

An Investigation of the Volatile Organic Compound Degradation Capacity of the Oligotrophic Surface Water Microbial Community of the Arctic Ocean

Susan McLatchie

A Thesis
In the Department of
Biology

Presented in Partial Fulfillment of the Requirements
For the Degree of
Master of Science (Biology)

at Concordia University
Montreal, Quebec, Canada

June 2023

©Susan McLatchie, 2023

**CONCORDIA UNIVERSITY
SCHOOL OF GRADUATE STUDIES**

This is to certify that the thesis prepared

By: Susan McLatchie

Entitled: An Investigation of the Volatile Organic Compound Degradation Capacity
of the Oligotrophic Surface Water Microbial Community of the Arctic
Ocean

and submitted in partial fulfillment of the requirements for the degree of

Master of Science (Biology)

Complies with the regulations of the University and meet the accepted standards with respect to
originality and quality

Signed by the final examining committee:

_____ External Examiner
Dr. Malcolm Whiteway

_____ Examiner
Dr. Selvadurai Dayanandan

_____ Examiner
Dr. Carly Ziter

_____ Thesis Supervisor
Dr. David Walsh

Approved by

Dr. Grant Brown, Graduate Program Director

June 6, 2023

Dr. Pascale Sicotte, Dean of Arts and Science

Abstract

An Investigation of the Volatile Organic Compound Degradation Capacity of the Oligotrophic Surface Water Microbial Community of the Arctic Ocean

Susan McLatchie

The summer surface waters of the Arctic Ocean (AO) are among the most oligotrophic marine ecosystems on Earth. The AO is also influenced by variable sea-ice and sunlight. Little is known about how marine bacteria have adapted to survive this extreme Arctic habitat. Volatile organic compounds (VOCs) are increasingly recognized as important resources for bacterial metabolism in oligotrophic oceans. Studies show that the AO is a net sink for VOCs leading to the hypothesis that Arctic Ocean bacteria consume VOCs. To test this hypothesis, we investigated a genome catalogue of AO prokaryotes for the presence of VOC-degradation pathways. A carboxylation pathway for acetone degradation dominated the surface waters. The key enzyme (acetone carboxylase) allows bacteria to access CO₂ through acetone carboxylation, suggesting a type of heterotrophic CO₂ fixation may operate in the AO. The acetone-consuming bacteria were mostly represented by a single population of *Porticoccus* bacteria (SP-01). The genome encoded few organic carbon transporters, but proteorhodopsin genes were identified indicating heterotrophic metabolism could be supplemented with energy from sunlight. Metatranscriptome analysis revealed that acetone carboxylase and proteorhodopsin were among the most highly expressed genes, indicating significant investment in the exploitation of acetone and sunlight in carbon and energy metabolism. Through 2004-2017 SP-01 had a 1-8% annual average abundance, peaking in 2017, a year distinguished by a longer ice-free period. The future relevance of bacteria, such as SP-01, adapted to AO oligotrophic surface water may increase with a warming Arctic, when ice-free conditions prevail, and ocean stratification and subsequent oligotrophication intensifies.

Acknowledgements

I am incredibly lucky to have had a supervisor as wonderful as David Walsh. I could not have asked for a better mentor to teach me how to properly conduct science and who was kind and challenged me to do things I was sometimes intimidated to do. I am sincerely grateful for your wisdom and support. I would also like to thank my committee members Selvadurai Dayanandan and Carly Ziter for your advice and feedback.

I am grateful for the funding I received during my master's from a Concordia Faculty of Arts and Science Award and NSERC CGS-M scholarship.

My experience as a graduate student was further enriched by having some fantastic lab mates. First, I would like to thank Arthi Ramachandran who made me feel welcome in my first undergraduate research experience in the Walsh Lab and who continued to be a good friend throughout my graduate studies. I would also like to thank past and present lab members (Thomas Grevesse, Susanne Kraemer, Charlène Lawruk-Desjardins and Vera Onana) for the laughs and words of encouragement during my studies. Although being a grad student from the kitchen table could be a bummer, returning to the lab and being able to laugh over vegan lunch and have some cathartic power walks with my lab mate, Rebecca Garner made up for lost time.

I would like to thank my college biochemistry instructor, Jamie Doran for inspiring my journey into research.

Finally, thank-you to my parents, Joanne, and Stuart and my grandparents Betty and Cyril for instilling curiosity in me when I was young and for your encouragement and support in my education from day one.

Dedication

For my mum, Joanne.

Contribution of Authors

Thomas Grevesse generated the metagenomic data and MAGs from the single metagenomes.

Sara Palestini generated the sea ice concentration heatmaps and conducted the statistical testing.

All authors reviewed the final manuscript and approved the contents.

Table of Contents

Introduction	1
Materials and methods	3
Sea water sampling	3
Metagenomic assembly, binning, dereplication and taxonomic classification.....	3
Fragment recruitment of Arctic Ocean and global ocean metagenomes	4
Gene expression activity in the MAG dataset	5
Functional annotation and assignment of VOC metabolism	5
Acetone carboxylase phylogeny	6
16S rRNA phylogeny	6
Comparative genomics	6
Metabolic reconstruction of SP-01 and <i>P. hydrocarbonoclasticus</i>	7
Microbial rhodopsin phylogeny and analysis of primary amino acid sequence of rhodopsin genes harboured by SP-01	7
Identification and statistical analysis of Port-CB9S-26 in a 16S rRNA ASV timeseries.....	7
Results	9
Generation and characteristics of the Canada Basin genome catalogue.....	9
Genomic evidence for VOC-degrading bacteria	12
Acetone degrading potential of the Canada Basin Arctic Ocean microbial communities.....	14
Phylogenomics of <i>Porticoccaceae</i>	18
Environmental drivers of SP-01 abundance in the Canada Basin in a 16S rRNA time series	28
Discussion	31
Oligotrophic Arctic Ocean surface specialist bacteria thrive on VOCs	31
Exploitation of varying light regimes of the Arctic Ocean.....	33
Nutrient acquisition adaptations of the acetone-degrading <i>Porticoccus</i> populations to the Arctic Ocean environment	34
Acetone-consuming bacteria regime shift associated with extended ice-free periods	35
Conclusions.....	35
References	37
Appendices	44

List of Figures

Figure 1. Vertical structuring of the microbial community in the Canada Basin, Arctic Ocean. (A) Map of sampling stations in the Canada Basin, Arctic Ocean. (B) Principal component analysis of the environmental characteristics of the metagenome sampling locations. (C) Principal coordinate analysis of the taxonomic variation among Arctic Ocean MAG assemblages based on TAD ₈₀ community fraction of the MAGs.	11
Figure 2. Identification and pattern of gene expression of VOC-consuming bacteria through the Arctic Ocean water column. (A) The pattern of summed TAD ₈₀ fraction of the metagenomic genome assemblages comprising MAGs harbouring VOC degradation marker genes. (B) The pattern of summed RPKM metatranscriptome read recruitment of VOC marker genes harboured by the MAGs.	14
Figure 3. Diversity and relative abundance of MAGs representing acetone-consuming bacteria. Taxonomy of the MAGs based on GTDB phylogenomic classification. Relative abundance of MAGs based on the TAD ₈₀ fraction of the metagenomic genome assemblages comprising the MAGs.	15
Figure 4. Phylogenetic relationships between acetone carboxylase <i>acxB</i> gene from MAGs representing acetone-consuming bacteria in the Canada Basin (red), Canada Basin metagenomes (blue) and experimentally validated acetone-consuming bacteria (bolded). A maximum likelihood tree was constructed using the JTT substitution model, gamma distribution (4 categories), and the nearest neighbor heuristic search model with 100 bootstrap replications. The tree was rooted using the alpha subunit gene of enzymes from the larger hydantoinase/oxoprolinase protein family.	17
Figure 5. Maximum likelihood phylogeny of marine <i>Gammaproteobacteria</i> based on 16S rRNA gene sequences. The blue taxon is from the Port-CB9S-26 MAG with the capacity for acetone carboxylation. The tree was constructed using the GTR + gamma distribution (4 categories) model of nucleotide substitution in MEGA v.11 with 100 bootstraps.	19
Figure 6. Genomic characteristics and gene overlap between SP-01 and <i>P. hydrocarbonoclasticus</i>	20
Figure 7. Metabolic reconstruction of the central carbon metabolism of SP-01 and <i>P. hydrocarbonoclasticus</i> . Black arrows indicate genes of pathways conserved between SP-01 and <i>P. hydrocarbonoclasticus</i> . Red arrows indicate genes of pathways possessed only by SP-01. Blue arrows indicate genes of pathways possessed only by <i>P. hydrocarbonoclasticus</i> . Grey dashed arrows indicate genes of pathways that are absent in SP-01 and <i>P. hydrocarbonoclasticus</i>	22
Figure 8. Phylogenetic analysis of rhodopsin genes and the proteorhodopsin biosynthesis operon within the SP-01 genome. (A) A maximum likelihood tree was constructed using the JTT substitution model, gamma distribution (4 categories), and the nearest neighbor heuristic search model with 100 bootstrap replications. Sequences from Port-CB9S-26 are highlighted in red. (B) Visualization of the proteorhodopsin and carotenoid biosynthesis genes within Port-CB9S-26.	23
Figure 9. Reconstruction of vitamin and inorganic nutrient transporters of SP-01 and <i>P. hydrocarbonoclasticus</i> . Black arrows indicate genes of pathways conserved between SP-01 and <i>P. hydrocarbonoclasticus</i> . Red arrows indicate genes of pathways possessed only by SP-01. Blue arrows indicate genes of pathways possessed only by <i>P. hydrocarbonoclasticus</i>	26
Figure 10. Top ten most expressed genes of SP-01 in the Canada Basin based on mean metatranscriptomic recruitment.	28

Figure 11. Average annual relative abundance of SP-01 in the Canada basin over 2004-2017. The relative abundance of SP-01 was averaged over the summer mixed layer (SML) (<15 m), upper Arctic water (UAW) (> 15m, < 31 PSU), and Pacific water (PW) (31-33 PSU) for each year.29

Figure 12. Sea ice concentration and relative abundance of SP-01 in the Canada Basin over 2004-2017 in the summer mixed layer and upper Arctic water. Relative abundance indicated with a black circle, white circle and grey circle in the SML, UAW, and PW, respectively. Black cross and white cross indicates SML not sampled, and UAW not sampled, respectively.30

Introduction

Large expanses of the ocean are oligotrophic, depleted in nutrients with low standing stocks of phytoplankton. Microbial processes control carbon, nutrient, and energy cycling in these systems. Bacteria are exquisitely adapted to these oligotrophic conditions (Dupont et al., 2012, Giovannoni et al., 2014, Giovannoni et al., 2017). Surface ocean nutrient depletion is expected to intensify with increased stratification of the upper ocean due to environmental change (Hutchins and Fu, 2017). While high productivity is associated with the Arctic Ocean due to its extensive continental shelf regions, the central Arctic Ocean has some of the most oligotrophic seas (Tremblay et al., 2012). In the Canada Basin, the circulation of the Beaufort Gyre retains freshwater from sea ice melt and river runoff, which leads to intensified stratification relative to other regions of the Arctic Ocean (Aagaard et al., 1989). Consequently, the surface waters of the Canada Basin remain highly oligotrophic through the summer to fall during open water conditions (McLaughlin et al., 2010, Nishino et al., 2020). Along with varying sea ice coverage, the Arctic Ocean faces intense seasonal differences in the number of daylight hours and is highly influenced by terrestrial land masses relative to other oceans. Distinct arctic-adapted ecotypes (Colatriano et al., 2018, Kraemer et al., 2020) and evolutionary adaptations (Colatriano et al., 2018, Ramachandran et al., 2021) of bacteria to the Arctic Ocean have been identified. Although lower latitude oligotrophic systems and the bacteria that inhabit them have been extensively investigated, little is known about the adaptations to the highly oligotrophic surface waters of the Arctic Ocean which are relatively fresh and face extremes in exposure to sunlight and ice cover.

Volatile organic compounds (VOCs) are a class of organic compounds that is being increasingly recognized in support growth in oligotrophic oceans (Duarte et al., 2013). VOCs account for a substantial proportion of the flux of dissolved organic carbon in the surface ocean (Wagner et al., 2020). VOCs, of which methanol, acetone, and acetaldehyde make up most of the atmospheric concentration (Singh et al., 1995), are ubiquitous carbon compounds that can enter the water column via atmospheric deposition (Wang et al., 2020) or can be produced in-situ by phytoplankton (Halsey et al., 2017) or via photodegradation of dissolved organic matter (de Bruyn et al., 2011). VOCs may support heterotrophic growth where labile phytoplankton derived carbon may be scarce (Beale et al., 2013). An investigation of the cycling of phytoplankton derived VOCs revealed that SAR11 bacteria metabolize VOCs at a rate high enough to consume

significant fractions of the marine VOC pool (Moore et al., 2022). Based on high undersaturation of VOCs in the water relative to the atmosphere, the Arctic Ocean is a net sink of methanol and acetone (Sjostedt et al., 2012, Wohl et al., 2022). In previous work, we found methanol-consuming bacteria to be common and active in the surface layer of the Arctic Ocean (Ramachandran et al., 2021). Furthermore, understanding VOC cycling has implications for climate as VOCs influence climate-active compounds in the atmosphere by increasing its oxidative capacity (Singh et al., 1995).

Given the highly oligotrophic conditions in the Canada Basin and the direction of VOC flux in the Arctic Ocean, we hypothesized that Arctic Ocean bacteria couple VOC degradation to their carbon and energy metabolism. To test this hypothesis, we investigated a metagenome-derived genome catalogue of Arctic Ocean prokaryotes for the presence of VOC metabolic pathways. We performed metatranscriptomics to assess the activity of VOC metabolism in Arctic Ocean prokaryotic carbon and energy metabolism. Metabolic reconstruction and comparative genomics with a lower latitude bacterial relative were carried out to investigate genomic adaptations to the oligotrophic surface waters of the Arctic Ocean. Finally, since VOC uptake and production in the ocean and Arctic Ocean stratification are influenced by sea-ice cover, we tracked the abundance of VOC-consuming bacteria over a multi-year time series to understand their relevance with the changing sea ice dynamics of the Arctic Ocean.

Materials and methods

Sea water sampling

Samples from 9 stations in the Canada Basin and Amundsen Gulf in the Western Arctic Ocean (CB4, CB6, CB9, CB11b, CB21, CB27, CBN3, AG5) (Fig. 1A) were collected in September 2017 on the *CCGS Louis S. St-Laurent* during the Joint Ocean Ice Study/Beaufort Gyre Exploration Project cruise. To explore the differences in environmental characteristics between the Canada Basin samples, we performed a principal component analysis using scaled environmental variables of temperature, salinity, oxygen, fluorescence, nitrate, phosphate, silicate, and chromophoric dissolved organic matter (CDOM). Eight water masses were sampled: two samples from the summer mixed layer (SML) (5m and 20m), one sample from the subsurface chlorophyll maximum (SCM), two samples the Pacific winter water (PWW) (32.3 PSU and 33.1 PSU), two samples from the Atlantic water (AW) (Tmax and 1000m), and one sample 10 or 100 m from the bottom (Grevesse et al., 2022). Fourteen litres of seawater for DNA samples and 7 litres for RNA samples were pre-filtered through a 3 µm pore size polycarbonate track etched filter and then filtered through a 0.22 µm pore size Sterivex filter to obtain the free-living microbial fraction (Grevesse et al., 2022). Thirty-eight samples total were collected and DNA and/or RNA was extracted as described in Grevesse et al. (2022) (Table S1).

Metagenomic assembly, binning, dereplication and taxonomic classification

Shotgun sequencing and assembly of single metagenomes were performed as described in Grevesse et al. (2022). Metagenome co-assemblies of samples from the Canada Basin and Amundsen Gulf (Table S1) were also generated to recover a greater taxonomic diversity of metagenome-assembled genomes (MAGs). Co-assemblies of 24 samples from the Canada Basin, 11 samples from Canada Basin surface, 7 samples from the Amundsen Gulf, and individual assemblies of all 31 samples from the Canada Basin and Amundsen Gulf were performed. All metagenome co-assemblies were generated using MEGAHIT v.1.2.7 (Li et al., 2015) with k-mer sizes of 27, 37, 47, 57, 67, 77, and 87. The input metagenome reads for each single and co-assembly were mapped to the assembly with BWA v.0.7.17 using the mem option (Li et al.,

2009). The mapping results were processed using `jgi_summarize_bam_contig_depths` from MetaBAT2 v.2.12.1 (Kang et al., 2019). Contig binning was performed for each of the assemblies using MetaBAT2 v.2.12.1 (Kang et al., 2019) with default settings, resulting in 4824 genome bins. Genome quality was evaluated using CheckM v.1.0.11 (Parks et al., 2015) with the `lineage_wf` workflow. The 924 bins with >50% completeness and <10% contamination and strain heterogeneity were considered at least medium-quality MAGs (Bowers et al., 2017). The medium-quality MAGs were dereplicated with dRep using 95% ANI cut-off to remove species level redundancy (Olm et al., 2017), resulting in 663 representative MAGs. Dereplicated MAGs with >50% completeness and <10% contamination and strain heterogeneity were taxonomically classified with the GTDB-tk v.1.3.0 using the `classify_wf` workflow (Chaumeil et al., 2019).

Fragment recruitment of Arctic Ocean and global ocean metagenomes

To determine the coverage of MAGs in the Canada Basin metagenomes (Table S1) and 129 TARA global oceans metagenomes (Karsenti et al. 2011) and 20 Southern Ocean metagenomes (Brown et al., 2012), trimmed metagenome reads were initially recruited to a concatenation of the MAGs at a minimum sequence identity of 98% using BMap v.35 (Bushnell, 2015). The mapping files were parsed to extract the mapping information for each individual MAG using a custom Unix script which implemented SAMtools (Li and Durbin, 2009). Mapping files for each MAG were converted from SAM to sorted BAM format files using SAMtools (Li and Durbin, 2009). Horizontal coverage was calculated following the method of Rodriguez-R et al. (2020) in which the average sequencing depth is truncated to the central 80% of the mapped positions (TAD_{80}) normalized by the number of genome equivalents. The number of genome equivalents in each metagenome was estimated using MicrobeCensus v.1.1.0 (Nayfach and Pollard, 2015). A principal coordinate analysis was performed to assess the taxonomic variation among medium-quality MAG assemblages based on the Bray-Curtis dissimilarities in TAD_{80} abundances using the function `pcoa()` in the R package `ape` (Paradis and Schliep, 2019).

