nature microbiology

Article

https://doi.org/10.1038/s41564-023-01322-0

Molecular hydrogen in seawater supports growth of diverse marine bacteria

Received: 26 August 2022

Accepted: 5 January 2023

Published online: 6 February 2023

Check for updates

Rachael Lappan[1,10], Guy Shelley[2,10], Zahra F. Islam[1,10], Pok Man Leung[1], Scott Lockwood[3], Philipp A. Nauer[4], Thanavit Jirapanjawat[1,5], Gaofeng Ni[1], Ya-Jou Chen[1,2], Adam J. Kessler[6], Timothy J. Williams[7], Ricardo Cavicchioli[7], Federico Baltar[8], Perran L. M. Cook[4], Sergio E. Morales[3] & Chris Greening[1,2,5,9]

Molecular hydrogen (H_2) is an abundant and readily accessible energy source in marine systems, but it remains unknown whether marine microbial communities consume this gas. Here we use a suite of approaches to show that marine bacteria consume H_2 to support growth. Genes for H_2-uptake hydrogenases are prevalent in global ocean metagenomes, highly expressed in metatranscriptomes and found across eight bacterial phyla. Capacity for H_2 oxidation increases with depth and decreases with oxygen concentration, suggesting that H_2 is important in environments with low primary production. Biogeochemical measurements of tropical, temperate and subantarctic waters, and axenic cultures show that marine microbes consume H_2 supplied at environmentally relevant concentrations, yielding enough cell-specific power to support growth in bacteria with low energy requirements. Conversely, our results indicate that oxidation of carbon monoxide (CO) primarily supports survival. Altogether, H_2 is a notable energy source for marine bacteria and may influence oceanic ecology and biogeochemistry.

Over the past decade, trace gases have emerged as major energy sources supporting the growth and survival of aerobic bacteria in terrestrial ecosystems. Two trace gases, molecular hydrogen (H_2) and carbon monoxide (CO), are particularly dependable substrates given their ubiquity, diffusibility and energy yields[1]. Bacteria oxidize these gases, including below atmospheric concentrations, using group 1 and 2 [NiFe]-hydrogenases and form I carbon monoxide dehydrogenases linked to aerobic respiratory chains[2–6]. Trace gas oxidation enables diverse organoheterotrophic bacteria to survive long-term starvation of their preferred organic growth substrates[7,8]. In addition, various microorganisms can grow mixotrophically by co-oxidizing trace gases with other organic or inorganic energy sources[7,9,10]. Thus far, bacteria from eight different phyla have been experimentally shown to consume H_2 and CO at ambient levels[1], with numerous other bacteria encoding the determinants of this process[6,11]. At the ecosystem scale, most bacteria in soil ecosystems harbour genes for trace gas oxidation and cell-specific rates of trace gas oxidation are theoretically sufficient to sustain their survival[12,13]. However, since most of these studies have focused on soil environments or isolates, the wider significance of trace gas oxidation remains largely unexplored.

[1]Department of Microbiology, Biomedicine Discovery Institute, Monash University, Melbourne, Victoria, Australia. [2]School of Biological Sciences, Monash University, Melbourne, Victoria, Australia. [3]Department of Microbiology and Immunology, University of Otago, Dunedin, New Zealand. [4]School of Chemistry, Monash University, Melbourne, Victoria, Australia. [5]Centre to Impact AMR, Monash University, Melbourne, Victoria, Australia. [6]School of Earth, Atmosphere, and Environment, Monash University, Melbourne, Victoria, Australia. [7]School of Biotechnology and Biomolecular Sciences, UNSW Sydney, Sydney, New South Wales, Australia. [8]Fungal and Biogeochemical Oceanography, Department of Functional and Evolutionary Ecology, University of Vienna, Vienna, Austria. [9]SAEF: Securing Antarctica's Environmental Future, Monash University, Melbourne, Victoria, Australia. [10]These authors contributed equally: Rachael Lappan, Guy Shelley, Zahra F. Islam. e-mail: chris.greening@monash.edu

https://doi.org/10.1038/s41564-023-01322-0

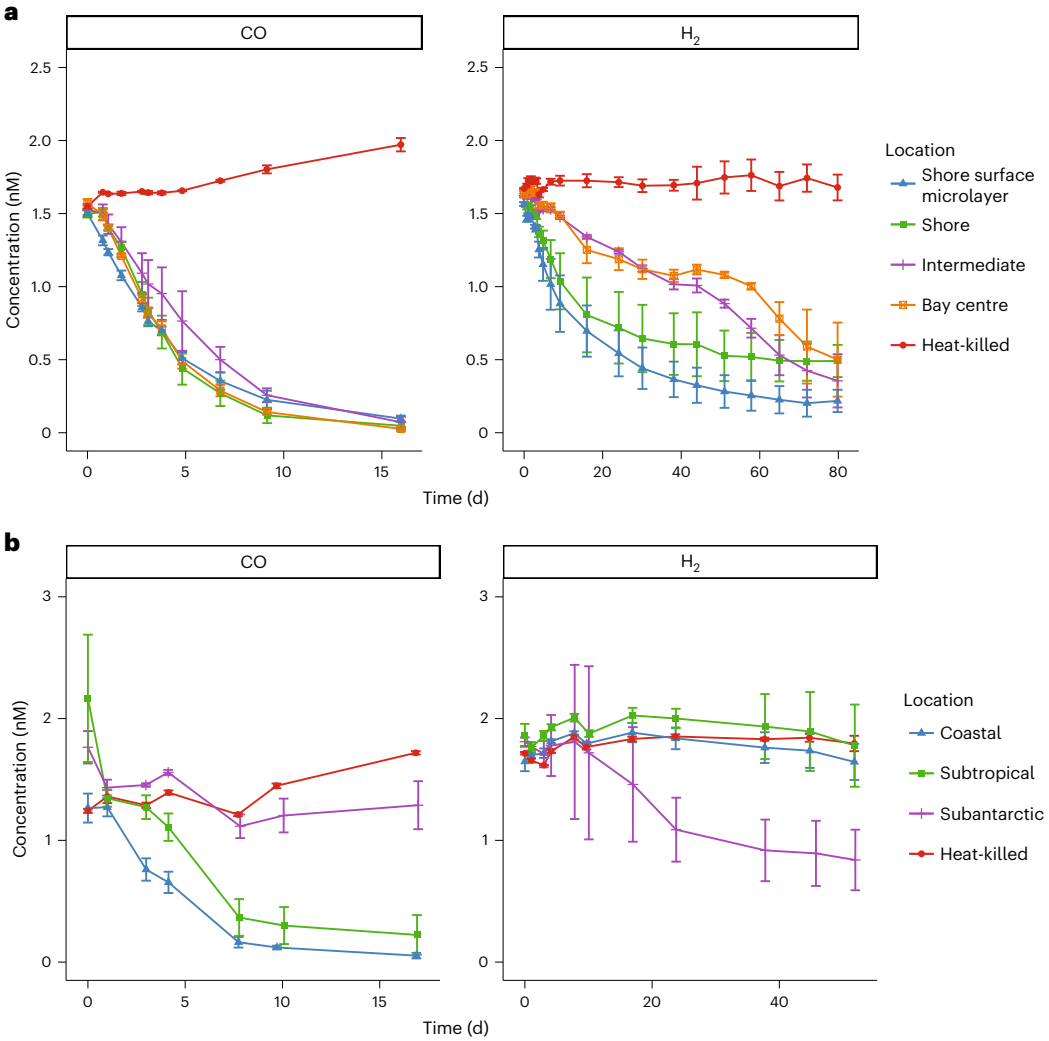


Fig. 1 | Ex situ oxidation of CO and H_2 by seawater communities. a,**b**, Results are shown for four samples in a transect at Port Phillip Bay, Victoria, Australia (**a**) and eight samples in the Munida transect off the coast of Otago, New Zealand (**b**). Each 120 ml sealed serum vial contained 60 ml of native seawater samples incubated in a 60 ml ambient-air headspace supplemented with ~2.5 ppmv H_2 or CO. At each timepoint, the mixing ratio of each gas in the headspace of each vial was measured on a gas chromatograph and converted to dissolved gas concentrations (nM). Data are presented as mean ± s.e.m. of three biologically independent samples.

Trace gases may be important energy sources for oceanic bacteria since they are generally available at elevated concentrations relative to the atmosphere, in contrast to most soils[1]. Surface layers of the world's oceans are generally supersaturated with H_2 and CO, typically by 2- to 5-fold (up to 15-fold) and 20- to 200-fold (up to 2,000-fold) relative to the atmosphere, respectively[14–17]. As a result, oceans contribute to net atmospheric emissions of these gases[18,19]. CO is mainly produced through photochemical oxidation of dissolved organic matter[20], whereas H_2 is primarily produced by cyanobacterial nitrogen fixation[21]. High concentrations of H_2 are also produced during fermentation in hypoxic sediments, and these high concentrations can diffuse into the overlying water column, especially in coastal waters[22]. For unresolved reasons, the distributions of these gases vary with latitude and exhibit opposite trends: while dissolved CO is highly supersaturated in polar waters, H_2 is often undersaturated[23–28]. These variations probably reflect differences in the relative rates of trace gas production and consumption in different climates.

Oceanic microbial communities have long been known to consume CO, although their capacity to use H_2 has not been systematically evaluated[29]. Approximately a quarter of bacterial cells in oceanic surface waters encode CO dehydrogenases in surface waters and these span a wide range of taxa, including the globally abundant family Rhodobacteraceae (previously known as the marine *Roseobacter* clade)[6,30–33]. Building on observations made for soil communities, CO oxidation potentially enhances the long-term survival of marine bacteria during periods of organic carbon starvation[6]; consistently, culture-based studies indicate that CO does not influence growth of marine isolates, but production of the enzymes responsible is strongly upregulated during starvation[34–37]. While aerobic and anaerobic oxidation of H_2 has been extensively described in benthic and hydrothermal vent communities[38–42], so far no studies have shown whether pelagic bacterial communities can use this gas. Several surveys have detected potential H_2-oxidizing hydrogenases in seawater samples and isolates[6,11,40,43]. Although Cyanobacteria are well-reported to oxidize H_2, including marine isolates such as *Trichodesmium*, this process is thought to be limited to the endogenous recycling of H_2 produced by the nitrogenase reaction[44,45].

In this study, we addressed these knowledge gaps by investigating the processes, distribution, mediators and potential roles of H_2 and CO oxidation by marine bacteria. To do so, we performed side-by-side

https://doi.org/10.1038/s41564-023-01322-0

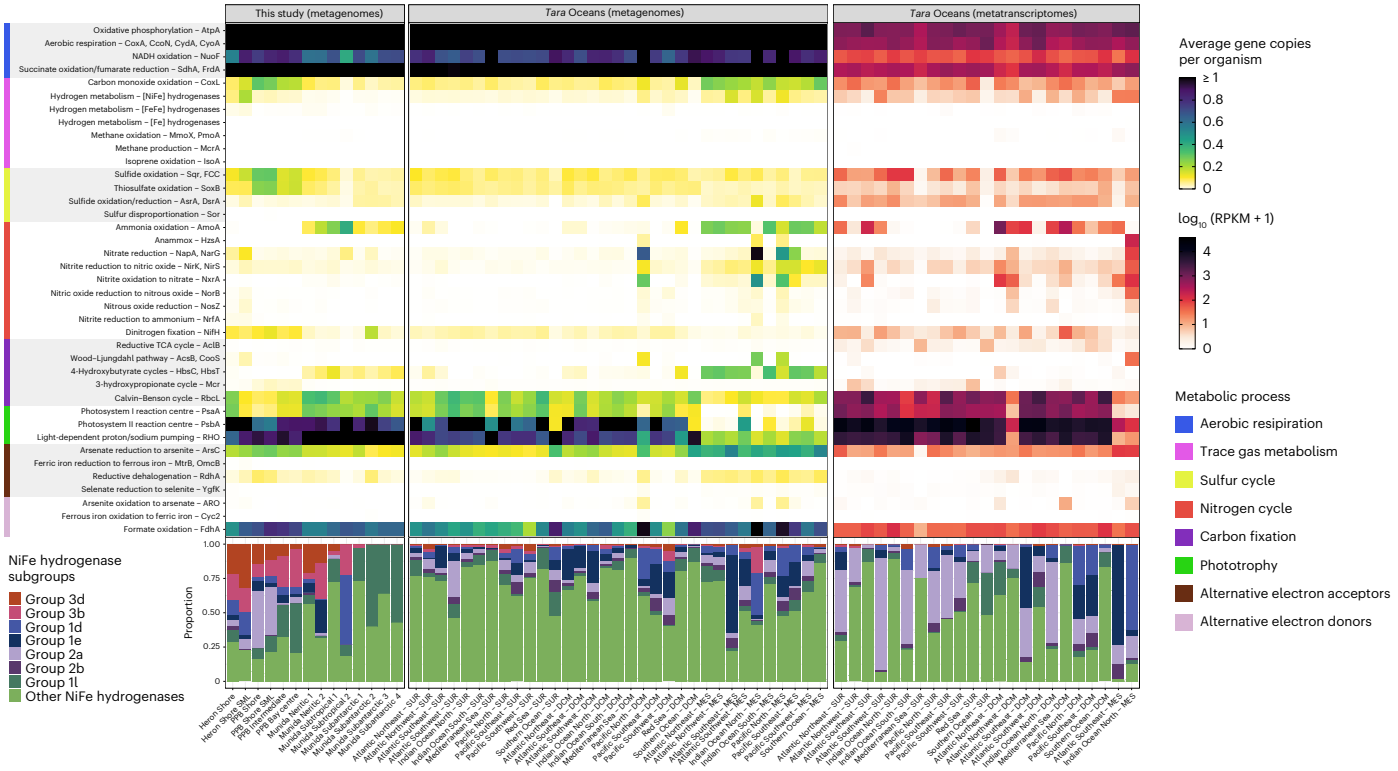


Fig. 2 | Abundance of metabolic genes encoded by marine communities. The abundance of metabolic marker genes is shown on the basis of the metagenomic short reads across the seawater sampled from the three study sites (left; $n = 14$), metagenomic short reads from the *Tara* Oceans dataset (middle; $n = 213$; replicates averaged) and metatranscriptomic short reads from the *Tara* Oceans dataset (right; $n = 89$; replicates averaged). Homology-based searches were used to calculate the relative abundance of marker genes as average gene copies per organism for the metagenomes (abundance relative to a set of universal single-copy marker genes; equivalent to the estimated proportion of the community encoding a given gene as a single copy) and RPKM for the metatranscriptomes. Where multiple marker genes are listed, values are summed. The bottom panels show the hydrogenase subgroups present in each sample. SUR, surface; DCM, deep chlorophyll maximum; MES, mesopelagic ocean layers.

metagenomic and biogeochemical profiling of 14 samples collected from a temperate oceanic transect, a temperate coastal transect and a tropical island, in addition to analysing the global *Tara* Oceans metagenomes and metatranscriptomes[46]. We also tested the capacity of three axenic marine bacterial isolates to aerobically consume atmospheric H_2. Altogether, we provide definitive ecosystem-scale and culture-based evidence that H_2 is an overlooked key energy source supporting growth of marine bacteria.

