrman JA, Rivero-Calle S, Gómez-Consarnau L. 2018. Proteorhodopsins dominate the expression of phototrophic mechanisms in seasonal and dynamic marine picoplankton communities. *PeerJ* 6:e5798. <https://doi.org/10.7717/peerj.5798>.
 36. Dubinsky V, Haber M, Burgsdorf I, Saurav K, Lehahn Y, Malik A, Sher D, Aharonovich D, Steindler L. 2017. Metagenomic analysis reveals unusually high incidence of proteorhodopsin genes in the ultraoligotrophic Eastern Mediterranean Sea. *Environ Microbiol* 19:1077–1090. <https://doi.org/10.1111/1462-2920.13624>.
 37. Shi Y, Tyson GW, Eppley JM, DeLong EF. 2011. Integrated metatranscriptomic and metagenomic analyses of stratified microbial assemblages in the open ocean. *ISME J* 5:999–1013. <https://doi.org/10.1038/ismej.2010.189>.
 38. Ottesen EA, Marin R, Preston CM, Young CR, Ryan JP, Scholin CA, DeLong EF. 2011. Metatranscriptomic analysis of autonomously collected and preserved marine bacterioplankton. *ISME J* 5:1881–1895. <https://doi.org/10.1038/ismej.2011.70>.
 39. Ottesen EA, Young CR, Gifford SM, Eppley JM, Marin R, Schuster SC, Scholin CA, DeLong EF. 2014. Multispecies diel transcriptional oscillations in open ocean heterotrophic bacterial assemblages. *Science* 345:207–212. <https://doi.org/10.1126/science.1252476>.
 40. Poretsky RS, Hewson I, Sun S, Allen AE, Zehr JP, Moran MA. 2009. Comparative day/night metatranscriptomic analysis of microbial communities in the North Pacific subtropical gyre. *Environ Microbiol* 11:1358–1375. <https://doi.org/10.1111/j.1462-2920.2008.01863.x>.
 41. Gómez-Consarnau L, González JM, Coll-Lladó M, Gourdon P, Pascher T, Neutze R, Pedrós-Alió C, Pinhassi J. 2007. Light stimulates growth of proteorhodopsin-containing marine Flavobacteria. *Nature* 445:210–213. <https://doi.org/10.1038/nature05381>.
 42. Lami R, Cottrell MT, Campbell BJ, Kirchman DL. 2009. Light-dependent growth and proteorhodopsin expression by Flavobacteria and SAR11 in experiments with Delaware coastal waters. *Environ Microbiol* 11:3201–3209. <https://doi.org/10.1111/j.1462-2920.2009.02028.x>.
 43. Giovannoni SJ. 2017. SAR11 bacteria: the most abundant plankton in the oceans. *Annu Rev Mar Sci* 9:231–255. <https://doi.org/10.1146/annurev-marine-010814-015934>.
 44. Ghiglione JF, Murray AE. 2012. Pronounced summer to winter differences and higher wintertime richness in coastal Antarctic marine bacterioplankton. *Environ Microbiol* 14:617–629. <https://doi.org/10.1111/j.1462-2920.2011.02601.x>.
 45. Rozema PD, Biggs T, Sprong PAA, Buma AGJ, Venables HJ, Evans C, Meredith MP, Bolhuis H. 2017. Summer microbial community composition governed by upper-ocean stratification and nutrient availability in northern Marguerite Bay, Antarctica. *Deep-Sea Res Part II Top Stud Oceanogr* 139:151–166. <https://doi.org/10.1016/j.dsr2.2016.11.016>.
 46. Trefault N, De la Iglesia R, Moreno-Pino M, Lopes dos Santos A, Géricas Ribeiro C, Parada-Pozo G, Cristi A, Marie D, Vulot D. 2021. Annual phytoplankton dynamics in coastal waters from Fildes Bay, Western Antarctic Peninsula. *Sci Rep* 11:1368. <https://doi.org/10.1038/s41598-020-80568-8>.
 47. Campbell BJ, Waidner LA, Cottrell MT, Kirchman DL. 2008. Abundant proteorhodopsin genes in the North Atlantic Ocean. *Environ Microbiol* 10:99–109. <https://doi.org/10.1111/j.1462-2920.2007.01436.x>.
 48. Hutchins DA, Fu F. 2017. Microorganisms and ocean global change. *Nat Microbiol* 2:17058. <https://doi.org/10.1038/nmicrobiol.2017.58>.
 49. Holmes RM, Aminot A, Kérouel R, Hooker BA, Peterson BJ. 1999. A simple and precise method for measuring ammonium in marine and freshwater ecosystems. *Can J Fish Aquat Sci* 56:1801–1808. <https://doi.org/10.1139/f99-128>.