Gene expression activity in the MAG dataset

Metatranscriptomic data were generated as described in Grevesse et al. (2022). To assess the gene expression activity of the MAGs, competitive fragment recruitment of the Canada Basin metatranscriptomes was carried out in the same manner as the metagenomic fragment recruitment. Metatranscriptomic RPKM were calculated for the gene-coding sequences to account for the differing sizes of the metatranscriptomes:

$$\left(\frac{\text{number of reads recruited}}{\text{length of gene in kbp}} \right) / \text{number of metatranscriptome reads in millions } .$$

Functional annotation and assignment of VOC metabolism

Single assembly metagenome gene prediction and functional annotation were performed as described in Grevesse et al. (2022). Protein-coding genes of the MAGs were annotated using Prokka v.1.12 (Seemann 2014) which implemented Prodigal v.2.6.3 for gene prediction (Hyatt 2010). Ribosomal RNA genes were annotated using Prokka v.1.12 which implemented Barrnap (Seemann 2014). Transfer and transfer-messenger RNA genes were annotated using Prokka v.1.12 which implemented ARAGORN v.1.2 (Laslett and Canback, 2004). MAG gene functions were annotated using KofamScan with default settings and a threshold score of 0.7 or higher and an e-value of 1×10^{-10} or lower (Aramaki et al., 2020).

VOC metabolism was identified in the microbial community by mining for acetaldehyde, methanol, and acetone marker genes in the MAG dataset. The marker genes for acetaldehyde degradation included *dmpF* (K00132), *adhE* (K04072), *mhpF* (K04073), and *bphJ* (K18366), each encoding acetaldehyde dehydrogenase. The marker gene for methanol degradation included *xoxF* (K23995), encoding methanol dehydrogenase. The marker genes for acetone degradation included *acxABC* (K10854, K10855, K10856), encoding acetone carboxylase and *acmA* (K18371), encoding acetone monooxygenase.

Acetone carboxylase phylogeny

Reference sequences including those from experimentally validated acetone-consuming bacteria, full-length sequences the acetone carboxylase-habouring MAGs, and sequences from the single metagenome assemblies (>500 bp in length) of the marker gene of acetone carboxylase (*acxB*) were included in a multiple sequence alignment using the MUSCLE algorithm in MEGA v.11 (Tamura et al., 2021). The maximum likelihood tree was constructed using the JTT substitution model, gamma distribution (4 categories), and the nearest neighbor heuristic search model in MEGA v.11 with 100 bootstrap replications (Tamura et al., 2021).

16S rRNA phylogeny

We performed a phylogenetic analysis using the 16S rRNA gene sequences from the Port-CB9-26 genome, reference sequences from the oligotrophic marine gammaproteobacteria group (Spring et al., 2015) and sequences from the *Porticoccaceae* group from biogeographic studies. A multiple-sequence alignment was generated using the MUSCLE algorithm as implemented in MEGA v.11 (Tamura et al., 2021). A maximum likelihood tree was constructed using the GTR + gamma distribution (4 categories) model of nucleotide substitution in MEGA v.11 with 100 bootstraps (Tamura et al., 2021).

Comparative genomics

Amino acid identity (AAI) was calculated between SP-01 and its closest relatives with a complete genome, *P. hydrocarbonoclasticus* (GCA_000744735.1) and HTCC2207 (GCA_000153445.1), using the Kostas Lab AAI calculator (Rodriguez-R et al., 2014). To explore differences in gene content between SP-01 and *P. hydrocarbonoclasticus*, the distributions of orthologous genes were analyzed by Proteinortho (Lechner et al., 2011).

Metabolic reconstruction of SP-01 and *P. hydrocarbonoclasticus*

The KEGG Mapper “Reconstruct” tool was used to identify central metabolic pathways harboured by SP-01 and *P. hydrocarbonoclasticus* (Kanehisa and Sato, 2020). The tool Gapmind for amino acids was used to identify possible amino acid auxotrophies of SP-01 and *P. hydrocarbonoclasticus* (Price et al., 2020). The tool Gapmind for small carbon sources was used to identify potential carbon sources of SP-01 and *P. hydrocarbonoclasticus* (Price et al., 2022).

Microbial rhodopsin phylogeny and analysis of primary amino acid sequence of rhodopsin genes harboured by SP-01

Employing the Prokka gene annotation tool (Seemann 2014) identified the rhodopsin genes harboured by SP-01 as blue light sensing proteorhodopsin (gene ID: FJFHIPCL_00744) and green light sensing proteorhodopsin (gene ID: FJFHIPCL_00745). Employing the KofamScan tool (Aramaki et al., 2020) annotated one rhodopsin gene (gene ID: FJFHIPCL_00744) as a sensory rhodopsin gene but did not annotate the other gene (gene ID: FJFHIPCL_00745). To verify the annotations, we performed a phylogenetic analysis to increase confidence in the biochemical function of the rhodopsin genes. Reference sequences for proteorhodopsin and sensory rhodopsin I and II were included in a multiple sequence alignment using the MUSCLE algorithm in MEGA v.11 (Tamura et al., 2021). The maximum likelihood tree was constructed using the JTT substitution model, gamma distribution (4 categories), and the nearest neighbor heuristic search model in MEGA v.11 with 100 bootstrap replications (Tamura et al., 2021). Finally, to gain additional confidence in the biochemical function of the proteorhodopsin genes, we analysed the primary amino acid sequence for the presence of proton pumping and light absorbing motifs (Wang et al., 2003).

Identification and statistical analysis of Port-CB9S-26 in a 16S rRNA ASV timeseries

We used blastn (Altschul et al., 1990) to identify the amplicon sequence variant (ASV) which represented SP-01 within the Arctic Ocean 16S rRNA ASV timeseries (Kraemer et al., 2023). We

explored whether extended ice-free periods influenced the abundance of SP-01. Sea ice concentrations for the coordinates of each station (CB29, CB15, CB21, CB29) resolved within 25 km², were obtained from the National Snow and Ice Data Center database. Sea ice concentrations for each date of sampling and 7, 14, 21 and 28 days before the earliest sampling date of each year were plotted in a heatmap using the R package MBA v.0.1.0. We grouped observations of SP-01 abundance based on whether in 3 weeks prior to sampling sea-ice concentration fell below 15%. We performed a Welch's two-sample t-test to validate the differences in the mean abundance under ice covered or ice-free conditions.

Results

Generation and characteristics of the Canada Basin genome catalogue

We generated 32 metagenomes selected to represent the environmentally distinct stratified water layers of the Western Arctic Ocean (Fig. 1A, B). The water layers were composed of the surface (summer mixed) layer, subsurface chlorophyll maximum, Pacific winter water masses (32.3 and 33.1 PSU) and the Atlantic water masses (1000 m and Tmax) (Fig. 1B). To maximize the quantity and diversity of MAGs, we generated MAGs from a collection of 3 co-assemblies and 32 single assemblies (Table S1). The selection of metagenomes for co-assembly was based on their oceanographic and environmental characteristics. Separate co-assemblies of Canada Basin metagenomes (CB assembly, 24 samples) (Table S1) and Amundsen Gulf (AG assembly, 7 samples) (Table S1) were performed based on their differing basin and coastal oceanography (Fig. 1A). A final co-assembly was performed with surface, 20 m and SCM euphotic zone metagenomes (CB surface assembly, 11 samples) (Table S1) of the Canada Basin due to their distinct lower salinity and nutrient conditions (Fig. 1B), which may be more immediately impacted by environmental change.

Contigs were binned based on tetranucleotide frequency and coverage. A total of 924 MAGs with estimated completeness greater or equal to 50% and contamination less than or equal to 10%, corresponding to at least medium-quality MAG standards (Bowers et al., 2017), were selected for further analyses (Fig. S1). The combination of co-assembly and single assembly strategies generated a greater diversity of at least medium-quality MAGs (Fig. S2). The CB assembly generated a much greater diversity of MAGs (493) relative to the AG assembly (72), CB surface assembly (79) and sum of all MAGs generated with the single assemblies (280) (Table S2). The MAG dataset was dereplicated at 95% ANI to generate 663 representative genomospecies (Table S2). Nearly all MAGs generated with the CB assembly either represented distinct genomospecies (<95% ANI) or the highest quality MAG within a genomospecies (>95% ANI). For example, the CB co-assembly only a 1.08-fold reduction in the number of genomospecies after dereplication in contrast to a 1.52, 1.41 and 2.69-fold reduction in the number of genomospecies for MAGs generated from the CB surface assembly, AG assembly and the sum of all MAGs from single

assemblies (Table S2). A large proportion of MAGs (389/493) generated from the CB assembly represented different genomospecies while MAGs generated from the single assemblies increased the microdiversity within genomospecies with a much lower proportion MAGs (52/280) representing different genomospecies (Fig. S2). The proportion of CB surface MAGs (28/79) and AG co-assembly MAGs (33/72) representing different genomospecies was intermediate to the number of MAGs generated from the CB assembly and the sum of MAGs generated from single assemblies. The MAG dataset represented a wide phylogenetic range, including 619 bacterial (Fig. S3) and 44 archaeal taxa (Fig. S4) of which the majority (93%) were novel at the species level (Fig. S5). Recovery of these MAGs provided a valuable resource for searching for VOC-degrading microbial populations in the Arctic Ocean.

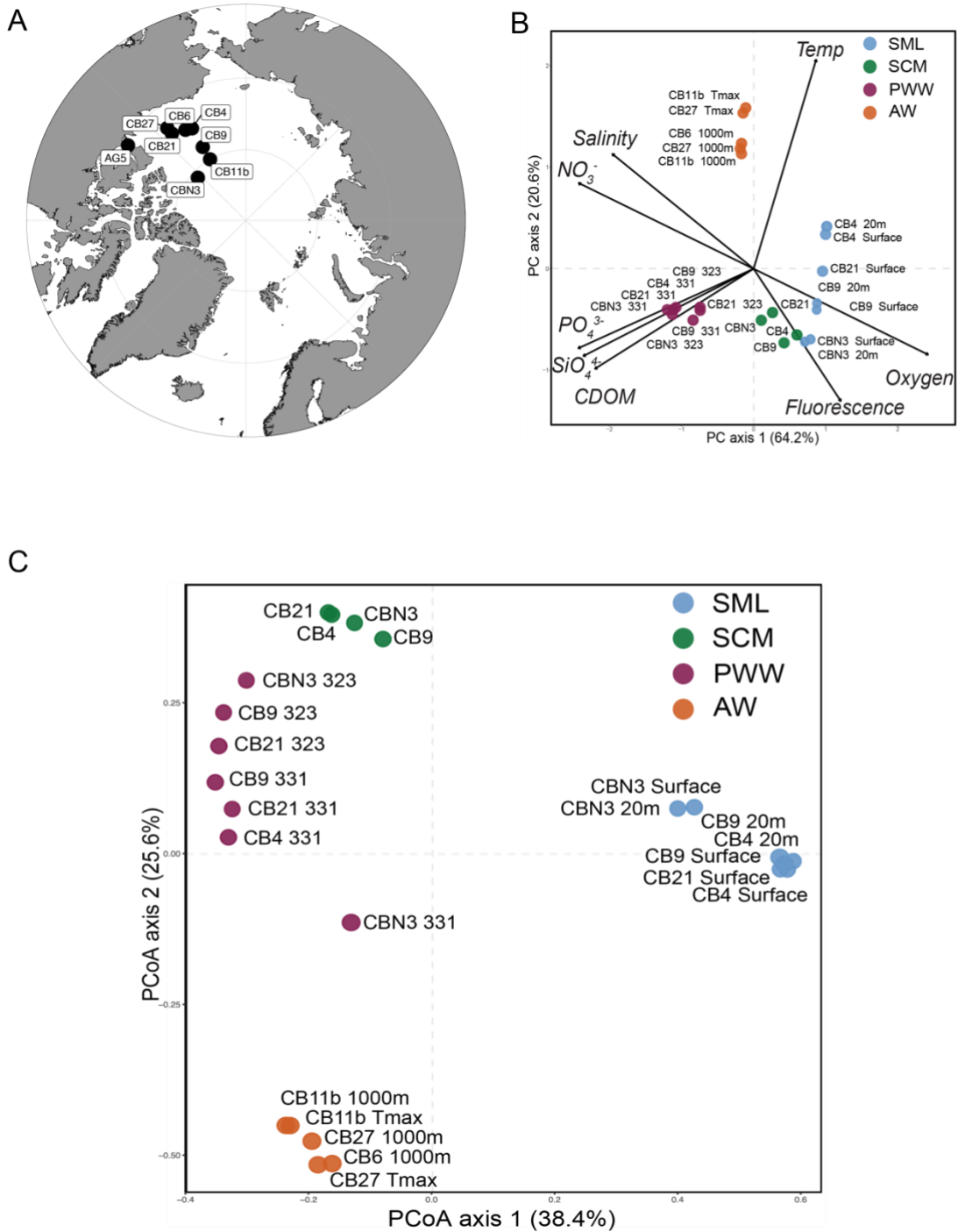


Figure 1. Vertical structuring of the microbial community in the Canada Basin, Arctic Ocean. (A) Map of sampling stations in the Canada Basin, Arctic Ocean. (B) Principal component analysis of the environmental characteristics of the metagenome sampling locations. (C) Principal coordinate analysis of the taxonomic variation among Arctic Ocean MAG assemblages based on TAD₈₀ community fraction of the MAGs.

Biogeography of MAGs across the Arctic and global oceans

Within the Arctic Ocean, read recruitment to the MAG dataset was lowest in the surface waters (8.34% read recruitment) and increased with depth with the highest mean read recruitment from the Pacific winter water (26.34%), indicating less genomic representation from the surface waters (Fig. S6). Correspondingly, MAG richness generally increased with depth (Fig. S7). The Arctic Ocean microbial communities demonstrated vertical structuring with the surface microbial community structure distinguished from that of the chlorophyll maximum, Pacific winter water and Atlantic water layers (Fig. 1C). In agreement with prior observations of the biogeography of prokaryotes from neritic regions of the Arctic Ocean (Royo-Llonch et al., 2021), the MAGs demonstrated a restricted distribution within the Arctic Ocean, with an 8.34 – 26.34% mean read recruitment of the MAGs from Arctic Ocean metagenomes compared to a 0.04 – 9.74% read recruitment of the MAGs from mid-latitude and Southern Ocean metagenomes (Fig. S6). As previously observed (Royo-Llonch et al., 2021) some Arctic Ocean prokaryotes may reach other regions of the global ocean via deep ocean circulation, as the mean metagenomic read recruitment of the MAGs increased with depth in the mid latitude oceans (0.04 – 1.60%) and the Southern Ocean (0.70 – 9.74%) (Fig. S6). Also in agreement with prior observations (Royo-Llonch et al., 2021), there was a bipolar distribution of some Arctic Ocean prokaryotes as evidenced by higher read recruitment (0.70% – 9.74%) in the Southern Ocean samples compared to mid-latitude oceans (0.04 – 1.60%) (Fig. S6).

Genomic evidence for VOC-degrading bacteria

Using a selection of functional gene markers, we identified six known VOC degradation pathways in the MAGs. We identified functional gene markers for two acetone degradation pathways, one methanol degradation pathway and four acetaldehyde degradation pathways (Fig. 2A). The functional gene markers for the carboxylating acetone degradation II pathway (*acxABC*) were most abundant in the surface layer MAG assemblages (8.10 – 12.23%) based on summed TAD₈₀ value (Fig. 2A). In agreement with a previous study (Ramachandran et al., 2021), the functional gene marker for methanol oxidation (*xoxF*) was identified in the surface layer MAG assemblages (0.01 – 1.03%) based on summed TAD₈₀ values (Fig. 2A). The functional gene

marker for the acetone degradation pathway I (*acmA*) was found in a small fraction of the MAG assemblages across the water column (0.002% – 0.2 (Fig. 2A). The functional gene marker for methanol oxidation (*xoxF*) was also identified in the Pacific winter water layer MAG assemblages (0.05 – 1.16%) MAGs (Fig 2A). The functional gene markers for acetaldehyde degradation via *mhpF* and *bphJ* were found in the largest fraction of the Pacific winter water layer MAG assemblages (0.07 – 0.83% and 0.02-0.8%, respectively) (Fig 2A). The functional gene marker for acetaldehyde degradation via *dmpF* was found in the largest fraction of the Atlantic water layer MAG assemblages (2.50 – 3.09%). The functional gene marker for acetaldehyde degradation via *adhE* was identified across the water column in a small fraction of the MAG assemblages (0.0004 –0.008%) (Fig. 2A).

We next mapped metatranscriptomes to the VOC-degrading MAGs. The functional gene markers for the carboxylating acetone degradation II pathway (*acxABC*) had the highest transcription levels based on summed RPKM values in the surface layer (162 – 1629 RPKM) (Fig 2B). In agreement with a previous study (Ramachandran 2021), the functional gene marker for methanol oxidation (*xoxF*) was most highly expressed in the surface layer (11 – 185 RPKM) (Fig. 2B). The functional gene marker for the acetone degradation pathway I (*acmA*) was most expressed in the SCM (0.4 – 1 RPKM) (Fig. 2B). The functional gene marker for acetaldehyde degradation, *bphJ*, was most expressed in the Pacific winter water layer (0.5 – 6 RPKM) (Fig. 2B). The functional gene markers for acetaldehyde degradation, *mhpF* and *dmpF*, were most expressed in the Atlantic water layer (2 – 4 RPKM and 2 – 3 RPKM) (Fig. 2B). The functional gene marker for acetaldehyde degradation, *adhE*, was expressed in the surface, SCM and Atlantic water layer (0.04 – 0.9 RPKM) (Fig. 2B). We focused further genomic analysis on acetone carboxylase dependent acetone-consuming bacteria due to the high abundance of MAGs harbouring *acxABC* and high expression of *acxABC* marker genes in the oligotrophic surface layer.