Results

Marine microbes consume H_2 slowly and CO rapidly

We measured in situ concentrations and ex situ oxidation rates of H_2 and CO in 14 surface seawater samples. The samples were collected from three locations (Supplementary Fig. 1): an oceanic transect spanning neritic, subtropical and subantarctic front waters (Munida transect off New Zealand coast; $n = 8$; Supplementary Fig. 2); a temperate urban bay (Port Phillip Bay, Australia; $n = 4$); and a tropical coral cay (Heron Island, Australia; $n = 2$). In line with global trends at these latitudes, both gases were supersaturated relative to the atmosphere in all samples. H_2 was supersaturated by 5.4-, 4.8- and 12.4-fold respectively in the oceanic transect (2.0 ± 1.2 nM), the temperate bay (1.8 ± 0.26 nM) and the tropical island (4.6 ± 0.3 nM). CO was moderately supersaturated in the oceanic transect (5.2-fold; 0.36 nM ± 0.07 nM), but highly oversaturated in both the temperate bay (123-fold; 8.5 ± 1.7 nM) and tropical island (118-fold; 8.2 ± 0.93 nM).

Microbial oxidation of trace gases was detected in all but one of the collected samples during ex situ incubations (Fig. 1). For the temperate bay, H_2 and CO were consumed in water samples collected from the shore, intermediary zone and bay centre (Fig. 1a). Based on in situ gas concentrations, bulk oxidation rates of CO were 18-fold faster than H_2 ($P < 0.0001$) (Supplementary Table 1). Bulk oxidation rates did not significantly differ between the surface microlayer (that is, the 1 mm interface between the atmosphere and ocean) and underlying waters. H_2 and CO oxidation was also evident in surface microlayer and underlying seawater samples collected from the tropical island (Supplementary Fig. 3). We similarly observed rapid CO and slower H_2 consumption across the multi-front Munida oceanic transect, although unexpectedly, these activities were mutually exclusive. Net CO oxidation occurred throughout the coastal and subtropical waters but was negligible in subantarctic waters. Conversely, net H_2 oxidation only occurred in the subantarctic waters (Fig. 1b). These divergent oxidation rates in water masses with contrasting physicochemical conditions may help explain the contrasting concentrations of H_2 and CO in global seawater[23–28], although wider sampling and in situ assays would be required to confirm this. It should be noted that these measurements probably underestimate rates and overestimate thresholds of H_2 oxidation since there will still be underlying endogenous production of H_2, primarily through nitrogen fixation, during the incubations. Nevertheless, they provide the first empirical report of H_2 oxidation in marine water columns.

Marine microbes express enzymes for CO and H_2 oxidation

To better understand the basis of these activities, we sequenced metagenomes of the 14 samples (Supplementary Tables 2 and 3) and used homology-based searches to determine the abundance of 50 metabolic marker genes in the metagenomic reads (Supplementary Table 3) and assemblies (Supplementary Table 4). In common with other surface seawater communities[47], analysis of community composition (Supplementary Fig. 4) and metabolic genes (Fig. 2) suggests that most

bacteria present are capable of aerobic respiration, organoheterotrophy and phototrophy via energy-converting rhodopsins. Capacity for aerobic CO oxidation was moderate: approximately 12% of bacterial and archaeal cells encoded the *coxL* gene (encoding the catalytic subunit of the form I CO dehydrogenase), although relative abundance decreased from an average of 25% in the temperate bay where CO oxidation was highly active to 5.1% in subantarctic waters (Fig. 2) where CO oxidation was negligible (Fig. 1). Diverse hydrogenases were also encoded by the community, including subgroups known to support hydrogenotrophic respiration, hydrogenotrophic carbon fixation, hydrogenogenic fermentation and H_2 sensing (Supplementary Table 3). Group 1d, 1l and 2a [NiFe]-hydrogenases (herein aerobic H_2-uptake hydrogenases), which enable cells to input electrons from H_2 into the aerobic respiratory chain[4,9,48,49], were by far the most abundant among the H_2-oxidizing enzymes (Fig. 2). Encoded by 1.0% of marine bacteria on average, the abundance of these hydrogenase subgroups was highest in the tropical island samples (average 3.5%) and declined to 0.11% in the neritic and subtropical samples from the oceanic transect (Fig. 2), in line with the contrasting H_2 oxidation rates between these samples (Fig. 1 and Supplementary Fig. 3). The dominant hydrogenase subgroups varied between the samples, namely group 1d in the tropical island samples, group 2a in the temperate shore and microlayer samples and group 1l in the subantarctic samples (Fig. 2). Relative abundance of H_2- and CO-oxidizing bacteria strongly predicted oxidation rates of each gas (R^2 of 0.55 and 0.88; P values of 0.0059 and <0.0001, respectively) (Supplementary Fig. 5), although it is likely that repression of gene expression contributes to the negligible activities of some samples.

To test whether these observations were globally representative, we determined the distribution and expression of the genes for H_2 and CO oxidation in the *Tara* Oceans dataset[47,50]. Similarly to our metagenomes, aerobic H_2-uptake hydrogenases were encoded by an average of 0.8% bacteria and archaea across the 213 *Tara* Oceans metagenomes, whereas form I CO dehydrogenases were encoded by 10.4%. These genes were observed in samples spanning all four oceans, as well as the Red Sea and Mediterranean Sea (Fig. 2). Despite their relatively low abundance based on the metagenomes, hydrogenase transcripts were highly numerous in the metatranscriptomes, with comparable levels to nitrogenase (*nifH*) transcripts (Fig. 2 and Supplementary Table 3). Expression ratios (average RNA:DNA ratios) of the aerobic H_2-uptake hydrogenases were high, that is, 2.2, 1.1 and 12.9 for the group 1d, 1l and 2a [NiFe]-hydrogenases, respectively (Supplementary Table 3); of the marker genes surveyed, only the determinants of phototrophy (*psaA*, *psbA*, energy-converting rhodopsins), nitrification (*amoA*, *nxrA*) and CO_2 fixation (*rbcL*) were expressed at higher ratios than the group 2a [NiFe]-hydrogenases. In contrast, expression levels were relatively low for the CO dehydrogenase (0.9), as well as the hydrogenases responsible for hydrogenotrophic carbon fixation, hydrogenogenic fermentation and H_2 sensing (average RNA/DNA <1 in all cases) (Supplementary Table 3). Together with the biogeochemical measurements (Fig. 1), these findings suggest that H_2-oxidizing bacteria can be highly active in seawater despite their relatively low abundance.

Eleven marine bacterial phyla encode H_2-oxidizing enzymes

We subsequently determined the distribution of the metabolic marker genes in 110 metagenome-assembled genomes (MAGs) constructed from the local dataset and 1,888 previously reported MAGs (Fig. 3a and Supplementary Fig. 6) from the *Tara* Oceans dataset (Fig. 3a). The three lineages of aerobic H_2-uptake hydrogenases were phylogenetically widespread, encoded by 75 (4.0%) of the bacterial MAGs, spanning 9 phyla and 26 orders, whereas CO dehydrogenases had a somewhat narrower distribution, that is, 70 (3.5%) MAGs, 6 phyla and 14 orders (Supplementary Table 5). Aerobic H_2-uptake hydrogenases and CO dehydrogenases were both encoded by MAGs within the Proteobacteria, Bacteroidota, Actinobacteriota, Chloroflexota, Myxococcota and candidate phylum SAR324, and hydrogenases were also present in MAGs from the Cyanobacteria, Planctomycetota and Eremiobacterota (Fig. 3a). Phylogenetic trees depict the evolutionary history and taxonomic distributions of the catalytic subunits of the H_2-oxidizing group 1 and 2 [NiFe]-hydrogenases (Fig. 3b and Supplementary Fig. 7), bidirectional group 3 and 4 [NiFe]-hydrogenases (Supplementary Fig. 8) and CO dehydrogenase (Supplementary Fig. 9).

Integrating genomic information with the wider literature, it is likely that H_2 and CO oxidation support a myriad of lifestyles in marine ecosystems. The group 1d [NiFe]-hydrogenase was typically co-encoded with both ribulose 1,5-bisphosphate carboxylase/oxygenase (RuBisCO) and the sensory group 2b [NiFe]-hydrogenase in the MAGs of multiple Rhodobacteraceae, Alteromonadaceae and other Proteobacteria (Fig. 3b and Supplementary Table 5); this suggests that this enzyme supports hydrogenotrophic growth in H_2-enriched waters, in line with the previously described roles of these hydrogenases in culture-based studies[11,38,51]. The group 1l [NiFe]-hydrogenase, recently shown to support persistence of a Bacteroidota isolate from Antarctic saline soils[4], was encoded by predicted organoheterotrophs from the Bacteroidota, SAR324 and, on the basis of cultured isolates, Proteobacteria (Fig. 3b and Supplementary Table 5). Group 2a [NiFe]-hydrogenases, known to support mixotrophic growth of diverse bacteria[9], were more phylogenetically diverse and taxonomically widespread; they were distributed in the MAGs of both predicted chemoorganoheterotrophs (Bacteroidota, Myxococcota, Proteobacteria) and photolithoautotrophs (Cyanobacteria) (Fig. 3b and Supplementary Table 5). CO dehydrogenases were mostly affiliated with Rhodobacteraceae and were also encoded by multiple MAGs from the classes Nanopelagicales S36-B12, Puniceispirillaceae, SAR324 NAC60-12 and Ilumatobacteraceae (Supplementary Fig. 9). Of *coxL*-encoding MAGs, 63% also encoded the genes for energy-converting rhodopsins or photosystem II, indicating that they can harvest energy concurrently or alternately from both CO and light, in support of previous culture-based findings[37]. While most of these MAGs are predicted to be obligate heterotrophs, 7% also encoded RuBisCO and hence are theoretically capable of carboxydotrophic growth (Supplementary Table 5). These findings support previous inferences that habitat generalists in marine waters benefit from metabolic flexibility, including consuming dissolved CO as a supplemental energy source[31,52].