 50. Strickland JHD, Parsons TR. 1972. A practical handbook of seawater analysis, 2nd ed. Fisheries Research Board of Canada, Ottawa, Canada.
 51. Tillett D, Neilan BA. 2000. Xanthogenate nucleic acid isolation from cultured and environmental cyanobacteria. *J Phycol* 36:251–258. <https://doi.org/10.1046/j.1529-8817.2000.99079.x>.
 52. Alarcón-Schumacher T, Guajardo-Leiva S, Antón J, Díez B. 2019. Elucidating viral communities during a phytoplankton bloom on the west Antarctic Peninsula. *Front Microbiol* 10:1014. <https://doi.org/10.3389/fmicb.2019.01014>.
 53. Bankevich A, Nurk S, Antipov D, Gurevich AA, Dvorkin M, Kulikov AS, Lesin VM, Nikolenko SI, Pham SON, Pribelski AD, Pyshtkin AV, Sirotkin AV, Vyahhi N, Tesler G, Alekseyev MA, Pevzner PA. 2012. SPAdes: a new genome assembly algorithm and its applications to single-cell sequencing. *J Comput Biol* 19:455–477. <https://doi.org/10.1089/cmb.2012.0021>.
 54. Hyatt D, Chen G, Locascio PF, Land ML, Larimer FW, Hauser LJ. 2010. Prodigal: prokaryotic gene recognition and translation initiation site identification. *BMC Bioinformatics* 11:119. <https://doi.org/10.1186/1471-2105-11-119>.
 55. Buchfink B, Xie C, Huson DH. 2015. Fast and sensitive protein alignment using DIAMOND. *Nat Methods* 12:59–60. <https://doi.org/10.1038/nmeth.3176>.
 56. Boeuf D, Audic S, Brillet-Guéguen L, Caron C, Jeanthon C. 2015. MicRhODE: a curated database for the analysis of microbial rhodopsin diversity and evolution. *Database* 2015:bav080. <https://doi.org/10.1093/database/bav080>.
 57. Finn RD, Coggill P, Eberhardt RY, Eddy SR, Mistry J, Mitchell AL, Potter SC, Punta M, Qureshi M, Sangrador-Vegas A, Salazar GA, Tate J, Bateman A. 2016. The Pfam protein families database: towards a more sustainable future. *Nucleic Acids Res* 44:D279–D285. <https://doi.org/10.1093/nar/gkv1344>.
 58. Finn RD, Attwood TK, Babbitt PC, Bateman A, Bork P, Bridge AJ, Chang HY, Dosztanyi Z, El-Gebali S, Fraser M, Gough J, Haft D, Holliday GL, Huang H, Huang X, Letunic I, Lopez R, Lu S, Marchler-Bauer A, Mi H, Mistry J, Natale DA, Necci M, Nuka G, Orengo CA, Park Y, Pesseat S, Piovesan D, Potter SC, Rawlings ND, Redaschi N, Richardson L, Rivoire C, Sangrador-Vegas A, Sigrist C, Sillitoe I, Smithers B, Squizzato S, Sutton G, Thanki N, Thomas PD, Tosatto SCE, Wu CH, Xenarios I, Yeh LS, Young SY, Mitchell AL. 2017. InterPro in 2017—beyond protein family and domain annotations. *Nucleic Acids Res* 45:D190–D199. <https://doi.org/10.1093/nar/gkw1107>.
 59. Langmead B, Salzberg SL. 2012. Fast gapped-read alignment with Bowtie 2. *Nat Methods* 9:357–359. <https://doi.org/10.1038/nmeth.1923>.
 60. Kopylova E, Noé L, Touzet H. 2012. Sequence analysis SortMeRNA: fast and accurate filtering of ribosomal RNAs in metatranscriptomic data. *Bioinformatics* 28:3211–3217. <https://doi.org/10.1093/bioinformatics/bts611>.
 61. Palovaara J, Akram N, Baltar F, Bunse C, Forsberg J, Pedrós-Alió C, González JM, Pinhassi J. 2014. Stimulation of growth by proteorhodopsin phototrophy involves regulation of central metabolic pathways in marine planktonic bacteria. *Proc Natl Acad Sci U S A* 111:E3650–E3659. <https://doi.org/10.1073/pnas.1402617111>.
 62. Booth MG, Hutchinson LA, Brumsted M, Aas P, Coffin RB, Downer RC, Kelley CA, Lyons MM, Pakulski JD, Holder Sandvik SL, Jeffrey WH, Miller RV. 2001. Quantification of recA gene expression as an indicator of repair

- potential in marine bacterioplankton communities of Antarctica. *Aquat Microb Ecol* 24:51–59. <https://doi.org/10.3354/ame024051>.