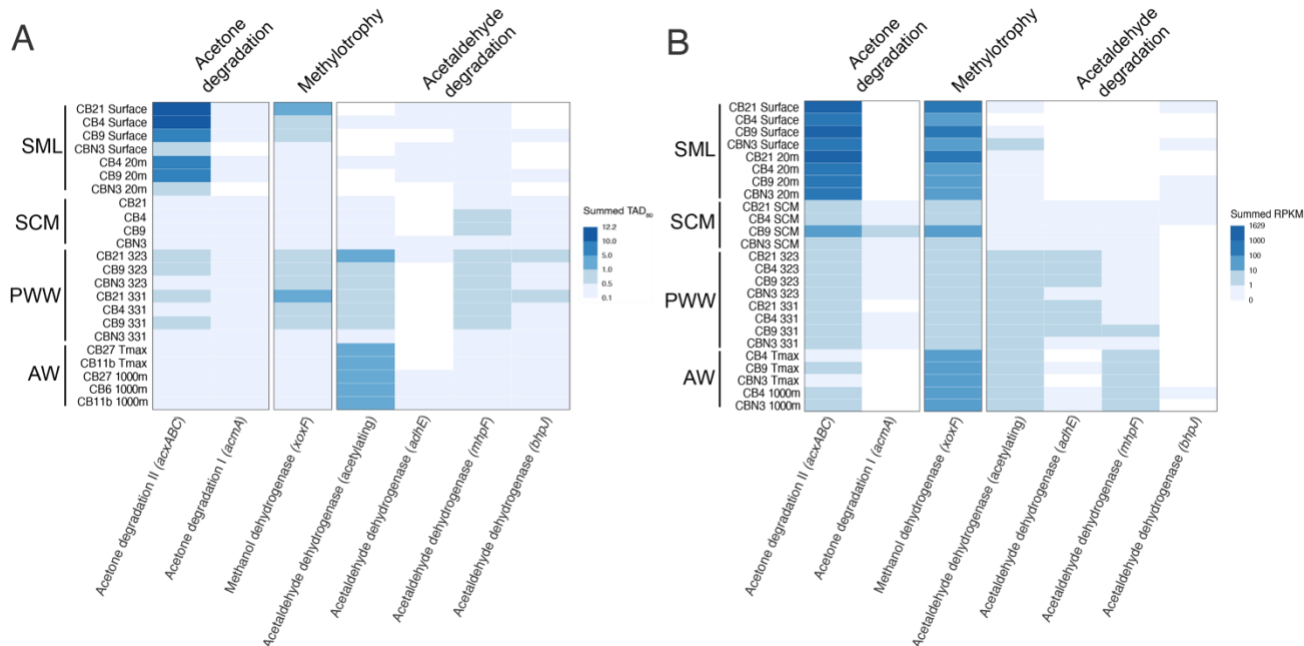


Figure 2. Identification and pattern of gene expression of VOC-consuming bacteria through the Arctic Ocean water column. (A) The pattern of summed TAD₈₀ fraction of the metagenomic genome assemblages comprising MAGs harbouring VOC degradation marker genes. (B) The pattern of summed RPKM metatranscriptome read recruitment of VOC marker genes harboured by the MAGs.

Acetone degrading potential of the Canada Basin Arctic Ocean microbial communities

Twelve MAGs contained the acetone carboxylase operon (*acxABC*) (Fig. 3). A single Canada Basin *Porticoccus* MAG (herein referred to as Port-CB9S-26) made up 0.8 – 12.2% of the metagenomic MAG assemblages based on the TAD₈₀ of the MAG in the surface waters (Fig. 3). Port-CB9S-26 made up >10% of the metagenomic MAG assemblages except for the ice-covered CBN3 MAG assemblages (Fig. 3). The remaining 11 MAGs were relatively rare (0.00002 – 0.3% of the metagenomic genome assemblages) and present in either the chlorophyll maximum and/or Pacific winter water and Atlantic water layers (Fig. 3).

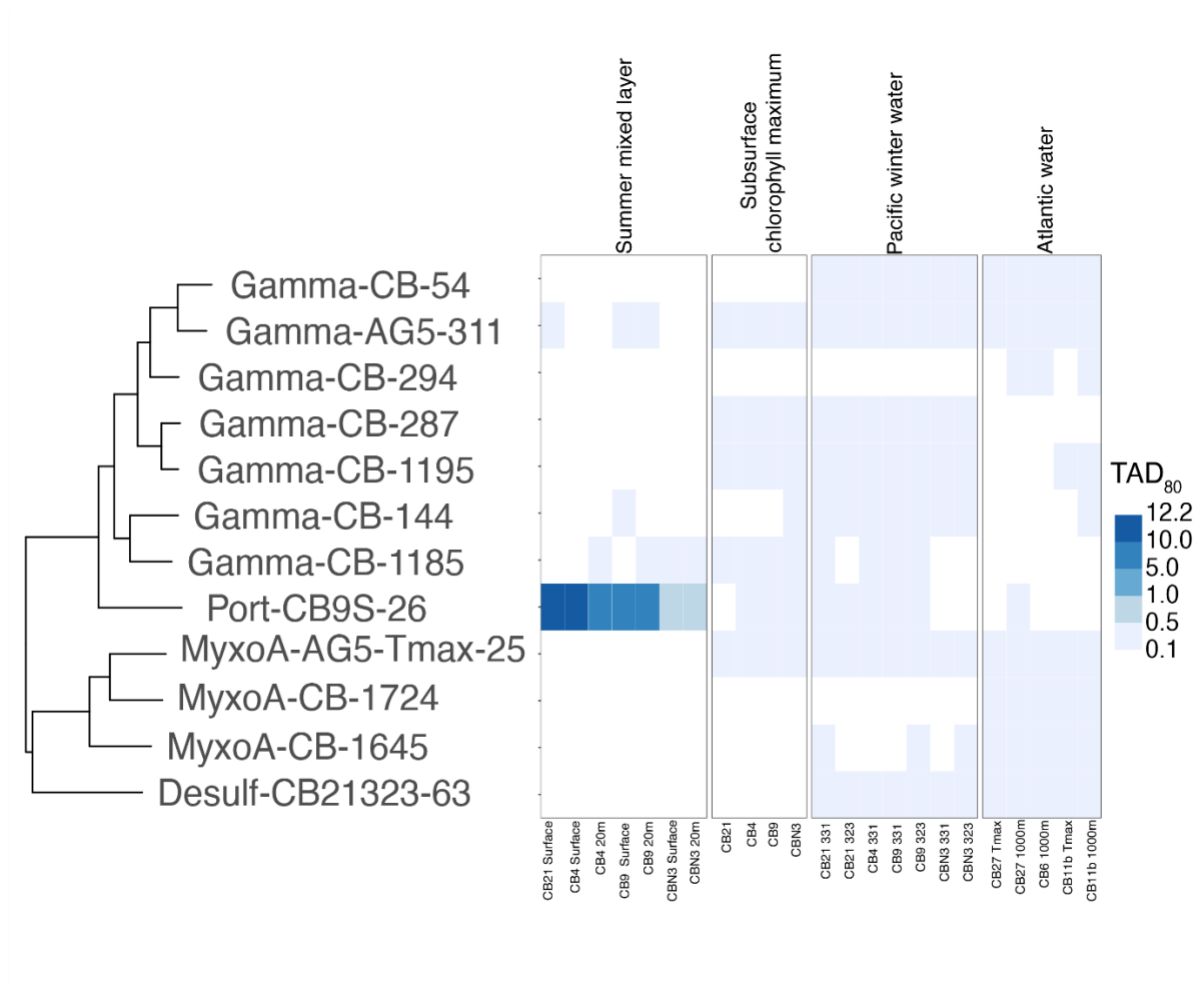


Figure 3. Diversity and relative abundance of MAGs representing acetone-consuming bacteria. Taxonomy of the MAGs based on GTDB phylogenomic classification. Relative abundance of MAGs based on the TAD₈₀ fraction of the metagenomic genome assemblages comprising the MAGs.

Acetone carboxylase is a member of the hydantoinase/oxoprolinase protein family (PF01968) which contains enzymes with different substrate specificities (Schuhle and Heider, 2012). To increase confidence that the genes we identified in the Arctic Ocean are involved in acetone carboxylation, we reconstructed the phylogenetic relationship between the acetone carboxylase alpha subunit (*acxB* gene product) of the MAGs and the corresponding subunits of hydantoinase and acetophenone carboxylase (Fig. 4). Two of twelve MAGs encoded homologs that are found in a clade of putative *acxB* genes that includes genes from experimentally validated acetone-consuming bacteria, increasing our confidence in their functional annotations. (Fig. 4). There was less confidence in acetone being a substrate of the enzyme represented by the homologs from the

other 11 deeper layer associated MAGs, as they fell outside the clade of *acxB* genes of experimentally validated acetone-consuming bacteria (Fig 4).

To explore the *acxB* diversity more deeply in the metagenomes, we recovered all *acxB* genes >500 bp in length from the metagenome assemblies, which likely capture more diversity than what is found only in the MAGs. The *acxB* genes from the assemblies that fell within the clade of *acxB* genes experimentally validated acetone-consuming bacteria are closely related to the *acxB* gene recovered from Port-CB9S-26 (Fig. 4). Therefore, the Port-CB9S-26 genome is an important representative of the acetone-carboxylating diversity of the microbial community in the surface layer.

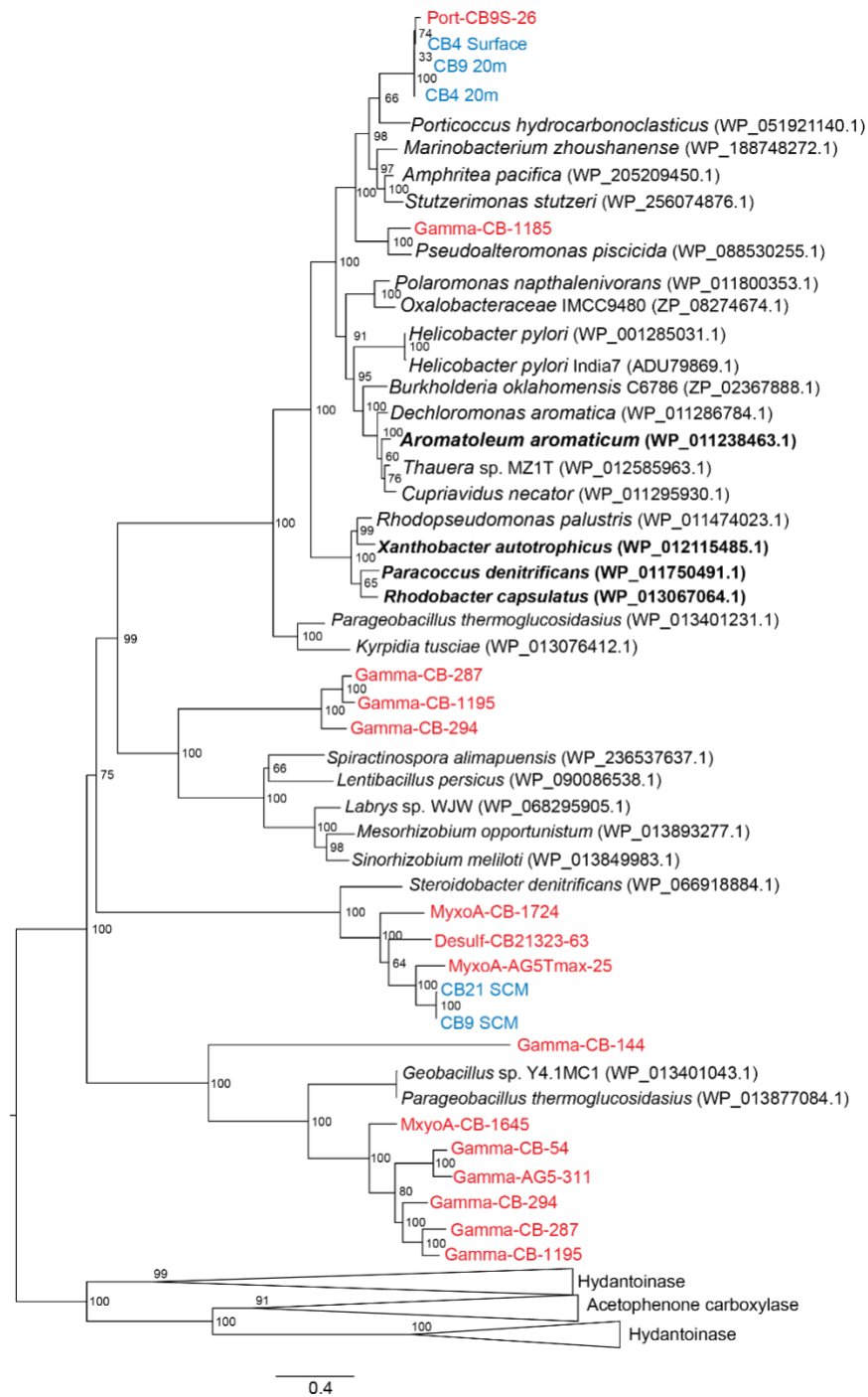


Figure 4. Phylogenetic relationships between acetone carboxylase *acxB* gene from MAGs representing acetone-consuming bacteria in the Canada Basin (red), Canada Basin metagenomes (blue) and experimentally validated acetone-consuming bacteria (bolded). A maximum likelihood tree was constructed using the JTT substitution model, gamma distribution (4 categories), and the nearest neighbor heuristic search model with 100 bootstrap replications. The tree was rooted using the alpha subunit gene of enzymes from the larger hydantoinase/oxoprolinase protein family.

Phylogenomics of *Porticoccaceae*

Port-CB9S-26 was the single dominant acetone-degrading population in the oligotrophic surface waters, prompting further analysis of genomic and metabolic characteristics. Port-CB9S-26 was 1.29 Mbp in length and had 41.6% GC content and estimated to be 92.5% complete with 0.37% contamination. The genome encoded 1,207 genes (including a single and complete rRNA operon) with a median intergenic spacer size of 54 bp.

To understand the phylogenetic placement of Port-CB9S-26 in the tree of life, we generated a 16S rRNA phylogeny of the *Porticoccaceae* and the wider oligotrophic marine Gamma Proteobacteria group (Fig. 5). Port-CB9S-26 and 16S rRNA sequences of *Porticoccus* bacteria from marine surface water formed a clade within the *Porticoccaceae* group, which we have named SP-01 (Fig. 5). Three other clades were identified within the *Porticoccaceae* group (Fig. 5). One clade (SP-02) included sequences from marine surface water, while the other (CP-01) included a coastal marine dinoflagellate-associated isolate *P. hydrocarbonoclasticus* and coastal marine isolate *P. litoralis* which we called CP-01. The third clade was the previously described SAR92 clade (Stingl et al., 2007) (Fig. 5).

Complete genome representatives from the *Porticoccaceae* group included *P. hydrocarbonoclasticus* (GCA_000744735.1) from the CP-01 clade and HTCC2207 (GCA_000153445.1) from the SAR92 clade. SP-01 and *P. hydrocarbonoclasticus* had 65.05% AAI and SP-01 and HTCC2207 had 57.05% AAI. *P. hydrocarbonoclasticus* and HTCC2207 had a genome size of 2.47 and 2.63 Mbp, respectively, both approximately twice the size of the SP-01 genome (1.29 Mbp). *P. hydrocarbonoclasticus* had a 53.1% GC content and HTCC2207 had a 49.1% GC content, 11.5% and 7.5% higher, respectively, than that of SP-01.

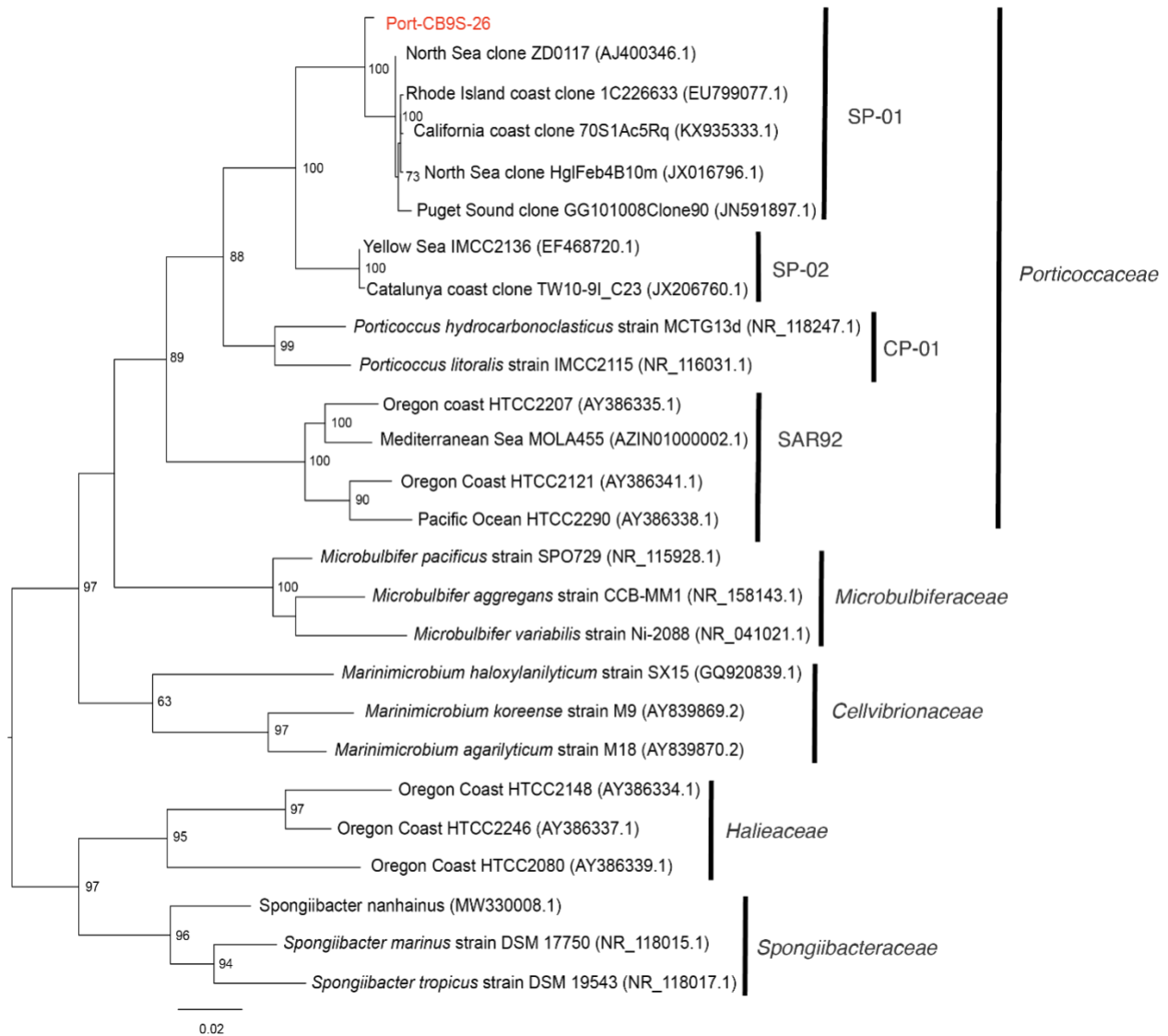


Figure 5. Maximum likelihood phylogeny of marine *Gammaproteobacteria* based on 16S rRNA gene sequences. The blue taxon is from the Port-CB9S-26 MAG with the capacity for acetone carboxylation. The tree was constructed using the GTR + gamma distribution (4 categories) model of nucleotide substitution in MEGA v.11 with 100 bootstraps.

Comparative genomics and metabolic reconstruction

Both Port-CB9S-26 and *P. hydrocarbonoclasticus* possessed an acetone carboxylase operon that was absent in HTCC2207. As SP-01 and *P. hydrocarbonoclasticus* were more closely related at the genome level and both genomes harboured the acetone carboxylase operon, we carried out

comparative genomic analyses with *P. hydrocarbonoclasticus* to better understand the origin and evolution of gene content in Port-CB9S-26.