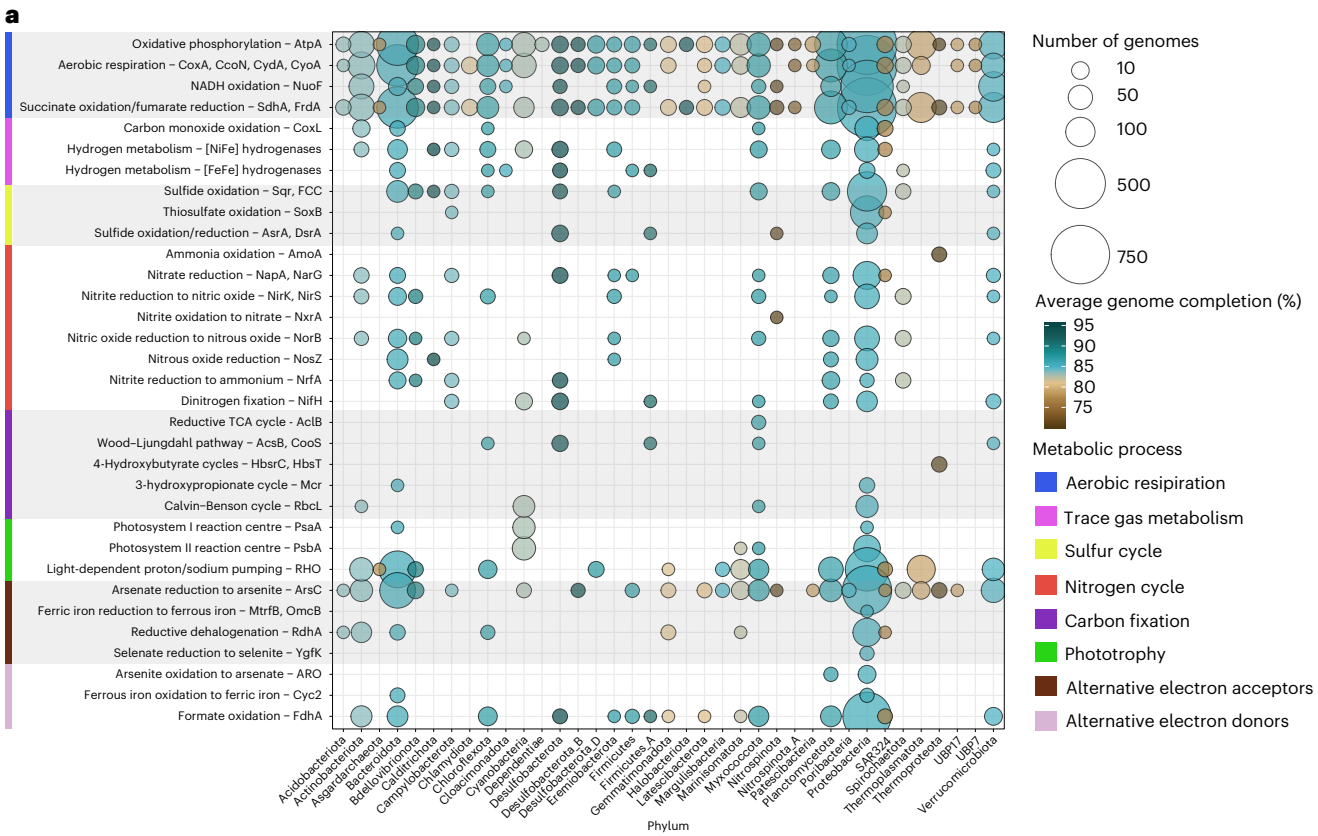
H_2 could support growth and survival of marine bacteria

We used two thermodynamic modelling calculations to estimate to what extent the measured rates of H_2 and CO oxidation sustain cellular growth or survival. First, assuming a median maintenance energy of 1.9×10^{-15} Watts (W) per cell based on measurements of mostly

Fig. 3 | Distribution of metabolic genes in marine bacteria. a, Bubble plot showing metabolic potential of the metagenome-assembled genomes constructed from the three study sites (110 MAGs) and previously reported for the *Tara* Oceans dataset (1,877 MAGs). MAGs are summarized at phylum level, with the size of the circle corresponding to the number of genomes in that phylum with a given gene, and the colour reflecting the percentage of genome completeness. Marker genes that were not detected in any MAG are omitted. **b**, Maximum-likelihood phylogenetic tree of the catalytic subunit of the group 1 and 2 [NiFe]-hydrogenases. Hydrogenase sequences retrieved from the new MAGs (coloured green) and *Tara* MAGs (coloured blue) are shown alongside representative reference sequences (coloured yellow), including the three cultured marine bacteria (names in red). Evolutionary history was inferred by using the JTT matrix-based model, the tree was bootstrapped using 50 replicates and the tree is rooted using the outgroup group 4a [NiFe]-hydrogenase sequence. The tree includes hydrogenase subgroups implicated in aerobic respiration (groups 1d, 1f, 1l, 2a), anaerobic respiration (groups 1a, 1b, 1c, 1e) and H_2 sensing (groups 2b and 2c).

https://doi.org/10.1038/s41564-023-01322-0

copiotrophic isolates[53], the measured oxidation rates would theoretically sustain an average of 2.0×10^7 H_2-oxidizing cells (range 1.4×10^6 to 8.3×10^7) and 6.1×10^7 CO-oxidizing cells (range 2.1×10^6 to 1.5×10^8) per litre at in situ dissolved gas concentrations (Supplementary Table 1). Second, we calculated the amount of power (that is, W per cell) generated on the basis of the observed rates of trace gas oxidation (Fig. 1 and Supplementary Table 1) and predicted number of trace gas oxidizers (Fig. 2 and Supplementary Table 1) in the sampled waters, with this



https://doi.org/10.1038/s41564-023-01322-0

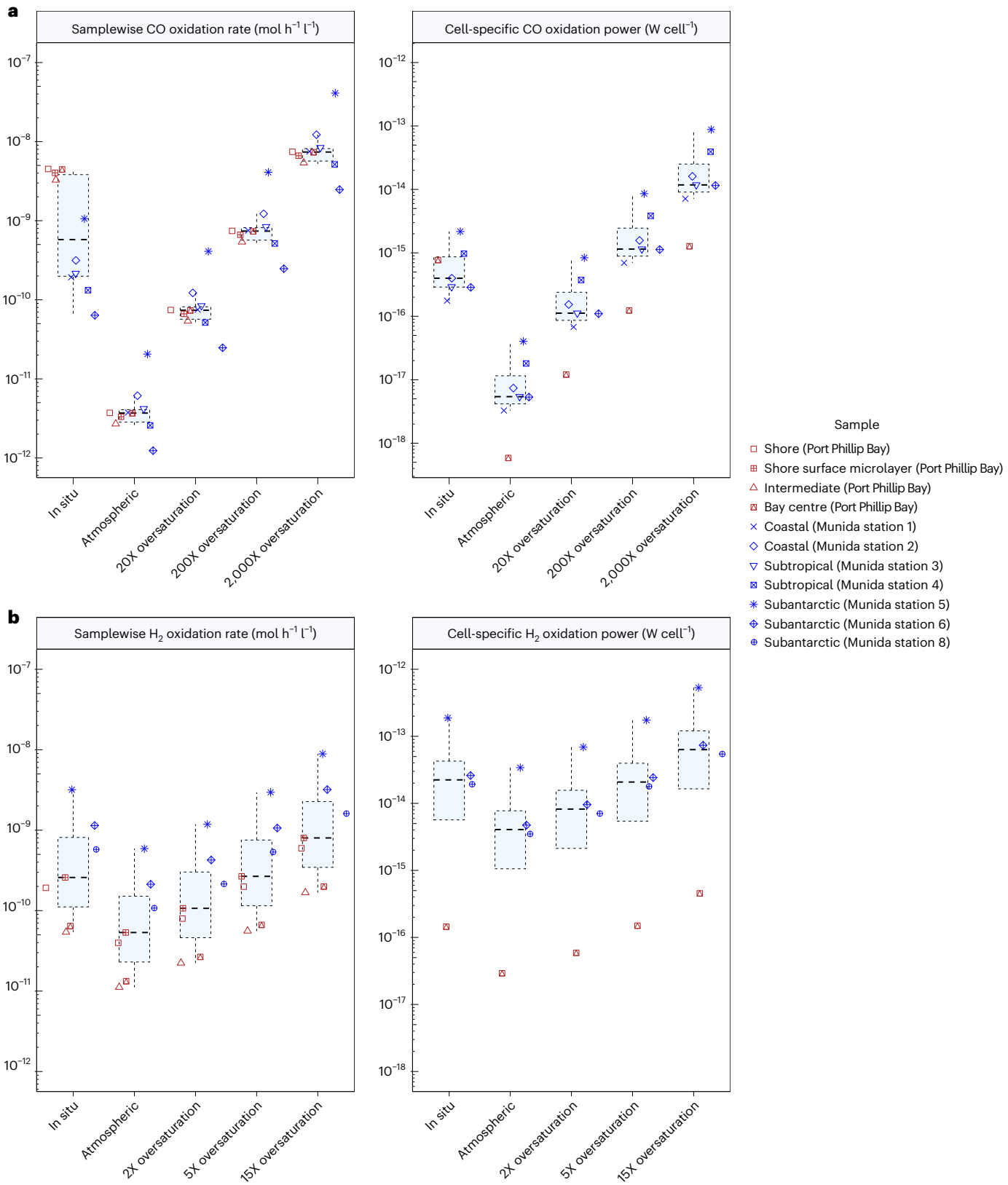


Fig. 4 | Thermodynamic modelling of H_2 and CO oxidation by marine bacteria. a,**b**, The results show the bulk oxidation rates (left) and power yields per cell (right) for CO oxidation (**a**) ($n = 10$ (rate) and $n = 7$ (power) biologically independent samples) and H_2 oxidation (**b**) ($n = 7$ (rate) and $n = 4$ (power) biologically independent samples). This analysis was only performed for samples where trace gas oxidation was measurable and cell-specific power was only calculated for samples where prokaryotic cell counts are available. Rates and power are shown on the basis of CO and H_2 concentrations at a range of environmentally relevant concentrations. Centre values show medians, boxes show upper and lower quartiles, and whiskers show maximum (upper quartile plus 1.5 times interquartile range) and minimum (lower quartile minus 1.5 times interquartile range) values.

https://doi.org/10.1038/s41564-023-01322-0

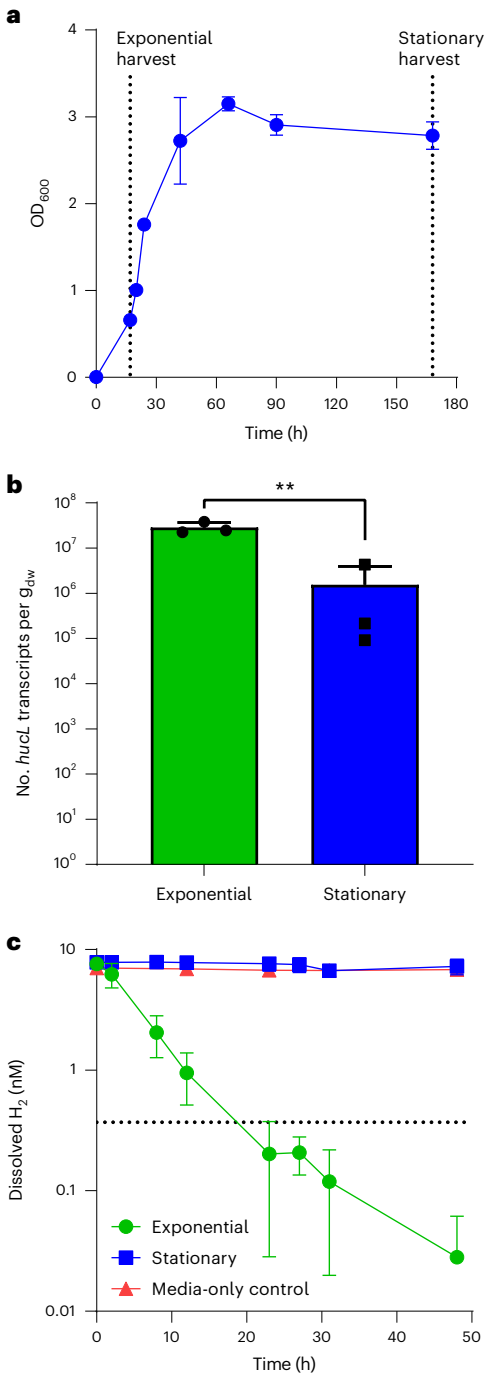


Fig. 5 | Hydrogenase expression and activity of *Sphingopyxis alaskensis*. **a**, Growth curve of *S. alaskensis* grown on Difco 2216 Marine Broth. Cultures were tested for gas consumption and collected for RT–qPCR in exponential phase (17 h, OD_{600} = 0.66) and stationary phase (168 h, 4 d post OD_{max}). Data are presented as mean ± s.d., n = 3 biologically independent samples. **b**, Number of transcripts of the group 2a [NiFe]-hydrogenase large subunit gene (*hucL*; locus Sala_3198) as measured by RT–qPCR in exponential and stationary phase cultures of *S. alaskensis*. Mean ± s.d. of three biological replicates (averaged from two technical duplicates) per condition. The comparison is statistically significant based on an unpaired two-sided t-test (**P = 0.0062). **c**, H_2 oxidation by exponential and stationary phase cultures of *S. alaskensis*. Mean ± s.d. of three biological replicates, with media-only vials monitored as negative controls. Dotted line shows the atmospheric concentration of hydrogen (0.53 ppmv).

analysis being limited to the samples where oxidation was observed and reliable cell counts are available. On average, oxidation of the measured in situ concentrations of CO and H_2 yields 7.2×10^{-16} W and 5.8×10^{-14} W per cell (Fig. 4). Together, these analyses suggest that the rates of CO oxidation are sufficient to sustain the survival, but not growth, of the numerous bacteria predicted to be capable of using this gas; this supports previous inferences that CO dehydrogenase primarily supports persistence in organoheterotrophic bacteria[6].

In contrast, marine H_2 oxidizers gain much power by oxidizing a relatively exclusive substrate at rapid cell-specific rates probably sufficient to support growth. The cell-specific power generated for the sample with the most active H_2 oxidizers (5.4×10^{-13} W; from the first subantarctic station) is within the range reported for cellular metabolic rates of bacterial isolates during growth (median: 2.6×10^{-14} W; range: 2.8×10^{-17} to 2.1×10^{-11} W), and is higher than that of copiotrophic marine isolates *Vibrio* sp. DW1 (3.2×10^{-14} W) and *V. anguillarum* (1.8×10^{-13} W)[53]. While estimation of cell-specific power from community data is less precise than estimates derived from axenic culture, these power per cell calculations are probably underestimates, given that they do not account for any internal cycling of trace gases, assume that all cells are equally active and do not consider relic DNA. It should also be noted that the power gained per cell will substantially increase when H_2 and CO become transiently highly elevated over space and time as depicted in Fig. 4. In combination with the genomic inferences that multiple MAGs encode hydrogenases known to support lithoautotrophic and lithoheterotrophic growth (Fig. 3), such thermodynamic modelling strongly suggests that a small proportion of bacteria in oceans can grow using H_2 as an electron donor for aerobic respiration and, in some cases, CO_2 fixation. By predominantly relying on energy derived from H_2 oxidation, marine bacteria could potentially allocate most organic carbon for biosynthesis rather than respiration, that is, adopting a predominantly lithoheterotrophic lifestyle.