63. Booth MG, Jeffrey WH, Miller RV. 2001. RecA expression in response to solar UVR in the marine bacterium *Vibrio natriegens*. *Microb Ecol* 42: 531–539. <https://doi.org/10.1007/s00248-001-1009-5>.
 64. Fu L, Niu B, Zhu Z, Wu S, Li W. 2012. CD-HIT: accelerated for clustering the next-generation sequencing data. *Bioinformatics* 28:3150–3152. <https://doi.org/10.1093/bioinformatics/bts565>.
 65. Li W, Godzik A. 2006. Cd-hit: a fast program for clustering and comparing large sets of protein or nucleotide sequences. *Bioinformatics* 22: 1658–1659. <https://doi.org/10.1093/bioinformatics/btl158>.
 66. Katoh K, Rozewicki J, Yamada KD. 2018. MAFFT online service: multiple sequence alignment, interactive sequence choice and visualization. *Brief Bioinform* 20:1160–1166. <https://doi.org/10.1093/bib/bbx108>.
 67. Capella-Gutiérrez S, Silla-Martínez JM, Gabaldón T. 2009. trimAl: a tool for automated alignment trimming in large-scale phylogenetic analyses. *Bioinformatics* 25:1972–1973. <https://doi.org/10.1093/bioinformatics/btp348>.
 68. Hoang DT, Chernomor O, von Haeseler A, Minh BQ, Vinh LS. 2018. UFBoot2: improving the ultrafast bootstrap approximation. *Mol Biol Evol* 35:518–522. <https://doi.org/10.1093/molbev/msx281>.
 69. Nguyen LT, Schmidt HA, Von Haeseler A, Minh BQ. 2015. IQ-TREE: a fast and effective stochastic algorithm for estimating maximum-likelihood phylogenies. *Mol Biol Evol* 32:268–274. <https://doi.org/10.1093/molbev/msu300>.
 70. Barbera P, Kozlov AM, Czech L, Morel B, Darriba D, Flouri T, Stamatakis A. 2019. EPA-ng: massively parallel evolutionary placement of genetic sequences. *Syst Biol* 68:365–369. <https://doi.org/10.1093/sysbio/syy054>.
 71. Czech L, Barbera P, Stamatakis A. 2020. Genesis and Gappa: processing, analyzing and visualizing phylogenetic (placement) data. *Bioinformatics* 36:3263–3265. <https://doi.org/10.1093/bioinformatics/btaa070>.
 72. Letunic I, Bork P. 2019. Interactive Tree of Life (iTOL) v4: recent updates and new developments. *Nucleic Acids Res* 47:W256–W259. <https://doi.org/10.1093/nar/gkz239>.
 73. Wheeler TJ, Clements J, Finn RD. 2014. Skyline: a tool for creating informative, interactive logos representing sequence alignments and profile hidden Markov models. *BMC Bioinformatics* 15:7. <https://doi.org/10.1186/1471-2105-15-7>.
 74. Kenwell A, Navarre-Sitchler A, Prugue R, Spear JR, Hering AS, Maxwell RM, Carroll RWH, Williams KH. 2016. Using geochemical indicators to distinguish high biogeochemical activity in floodplain soils and sediments. *Sci Total Environ* 563-564:386–395. <https://doi.org/10.1016/j.scitotenv.2016.04.014>.
 75. R Development Core Team. 2020. R: a language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria.
 76. Oksanen J, Blanchet FG, Friendly M, Kindt R, Legendre P, Mcglinn D, Minchin PR, O'Hara RB, Simpson GL, Solymos P, Stevens MHH, Szoecs E, Wagner H. 2019. vegan: community ecology package. R package version 2.4-2. Community ecology package 2.5-6.
 77. Harrel F. 2015. Package "Hmisc" (v4.0-0). <https://cran.r-project.org/web/packages/Hmisc/index.html>.
 78. Wei T, Simko V. 2016. The corplot package. R Core Team.
 79. Moritz S, Bartz-Beielstein T. 2017. imputeTS: time series missing value imputation in R. *R J* 9:207–218.
 80. Kendall MG. 1975. Rank correlation methods, 4th ed. Charles Griffin, San Francisco, CA.
 81. Benjamini Y, Hochberg Y. 1995. Controlling the false discovery rate: a practical and powerful approach to multiple testing. *J R Stat Soc B* 57: 289–300. <https://doi.org/10.1111/j.2517-6161.1995.tb02031.x>.
 82. Anderson MJ. 2001. A new method for non-parametric multivariate analysis of variance. *Austral Ecol* 26:32–46. <https://doi.org/10.1111/j.1442-9993.2001.01070.pp.x>.
 83. Schmieder R, Edwards R. 2011. Quality control and preprocessing of metagenomic datasets. *Bioinformatics* 27:863–864. <https://doi.org/10.1093/bioinformatics/btr026>.