SP-01 and *P. hydrocarbonoclasticus* shared 1,093 genes while 114 genes were unique to Port-CB9S-26 and 1,206 genes were unique to *P. hydrocarbonoclasticus* (Fig. 6). Metabolic reconstructions of SP-01 and *P. hydrocarbonoclasticus* revealed a minimal central carbon metabolism which relied on acetone, pyruvate, and oligopeptides for carbon and energy sources (Fig. 7). *P. hydrocarbonoclasticus* could also transport TCA cycle intermediates (succinate, fumarate, L-malate, and an uncharacterised dicarboxylate) and amino acids (L-alanine, L-glutamate, L-proline, and L-aspartate) for use as carbon and energy sources (Fig. 7). SP-01 and *P. hydrocarbonoclasticus* also harboured six orthologous uncharacterised ABC transporters (Table S3). *P. hydrocarbonoclasticus* harboured 6 additional uncharacterised ABC transporters to SP-01 (Table S3).

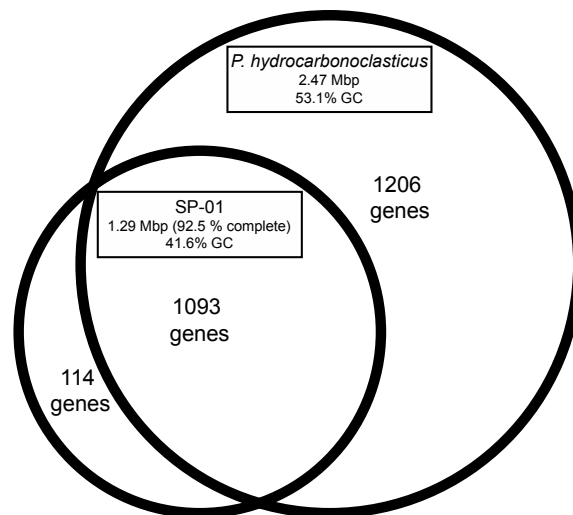


Figure 6. Genomic characteristics and gene overlap between SP-01 and *P. hydrocarbonoclasticus*.

The reduced set of organic carbon transporters of SP-01 indicated that acetone may be a primary carbon source of SP-01. Acetone carboxylase catalyses the carboxylation of acetone to

acetoacetate with CO₂ or bicarbonate (Sluis and Ensign, 1997). Both SP-01 and harboured a cytoplasmic carbonic anhydrase (*can*), facilitating the interconversion of CO₂ and bicarbonate to possibly facilitate the fixation of inorganic carbon in acetone degradation (Fig. 7). *P. hydrocarbonoclasticus* harboured an inorganic carbon transporter (*sulP*) indicating that inorganic carbon could be acquired from the environment for fixation during acetone degradation (Fig. 7).

Acetone-dependent CO₂ fixation leads to the production of acetoacetate (Sluis and Ensign 1997). The degradation of acetoacetate to acetoacetyl-CoA proceeds via a CoA transferase enzyme with either succinyl-CoA or acetyl-CoA as the donor (Platen and Schink 1990, Birks and Kelly, 1997). In denitrifying acetone-consuming bacteria, a succinyl-CoA: acetoacetate transferase catalyses the reaction (Platen and Schink 1990). In butyrate-consuming bacteria, an acetyl-CoA:acetoacetate-CoA transferase catalyses the reaction (Pauli and Overath, 1972). An acetyl-CoA acetyltransferase enzyme catalyzes the generation of 2 acetyl-CoA from acetoacetyl-CoA (Platen and Schink 1990, Pauli and Overath, 1972). SP-01 and *P. hydrocarbonoclasticus* harboured the genes for succinyl-CoA: acetoacetate transferase (*scoAB*) and the gene for an acetyl-CoA acetyltransferase enzyme (*atoD*). SP-01 and *P. hydrocarbonoclasticus* therefore may carry out acetone degradation in a same manner to known acetone-consuming bacteria and use succinyl-CoA as a CoA donor for activation of acetoacetate.

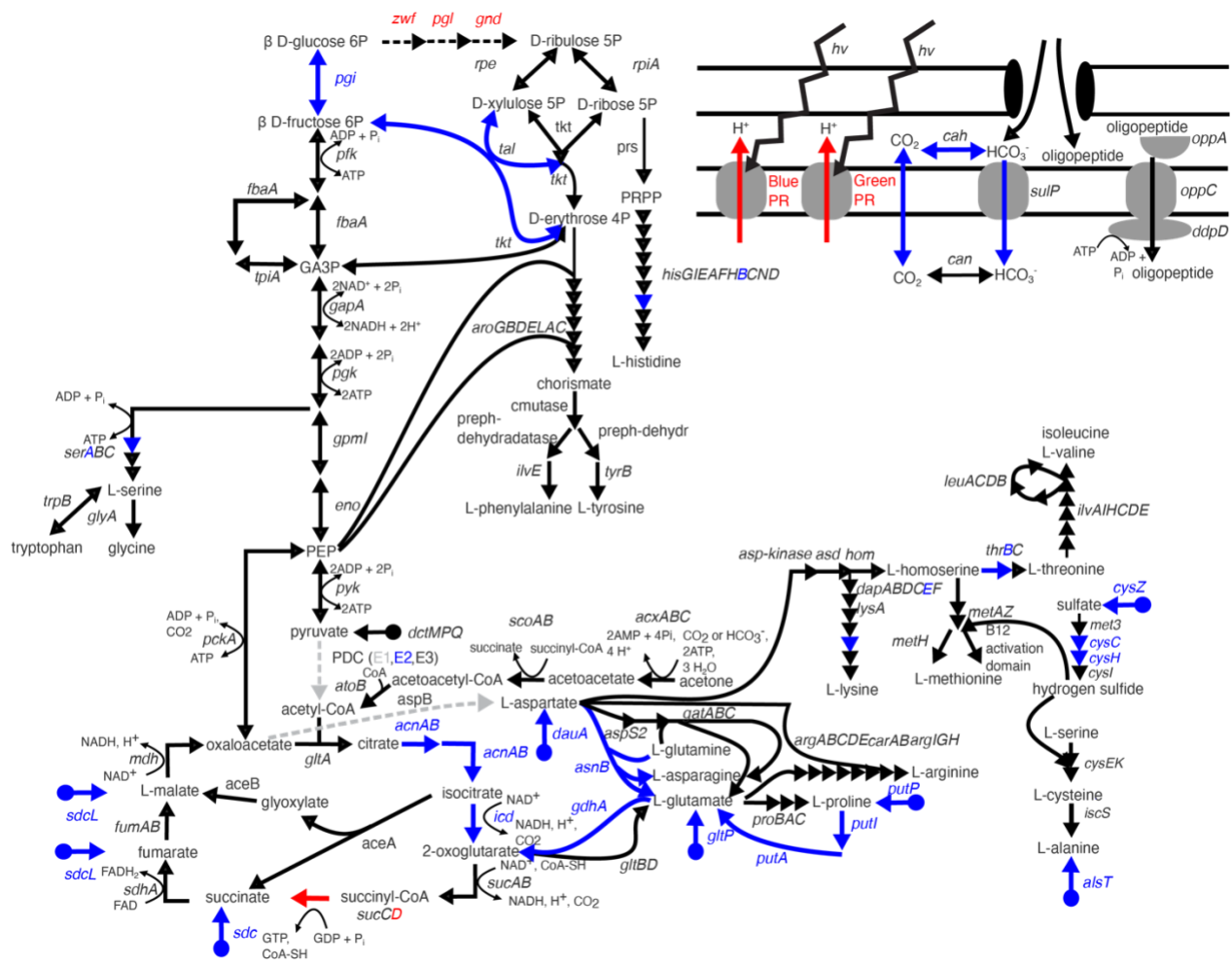


Figure 7. Metabolic reconstruction of the central carbon metabolism of SP-01 and *P. hydrocarbonoclasticus*. Black arrows indicate genes of pathways conserved between SP-01 and *P. hydrocarbonoclasticus*. Red arrows indicate genes of pathways possessed only by SP-01. Blue arrows indicate genes of pathways possessed only by *P. hydrocarbonoclasticus*. Grey dashed arrows indicate genes of pathways that are absent in SP-01 and *P. hydrocarbonoclasticus*.

SP-01 appeared to be able to exploit the continuous daylight of the Arctic summer by obtaining energy from sunlight. Two rhodopsin genes were identified in SP-01 and absent in *P. hydrocarbonoclasticus* (Table S3). Both rhodopsin genes are predicted proton pumps and therefore energy generating. The evidence for this is their positioning within the proteorhodopsin clade (Fig. 8B) and an amino acid motif of aspartate, threonine, and glutamate at positions 97, 101, 108, respectively (Wang et al., 2003). The proteorhodopsin genes were adjacent to each other in the genome (Fig. 8C). Interestingly, the upstream rhodopsin was blue light absorbing, and the downstream rhodopsin was green light absorbing, based on the glutamine and leucine, at

position 105 (Wang et al., 2003). A complete carotenoid biosynthesis pathway was located immediately downstream of the proteorhodopsin genes, (Table S3, Fig 8C) (Misawa et al., 1995, Sabehi et al., 2005).

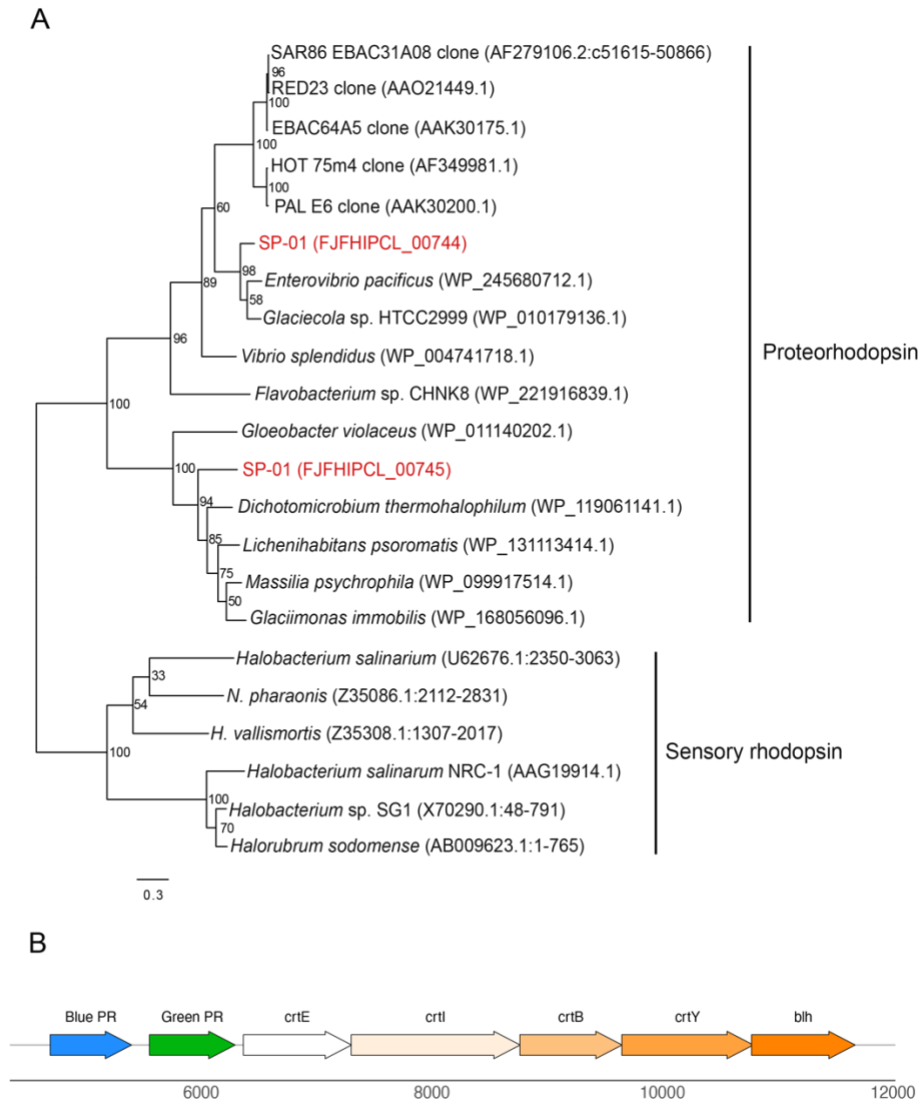


Figure 8. Phylogenetic analysis of rhodopsin genes and the proteorhodopsin biosynthesis operon within the SP-01 genome. (A) A maximum likelihood tree was constructed using the JTT substitution model, gamma distribution (4 categories), and the nearest neighbor heuristic search model with 100 bootstrap replications. Sequences from Port-CB9S-26 are highlighted in red. (B) Visualization of the proteorhodopsin and carotenoid biosynthesis genes within Port-CB9S-26.

Amino acids may represent an additional carbon source for SP-01 and *P. hydrocarbonoclasticus*, however SP-01 relies more on the acquisition of amino acids from the environment for biosynthesis purposes as genes in pathways L-threonine (*thrB*) (and consequently, isoleucine and valine), L-lysine (*dapE*), L-cysteine (*cysCH*), L-glutamine (*asnB*), L-aspartate (*aspB*) (and as a consequence, L-asparagine) are absent (Fig. 7). In contrast, the only amino acid auxotrophy of *P. hydrocarbonoclasticus*, was L-aspartate as aspartate aminotransferase (*aspB*), is absent in the genome (Fig. 7). *P. hydrocarbonoclasticus* may acquire L-aspartate via *dauA* C4 dicarboxylic acid transporter (Fig. 7). SP-01 harbours an oligopeptide transporter (*opp*) (Fig. 7) which may allow for the acquisition of amino acids.

Metabolic reconstruction of SP-01 and *P. hydrocarbonoclasticus* revealed a minimal central carbon metabolism tuned to the degradation of oxidized carbon sources. Conspicuously, the genes for the E1 component of the pyruvate dehydrogenase complex (pyruvate dehydrogenase E1 component alpha subunit, *pdhA*, pyruvate dehydrogenase E1 component beta subunit, *pdhB*, and pyruvate dehydrogenase E1 component, *aceE*) were absent in both genomes (Fig. 7). The gene for the E3 component of the pyruvate dehydrogenase complex (dihydrolipoamide dehydrogenase, *pdhD*) was present in both genomes and the gene for the E2 component (pyruvate dehydrogenase E2 component, *pdhC*) was identified in *P. hydrocarbonoclasticus*. Therefore, consistent with the lack of sugar transporters, both genomes did not have the capacity to degrade more reduced sugars via glycolysis nor have the capacity for the degradation of sugars to eventually enter the TCA cycle through the oxidation of pyruvate to acetyl-CoA (Fig. 7).

The genes for the gluconeogenesis pathway to β D-glucose 6P and β D-fructose 6P were present in the *P. hydrocarbonoclasticus* and SP-01 genomes (Fig. 7). In both genomes, the oxidative, NADPH generating portion of the pentose phosphate pathway from β D-Glucose 6P was absent, while a complete and near complete (transaldolase, *tal*, absent), non-oxidative portion of the pentose phosphate pathway was present in both *P. hydrocarbonoclasticus* and SP-01 (Fig. 7). The lack of an oxidative pentose phosphate pathway is perhaps consistent with the oxidized carbon compounds being a carbon and energy source of SP-01 and *P. hydrocarbonoclasticus* as β D-glucose 6P would be solely a product of gluconeogenesis. A near complete TCA cycle was present in the *P. hydrocarbonoclasticus* genome (succinyl-CoA synthetase alpha subunit, *sucD*

was absent). A partially complete TCA cycle was present in SP-01 (aconitate hydratase, *acnA*, aconitate hydratase 2, *acnB*, and isocitrate dehydrogenase, *icd* were absent). A complete glyoxylate pathway was present in *P. hydrocarbonoclasticus* (Fig. 7). A partially complete glyoxylate pathway was present in SP-01 (aconitate hydratase subunit, *acnA*, and aconitate hydratase 2, *acnB* was absent) (Fig. 7).

SP-01 had a reduced genetic repertoire compared to *P. hydrocarbonoclasticus* with respect to inorganic nutrient acquisition and metabolism (Fig. 9). The ability of SP-01 to acquire diverse inorganic nitrogen compounds is reduced compared to *P. hydrocarbonoclasticus* (Fig. 9). An ammonium transporter (*amt*) was present in both genomes while *P. hydrocarbonoclasticus* also harboured a nitrate ABC transporter (*nrtABC*), assimilatory nitrate reduction genes (*nasAB* and *nirBD*) and a cyanase gene (*cynS*) (Fig. 9). SP-01 had the ability to acquire the same phosphorus compounds with the conservation of a phosphate ABC transporter (*pstABCS*) and polyphosphate/phosphate selective porin (*oprO/P*), however, *P. hydrocarbonoclasticus* possessed a higher diversity of genes for acquiring phosphate. *P. hydrocarbonoclasticus* harboured an inorganic phosphate transporter (PiT family), which was absent in SP-01 (Fig. 9). With respect to acquisition of phosphorus from organic matter, *P. hydrocarbonoclasticus* has an alkaline phosphatase gene (*phoA*) which was absent in SP-01 (Fig. 9). SP-01 had a reduced number of genes associated with polyphosphate metabolism (Fig. 9). An exopolyphosphatase gene (*ppx*) was conserved in both genomes, although *P. hydrocarbonoclasticus* retained the full-length gene (Fig. 9). Polyphosphate kinase 1 and 2 genes were absent in the SP-01 genome but were retained in the *P. hydrocarbonoclasticus* genome (Fig. 9).