A marine isolate uses atmospheric H_2 mixotrophically

To better understand the mediators and roles of marine H_2 oxidation, we investigated H_2 uptake by three heterotrophic marine isolates encoding uptake hydrogenases closely related to those in the MAGs (Fig. 3). Two strains, *Robiginitalea biformata* DSM-15991 (Flavobacteriaceae)[54] and *Marinovum algicola* FF3 (Rhodobacteraceae)[55], did not substantially consume H_2 over a 3 week period across a range of conditions despite encoding group 1l [NiFe]-hydrogenases. It is unclear whether hydrogenases have become non-functional in these fast-growing laboratory-adapted isolates or whether they are instead only active under very specific conditions. *Sphingopyxis alaskensis* RB2256 (Sphingomonadaceae)[56,57], which encodes a plasmid-borne group 2a [NiFe]-hydrogenase, aerobically consumed H_2 in a first-order kinetic process to sub-atmospheric levels (Fig. 5). Abundant in oligotrophic polar waters, *S. alaskensis* requires minimal resources to replicate since it forms extremely small cells (<0.1 μm^3) and has a streamlined genome[57–60]. Previously thought to be an obligate organoheterotroph[61], the discovery that this oligotrophic, exceptionally small bacterium (ultramicrobacterium)[62] uses an abundant reduced gas as an energy source further rationalizes its ecological success. This is presumably the first report of atmospheric H_2 oxidation by a marine bacterium.

We then determined whether *S. alaskensis* uses H_2 oxidation primarily to support mixotrophic growth or survival. Expression of its hydrogenase large subunit gene (*hucL*) was quantified by reverse transcription quantitative PCR (RT–qPCR). Under ambient conditions, this gene was expressed at significantly higher levels (P = 0.006) during aerobic growth on organic carbon sources (mid-exponential phase; av. 2.9×10^7 copies per g_{dw}) than during survival (4 d in stationary phase; av. 1.5×10^6 copies per g_{dw}; P = 0.006) (Fig. 5a,b). This expression pattern is similar to other organisms possessing a group 2a [NiFe]-hydrogenase[9] and is antithetical to that of the groups 1h and 1l [NiFe]-hydrogenases that are typically induced by starvation[1]. The activity of the hydrogenase was monitored under the same two conditions by monitoring depletion of headspace H_2 mixing ratios over time by gas chromatography. H_2 was rapidly oxidized by exponentially

https://doi.org/10.1038/s41564-023-01322-0

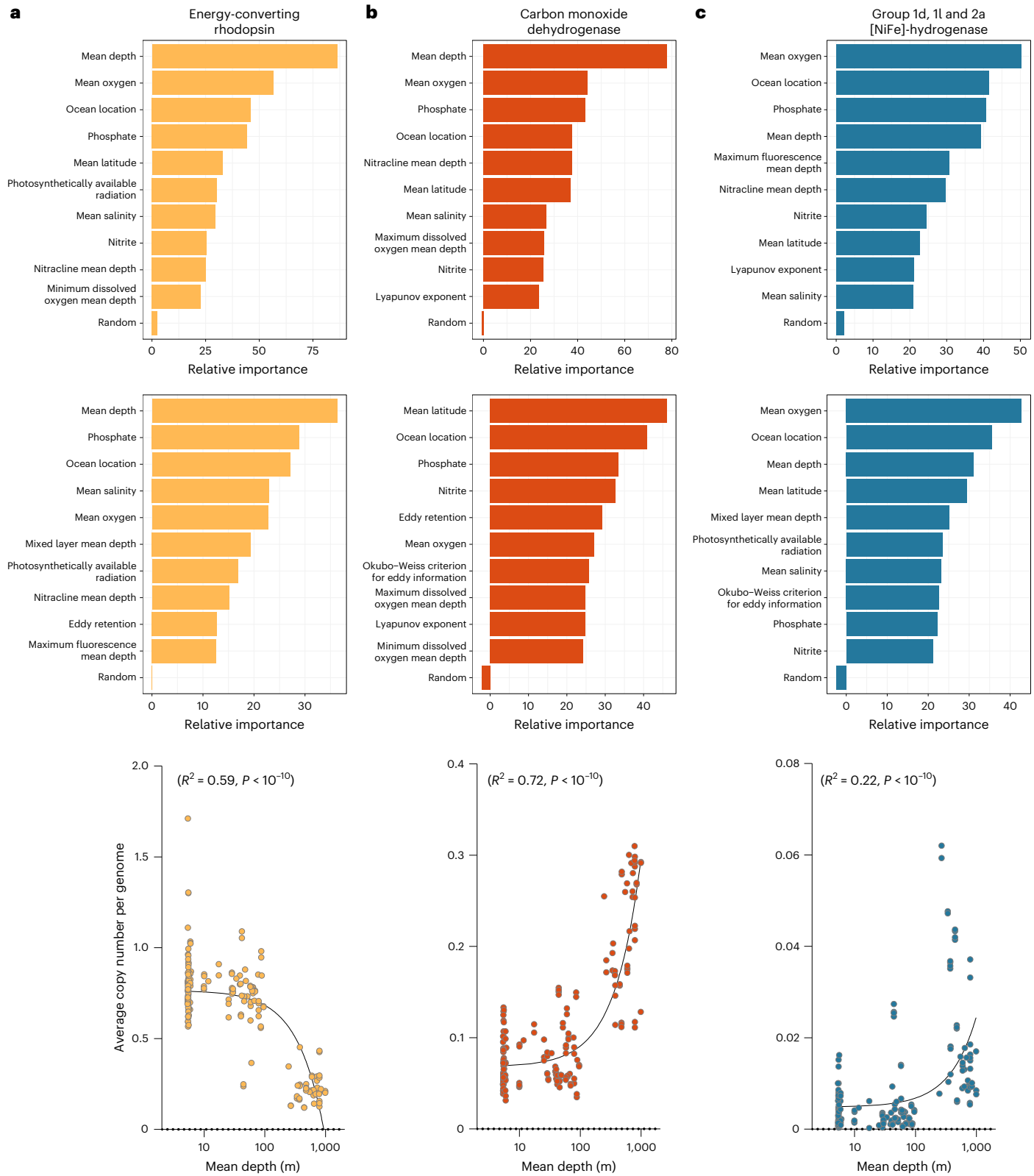


Fig. 6 | Drivers of the abundance and expression of metabolic genes in *Tara* Oceans metagenomes. a–**c**, This analysis is visualized for energy-converting rhodopsins (**a**), CO dehydrogenases (**b**) and aerobic H_2-uptake hydrogenases (**c**). The top and middle panels show random forest modelling of the environmental variables that best predict marker gene abundance in metagenomes and metatranscriptomes, respectively. The relative importance (percentage increase in mean squared error, %IncMSE, as a measure of decrease in model accuracy) of the top ten most important variables for each model is shown in addition to a randomized variable used to benchmark importance. The bottom panel shows simple linear correlations between the metagenomic abundance of each gene and water depth. For each gene, Pearson's R^2 values show goodness of fit and *P* values confirm that each slope significantly deviates from zero.

growing cultures to sub-atmospheric concentrations within a period of 30 h, whereas negligible consumption occurred in stationary phase cultures (Fig. 5c). Together, these findings suggest that *S. alaskensis* can grow mixotrophically in marine waters by simultaneously consuming dissolved H_2 with available organic substrates. These findings align closely with that observed for other organisms harbouring group 2a [NiFe]-hydrogenases[9,10] and support the inferences from thermodynamic modelling (Fig. 4) that H_2 probably supports growth of some marine bacteria.

H_2 and CO oxidation capacity changes with water depth

Finally, we investigated the environmental correlates of the abundance and expression of trace gas oxidation genes in the *Tara* Oceans datasets (Fig. 6 and Supplementary Table 6). Linear correlation analysis confirmed that genes encoding the aerobic H_2-uptake hydrogenase ($R^2 = 0.22$, $P < 0.0001$) and the CO dehydrogenase ($R^2 = 0.72$, $P < 0.0001$) both significantly increased with depth (Fig. 6), as illustrated by their increased abundance in the metagenomes from mesopelagic waters (Fig. 2). This contrasts with the sharp decreases in the genes responsible for phototrophy, such as energy-converting rhodopsins ($R^2 = 0.59$, $P < 0.0001$), with depth (Figs. 2 and 6). This pattern was consistent across sites in the Atlantic, Indian, Pacific and Southern Oceans. These findings suggest that as light and hence energy availability decreases, there is a greater selective advantage for bacteria that use trace gases (lithoheterotrophy) rather than photosynthesis (photoheterotrophy).

These inferences were nuanced, after accounting for co-correlated variables (Supplementary Fig. 10), by random forest modelling (Fig. 6 and Supplementary Figs. 11 and 12). Depth was among the top three strongest predictors of the abundance of group 1l and 2a [NiFe]-hydrogenases, CO dehydrogenase and energy-converting rhodopsins (Fig. 6 and Supplementary Fig. 11). Latitude proved to be a strong predictor of the expression of the group 1l [NiFe]-hydrogenases and CO dehydrogenases, the latter peaking in the tropics (Fig. 6 and Supplementary Figs. 11–13). One explanation for the latter is that in tropical waters, increased photochemical and thermochemical CO production enhances substrate availability for CO oxidizers. These observations are consistent with the inverse CO and H_2 oxidation rates observed across the Munida transect (Fig. 1), as well as previously reported latitudinal variations in seawater concentrations of these gases[23–28]. In contrast, group 1d [NiFe]-hydrogenase gene abundance and expression levels were highest in hypoxic waters (Fig. 6 and Supplementary Fig. 14); this suggests that in contrast to its high-affinity oxygen-insensitive counterparts, this hydrogenase will be most transcribed when H_2 levels are elevated due to hypoxic fermentation (resulting in activation of the sensory hydrogenase) and most active when O_2 levels are low enough to minimize active site inhibition[38,51]. Collectively, our analyses suggest that there are complex environmental controls on the abundance and activities of marine trace gas oxidizers, and that the three H_2-uptake hydrogenases are ecophysiologically distinct.

Discussion

Through an integrative approach, we provide presumably the first demonstration that H_2 is an important energy source for seawater communities. The biogeochemical, metagenomic and thermodynamic modelling analyses together suggest that H_2 is oxidized by a diverse but small proportion of community members, but at sufficiently fast cell-specific rates to enable lithotrophic growth. These findings are supported by experimental observations that the ultramicrobacterium *S. alaskensis* consumes H_2 during heterotrophic growth. Marine bacteria with the capacity to oxidize H_2 probably gain a major competitive advantage from being able to consume this abundant, diffusible, high-energy gas. H_2-oxidizing marine microorganisms are globally distributed, although activity measurements and hydrogenase distribution profiles suggest complex controls on their activity and that they may be particularly active in low-chlorophyll waters. In contrast, our findings support that CO oxidation is a widespread trait that enhances the flexibility and likely primarily survival of habitat generalists[30,31], especially in high-chlorophyll waters. At the biogeochemical scale, our findings indicate that marine bacteria mitigate atmospheric H_2 emissions[19] and potentially account for undersaturation of H_2 in Antarctic waters[28].

Yet a major enigma remains. H_2 and CO are among the most dependable energy sources in the sea given their relatively high concentrations and energy yields. So why do relatively few bacteria harness them? By comparison, soils are net sinks for these trace gases given that the numerous bacteria present rapidly consume them[12]. We propose the straightforward explanation that the resource investment required to make the metalloenzymes to harness these trace gases may not always be justified by the energy gained. In the acutely iron-limited ocean, hydrogenases (containing 12–13 Fe atoms per protomer[11]) and to a lesser extent CO dehydrogenases (containing 4 Fe atoms per protomer[63]) are a major investment. This trade-off is likely to be most pronounced in the surface ocean, where solar energy can be harvested using minimal resources through energy-converting rhodopsins. However, the iron investment required to consume H_2 and CO is likely to be justified in energy-limited waters at depths and regions or seasons where primary production is low. This is consistent with the observed enrichment of hydrogenases and CO dehydrogenases in metagenomes from mesopelagic waters, as well as increased H_2 oxidation observed in subantarctic waters. Moreover, iron availability is typically higher in deeper circulating waters and around continental shelves (due to both deep water upwelling and terrestrial inputs), where high hydrogenase expression and activity were observed[64]. Thus, oceans continue to be a net source of H_2 and CO despite the importance of these energy sources for diverse marine bacteria.

Methods

Sample collection and characteristics

To determine the ability of marine microbial communities to oxidize trace gases, a total of 14 marine surface water samples were collected from three different locations (Supplementary Fig. 1). Eight samples were collected from across the Munida Microbial Observatory Time-Series transect (Otago, New Zealand)[65] on 23 July 2019 in calm weather on the RV *Polaris II*. This marine transect begins off the coast of Otago, New Zealand and extends through neritic, subtropical and subantarctic waters[65]. Eight equidistant stations were sampled travelling east, ranging from approximately 15 km to 70 km from Taiaroa Head. At each station, water was collected at 1 m depth using Niskin bottles and stored in two 1 l autoclaved bottles. One bottle was reserved for DNA filtration and extraction, whereas the other was used for microcosm incubation experiments. The vessel measured changes in salinity and temperature to determine the boundaries of each water mass (Supplementary Fig. 2).

Four samples were also collected from the temperate Port Phillip Bay at Carrum Beach (Victoria, Australia) on 20 March 2019 and two were collected from the tropical Heron Island (Queensland, Australia) on 9 July 2019. At both sites, near-shore surface microlayer and surface water samples were collected in the subtidal zone (water depth ca. 1 m). At Port Phillip Bay, two samples were also collected at 7.5 km and 15 km east of the mouth of the Patterson River, labelled 'Intermediate' and 'Centre' respectively. In all cases, surface water samples of 3 l were collected with a sterile Schott bottle from approximately 20 cm depth and aliquoted for microcosm incubation and DNA extraction. Surface microlayer samples were collected using a manual glass-plate sampler of 1,800 cm^2 surface area[66]. A total of 520–580 ml was collected in 150–155 dips, resulting in an average sampling thickness of 20 µm. For the surface microlayer samples, 180 ml was reserved for microcosm incubations, with the remaining volume used for DNA extraction. From all transects, each sample reserved for DNA extraction was vacuum-filtered using 0.22 µm polycarbonate filters and then stored at −80 °C until extraction.

Measurement of dissolved H_2 and CO

Dissolved gases were also sampled in situ at each transect to measure dissolved concentrations of CO and H_2. Serum vials (160 ml) were filled with seawater using a gas-tight tube, allowing approximately 300 ml to overflow. The vial was then sealed with a treated lab-grade butyl rubber stopper, avoiding the introduction of gas to the vial. An ultra-pure N_2 headspace (20 ml) was introduced to the vial by concurrently removing 20 ml of liquid using two gas-tight syringes. The vials were then shaken vigorously for 2 min before being equilibrated for 5 min to allow dissolved gases to enter the headspace. Of the headspace, 17 ml was then collected into a syringe flushed with N_2 by returning the removed liquid to the vial, and 2 ml was purged to flush the stopcock and needle before injecting the remaining 15 ml into a N_2-flushed and evacuated silicone-closed Exetainer[67] for storage. Exetainers were sealed with a stainless-steel bolt and O-ring and stored until measurement. H_2 and CO concentrations in the Exetainers were analysed by gas chromatography using a pulse discharge helium ionization detector (model TGA-6792-W-4U-2, Valco Instruments), as previously described[68], calibrated against standard CO and H_2 gas mixtures of known concentrations.

Ex situ activity assays

To determine the ability of these marine microbial communities to oxidize CO and H_2, the seawater samples were incubated with these gases under laboratory conditions and their concentration over time was measured using gas chromatography. For each sample, triplicate microcosms were setup in which seawater was transferred into foil-wrapped serum vials (60 ml seawater in 120 ml vials for Munida transect and Port Phillip Bay; 80 ml seawater in 160 ml vials for Heron Island) and sealed with treated lab-grade butyl rubber stoppers[67]. For each sampling location, one set of triplicates was also autoclaved and used as a control. The ambient-air headspace of each vial was spiked with H_2 and CO so that they reached initial headspace mixing ratios of either 2 ppmv (Munida transect and Port Phillip Bay) or 10 ppmv (Heron Island). Microcosms were continuously agitated at 20 °C on a shaker table at 100 r.p.m. For Munida and Port Phillip Bay samples, 1 ml samples were extracted daily from the headspace and their content was measured by gas chromatography as described above. For Heron Island samples, at each timepoint, 6 ml gas was extracted and stored in 12 ml UHP-He-flushed conventional Exetainers (2018) or pre-evacuated 3 ml silicone-sealed Exetainers[67].

Calculation of dissolved gas concentrations

The concentrations of dissolved gases in seawater at equilibrium state and at 1 atmospheric pressure were calculated according to the Sechenov relation for mixed electrolyte solutions, as described in ref. [69]:

$$\log\left(\frac{k_{G,0}}{k_G}\right) = \sum(h_i + h_G)\,c_i \quad (1)$$

where $k_{G,0}$ and k_G denote the gas solubility (or Henry's law constant in equivalent) in water and the mixed electrolyte solution, respectively, h_i is a constant specific to the dissolved ion i (m^3 $kmol^{-1}$), h_G is a gas-specific parameter (m^3 $kmol^{-1}$) and c_i represents the concentration of the dissolved ion i in solution (kmol m^{-3}). The gas-specific constant, h_G, at temperature T (in K) follows the equation:

$$h_G = h_{G,0} + h_T\,(T - 298.15) \quad (2)$$

where $h_{G,0}$ represents the value of h_G at 298.15 K and h_T is a gas-specific parameter for the temperature effect (m^3 $kmol^{-1}$ K^{-1}). The gas solubility parameter $k_{G,0}$ at temperature T follows combined Henry's law and van 't Hoff equation:

$$k_{G,0} = k'_{G,0} \times e^{\frac{-\Delta_{soln}H}{R}\left(\frac{1}{T} - \frac{1}{298.15}\right)} \quad (3)$$

where $k'_{G,0}$ denotes Henry's law constant of the gas at 298.15 K, $\Delta_{soln}H$ is the enthalpy of solution and R is the ideal gas law constant.

The concentrations of dissolved gases at equilibrium with the headspace gas phase at 1 atmospheric pressure and incubation temperature of 20 °C were calculated on the basis of a mean seawater composition as reported in ref. [70]. The salinity correcting constants h_i, $h_{G,0}$ and h_T were adopted from ref. [69], while the temperature correcting constants $k'_{G,0}$ and $\frac{-\Delta_{soln}H}{R}$ were obtained from ref. [71].

Kinetic analysis and thermodynamic modelling

For kinetic analysis, measurement timepoints of up to 30 d of incubation time were used. The gas consumption pattern was fitted with both an exponential model and a linear model. The former showed a lowest overall Akaike information criterion value for both H_2 and CO consumption (Supplementary Table 1). As such, first-order reaction rate constants were calculated and used for the kinetic modelling. In addition, only samples having at least two replicates with a positive rate constant were deemed to have a confident gas consumption. Bulk atmospheric gas oxidation rates for each sample were calculated with respect to the mean atmospheric mixing ratio of the corresponding trace gases (H_2: 0.53 ppmv; CO: 0.09 ppmv; CH_4: 1.9 ppmv). To estimate the cell-specific gas oxidation rate, the average direct cell count values reported for surface seawaters at Port Phillip Bay centre[72] and the eight stations along the Munida transect were used[65,73]. Assuming that all cells are viable and active, cell-specific gas oxidation rates were then inferred by multiplying the estimated relative abundance of trace gas oxidizers derived from the metagenomic short reads (the average gene copy number, assuming one copy per organism; see 'Metabolic annotation' below) by the cell counts to obtain the number of trace gas oxidizers.

To estimate the energetic contributions of H_2 and CO oxidation to the corresponding marine trace gas oxidizers, we performed thermodynamic modelling to calculate their respective theoretical energy yields according to the first-order kinetics of each sample estimated above. Power (Gibbs energy per unit time per cell) P follows the equation:

$$P = \frac{\upsilon \times \Delta G_r}{B} \quad (4)$$

where υ denotes the rate of substrate consumption per litre of seawater (mol l^{-1} s^{-1}) and B is the number of microbial cells (cells l^{-1}) performing the reactions $H_2 + 0.5\,O_2 \rightarrow H_2O$ (dihydrogen oxidation) and $CO + 0.5\,O_2 \rightarrow CO_2$ (carbon monoxide oxidation). ΔG_r represents the Gibbs free energy of the reaction at the experimental conditions (J mol^{-1}) and follows the equation:

$$\Delta G_r = \Delta G_r^0 + RT \ln Q_r \quad (5)$$

where ΔG_r^0 denotes the standard Gibbs free energy of the reaction, Q_r denotes the reaction quotient, R represents the ideal gas constant and T represents temperature in Kelvin. Values of ΔG_r^0 of the hydrogen oxidation and carbon monoxide oxidation were obtained from ref. [74]. Values of Q_r for each reaction were calculated using:

$$Q_r = \prod a_g^{n_i} \quad (6)$$

where a_g and n_i denote the dissolved concentration of the ith species in seawater and the stoichiometric coefficient of the ith species in the reaction of interest, respectively. Gibbs free energies were calculated for oxidation of hydrogen and carbon monoxide at atmospheric pressure and 20 °C incubation temperature. To contextualize cellular power yield from H_2 and CO oxidation in relation to reported cellular energy requirements, a comprehensive list of maintenance (endogenous

rate) and growth (active rate) power requirements of 121 organoheterotrophic bacteria at 20 °C reported in ref. [53] was used as the primary reference. A median maintenance energy of 1.9×10^{-15} W per cell was derived from the bacterial endogenous rates obtained in the supporting information sd01 of the above reference.

Metagenomic sequencing and assembly

DNA was extracted from the sample filters using the DNeasy PowerSoil kit (QIAGEN) following the manufacturer's instructions. Sample libraries, including an extraction blank control, were prepared with the Nextera XT DNA Sample Preparation kit (Illumina) and sequenced on an Illumina NextSeq500 platform (2 × 151 bp) at the Australian Centre for Ecogenomics (University of Queensland). An average of 20,122,526 read pairs were generated per sample, with 827,868 read pairs sequenced in the negative control (Supplementary Table 2). Raw metagenomic data were quality controlled with the BBTools suite v38.90 (https://sourceforge.net/projects/bbmap/), using BBDuk to remove the 151st base, trim adapters, filter PhiX reads, trim the 3' end at a quality threshold of 15 and discard reads below 50 bp in length. Reads detected in the extraction blank were additionally removed with BBMap v38.90, leaving a total of 97.7% of raw sample reads for further analysis. Taxonomy was profiled from high-quality short reads by assembling and classifying 16S rRNA and 18S rRNA genes with PhyloFlash v3.4 (ref. [75]). Short reads were assembled individually with metaSPAdes v3.14.1 (ref. [76]) and collectively (all samples together, and by location) with MEGAHIT v1.2.9 (ref. [77]). Coverage profiles for each contig were generated by mapping the short reads to the assemblies with BBMap v38.90 (ref. [78]).

Genome binning was performed with MetaBAT2 v2.15.5 (ref. [79]), MaxBin 2 v2.2.7 (ref. [80]) and CONCOCT v1.1.0 (ref. [81]) after setting each tool to retain only contigs ≥2,000 bp in length. For each assembly, resulting bins were dereplicated across binning tools with DAS_Tool v1.1.3 (ref. [82]). All bins were refined with RefineM v0.1.2 (ref. [83]) and consolidated into a final set of non-redundant metagenome-assembled-genomes (MAGs) at the default 99% average nucleotide identity using dRep v3.2.2 (ref. [84]). The completeness, contamination and strain heterogeneity of each MAG were calculated with CheckM v1.1.3 (ref. [85]), resulting in a total of 21 high-quality (>90% completeness, <5% contamination[86]) and 89 medium-quality (>50% completeness, <10% contamination[86]) MAGs. Taxonomy was assigned to each MAG with GTDB-Tk v1.6.0 (ref. [87]) (using GTDB release 202)[88] and open reading frames were predicted from each MAG and additionally across all contigs (binned and unbinned) with Prodigal v2.6.3 (ref. [89]). CoverM v0.6.1 (https://github.com/wwood/CoverM) 'genome' was used to calculate the relative abundance of each MAG in each sample (–min-read-aligned-percent 0.75, –min-read-percent-identity 0.95, –min-covered-fraction 0) and the mean read coverage per MAG across the dataset (-m mean, –min-covered-fraction 0).

For global comparisons, raw metagenome (PRJEB1787) and metatranscriptome (PRJEB6608) data from the *Tara* Oceans global dataset were downloaded from the European Nucleotide Archive[47,50]. In addition, 1,888 bacterial and archaeal MAGs generated in ref. [90] were downloaded (via https://www.genoscope.cns.fr/tara/).

Metabolic annotation

For both the metagenomes generated in this study and those from the *Tara* Oceans dataset, high-quality short reads and predicted proteins from assemblies and MAGs underwent metabolic annotation using DIAMOND v2.0.9 (–max-target-seqs 1, –max-hsps 1)[91] for alignment against a custom set of 50 metabolic marker protein databases. The marker proteins (https://doi.org/10.26180/c.5230745) cover the major pathways for aerobic and anaerobic respiration, energy conservation from organic and inorganic compounds, carbon fixation, nitrogen fixation and phototrophy[4]. Gene hits were filtered as follows: alignments were filtered to retain only those either at least 40 amino acids in length (150 bp metagenomes from the current study), 32 amino acids in length (100 bp *Tara* metagenomes and metatranscriptomes) or with at least 80% query or 80% subject coverage (predicted proteins from assemblies and MAGs). Alignments were further filtered by a minimum percentage identity score by protein: for short reads, this was 80% (PsaA), 75% (HbsT), 70% (PsbA, IsoA, AtpA, YgfK and ARO), 60% (CoxL, MmoA, AmoA, NxrA, RbcL, NuoF, FeFe hydrogenases and NiFe Group 4 hydrogenases) or 50% (all other genes). For predicted proteins, the same thresholds were used except for AtpA (60%), PsbA (60%), RdhA (45%), Cyc2 (35%) and RHO (30%).