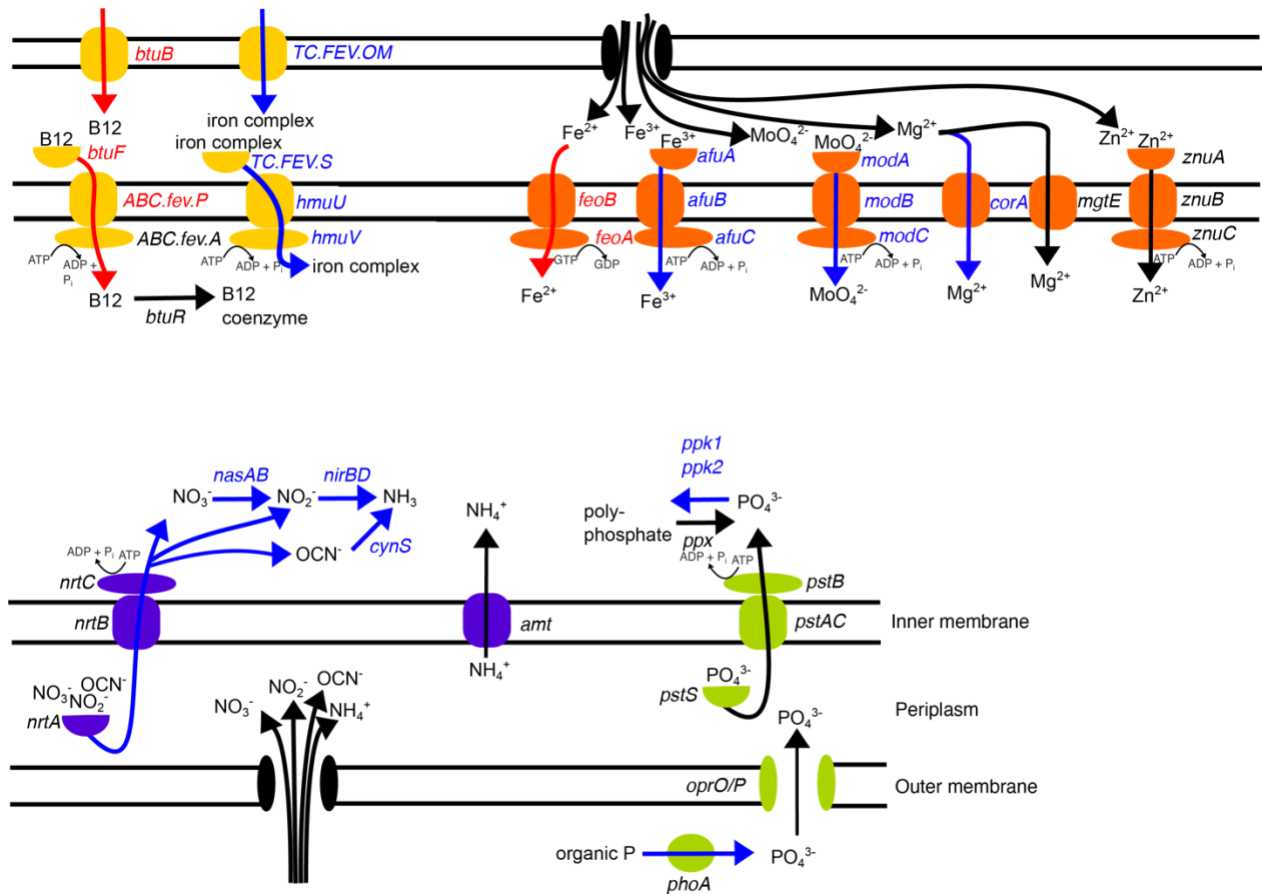


Figure 9. Reconstruction of vitamin and inorganic nutrient transporters of SP-01 and *P. hydrocarbonoclasticus*. Black arrows indicate genes of pathways conserved between SP-01 and *P. hydrocarbonoclasticus*. Red arrows indicate genes of pathways possessed only by SP-01. Blue arrows indicate genes of pathways possessed only by *P. hydrocarbonoclasticus*.

SP-01 and *P. hydrocarbonoclasticus* diverged in their vitamin and metal acquisition capacity. SP-01 harboured an operon for acquisition of vitamin B12 (B12 outermembrane receptor and permease protein, *btuBF* and an ABC transporter, *ABC.fev.AP*) (Fig. 9). *P. hydrocarbonoclasticus* harboured an operon for the acquisition of heme (iron complex outer membrane receptor, *TC.fev.OM*, and substrate binding protein, *TC.fev.S*, and heme ABC transporter, *hmuUV*) (Fig. 9). All genes in both operons, except the gene for the outer membrane receptor, were identified as orthologs indicating functions of the genes may be more conserved than the annotations suggest. Consistent with this possibility, *P. hydrocarbonoclasticus* also harboured a cobalamin adenosyltransferase (*btuR*) for synthesis of the B12 coenzyme from vitamin B12 within the heme transport operon. Neither genome harboured genes for vitamin B12

biosynthesis pathway, indicating that both SP-01 and *P. hydrocarbonoclasticus* would rely on B12 transport. Both genomes harboured genes for magnesium (*mgtE*) and zinc (*znuABC*) transporters (Fig. 9). *P. hydrocarbonoclasticus* also harboured an additional gene for magnesium transport (*corA*). *P. hydrocarbonoclasticus* harboured genes for a Fe³⁺ transporter (*afuABC*) and molybdenum transporter (*modABC*) which were absent in SP-01 (Fig. 9). SP-01 harboured genes for a Fe²⁺ transporter (*feoAB*) which were absent in *P. hydrocarbonoclasticus* (Fig. 9).

We investigated the gene expression activity of SP-01 and found that in the surface layer nearly all genes (90.0-99.8%) within the genome were expressed. Most genes (945/1207, 78.2%) had an RPKM between 1 and 10 (Fig. S9). Two of the most highly expressed genes were post-transcriptional regulators, transfer-messenger RNA and cold-shock protein (*cspA*) genes (Fig. 8A). The acetone carboxylase genes and a blue-light sensing proteorhodopsin were also some of most highly expressed genes in the genome (Fig. 8A). While the blue light sensing proteorhodopsin gene was one of the most highly expressed in the genome (Fig. 8A), the green light sensing proteorhodopsin was less expressed (~64-fold reduction).

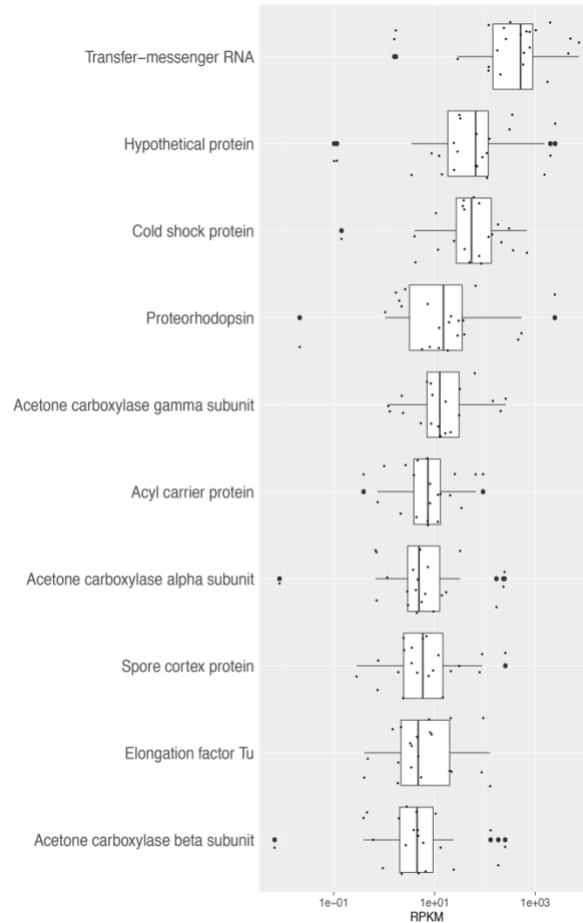


Figure 10. Top ten most expressed genes of SP-01 in the Canada Basin based on mean metatranscriptomic recruitment.

Environmental drivers of SP-01 abundance in the Canada Basin in a 16S rRNA time series

We were fortunate that the Port-CB9S-26 genome possessed a 16S rRNA gene as it allowed us to identify the population in a 16S rRNA time series dataset of Canada Basin summer waters (2004-2017) (Kraemer et al., 2023). A single amplicon sequence variant (ASV) was identified which exhibited 100% identity to the 16S rRNA gene in the Port-CB9S-26 genome. Agreeing with the estimation of community fraction using the metagenomes, the average relative abundance of SP-01 was 8% in 2017 (Fig. 3, Fig. 10). Before 2017, the average relative abundance fluctuated, but remained below 2.5% (Fig. 10). Throughout the timeseries, SP-01 was restricted to the summer mixed layer (<15 m) and upper Arctic water (> 15m and below 31 PSU) and was absent in

Pacific water below the surface except in 2010 and in 2012 when the average abundance in the Pacific water was 0.01% (Fig. 10).

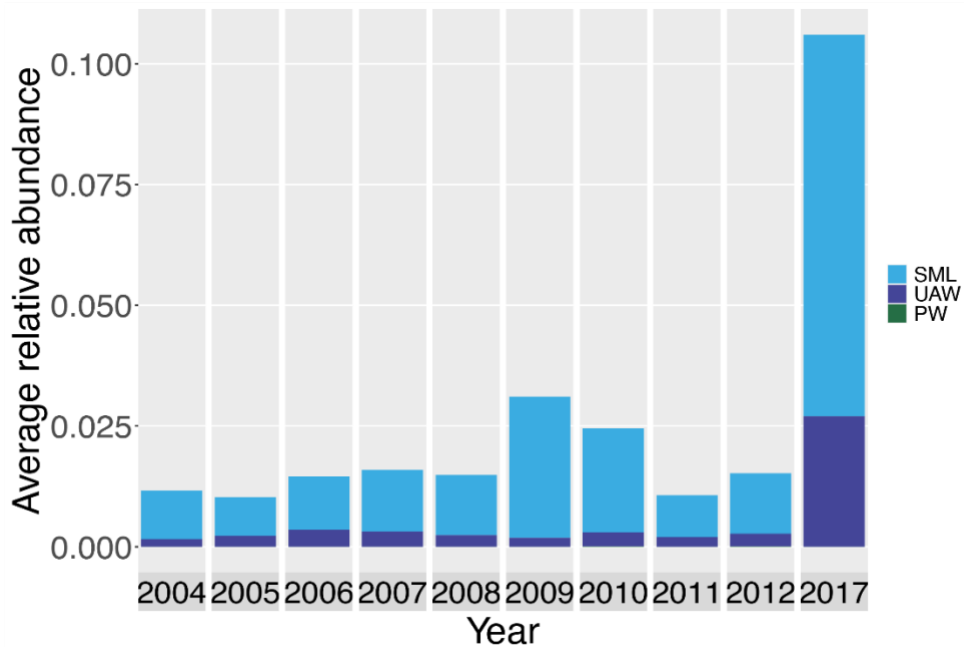


Figure 11. Average annual relative abundance of SP-01 in the Canada basin over 2004-2017. The relative abundance of SP-01 was averaged over the summer mixed layer (SML) (<15 m), upper Arctic water (UAW) (> 15m, < 31 PSU), and Pacific water (PW) (31-33 PSU) for each year.

The peak in abundance of SP-01 in 2017 was distinguished by being one of three years (2008 at CB29, 2012 at CB29 and CB21 and 2017 at CB21 and CB9) over the time series when an ice free (< %15 sea ice concentration) period lasted at least one month (Fig. 11). Therefore, we explored whether ice free conditions could be the driver of the abundance of the acetone-consuming bacteria. The median relative abundance of the bacteria (0.6%) during ice-free periods was higher than during ice-covered periods (1.9%) (Fig. S10), however the difference was not statistically significant ($p > 0.05$, Welch two sample t-test).

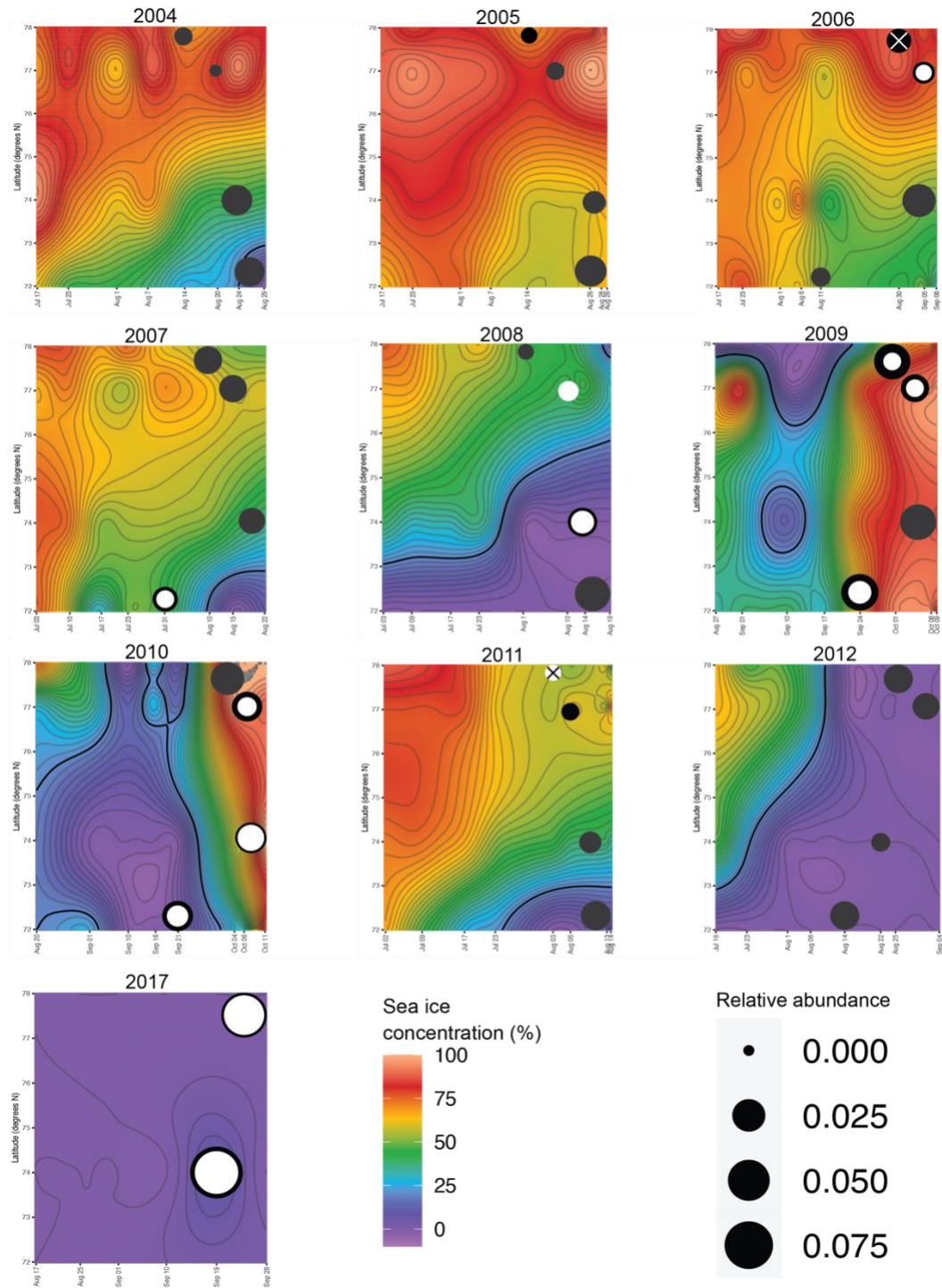


Figure 12. Sea ice concentration and relative abundance of SP-01 in the Canada Basin over 2004-2017 in the summer mixed layer and upper Arctic water. Relative abundance indicated with a black circle, white circle and grey circle in the SML, UAW, and PW, respectively. Black cross and white cross indicates SML not sampled, and UAW not sampled, respectively.

Discussion

Oligotrophic Arctic Ocean surface specialist bacteria thrive on VOCs

VOCs are increasingly recognized as a carbon and energy source for bacterial growth under seasonal oligotrophic regimes (Dixon et al., 2014, Beale et al., 2015) and in oligotrophic seas (Dixon et al., 2011, Beale et al., 2013). Correspondingly, in the oligotrophic surface waters of the Arctic Ocean, we identified active methanol-consuming bacteria as previously reported (Ramachandran et al., 2021), and active acetone-consuming bacteria. Remarkably, the acetone-consuming bacteria made up over 10% of the microbial community in the surface water and were almost completely represented by a single population of *Porticococcus* bacteria.

VOCs may be an important carbon and energy source for oligotrophic bacteria as their passive uptake does not rely on the expression of transporters nor energy input. In the metabolism of acetone, acetone carboxylase would allow bacteria to access inorganic carbon (CO_2 or HCO_3^-), which is readily available for the heterotrophic metabolism of VOCs. This provides further evidence of heterotrophic inorganic carbon fixation being a widespread strategy for survival of Arctic Ocean bacteria in oligotrophic conditions (Alonso-Sáez et al., 2010). Genome reduction of SP-01 relative to *P. hydrocarbonoclasitus* and the SAR92 clade suggests that VOC-based carbon metabolism may be a metabolic strategy that promotes genome reduction which would allow oligotrophic bacteria to remain competitive in their resource-scarce environment.

Acetone degradation in the surface water appears to be a highly specialized metabolism only found in a single lineage of bacteria in the Arctic Ocean. The low diversity of acetone-consuming bacteria could arise due to periodic availability of acetone in surface waters and/or competition with degraders of more labile carbon compounds. Acetone may enter the water column through photodegradation of chromophoric dissolved organic matter (DeBruyn et al., 2011) or through air-sea exchange, both of which would require reduced sea-ice cover for exposure to light or the atmosphere, respectively (Wohl et al., 2022). During ice free periods when intensified stratification of the water column occurs, labile carbon sources may be scarce and acetone-degrading populations of bacteria may increase in abundance. After phytoplankton blooms in the

spring, organisms which use other more energy-rich carbon sources may outcompete acetone-degrading bacteria resulting in the reduction of their population size. The metabolic specialization may be driven by the significant investment in cellular resources necessary to carry out acetone carboxylase dependent acetone degradation where ~25% of the proteome is composed of acetone carboxylase (Sluis and Ensign, 1997). Therefore, for acetone degradation to be a productive metabolism it would likely have a strong effect on the solvent capacity of the cell. The promotion of metabolic specialization has been proposed to arise from biochemical conflicts, such as cellular near saturation of solvent capacity (Johnson et al., 2012). Because the solvent capacity is near saturation, the expression of acetone carboxylase to high levels may concurrently reduce the production of enzymes in more productive pathways. Therefore, a potentially infrequent scenario of greater supply of acetone over other substrates and large investment in the production of acetone carboxylase would need to co-occur to provide a selective advantage for acetone metabolism in bacteria and lateral transfer to other lineages.

Another strategy of oligotrophic marine bacteria to be responsive to a patchy environment is to reduce transcriptional level control on gene expression (Noell et al., 2023). Overall, SP-01 appeared to have little transcriptional gene expression regulation as nearly all genes in the genome were active in the surface waters and most had a similar level of activity. Some of the most highly expressed genes, transfer-messenger RNA and cold-shock protein genes could be involved in post-transcriptional gene regulation. Transfer-messenger RNA is a mechanism to rescue stalled ribosomes which could arise in the absence of appropriate charged tRNA, for example, during amino-acid starvation (Withey and Friedman, 2003). Cold-shock proteins could modulate translation of mRNA by acting as an RNA chaperone. Although translational level control on gene expression appears to be important, some gene regulation may occur at the transcriptional level in SP-01 as evidenced by the larger median intergenic spacer regions in the genome and by the dramatic difference in expression level of some genes, for example the proteorhodopsin genes which differ in their light absorption optima. In the oligotrophic OM43 clade, some of the strongest changes in transcriptional response to different nutrient conditions were observed in carbon metabolism associated genes (including methanol dehydrogenase and proteorhodopsin) (Gifford et al., 2016). The strategy of gene expression regulation at the transcriptional level of carbon metabolism associated genes (acetone carboxylase and

proteorhodopsin) while having little gene expression regulation at the transcriptional level may also be carried out in SP-01.

Exploitation of varying light regimes of the Arctic Ocean

The ability to use sunlight to supplement heterotrophic energy generation is ubiquitous within microbes inhabiting the oligotrophic ocean photic zone (Finkel et al., 2013) and could be an especially advantageous trait in the summer surface waters of the Canada Basin which receive up to 24 hours of daylight. The presence of two proteorhodopsin genes that differ in their light-absorbing capacities may indicate the importance of proteorhodopsin in the survival in the summer surface water of the Arctic Ocean. There is evidence for the retention of multiple proteorhodopsin homologs in prokaryotic genomes (Olson et al., 2018). However, the maintenance of both blue and green light absorbing proteorhodopsins in genomes of marine prokaryotes has not been described previously in the literature to our knowledge and therefore may be an uncommon strategy in marine lineages of prokaryotes. Differing light-absorbing capacities are a driving force of vertical niche differentiation within lineages of marine prokaryotes (Prochlorococcus, SAR11, MGII) (Frigaard et al., 2006). For example, only single homologs of proteorhodopsin genes are maintained in genomes of the Marine Group II Archaea, possibly to allow for the quick transition to exploit the wavelength of light at different depths (Rinke et al., 2018). In the sea-ice influenced Arctic Ocean, in contrast, having a wider range of type of light to exploit for energy generation would be a beneficial adaptation to the varying light regime at the same depth of the Arctic Ocean due to dynamic ice-cover and wide range of angle of incidence of the sunlight on water at the Pole. On the other hand, it is possible that we captured the spectral tuning of the *Porticoccus* genome in action as the blue light sensing proteorhodopsin was much more active compared to the green-light sensing proteorhodopsin.

Proteorhodopsin may have additional roles in survival in the surface waters of the Arctic Ocean. In addition to energy generation to support growth, proteorhodopsin expression may serve to enhance the long-term viability of cells by stabilizing the cell membrane (Song et al., 2020). This survival strategy may explain the observations that proteorhodopsin continues to be transcribed through the polar night into the winter (Nguyen et al., 2015). Expression of proteorhodopsin may

supplement heterotrophic energy metabolism during spring and summer ice-free periods while promoting long term survival of bacteria during the polar night.

Nutrient acquisition adaptations of the acetone-degrading *Porticoccus* populations to the Arctic Ocean environment

The evolutionary adaptation of SP-01 to the oligotrophic waters of the Arctic Ocean is reflected in the reduced nutrient transporter content compared to *P. hydrocarbonoclasticus*. Post-bloom and into the autumn, the surface waters of the Arctic Ocean are depleted in nitrate while concentrations of ammonia increase (Simpson et al., 2008) due to nitrogen remineralization and concurrent inhibition of nitrification in the sunlit surface waters. The absence of a nitrate transporter in the Port-CB9S-26, but presence of an ammonia transporter, may be a consequence of the absence of nitrate but increasingly available ammonia. The absence of nitrate transporter was also identified in methanol-consuming bacteria active in the surface waters (Ramachandran et al., 2021) indicating that this may be a more general strategy of genome reduction in Arctic Ocean surface bacteria.

High affinity phosphate ABC transporters are important for the uptake of phosphate in nutrient limiting systems (Sowell et al., 2009, Ustick et al., 2021). We identified a phosphate ABC transporter (*PstS*) within the genome of Port-CB9S-26 and previously in methanol-consuming populations of bacteria active in the Arctic Ocean surface waters (Ramachandran et al., 2021). Port-CB9S-26 may remain competitive in the oligotrophic environment by further increasing the rate of phosphate uptake with the expression of polyphosphate/phosphate selective porins (*OprO* or *OprR*) compared to passive diffusion of phosphate through general porins into the periplasm. Unlike *P. hydrocarbonoclasticus* and the methanol-consuming Arctic Ocean surface populations, SP-01 appears to lack the capacity for phosphate storage as polyphosphate, retaining only a partial exopolyphosphatase. There is little evidence for exopolyphosphatase function as the gene does not include the active site of the enzyme, although curiously it appears to be transcribed. Perhaps the absence of polyphosphate metabolism is related to the increased phosphate rate of uptake via phosphate selective porins which may reduce the reliance in storing phosphate in the

form of polyphosphate as phosphate be taken up more quickly when nitrogen is no longer limiting growth.

Alkaline phosphatase is a genomic adaptation to acquire phosphate from dissolved organic matter in phosphate depleted environments of *Cyanobacteria* (Ustick et al., 2021). Although present in the genome of phytoplankton-associated *P. hydrocarbonoclasticus*, alkaline phosphatase gene is absent within the Port-CB9S-26 genome. In the oligotrophic post bloom surface waters of the Arctic Ocean SP-01 would have less access to fresher phytoplankton derived organic matter for phosphate acquisition.

Acetone-consuming bacteria regime shift associated with extended ice-free periods

The microbial community of the Arctic Ocean must contend with some of the most oligotrophic seas on earth. Along with the intensified oligotrophication that results from sea-ice melt, the onset of the polar day permits photochemical degradation of organic matter in the atmosphere and in melting snowpack in the Arctic region (Gao et al., 2012, Pernov et al., 2021). The reduced sea ice cover may increase air-sea exchange and/or photochemical degradation of organic matter within the water column resulting in the input of acetone in the surface waters (Sjostedt et al., 2012, Wohl et al., 2022). The peak in abundance of SP-01 in 2017 coincided with one of lowest overall sea-ice concentrations in the Beaufort Sea on record (Babb et al., 2020). This indicates that VOCs may make convenient carbon sources for oligotrophs and therefore bacteria such as VOC-consuming bacteria Arctic Ocean surface specialist bacteria may increase in relevance as oligotrophication intensifies with environmental change.

Conclusions

This work expanded our understanding of the bacterial metabolism in the oligotrophic surface waters of the Arctic Ocean, a little studied ecosystem which is undergoing intensified oligotrophication, along with large regions of the global ocean. The generation of a metagenomic genome catalogue provides an additional resource for studying the microbial community of the stratified water column of an oligotrophic region of the Arctic Ocean. The identification of

bacteria with the capacity to consume acetone in the surface waters of the Arctic Ocean is consistent with observations of acetone uptake by microbes in oligotrophic gyres and winter coastal ecosystems. We provide evidence that acetone may be degraded through heterotrophic CO₂ fixation in oligotrophic marine surface waters. We find evidence for genome reduction of putative acetone-degrading bacteria related to central carbon metabolism and nutrient transport. The investigation of VOC degradation capacity the Arctic Ocean microbial community provides evidence that VOCs degradation may be an important carbon and energy sources of bacteria in the oligotrophic surface waters of the Arctic Ocean. To further characterise the role of acetone-consuming bacteria in the cycling of VOCs in oligotrophic surface waters of the Arctic Ocean, targeted cultivation of acetone-consuming bacteria could be carried out in order to quantify rates of acetone consumption. Additionally, future work characterising the global ecology of acetone carboxylase would help us understand the distributions of the metabolism of acetone via acetone carboxylase and whether the niche of the organisms possessing this metabolism is restricted to oligotrophic waters globally.

References

1. Aagaard K, Carmack EC. The role of sea ice and other fresh water in the Arctic circulation. *J Geophys Res Oceans* 1989; 94: 14485–14498.
2. Alonso-Sáez L, Galand PE, Casamayor EO, Pedrós-Alió C, Bertilsson S. High bicarbonate assimilation in the dark by Arctic bacteria. *Isme J* 2010; 4: 1581–1590.
3. Altshul SF, Gish W, Miller W, Myers EW, Lipman DJ. Basic local alignment search tool. *J Mol Biol* 1990; 215(3): 403-410.
4. Aramaki T, Blanc-Mathieu R, Endo H, Ohkubo K, Kanehisa M, Goto S, et al. KofamKOALA: KEGG ortholog assignment based on profile HMM and adaptive score threshold. *Bioinformatics* 2019; 36: 2251–2252.
5. Babb DG, Landy JC, Lukovich JV, Haas C, Hendricks S, Barber DG, et al. The 2017 Reversal of the Beaufort Gyre: Can Dynamic Thickening of a Seasonal Ice Cover During a Reversal Limit Summer Ice Melt in the Beaufort Sea? *J Geophys Res Oceans* 2020; 125.
6. BBMap – Bushnell B. – sourceforge.net/projects/bbmap/ (2015)
7. Beale R, Dixon JL, Arnold SR, Liss PS, Nightingale PD. Methanol, acetaldehyde, and acetone in the surface waters of the Atlantic Ocean. *J Geophys Res Oceans* 2013; 118: 5412–5425.
8. Beale R, Dixon JL, Smyth TJ, Nightingale PD. Annual study of oxygenated volatile organic compounds in UK shelf waters. *Mar Chem* 2015; 171: 96–106.
9. Birks SJ and Kelly DJ. Assay and properties of acetone carboxylase, a novel enzyme involved in acetone-dependent growth and CO₂ fixation in *Rhodobacter capsulatus* and other photosynthetic and denitrifying bacteria. *Microbiology* 1997; 143, 755-766.
10. Bowers RM, Kyrpides NC, Stepanauskas R, Harmon-Smith M, Doud D, Reddy TBK, et al. Minimum information about a single amplified genome (MISAG) and a metagenome-assembled genome (MIMAG) of bacteria and archaea. *Nat Biotechnol* 2017; 35: 725–731.
11. Brown MV, Lauro FM, DeMaere MZ, Muir L, Wilkins D, Thomas T, et al. Global biogeography of SAR11 marine bacteria. *Mol Syst Biol* 2012; 8: 595–595.

12. Bruyn WJ de, Clark CD, Pagel L, Takehara C. Photochemical production of formaldehyde, acetaldehyde and acetone from chromophoric dissolved organic matter in coastal waters. *J Photochem Photobiology Chem* 2011; 226: 16–22.
13. Chaumeil P-A, Mussig AJ, Hugenholtz P, Parks DH. GTDB-Tk: a toolkit to classify genomes with the Genome Taxonomy Database. *Bioinformatics* 2019.
14. Colatriano D, Tran PQ, Guéguen C, Williams WJ, Lovejoy C, Walsh DA. Genomic evidence for the degradation of terrestrial organic matter by pelagic Arctic Ocean Chloroflexi bacteria. *Commun Biology* 2018; 1: 90.
15. Dixon JL, Beale R, Nightingale PD. Rapid biological oxidation of methanol in the tropical Atlantic: significance as a microbial carbon source. *Biogeosciences* 2011; 8: 2707–2716.
16. Dixon JL, Beale R, Sargeant SL, Tarran GA, Nightingale PD. Microbial acetone oxidation in coastal seawater. *Front Microbiol* 2014; 5: 243.
17. Duarte CM, Regaudie-de-Gioux A, Arrieta JM, Delgado-Huertas A, Agustí S. The Oligotrophic Ocean Is Heterotrophic*. *Mar Sci* 2013; 5: 551–569.
18. Dupont CL, Rusch DB, Yooseph S, Lombardo M-J, Richter RA, Valas R, et al. Genomic insights to SAR86, an abundant and uncultivated marine bacterial lineage. *Isme J* 2012; 6: 1186–1199.
19. Ensign SA, Small FJ, Allen JR, Sluis MK. New roles for CO₂ in the microbial metabolism of aliphatic epoxides and ketones. *Arch Microbiol* 1998; 169: 179–187.
20. Finkel OM, Béjà O, Belkin S. Global abundance of microbial rhodopsins. *Isme J* 2013; 7: 448–451.
21. Frigaard N-U, Martinez A, Mincer TJ, DeLong EF. Proteorhodopsin lateral gene transfer between marine planktonic Bacteria and Archaea. *Nature* 2006; 439: 847–850.
22. Gao SS, Sjostedt SJ, Sharma S, Hall SR, Ullmann K, Abbatt JPD. PTR-MS observations of photo-enhanced VOC release from Arctic and midlatitude snow. *J Geophys Res Atmospheres* 2012; 117: n/a-n/a.
23. Gifford SM, Becker JW, Sosa OA, Repeta DJ, DeLong EF. Quantitative Transcriptomics Reveals the Growth and- Nutrient-Dependent Response of a Streamlined Marine Methylotroph to Methanol and Naturally Occuring Dissolved Organic Matter. *Mbio* 2016; 7:e01279-16.

24. Giovannoni SJ, Thrash JC, Temperton B. Implications of streamlining theory for microbial ecology. *Isme J* 2014; **8**: 1553–1565.
25. Giovannoni SJ. SAR11 Bacteria: The Most Abundant Plankton in the Oceans. *Annu Rev Mar Sci* 2017; **9**: 231–255.
26. Grevesse T, Guéguen C, Onana VE, Walsh DA. Degradation pathways for organic matter of terrestrial origin are widespread and expressed in Arctic Ocean microbiomes. *Microbiome* 2022; **10**: 237.
27. Halsey KH, Giovannoni SJ, Graus M, Zhao Y, Landry Z, Thrash JC, et al. Biological cycling of volatile organic carbon by phytoplankton and bacterioplankton. *Limnol Oceanogr* 2017; **62**: 2650–2661.
28. Hutchins DA, Fu F. Microorganisms and ocean global change. *Nat Microbiol* 2017; **2**: 17058.
29. Hyatt D, Chen G-L, LoCascio PF, Land ML, Larimer FW, Hauser LJ. Prodigal: prokaryotic gene recognition and translation initiation site identification. *Bmc Bioinformatics* 2010; **11**: 119
30. Johnson DR, Goldschmidt F, Lilja EE, Ackermann M. Metabolic specialization and the assembly of microbial communities. *Isme J* 2012; 1985-1991.
31. Kanehisa M, Sato Y. KEGG Mapper for inferring cellular functions from protein sequences. *Protein Sci* 2020; **29**: 28–35.
32. Kang DD, Li F, Kirton E, Thomas A, Egan R, An H, et al. MetaBAT 2: an adaptive binning algorithm for robust and efficient genome reconstruction from metagenome assemblies. *Peerj* 2019; **7**: e7359.
33. Karsenti E, Acinas SG, Bork P, Bowler C, Vargas CD, Raes J, et al. A Holistic Approach to Marine Eco-Systems Biology. *Plos Biol* 2011; **9**: e1001177.
34. Kraemer S, Ramachandran A, Colatriano D, Lovejoy C, Walsh DA. Diversity and biogeography of SAR11 bacteria from the Arctic Ocean. *Isme J* 2020; **14**: 79–90.
35. Kraemer S, Ramachandran A, Onana VE, Li WKW, Walsh DA. Change in bacterial and archaeal diversity linked to sea ice loss in the Arctic Ocean. (Unpublished)
36. Laslett D, Canback B. ARAGORN, a program to detect tRNA genes and tmRNA genes in nucleotide sequences. *Nucleic Acids Res* 2004; **32**: 11-16

37. Lechner M, Findeiß S, Steiner L, Marz M, Stadler PF, Prohaska SJ. Proteinortho: Detection of (Co-)orthologs in large-scale analysis. *Bmc Bioinformatics* 2011; **12**: 124–124.
38. Li D, Liu C-M, Luo R, Sadakane K, Lam T-W. MEGAHIT: an ultra-fast single-node solution for large and complex metagenomics assembly via succinct de Bruijn graph. *Bioinformatics* 2015; **31**: 1674–1676.
39. Li G, Cheng L, Zhu J, Trenberth KE, Mann ME, Abraham JP. Increasing ocean stratification over the past half-century. *Nat Clim Change* 2020; **10**: 1116–1123.
40. Li H, Handsaker B, Wysoker A, Fennell T, Ruan J, Homer N, et al. The Sequence Alignment/Map format and SAMtools. *Bioinformatics* 2009; **25**: 2078-2079.
41. Li H, Durbin R. Fast and accurate short read alignment with Burrows–Wheeler transform. *Bioinformatics* 2009; **25**: 1754–1760.
42. McLaughlin FA, Carmack EC. Deepening of the nutricline and chlorophyll maximum in the Canada Basin interior, 2003–2009. *Geophys Res Lett* 2010; **37**: n/a-n/a.
43. Misawa N, Satomi Y, Kondo K, Yokoyama A, Kajiwara S, Saito T, et al. Structure and functional analysis of a marine bacterial carotenoid biosynthesis gene cluster and astaxanthin biosynthetic pathway proposed at the gene level. *J Bacteriol* 1995; **177**: 6575–6584.
44. Moore ER, Weaver AJ, Davis EW, Giovannoni SJ, Halsey KH. Metabolism of key atmospheric volatile organic compounds by the marine heterotrophic bacterium *Pelagibacter* HTCC1062 (SAR11). *Environ Microbiol* 2022; **24**: 212–222.
45. Nayfach S, Pollard KS. Average genome size estimation improves comparative metagenomics and sheds light on the functional ecology of the human microbiome. *Genome Biol* 2015; **16**:51.
46. Nguyen D, Maranger R, Balagué V, Coll-Lladó M, Lovejoy C, Pedrós-Alió C. Winter diversity and expression of proteorhodopsin genes in a polar ocean. *Isme J* 2015; **9**: 1835–1845.
47. Noell SE, Hellweger FL, Temperton B, Giovannoni SJ. A Reduction of Transcriptional Regulation in Aquatic Oligotrophic Microorganisms Enhances Fitness in Nutrient-Poor Environments. *Microbiol Mol Biol R* 2023; e00124-22

48. Nishino S, Kawaguchi Y, Inoue J, Yamamoto-Kawai M, Aoyama M, Harada N, et al. Do Strong Winds Impact Water Mass, Nutrient, and Phytoplankton Distributions in the Ice-Free Canada Basin in the Fall? *J Geophys Res Oceans* 2020; 125.
49. Olm MR, Brown CT, Brooks B, Banfield JF. dRep: a tool for fast and accurate genomic comparisons that enables improved genome recovery from metagenomes through de-replication. *Isme J* 2017; **11**: 2864–2868.
50. Olson DK, Yoshizawa S, Boeuf D, Iwasaki W, DeLong EF. Proteorhodopsin variability and distribution in the North Pacific Subtropical Gyre. *Isme J* 2018; **12**: 1047–1060.
51. Paradis E, Schliep K. ape 5.0: an environment for modern phylogenetics and evolutionary analyses in R. *Bioinformatics* 2018; **35**: 526-528
52. Parks DH, Imelfort M, Skennerton CT, Hugenholtz P, Tyson GW. CheckM: assessing the quality of microbial genomes recovered from isolates, single cells, and metagenomes. *Genome Res* 2015; **25**: 1043–1055.
53. Pauli G, Overath P. ato Operon: a Highly Inducible System for Acetoacetate and Butyrate Degradation in *Escherichia coli*. *Eur J Biochem* 1972; **29**: 553–562.
54. Pernov JB, Bossi R, Lebourgeois T, Nøjgaard JK, Holzinger R, Hjorth JL, et al. Atmospheric VOC measurements at a High Arctic site: characteristics and source apportionment. *Atmos Chem Phys* 2021; **21**: 2895–2916.
55. Platen H, Schink B. Enzymes involved in anaerobic degradation of acetone by a denitrifying bacterium. *Biodegradation* 1990; **1**: 243–251.
56. Polovina JJ, Howell EA, Abecassis M. Ocean's least productive waters are expanding. *Geophys Res Lett* 2008; **35**.
57. Price MN, Deutschbauer AM, Arkin AP. GapMind: Automated Annotation of Amino Acid Biosynthesis. *Msystems* 2020; **5**: e00291-20.
58. Price MN, Deutschbauer AM, Arkin AP. Filling gaps in bacterial catabolic pathways with computation and high-throughput genetics. *Plos Genet* 2022; **18**: e1010156.
59. Ramachandran A, McLatchie S, Walsh DA. A Novel Freshwater to Marine Evolutionary Transition Revealed within Methylophilaceae Bacteria from the Arctic Ocean. *Mbio* 2021; **12**: e01306-21.