For short reads, gene abundance in the community was estimated as 'average gene copies per organism' by dividing the abundance of the gene (in reads per kilobase million, RPKM) by the mean abundance of 14 universal single-copy ribosomal marker genes (in RPKM, obtained from the SingleM v0.13.2 package, https://github.com/wwood/singlem). For single-copy metabolic genes, this corresponds to the proportion of community members that encode the gene. A linear correlation analysis, performed in GraphPad Prism 9, was used to determine how metagenomic gene abundance correlated with ex situ H_2 and CO oxidation rates. For the *Tara* Oceans dataset, the RNA:DNA ratio was calculated by dividing gene abundance in the metatranscriptome (in RPKM) by the gene abundance in the corresponding metagenome (RPKM) to examine gene expression relative to abundance. Where replicate metagenomes or metatranscriptomes were present, RPKM values were averaged by sample.

Phylogenetic analysis

Phylogenetic trees were constructed to understand the distribution and diversity of marine microorganisms capable of H_2 and CO oxidation. Trees were constructed for the catalytic subunits of the groups 1 and 2 [NiFe]-hydrogenases, groups 3 and 4 [NiFe]-hydrogenases, and the form I CO dehydrogenase (CoxL). In all cases, protein sequences retrieved from the MAGs by homology-based searches were aligned against a subset of reference sequences from custom protein databases[6,49] using ClustalW in MEGA11 (ref. [92]). In brief, evolutionary relationships were visualized by constructing a maximum-likelihood phylogenetic tree; specifically, initial trees for the heuristic search were obtained automatically by applying Neighbour-Join and BioNJ algorithms to a matrix of pairwise distances estimated using a Jones-Taylor-Thornton (JTT) model, and then selecting the topology with superior log likelihood value within MEGA11. All residues were used and trees were bootstrapped with 50 replicates. Phylogenetic tree annotation and visualization were performed using iTOL (v6.6).

Environmental driver analysis

Random forest models, Pearson correlations and Spearman correlations were generated for the *Tara* Oceans dataset to identify significant correlations between sample environmental metadata and the normalized abundance of carbon monoxide dehydrogenase, rhodopsin and [NiFe] groups 1d, 1e, 1l, 2a, 3b and 3d hydrogenase genes (shown as copies per organism for metagenomes, $\log_{10}$(RPKM + 1) for metatranscriptomes). To account for collinearity, where environmental variables were highly correlated (Pearson coefficient > |0.7|, Supplementary Fig. 10), one was excluded from the random forest models to avoid the division of variable importance across those features. These excluded variables were selected at random, unless they were highly correlated with depth (which was kept). Then, using imputed values where data were missing (function rfImpute()), a random forest model was generated for each gene above using the environmental variables marked in Supplementary Table 6 as predictors (importance = TRUE, ntree = 3,000), using the R package randomForest[93]. All combinations of the above genes and environmental variables were additionally correlated with Pearson's and Spearman's rank correlations, omitting missing values and adjusting all *P* values with the false discovery rate correction.

Culture-based growth and gas consumption analysis

Axenic cultures of three bacterial strains were analysed in this study: *Sphingopyxis alaskensis* (RB2256)[56,57] obtained from UNSW Sydney,

Robiginitalea biformata DSM-15991 (ref. [54]) imported from DSMZ and *Marinovum algicola* FF3 (Rhodobacteraceae)[55] imported from DSMZ. Cultures were maintained in 120 ml glass serum vials containing a headspace of ambient air (H_2 mixing ratio ~0.5 ppmv) sealed with treated lab-grade butyl rubber stoppers[67]. Broth cultures of all three species were grown in 30 ml of Difco 2216 Marine Broth media and incubated at 30 °C at an agitation speed of 150 r.p.m. in a Ratek orbital mixer incubator with access to natural day/night cycles. Growth was monitored by determining the optical density (OD_{600}) of periodically sampled 1 ml extracts using an Eppendorf BioSpectrophotometer. The ability of the three cultures to oxidize H_2 was measured by gas chromatography. Cultures in biological triplicate were opened, equilibrated with ambient air (1 h) and resealed. These re-aerated vials were then amended with H_2 (via 1% v/v H_2 in N_2 gas cylinder, 99.999% pure) to achieve final headspace concentrations of ~10 ppmv. Headspace mixing ratios were measured immediately after closure and at regular intervals thereafter until the limit of quantification of the gas chromatograph was reached (42 ppbv H_2). This analysis was performed for both exponential (OD_{600} 0.67 for *S. alaskensis*) and stationary phase cultures (~72 h post OD_{max} for *S. alaskensis*).

RT–qPCR analysis

Quantitative reverse transcription PCR (RT–qPCR) was used to determine the expression levels of the group 2a [NiFe]-hydrogenase large subunit gene (*hucL*; locus Sala_3198) in *S. alaskensis* during growth and survival. For RNA extraction, triplicate 30 ml cultures of *S. alaskensis* were grown synchronously in 120 ml sealed serum vials. Cultures were grown to either exponential phase (OD_{600} 0.67) or stationary phase (48 h post OD_{max} ~3.2). Cells were then quenched using a glycerol-saline solution (−20 °C, 3:2 v/v), collected by centrifugation (20,000 × *g*, 30 min, −9 °C), resuspended in 1 ml cold 1:1 glycerol:saline solution (−20 °C) and further centrifuged (20,000 × *g*, 30 min, −9 °C). Briefly, resultant cell pellets were resuspended in 1 ml TRIzol reagent (Thermo Fisher), mixed with 0.1 mm zircon beads (0.3 g) and subjected to bead beating (three 30 s on/30 s off cycles, 5,000 r.p.m.) in a Precellys 24 homogenizer (Bertin Technologies) before centrifugation (12,000 × *g*, 10 min, 4 °C). Total RNA was extracted using the phenol-chloroform method following the manufacturer's instructions (TRIzol reagent user guide, Thermo Fisher) and resuspended in diethylpyrocarbonate-treated water. RNA was treated using the TURBO DNA-free kit (Thermo Fisher) following the manufacturer's instructions. RNA concentration and purity were confirmed using a NanoDrop ND-1000 spectrophotometer.

Complementary DNA was synthesized using a SuperScript III First-Strand Synthesis System kit for RT–qPCR (Thermo Fisher) with random hexamer primers, following the manufacturer's instructions. RT–qPCR was performed in a QuantStudio 7 Flex Real-Time PCR System (Applied Biosystems) using a LightCycler 480 SYBR Green I Master Mix (Roche) in 96-well plates according to the manufacturer's instructions. Primers were designed using Primer3 (ref. [94]) to target the *hucL* gene (HucL_fw: AGCTACACAAACCCTCGACA; HucL_rvs: AGTCGATCATGAACAGGCCA) and the 16S rRNA gene as a housekeeping gene (16S_fwd: AACCCTCATCCCTAGTTGCC; 16S_rvs: GGTTAGAGCATTGCCTTCGG). Copy numbers for each gene were interpolated from standard curves of each gene created from threshold cycle (C_T) values of amplicons that were serially diluted from 10^8 to 10 copies ($R^2 > 0.98$). Hydrogenase expression data were then normalized to the housekeeping gene in exponential phase. All biological triplicate samples, standards and negative controls were run in technical duplicate. A Student's *t*-test in GraphPad Prism 9 was used to compare *hucL* expression levels between exponential and stationary phases.

Reporting summary

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

Data availability

All raw metagenomes and metagenome-assembled genomes are deposited to the NCBI Sequence Read Archive under the BioProject accession number PRJNA801081. Raw metagenome (PRJEB1787) and metatranscriptome (PRJEB6608) data from the Tara Oceans global dataset were downloaded from the European Nucleotide Archive (https://www.ebi.ac.uk/ena/browser/view/PRJEB402). Bacterial and archaeal MAGs (1,888) generated in ref. [90] were downloaded from https://www.genoscope.cns.fr/tara/. The metabolic marker protein database used in this study, which includes reference hydrogenase and carbon monoxide dehydrogenase sequences, can be obtained from https://bridges.monash.edu/collections/_/5230745.

Code availability

Metagenomics analysis scripts are publicly available at https://github.com/greeninglab/MarineOxidationManuscript.

References

1. Greening, C. & Grinter, R. Microbial oxidation of atmospheric trace gases. *Nat. Rev. Microbiol*. https://doi.org/10.1038/s41579-022-00724-x (2022).
2. Greening, C., Berney, M., Hards, K., Cook, G. M. & Conrad, R. A soil actinobacterium scavenges atmospheric H_2 using two membrane-associated, oxygen-dependent [NiFe] hydrogenases. *Proc. Natl Acad. Sci. USA* **111**, 4257–4261 (2014).
3. Myers, M. R. & King, G. M. Isolation and characterization of *Acidobacterium ailaaui* sp. nov., a novel member of Acidobacteria subdivision 1, from a geothermally heated Hawaiian microbial mat. *Int. J. Syst. Evol. Microbiol*. **66**, 5328–5335 (2016).
4. Ortiz, M. et al. Multiple energy sources and metabolic strategies sustain microbial diversity in Antarctic desert soils. *Proc. Natl Acad. Sci. USA* **118**, e2025322118 (2021).
5. King, G. M. Molecular and culture-based analyses of aerobic carbon monoxide oxidizer diversity. *Appl. Environ. Microbiol*. **69**, 7257–7265 (2003).
6. Cordero, P. R. F. et al. Atmospheric carbon monoxide oxidation is a widespread mechanism supporting microbial survival. *ISME J.* **13**, 2868–2881 (2019).
7. Greening, C., Villas-Bôas, S. G., Robson, J. R., Berney, M. & Cook, G. M. The growth and survival of *Mycobacterium smegmatis* is enhanced by co-metabolism of atmospheric H_2. *PLoS ONE* **9**, e103034 (2014).
8. Liot, Q. & Constant, P. Breathing air to save energy – new insights into the ecophysiological role of high-affinity [NiFe]-hydrogenase in *Streptomyces avermitilis*. *Microbiologyopen* **5**, 47–59 (2016).
9. Islam, Z. F. et al. A widely distributed hydrogenase oxidises atmospheric H_2 during bacterial growth. *ISME J.* **14**, 2649–2658 (2020).
10. Leung, P. M. et al. A nitrite-oxidising bacterium constitutively oxidises atmospheric H2. *ISME J.* https://doi.org/10.1038/s41396-022-01265-0 (2022).
11. Greening, C. et al. Genomic and metagenomic surveys of hydrogenase distribution indicate H_2 is a widely utilised energy source for microbial growth and survival. *ISME J.* **10**, 761–777 (2016).
12. Bay, S. K. et al. Trace gas oxidizers are widespread and active members of soil microbial communities. *Nat. Microbiol.* **6**, 246–256 (2021).
13. Xu, Y. et al. Genome-resolved metagenomics reveals how soil bacterial communities respond to elevated H_2 availability. *Soil Biol. Biochem.* **163**, 108464 (2021).
14. Schmidt, U. Molecular hydrogen in the atmosphere. *Tellus* **26**, 78–90 (1974).

15. Walter, S. et al. Isotopic evidence for biogenic molecular hydrogen production in the Atlantic Ocean. *Biogeosciences* **13**, 323–340 (2016).
16. Moore, R. M. et al. Extensive hydrogen supersaturations in the western South Atlantic Ocean suggest substantial underestimation of nitrogen fixation. *J. Geophys. Res. Oceans* **119**, 4340–4350 (2014).
17. Conte, L., Szopa, S., Séférian, R. & Bopp, L. The oceanic cycle of carbon monoxide and its emissions to the atmosphere. *Biogeosciences* **16**, 881–902 (2019).
18. Khalil, M. A. K. & Rasmussen, R. A. The global cycle of carbon monoxide: trends and mass balance. *Chemosphere* **20**, 227–242 (1990).
19. Ehhalt, D. H. & Rohrer, F. The tropospheric cycle of H_2: a critical review. *Tellus B* **61**, 500–535 (2009).
20. Miller, W. L. & Zepp, R. G. Photochemical production of dissolved inorganic carbon from terrestrial organic matter: significance to the oceanic organic carbon cycle. *Geophys. Res. Lett.* **22**, 417–420 (1995).
21. Moore, R. M., Punshon, S., Mahaffey, C. & Karl, D. The relationship between dissolved hydrogen and nitrogen fixation in ocean waters. *Deep Sea Res. I* **56**, 1449–1458 (2009).
22. Kessler, A. J. et al. Bacterial fermentation and respiration processes are uncoupled in permeable sediments. *Nat. Microbiol.* **4**, 1014–1023 (2019).
23. Swinnerton, J. W., Linnenbom, V. J. & Lamontagne, R. A. The ocean: a natural source of carbon monoxide. *Science* **167**, 984–986 (1970).
24. Swinnerton, J. W. & Lamontagne, R. A. Carbon monoxide in the South Pacific Ocean. *Tellus* **26**, 136–142 (1974).
25. Herr, F. L., Scranton, M. I. & Barger, W. R. Dissolved hydrogen in the Norwegian Sea: mesoscale surface variability and deep-water distribution. *Deep Sea Res. A* **28**, 1001–1016 (1981).
26. Herr, F. L. Dissolved hydrogen in Eurasian Arctic waters. *Tellus B* **36**, 55–66 (1984).
27. Conrad, R., Seiler, W., Bunse, G. & Giehl, H. Carbon monoxide in seawater (Atlantic Ocean). *J. Geophys. Res. Oceans* **87**, 8839–8852 (1982).
28. Conrad, R. & Seiler, W. Methane and hydrogen in seawater (Atlantic Ocean). *Deep Sea Res. A* **35**, 1903–1917 (1988).
29. Conrad, R. & Seiler, W. Photooxidative production and microbial consumption of carbon monoxide in seawater. *FEMS Microbiol. Lett.* **9**, 61–64 (1980).
30. Tolli, J. D., Sievert, S. M. & Taylor, C. D. Unexpected diversity of bacteria capable of carbon monoxide oxidation in a coastal marine environment, and contribution of the *Roseobacter*-associated clade to total CO oxidation. *Appl. Environ. Microbiol.* **72**, 1966–1973 (2006).
31. Mou, X., Sun, S., Edwards, R. A., Hodson, R. E. & Moran, M. A. Bacterial carbon processing by generalist species in the coastal ocean. *Nature* **451**, 708–711 (2008).
32. Cunliffe, M. Correlating carbon monoxide oxidation with cox genes in the abundant marine Roseobacter clade. *ISME J.* **5**, 685–691 (2011).
33. Royo-Llonch, M. et al. Compendium of 530 metagenome-assembled bacterial and archaeal genomes from the polar Arctic Ocean. *Nat. Microbiol.* **6**, 1561–1574 (2021).
34. Cunliffe, M. Physiological and metabolic effects of carbon monoxide oxidation in the model marine bacterioplankton *Ruegeria pomeroyi* DSS-3. *Appl. Environ. Microbiol.* **79**, 738–740 (2013).
35. Christie-Oleza, J. A., Fernandez, B., Nogales, B., Bosch, R. & Armengaud, J. Proteomic insights into the lifestyle of an environmentally relevant marine bacterium. *ISME J.* **6**, 124–135 (2012).
36. Muthusamy, S. et al. Comparative proteomics reveals signature metabolisms of exponentially growing and stationary phase marine bacteria. *Environ. Microbiol.* **19**, 2301–2319 (2017).
37. Giebel, H.-A., Wolterink, M., Brinkhoff, T. & Simon, M. Complementary energy acquisition via aerobic anoxygenic photosynthesis and carbon monoxide oxidation by *Planktomarina temperata* of the *Roseobacter* group. *FEMS Microbiol. Ecol.* **95**, fiz050 (2019).
38. Schwartz, E., Fritsch, J. & Friedrich, B. *H_2-Metabolizing Prokaryotes* (Springer, 2013).
39. Adam, N. & Perner, M. Microbially mediated hydrogen cycling in deep-sea hydrothermal vents. *Front. Microbiol.* **9**, 2873 (2018).
40. Anantharaman, K., Breier, J. A., Sheik, C. S. & Dick, G. J. Evidence for hydrogen oxidation and metabolic plasticity in widespread deep-sea sulfur-oxidizing bacteria. *Proc. Natl Acad. Sci. USA* **110**, 330–335 (2013).
41. Kleiner, M. et al. Use of carbon monoxide and hydrogen by a bacteria–animal symbiosis from seagrass sediments. *Environ. Microbiol.* **17**, 5023–5035 (2015).
42. Petersen, J. M. et al. Hydrogen is an energy source for hydrothermal vent symbioses. *Nature* **476**, 176–180 (2011).
43. Barz, M. et al. Distribution analysis of hydrogenases in surface waters of marine and freshwater environments. *PLoS ONE* **5**, e13846 (2010).
44. Eichner, M. J., Basu, S., Gledhill, M., de Beer, D. & Shaked, Y. Hydrogen dynamics in *Trichodesmium* colonies and their potential role in mineral iron acquisition. *Front. Microbiol.* **10**, 1565 (2019).
45. Bothe, H., Schmitz, O., Yates, M. G. & Newton, W. E. Nitrogen fixation and hydrogen metabolism in cyanobacteria. *Microbiol. Mol. Biol. Rev.* **74**, 529–551 (2010).
46. Sunagawa, S. et al. Tara Oceans: towards global ocean ecosystems biology. *Nat. Rev. Microbiol.* **18**, 428–445 (2020).
47. Sunagawa, S. et al. Structure and function of the global ocean microbiome. *Science* **348**, 1261359 (2015).
48. Podzuweit, H. G., Arp, D. J., Schlegel, H. G. & Schneider, K. Investigation of the H_2-oxidizing activities of *Alcaligenes eutrophus* H16 membranes with artificial electron acceptors, respiratory inhibitors and redox-spectroscopic procedures. *Biochimie* **68**, 103–111 (1986).
49. Søndergaard, D., Pedersen, C. N. S. & Greening, C. HydDB: a web tool for hydrogenase classification and analysis. *Sci. Rep.* **6**, 34212 (2016).
50. Salazar, G. et al. Gene expression changes and community turnover differentially shape the global ocean metatranscriptome. *Cell* **179**, 1068–1083 (2019).
51. Lenz, O. & Friedrich, B. A novel multicomponent regulatory system mediates H_2 sensing in *Alcaligenes eutrophus*. *Proc. Natl Acad. Sci. USA* **95**, 12474–12479 (1998).
52. Chen, Y. J. et al. Metabolic flexibility allows bacterial habitat generalists to become dominant in a frequently disturbed ecosystem. *ISME J.* https://doi.org/10.1038/s41396-021-00988-w (2021).
53. DeLong, J. P., Okie, J. G., Moses, M. E., Sibly, R. M. & Brown, J. H. Shifts in metabolic scaling, production, and efficiency across major evolutionary transitions of life. *Proc. Natl Acad. Sci. USA* **107**, 12941–12945 (2010).
54. Cho, J.-C. & Giovannoni, S. J. *Robiginitalea biformata* gen. nov., sp. nov., a novel marine bacterium in the family Flavobacteriaceae with a higher G+C content. *Int. J. Syst. Evol. Microbiol.* **54**, 1101–1106 (2004).
55. Lafay, B. et al. *Roseobacter algicola* sp. nov., a new marine bacterium isolated from the phycosphere of the toxin-producing dinoflagellate *Prorocentrum lima*. *Int. J. Syst. Evol. Microbiol.* **45**, 290–296 (1995).

56. Schut, F. et al. Isolation of typical marine bacteria by dilution culture: growth, maintenance, and characteristics of isolates under laboratory conditions. *Appl. Environ. Microbiol.* **59**, 2150–2160 (1993).
57. Schut, F., Gottschal, J. C. & Prins, R. A. Isolation and characterisation of the marine ultramicrobacterium *Sphingomonas* sp. strain RB2256. *FEMS Microbiol. Rev.* **20**, 363–369 (1997).
58. Vancanneyt, M. et al. *Sphingomonas alaskensis* sp. nov., a dominant bacterium from a marine oligotrophic environment. *Int. J. Syst. Evol. Microbiol.* **51**, 73–79 (2001).
59. Eguchi, M. et al. *Sphingomonas alaskensis* strain AFO1, an abundant oligotrophic ultramicrobacterium from the North Pacific. *Appl. Environ. Microbiol.* **67**, 4945–4954 (2001).
60. Cavicchioli, R., Ostrowski, M., Fegatella, F., Goodchild, A. & Guixa-Boixereu, N. Life under nutrient limitation in oligotrophic marine environments: an eco/physiological perspective of *Sphingopyxis alaskensis* (formerly *Sphingomonas alaskensis*). *Microb. Ecol.* **45**, 203–217 (2003).
61. Williams, T. J., Ertan, H., Ting, L. & Cavicchioli, R. Carbon and nitrogen substrate utilization in the marine bacterium *Sphingopyxis alaskensis* strain RB2256. *ISME J.* **3**, 1036–1052 (2009).
62. Lauro, F. M. et al. The genomic basis of trophic strategy in marine bacteria. *Proc. Natl Acad. Sci. USA* **106**, 15527–15533 (2009).
63. Dobbek, H., Gremer, L., Meyer, O. & Huber, R. Crystal structure and mechanism of CO dehydrogenase, a molybdo iron-sulfur flavoprotein containing S-selanylcysteine. *Proc. Natl Acad. Sci. USA* **96**, 8884–8889 (1999).
64. Laës, A. et al. Deep dissolved iron profiles in the eastern North Atlantic in relation to water masses. *Geophys. Res. Lett.* **30**, 1902 (2003).
65. Baltar, F., Stuck, E., Morales, S. & Currie, K. Bacterioplankton carbon cycling along the subtropical frontal zone off New Zealand. *Prog. Oceanogr.* **135**, 168–175 (2015).
66. Cunliffe, M. & Wurl, O. *Guide to Best Practices to Study the Ocean's Surface* (Marine Biological Association of the United Kingdom for SCOR, 2014).
67. Nauer, P. A., Chiri, E., Jirapanjawat, T., Greening, C. & Cook, P. L. M. Technical note: inexpensive modification of Exetainers for the reliable storage of trace-level hydrogen and carbon monoxide gas samples. *Biogeosciences* **18**, 729–737 (2021).
68. Islam, Z. F. et al. Two Chloroflexi classes independently evolved the ability to persist on atmospheric hydrogen and carbon monoxide. *ISME J.* **13**, 1801–1813 (2019).
69. Weisenberger, S. & Schumpe, A. Estimation of gas solubilities in salt solutions at temperatures from 273 K to 363 K. *AIChE J.* **42**, 298–300 (1996).
70. Dickson, A. G. & Goyet, C. *Handbook of Methods for the Analysis of the Various Parameters of the Carbon Dioxide System in Sea Water* (Oak Ridge National Laboratory, 1994).
71. Sander, R. Compilation of Henry's law constants (version 4.0) for water as solvent. *Atmos. Chem. Phys.* **15**, 4399–4981 (2015).
72. Sharma, R. V., Edwards, R. T. & Beckett, R. Analysis of bacteria in aquatic environments using sedimentation field-flow fractionation: (II) physical characterization of cells. *Water Res.* **32**, 1508–1514 (1998).
73. Wenley, J. et al. Seasonal prokaryotic community linkages between surface and deep ocean water. *Front. Mar. Sci.* **8**, 777 (2021).
74. Thauer, R. K., Jungermann, K. & Decker, K. Energy conservation in chemotrophic anaerobic bacteria. *Bacteriol. Rev.* **41**, 100–180 (1977).
75. Gruber-Vodicka, H. R., Seah, B. K. B. & Pruesse, E. PhyloFlash: rapid small-subunit rRNA profiling and targeted assembly from metagenomes. *mSystems* **5**, e00920 (2019).
76. Nurk, S., Meleshko, D., Korobeynikov, A. & Pevzner, P. A. metaSPAdes: a new versatile metagenomic assembler. *Genome Res.* **27**, 824–834 (2017).
77. Li, D. H. et al. MEGAHIT v1.0: a fast and scalable metagenome assembler driven by advanced methodologies and community practices. *Methods* **102**, 3–11 (2016).
78. Bushnell, B. BBMap: a fast, accurate, splice-aware aligner (2015). In *Conference: 9th Annual Genomics of Energy & Environment Meeting* (USDOE Office of Science, 2014).
79. Kang, D. et al. MetaBAT 2: an adaptive binning algorithm for robust and efficient genome reconstruction from metagenome assemblies. *PeerJ* **7**, e7359 (2019).
80. Wu, Y.-W., Simmons, B. A. & Singer, S. W. MaxBin 2.0: an automated binning algorithm to recover genomes from multiple metagenomic datasets. *Bioinformatics* **32**, 605–607 (2015).
81. Alneberg, J. et al. Binning metagenomic contigs by coverage and composition. *Nat. Methods* **11**, 1144–1146 (2014).
82. Sieber, C. M. K. et al. Recovery of genomes from metagenomes via a dereplication, aggregation and scoring strategy. *Nat. Microbiol.* **3**, 836–843 (2018).
83. Parks, D. H. et al. Recovery of nearly 8,000 metagenome-assembled genomes substantially expands the tree of life. *Nat. Microbiol.* **2**, 1533–1542 (2017).
84. Olm, M. R., Brown, C. T., Brooks, B. & Banfield, J. F. dRep: a tool for fast and accurate genomic comparisons that enables improved genome recovery from metagenomes through de-replication. *ISME J.* **11**, 2864–2868 (2017).
85. Parks, D. H., Imelfort, M., Skennerton, C. T., Hugenholtz, P. & Tyson, G. W. CheckM: assessing the quality of microbial genomes recovered from isolates, single cells, and metagenomes. *Genome Res.* **25**, 1043–1055 (2015).
86. Bowers, R. M. et al. Minimum information about a single amplified genome (MISAG) and a metagenome-assembled genome (MIMAG) of bacteria and archaea. *Nat. Biotechnol.* **35**, 725–731 (2017).
87. Chaumeil, P.-A., Mussig, A. J., Hugenholtz, P. & Parks, D. H. GTDB-Tk: a toolkit to classify genomes with the Genome Taxonomy Database. *Bioinformatics* **36**, 1925–1927 (2020).
88. Parks, D. H. et al. A standardized bacterial taxonomy based on genome phylogeny substantially revises the tree of life. *Nat. Biotechnol.* **36**, 996–1004 (2018).
89. Hyatt, D. et al. Prodigal: prokaryotic gene recognition and translation initiation site identification. *BMC Bioinformatics* **11**, 119 (2010).
90. Delmont, T. O. et al. Heterotrophic bacterial diazotrophs are more abundant than their cyanobacterial counterparts in metagenomes covering most of the sunlit ocean. *ISME J.* **16**, 927–936 (2022).
91. Buchfink, B., Xie, C. & Huson, D. H. Fast and sensitive protein alignment using DIAMOND. *Nat. Methods* **12**, 59–60 (2014).
92. Tamura, K., Stecher, G. & Kumar, S. MEGA11: molecular evolutionary genetics analysis version 11. *Mol. Biol. Evol.* **38**, 3022–3027 (2021).
93. Liaw, A. & Wiener, M. Classification and regression by randomForest. *R News* **2**, 18–22 (2002).
94. Untergasser, A. et al. Primer3—new capabilities and interfaces. *Nucleic Acids Res.* **40**, e115 (2012).