60. Rinke C, Rubino F, Messer LF, Youssef N, Parks DH, Chuvochina M, et al. A phylogenomic and ecological analysis of the globally abundant Marine Group II archaea (Ca. Poseidoniales ord. nov.). *Isme J* 2019; **13**: 663–675.
61. Rodriguez-R LM, Konstantinidis KT. Bypassing Cultivation To Identify Bacterial Species: Culture-independent genomic approaches identify credibly distinct clusters, avoid cultivation bias, and provide true insights into microbial species. *Microbe Mag* 2014; **9**: 111–118.
62. Rodriguez-R LM, Tsementzi D, Luo C, Konstantinidis KT. Iterative subtractive binning of freshwater chronoserries metagenomes identifies over 400 novel species and their ecologic preferences. *Environ Microbiol* 2020; **22**: 3394–3412.
63. Royo-Llonch M, Sánchez P, Ruiz-González C, Salazar G, Pedrós-Alió C, Sebastián M, et al. Compendium of 530 metagenome-assembled bacterial and archaeal genomes from the polar Arctic Ocean. *Nat Microbiol* 2021; 1–14.
64. Schuhle K, Heider J. Acetone and Butanone Metabolism of the Denitrifying Bacterium “*Aromatoleum aromaticum*” Demonstrates Novel Biochemical Properties of an ATP-Dependent Aliphatic Ketone Carboxylase. *J Bacteriol* 2011; **194**: 131–141.
65. Seemann T. Prokka: rapid prokaryotic genome annotation. *Bioinformatics* 2014; **30**: 2068–2069.
66. Simpson KG, Tremblay J, Gratton Y, Price NM. An annual study of inorganic and organic nitrogen and phosphorus and silicic acid in the southeastern Beaufort Sea. *J Geophys Res Oceans* 2008; **113**.
67. Sjostedt SJ, Leitch WR, Levasseur M, Scarratt M, Michaud S, Motard-Côté J, et al. Evidence for the uptake of atmospheric acetone and methanol by the Arctic Ocean during late summer DMS-Emission plumes. *J Geophys Res Atmospheres* 2012; <https://doi.org/10.1029/2011JD017086>.
68. Singh HB, Kanakidou M, Crutzen PJ, Jacob DJ. High concentrations and photochemical fate of oxygenated hydrocarbons in the global troposphere. *Nature* 1995; **378**: 50–54.
69. Stingl U, Desiderio RA, Cho J-C, Vergin KL, Giovannoni SJ. The SAR92 Clade: an Abundant Coastal Clade of Culturable Marine Bacteria Possessing Proteorhodopsin. *Appl Environ Microb* 2007; **73**: 2290–2296.

70. Sluis MK, Ensign SA. Purification and characterization of acetone carboxylase from *Xanthobacter* strain Py2. *Proc National Acad Sci* 1997; 94: 8456–8461.
71. autotrophicus Strain Py2 and *Rhodobacter capsulatus* Strain B10. *J Bacteriol* 2002; **184**: 2969–2977.
72. Song Y, Cartron ML, Jackson PJ, Davison PA, Dickman MJ, Zhu D, et al. Proteorhodopsin Overproduction Enhances the Long-Term Viability of *Escherichia coli*. *Appl Environ Microb* 2019; **86**: e02087-19.
73. Spring S, Scheuner C, Göker M, Klenk H-P. A taxonomic framework for emerging groups of ecologically important marine gammaproteobacteria based on the reconstruction of evolutionary relationships using genome-scale data. *Front Microbiol* 2015; **6**: 281.
74. Sowell SM, Wilhelm LJ, Norbeck AD, Lipton MS, Nicora CD, Barofsky DF, et al. Transport functions dominate the SAR11 metaproteome at low-nutrient extremes in the Sargasso Sea. *Isme J* 2009; **3**: 93–105.
75. Tamura K, Stecher G, Kumar S. MEGA11: Molecular Evolutionary Genetics Analysis version 11. *Mol Biol Evol* 2021; **38**: msab120-.
76. Ustick LJ, Larkin AA, Garcia CA, Garcia NS, Brock ML, Lee JA, et al. Metagenomic Analysis reveals global-scale patterns of nutrient limitation. *Science* 2021; 372: 287-291.
77. Wagner S, Schubotz F, Kaiser K, Hallmann C, Waska H, Rossel PE, et al. Soothsaying DOM: A Current Perspective on the Future of Oceanic Dissolved Organic Carbon. *Frontiers Mar Sci* 2020; **7**: 341.
78. Wang S, Apel EC, Schwantes RH, Bates KH, Jacob DJ, Fischer EV, et al. Global Atmospheric Budget of Acetone: Air-Sea Exchange and the Contribution to Hydroxyl Radicals. *J Geophys Res Atmospheres* 2020; 125.
79. Withey JH, Friedman DI. A SALVAGE PATHWAY FOR PROTEIN SYNTHESIS: tmRNA and Trans-Translation. *Annu Rev Microbiol* 2003; **57**: 101–123.
80. Wohl C, Jones AE, Sturges WT, Nightingale PD, Else B, Butterworth BJ, et al. Sea ice concentration impacts dissolved organic gases in the Canadian Arctic. *Biogeosciences* 2022; 19: 1021–1045.

Appendices

Table S1.

Metagenome	CB co-assembly	CB surface co-assembly	AG5 co-assembly	Single assemblies	Metagenomes used in fragment recruitment analysis	Metatranscriptomes used in gene expression analysis
CB4 Surface						
CB9 Surface						
CB21 Surface						
CBN3 Surface						
CB4 20m						
CB9 20m						
CB21 20m						
CBN3 20m						
CB4 SCM						
CB9 SCM						
CB21 SCM						
CBN3 SCM						
CB4 32.3						
CB9 32.3						
CB21 32.3						
CBN3 32.3						
CB4 33.1						
CB9 33.1						
CB21 33.1						
CBN3 33.1						
CB11b Tmax						
CB4 Tmax						
CB9 Tmax						
CBN3 Tmax						
CB27 Tmax						
CB4 AW						
CBN3 AW						
CB11b AW						
CB27 AW						
CB6 AW						
CB6 Bottom						
AG5 surface						
AG5 20m						
AG5 SCM						
AG5 32.3						
AG5 33.1						
AG5 Tmax						
AG5 Bottom						
Number of MAGs generated	493	79	72	280	N/A	N/A

Table S2.

Assembly	Selected MAGs^e	Genomospecies^f	Fold reduction into genomospecies
CB^b	493	456	1.08
CB surface^c	79	52	1.52
AG^d	72	51	1.41
single^a	280	104	2.69
total	924	663	1.39

^aindividual assemblies of 32 samples

^bcombined assembly of 24 samples

^ccombined assembly of 11 samples

^dcombined assembly of 7 samples

^eselected based on $\geq 50\%$ completeness and $\leq 10\%$ contamination and strain heterogeneity as calculated by checkM

^fMAGs were clustered at $\geq 95\%$ ANI and $\geq 10\%$ alignment to produce genomospecies

Table S3.

		<i>P. hyd</i>	Port- CB9S- 26
C metabolism	acetone oxidation	1	1
	C4-dicarboxylate TRAP transporter	1	1
	uncharacterised TRAP transporter	2	2
	dicarboxylate/amino acid cation symporter	1	
	C4-dicarboxylic acid transporter	1	
Light energy metabolism	proteorhodopsin		2
Amino acid transport	dipeptide ABC transporter	1	
	oligopeptide ABC transporter	2	2
N metabolism	nitrate ABC transporter	1	
	ass. nitrate reduction	1	
	ammonium transporter	1	1
	cyanate degradation	1	
P metabolism	phosphate ABC transporter	1	1
	inorganic phosphate transporter (PiT family)	1	
	phosphate/Na ⁺ symporter	1	
	alkaline phosphatase	1	
	polyphosphate/phosphate selective porin (OprO/OprP)	1	2
	polyphosphate kinase 1	1	
	polyphosphate kinase 2	1	
	exopolyphosphatase	1	*partial
Motility/extracellular structures	flagella		
	type IV pili		
Metal and vitamin transport	Mo ABC transporter	1	
	Mg transporter	3	1
	Mg/Co transporter	1	1
	Fe(2 ⁺) transporter		1
	Fe(3 ⁺) ABC transporter	1	
	heme ABC transporter	2	1
	Zn ABC transporter	1	1
	vitamin B12 transporter	1	2
Misc. transporters	sulfate transporter	1	
	uncharacterised ABC transporter	12	6
Other pathways	carotenoid biosynthesis		1

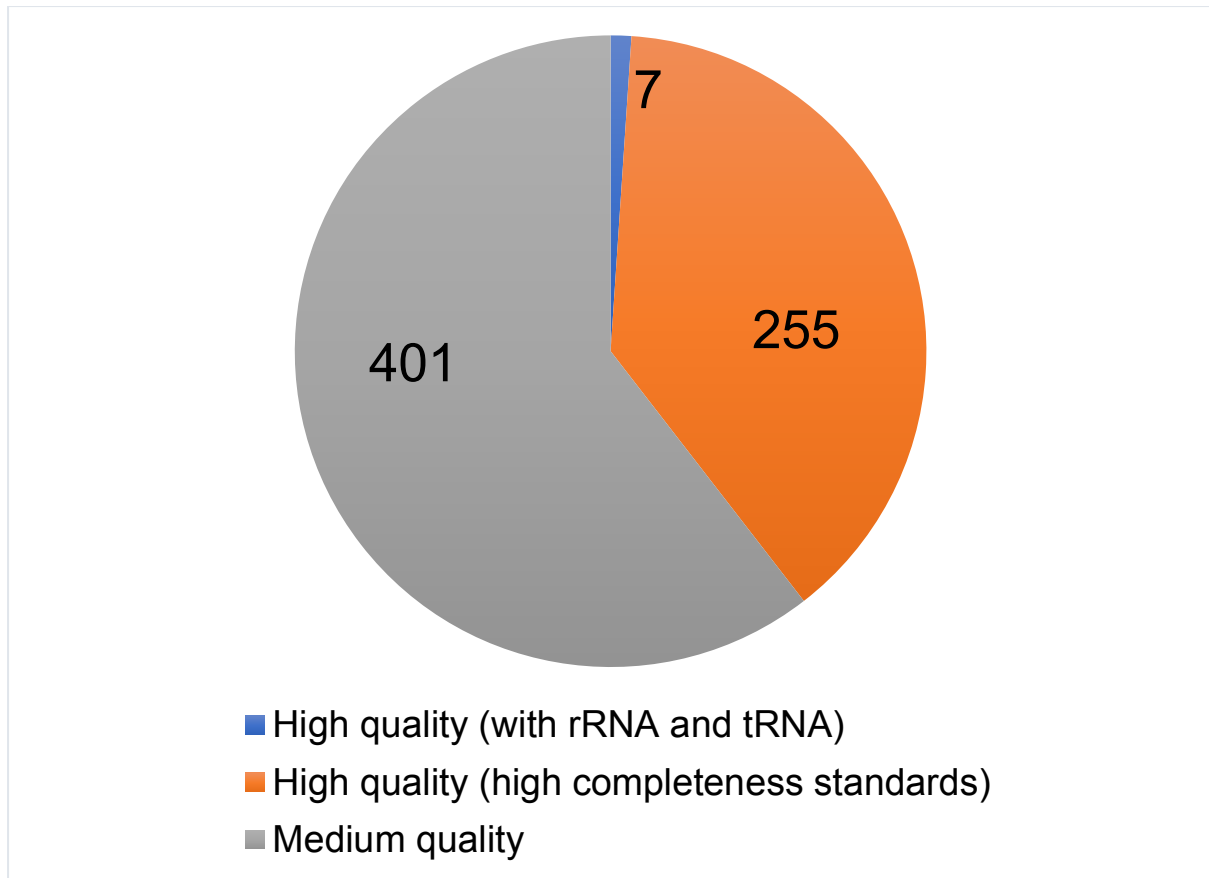


Figure S1.

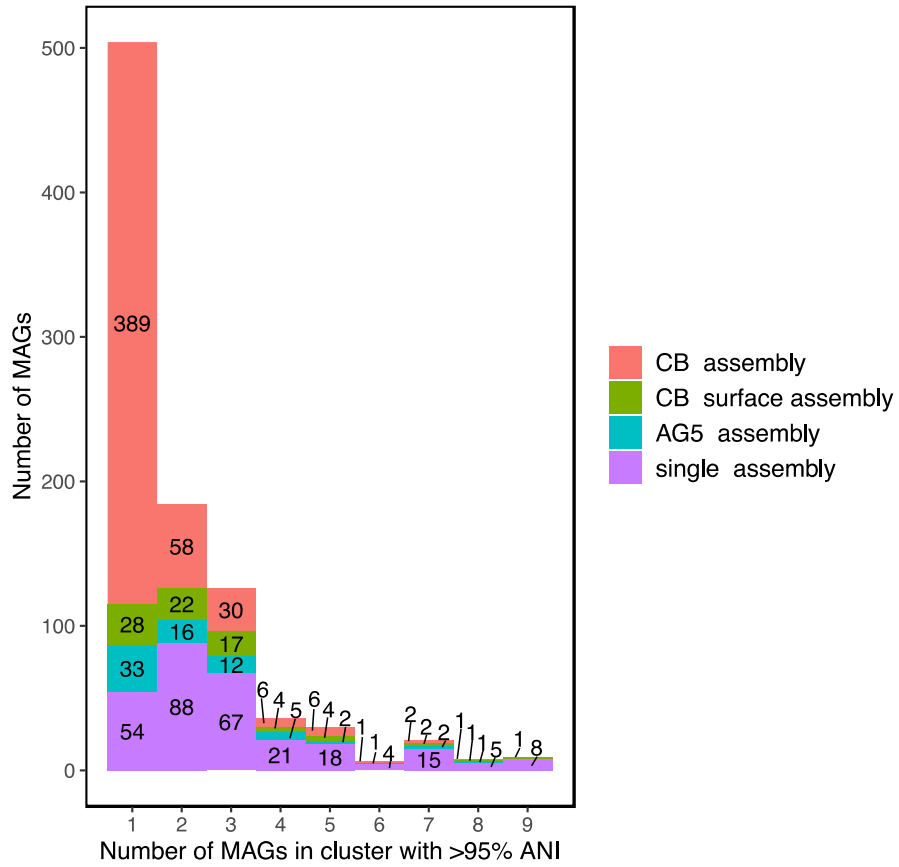


Figure S2.

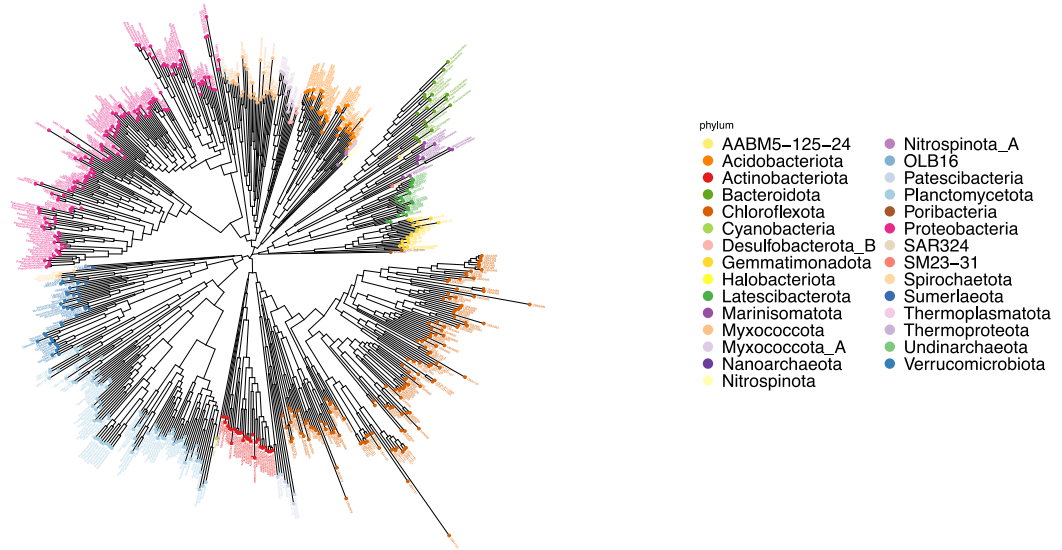


Figure S3.

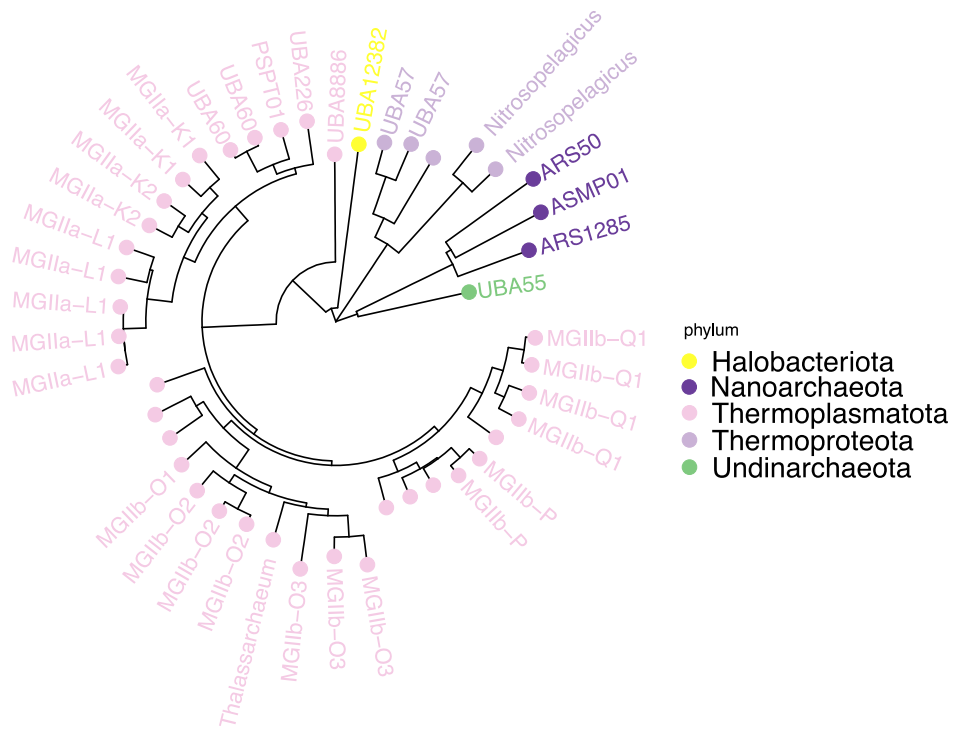


Figure S4.

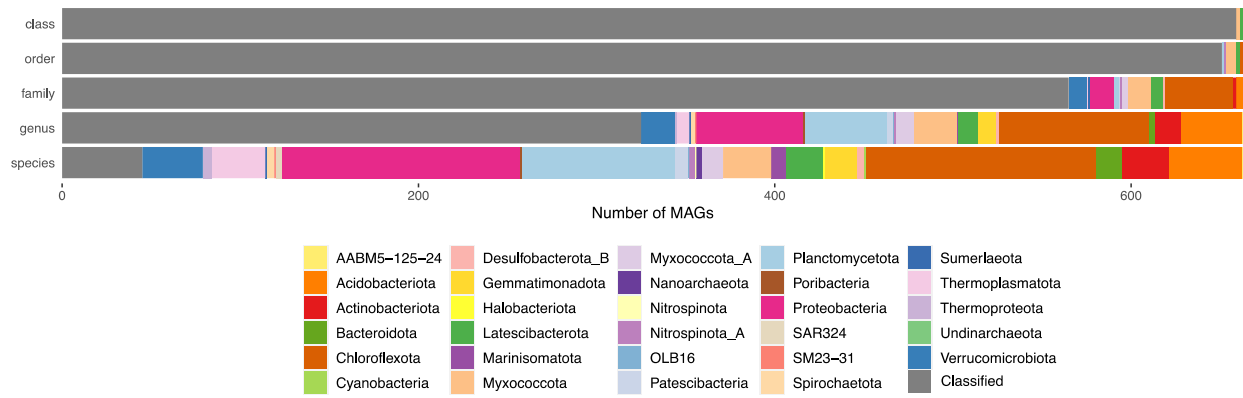


Figure S5.

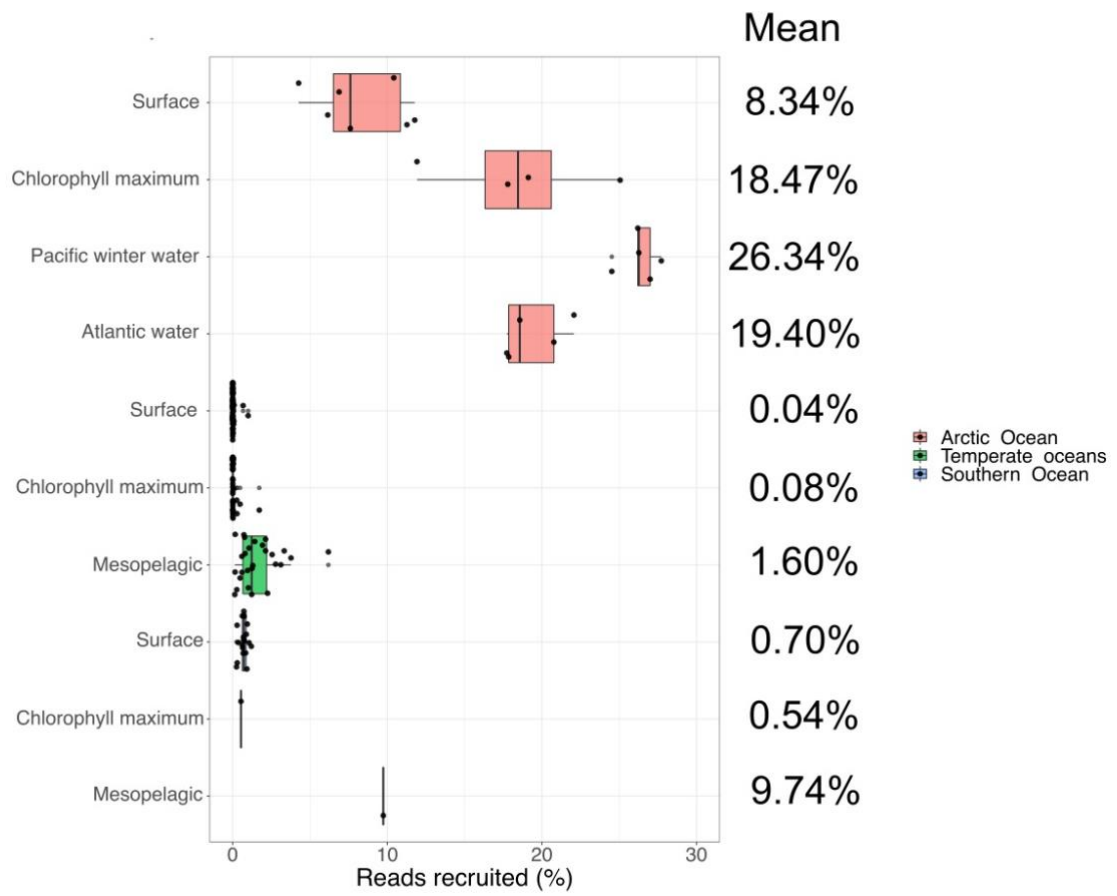


Figure S7.

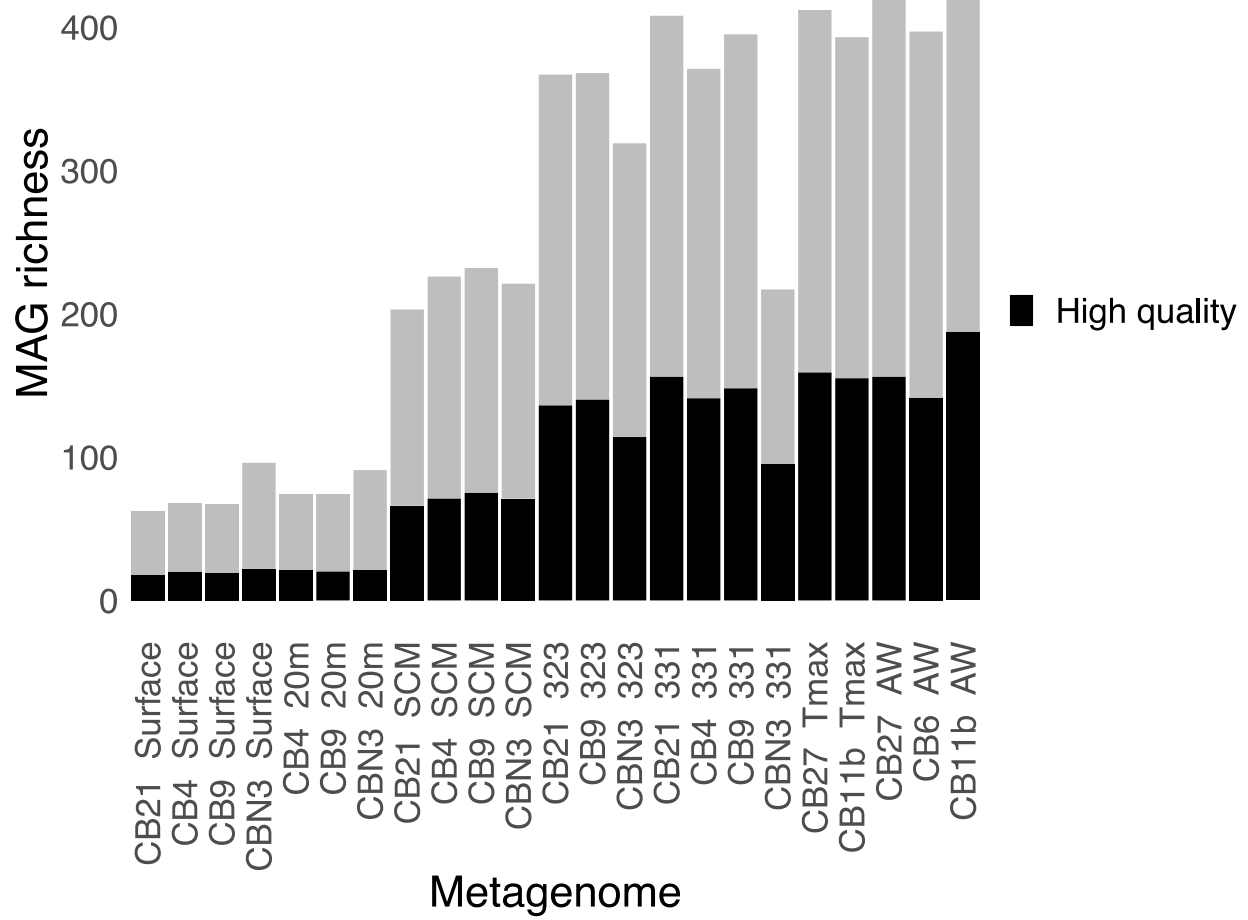


Figure S7.

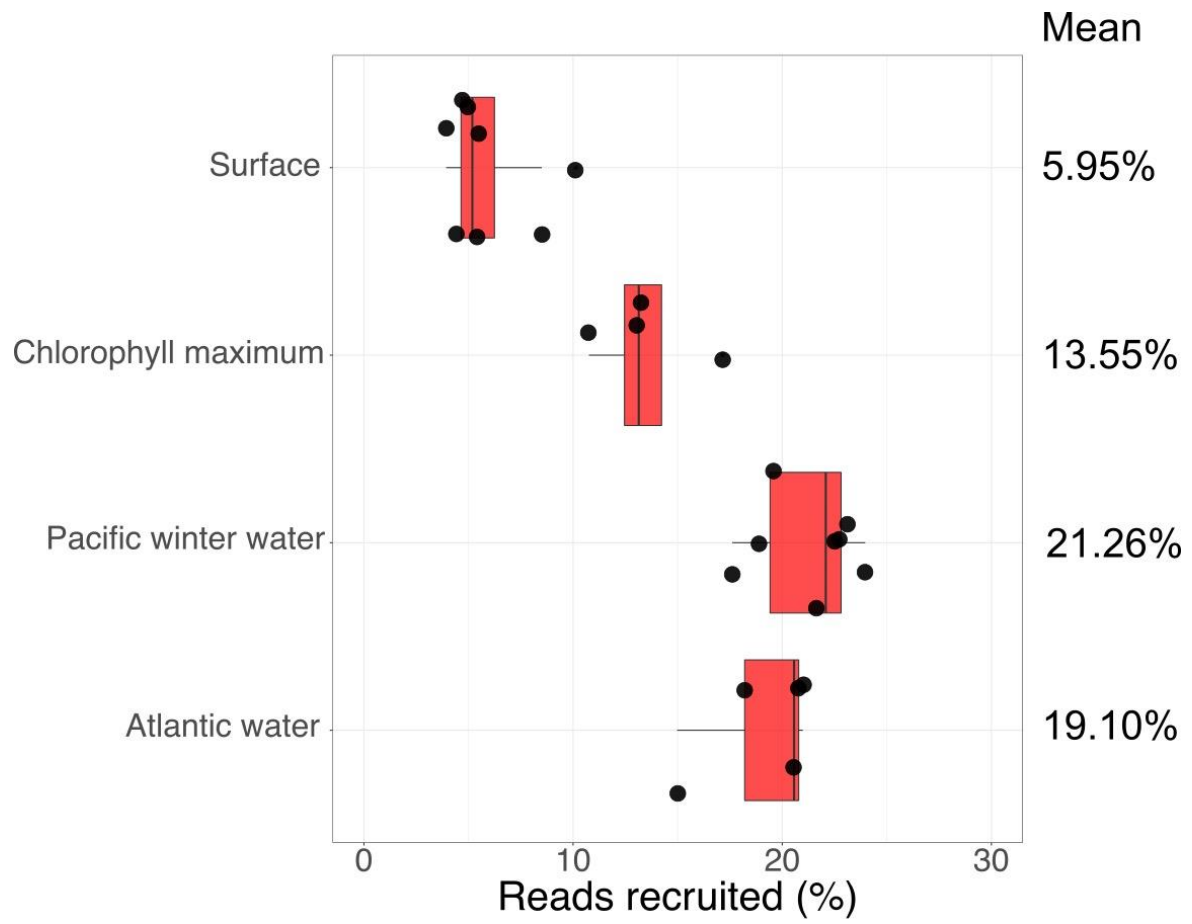


Figure S8.

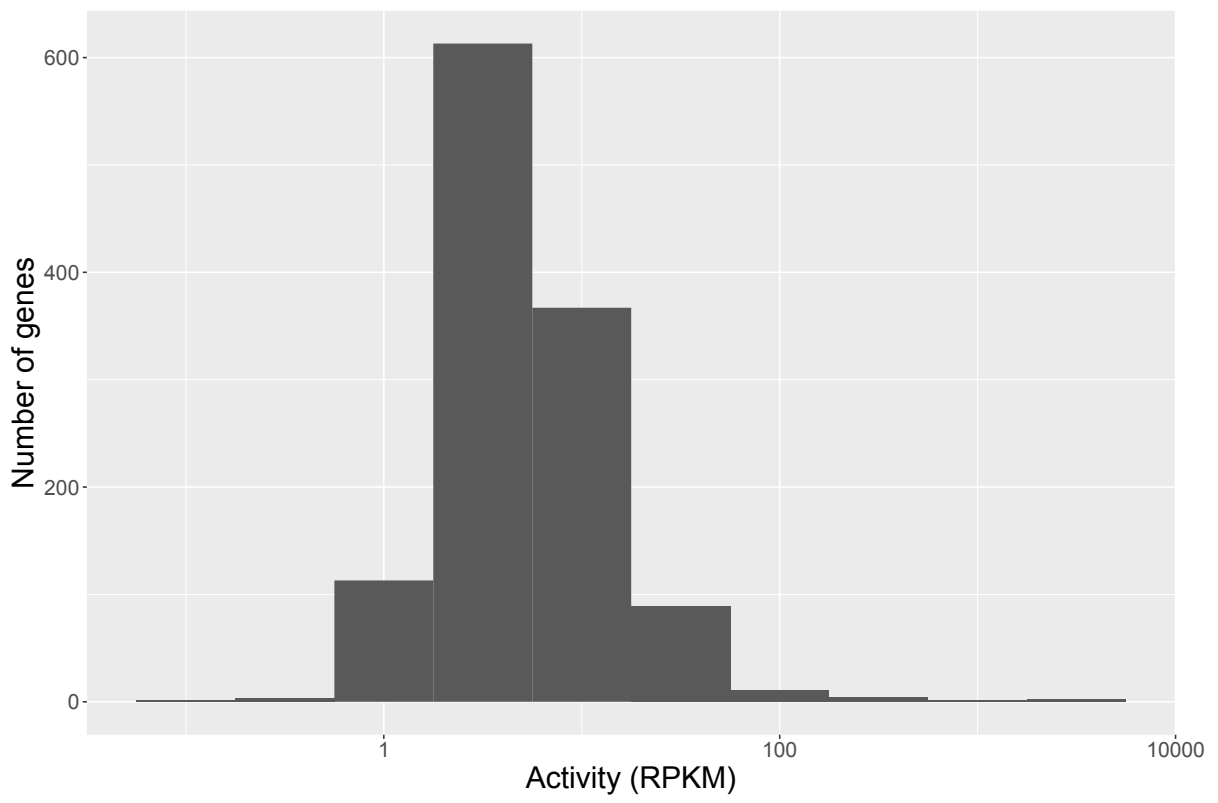


Figure S9.

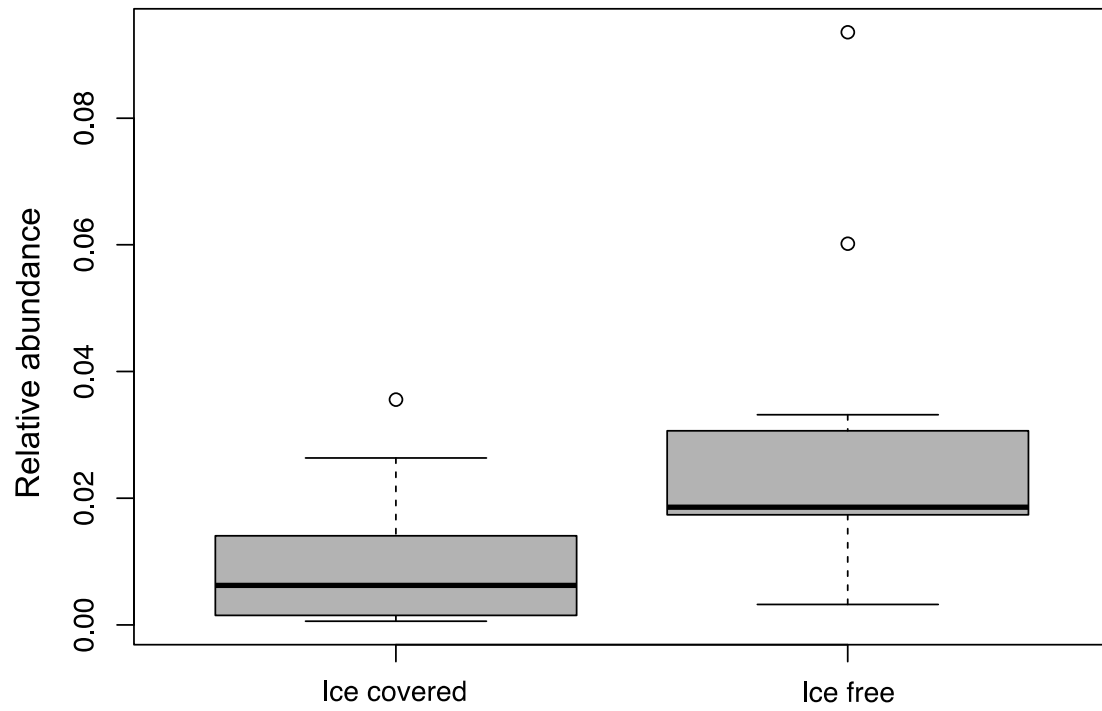


Figure S10.