Acknowledgements

This study was supported by ARC Discovery Project grants (DP180101762 and DP210101595 both awarded to P.L.M.C. and C.G.; DP150100244 awarded to R.C.), an ARC DECRA Fellowship (DE170100310; salary for C.G.), an NHMRC EL2 Fellowship (APP1178715; salary for C.G.), an Australian Government Research Training Stipend

Scholarship (awarded to P.M.L.), Monash International Tuition Scholarships (awarded to P.M.L. and Y.-J.C.) and Monash Postgraduate Publications Awards (awarded to Z.F.I. and Y.-J.C.).

Author contributions

C.G. conceived and supervised this study. C.G., G.S., S.E.M., P.L.M.C., R.L. and Z.F.I. designed experiments. G.S., S.L., S.E.M., P.A.N., Y.-J.C., A.J.K. and P.L.M.C. contributed to field work. R.L., G.S., S.L. and C.G. contributed to metagenome analysis. C.G. and G.N. contributed to phylogenetic analysis. G.S., P.M.L., P.A.N. and C.G. contributed to biogeochemical analysis. P.M.L., C.G., F.B. and P.M.L.C. contributed to thermodynamic modelling. Z.F.I., T.J., G.S., T.J.W., R.C. and C.G. contributed to culture-based work. R.L. and C.G. contributed to environmental driver analysis. C.G., R.L. and Z.F.I. wrote the paper with input from all authors.

Competing interests

The authors declare no conflict of interest.

Additional information

Supplementary information The online version contains supplementary material available at https://doi.org/10.1038/s41564-023-01322-0.

Correspondence and requests for materials should be addressed to Chris Greening.

Peer review information *Nature Microbiology* thanks the anonymous reviewers for their contribution to the peer review of this work.

nature portfolio

Corresponding author(s): Chris Greening

Last updated by author(s): 16/12/22

Reporting Summary

Nature Portfolio wishes to improve the reproducibility of the work that we publish. This form provides structure for consistency and transparency in reporting. For further information on Nature Portfolio policies, see our Editorial Policies and the Editorial Policy Checklist.

Statistics

For all statistical analyses, confirm that the following items are present in the figure legend, table legend, main text, or Methods section.

n/a	Confirmed	
☐	☒	The exact sample size (*n*) for each experimental group/condition, given as a discrete number and unit of measurement
☐	☒	A statement on whether measurements were taken from distinct samples or whether the same sample was measured repeatedly
☐	☒	The statistical test(s) used AND whether they are one- or two-sided *Only common tests should be described solely by name; describe more complex techniques in the Methods section.*
☒	☐	A description of all covariates tested
☒	☐	A description of any assumptions or corrections, such as tests of normality and adjustment for multiple comparisons
☐	☒	A full description of the statistical parameters including central tendency (e.g. means) or other basic estimates (e.g. regression coefficient) AND variation (e.g. standard deviation) or associated estimates of uncertainty (e.g. confidence intervals)
☐	☒	For null hypothesis testing, the test statistic (e.g. *F*, *t*, *r*) with confidence intervals, effect sizes, degrees of freedom and *P* value noted *Give P values as exact values whenever suitable.*
☒	☐	For Bayesian analysis, information on the choice of priors and Markov chain Monte Carlo settings
☒	☐	For hierarchical and complex designs, identification of the appropriate level for tests and full reporting of outcomes
☐	☒	Estimates of effect sizes (e.g. Cohen's *d*, Pearson's *r*), indicating how they were calculated

Our web collection on statistics for biologists contains articles on many of the points above.

Software and code

Policy information about availability of computer code

Data collection	Metagenomics analysis scripts are publicly available at https://github.com/greeninglab/MarineOxidationManuscript.
Data analysis	Databases used to analyse data or against which data was compared to include: - UniProtKB (https://www.uniprot.org/help/uniprotkb) - HydDB (https://services.birc.au.dk/hyddb/) - Curated metabolic marker databases (https://bridges.monash.edu/collections/_/5230745) - GTDB (release 202; https://gtdb.ecogenomic.org/) Data were analysed using the following published software: - BBTools suite v38.90 - BBMap v38.90 - PhyloFlash v3.4 - MEGAHIT v1.2.9 - MetaBAT2 v2.15.5 - MaxBin 2 v2.2.7 - CONCOCT v.1.1.0 - DAS_Tool v1.1.3 - RefineM v0.1.2 - CheckM v1.1.3 - GTDB-Tk v1.6.0 - Prodigal v2.6.3

- CoverM v0.6.1
- DIAMOND v2.0.9
- SingleM v0.13.2
- MEGA 11 v11.0.13 (including the following parameters: ClustalW (v2.0), Jones-Taylor-Thornton (JTT) model, Neighbour-Join statistical algorithm and BioNJ heuristic inference)
- Fastqc v0.11.7
- R v4.2.0 (including the following packages: randomForest, tidyr, dplyr, purrr, ggplot2, MetBrewer and ggcorrplot)
- GraphPad Prism 9
- dRep v3.2.2
- iTOL v6.6

Code availability:
Bash scripts used to process metagenomes, metatranscriptomes and MAGs are publicly available here: https://github.com/GreeningLab/MarineOxidationManuscript.

For manuscripts utilizing custom algorithms or software that are central to the research but not yet described in published literature, software must be made available to editors and reviewers. We strongly encourage code deposition in a community repository (e.g. GitHub). See the Nature Portfolio guidelines for submitting code & software for further information.

Data

Policy information about availability of data

All manuscripts must include a data availability statement. This statement should provide the following information, where applicable:

- Accession codes, unique identifiers, or web links for publicly available datasets
- A description of any restrictions on data availability
- For clinical datasets or third party data, please ensure that the statement adheres to our policy

All raw metagenomes and metagenome-assembled genomes are deposited to the NCBI Sequence Read Archive under the BioProject accession number PRJNA801081. Raw metagenome and metatranscriptome data from the Tara Oceans global dataset were downloaded from the European Nucleotide Archive (https://www.ebi.ac.uk/ena/browser/home), under the following study numbers: PRJEB1787 (metagenomes) and PRJEB6608 (metatranscriptomes). 1888 bacterial and archaeal MAGs generated by Delmont et al. (2021) were downloaded from https://www.genoscope.cns.fr/tara/. The metabolic marker protein database used in this study, which includes reference hydrogenase and carbon monoxide dehydrogenase sequences can be obtained from https://bridges.monash.edu/collections/_/5230745.

Human research participants

Policy information about studies involving human research participants and Sex and Gender in Research.

Reporting on sex and gender	This information was not collected as there were no human research participants involved in this study.
Population characteristics	This information was not collected as there were no human research participants.
Recruitment	This information was not collected as there were no human research participants.
Ethics oversight	This information was not collected as there were no human research participants.

Note that full information on the approval of the study protocol must also be provided in the manuscript.

Field-specific reporting

Please select the one below that is the best fit for your research. If you are not sure, read the appropriate sections before making your selection.

☐ Life sciences ☐ Behavioural & social sciences ☒ Ecological, evolutionary & environmental sciences

For a reference copy of the document with all sections, see nature.com/documents/nr-reporting-summary-flat.pdf

Ecological, evolutionary & environmental sciences study design

All studies must disclose on these points even when the disclosure is negative.

Study description	This study describes the significance of marine microorganisms utilising hydrogen as an alternate energy source and reveals new mechanisms by which trace gas oxidising marine microorganisms influence the biogeochemistry of global oceans. This study combines culture-dependent and culture-independent based analyses, including genome-resolved metagenomics and thermodynamic modelling to show that diverse marine bacteria consume hydrogen to support growth. Specifically, we demonstrated that microbial communities spanning tropical, temperate and subantarctic waters consume H2 at rates sufficient to support the growth of bacteria with low energy requirements and at environmentally relevant concentrations. Using axenic cultures of the ultramicrobium Sphingopyxis alaskensis, we provide the first demonstration of atmospheric H2 oxidation by a marine bacterium. Concomitantly, using the Tara Oceans dataset, we revealed that the capacity for H2 oxidation as well as carbon monoxide oxidation is

widespread amongst marine bacteria, and of increasing importance in the deeper ocean and in subantarctic waters, with H2 likely supporting mixotrophic growth, and CO supporting long-term bacterial survival.

Metagenomes (n = 14), ex-situ activity (n = 14) and dissolved H2 and CO (n = 14) measurements were prepared using subsamples of homogenised surface water and surface microlayer samples collected from each of the sampling locations (Munida Transect [8], Port Phillip Bay [4] and Heron Island [2]). Growth curves and gas consumption analysis was performed on exponential and stationary phase cultures of Sphingopyxis alaskensis (n = 3), Robiginitalea biformata (n = 3) and Marinovum algicola (n = 3). Quantitative RT-PCR was performed on S. alakensis harvested during exponential (n = 3) and stationary phase (n = 3), performed in technical duplicate.

Research sample

Surface water and surface microlayer samples were obtained from three different sampling locations, eight from the Munida Microbial Observatory Time-Series transect, four from Carrum Beach in Port Phillip Bay, Victoria and two from Heron Island in Queensland. These fourteen samples are representative of diverse marine ecosystems, spanning tropical, temperate, neritic, subtropical and subantarctic waters.

The existing Tara Oceans global dataset was analysed to provide global comparisons and to further understand the environmental drivers of trace gas oxidation by marine bacteria.It can be obtained from the European Nucleotide Archive. In addition, 1888 bacterial and archaeal MAGs generated by Delmont et al. (2021) were downloaded from https://www.genoscope.cns.fr/tara/. This represents a high quality set of bacterial and archaeal MAGs obtained from the 0.8-2000 µm planktonic cellular size fractions in surface oceans and seas collected by the Tara Oceans expeditions.

Three distinct marine bacterial species, Sphingopyxis alaskensis, Robiginitalea biformata and Marinovum algicola, were chosen due to the species harbouring hydrogenase classes resembling those identified in the MAGs (namely the group 1l and 2a [NiFe]-hydrogenases). These bacterial species are readily attainable from the authors or via the commercial supplier DSMZ, and are able to be easily grown under laboratory conditions using commercially available growth media.

Sampling strategy

For seawater samples, 1 L water samples were obtained from approximately 0 to 20 cm below the surface using sterile Schott bottles or using a manual glass plate sampler. These samples were then homogenised prior to aliquoting into serum vials in triplicate for ex-situ analysis of dissolved H2 and CO, and microcosm activity experiments. For each sampling location, one set of triplicate samples were autoclaved and used as a control. In addition, a portion of these samples were also used for DNA extraction for metagenomic analysis. For axenic culture experiments, cultures were also grown and analysed in triplicate for each condition. For qRT-PCR experiments, all biological triplicate samples, standards and negative controls were run in technical duplicate. A replication level of three is the standard sample size for environmental microbiological studies, and the results indicate consistency between replicates, with error bars included on all plots to illustrate any variation between replicates No statistical methods were used to predetermine sample size.

Data collection

Samples were collected by different authors as described within the author contribution statement (copied below). All data was recorded digitally and was analysed computationally as follows:
- Concentrations of H2 and CO gas for both microcosm and axenic culture experiments were measured using a Valco TGA-6791-W-4U-2 Trace Gas Analyzer
- DNA for shotgun metagenomic sequencing was extracted using a DNeasy PowerSoil kit and sample libraries were prepared using a Nextera XT DNA Sample Preparation Kit.
- Shotgun metagenomic sequencing was performed on an Illumina NextSeq500 platform at the Australian Centre for Ecogenomics.
- Metagenomes were quality controlled using BBDuk and BBMap
- Taxonomic profiles were generated from the metagenomes using PhyloFLASH
- Metagenome assembled genomes were assembled, binned and analysed using MEGAHIT, MetaBAT2, MaxBin2, CONCOCT, DAS_Tool, RefineM, CheckM, GTDB-Tk, Prodigal and CoverM.
- Metabolic annotations were completed using DIAMOND and normalized using the single copy ribosomal proteins available via SingleM.
- Phylogenetic analysis was carried out using ClustalW in MEGA11
- Environmental driver analysis was carried out in R using the package randomForest
- Growth of Sphingopyxis alaskensis was determined using optical density using an Eppendorf BioSpectrophotometer and analysed using GraphPad Prism 9
- Quantitative RT-PCR was conducted using a QuantStudio 7 Flex Real-Time PCR system and analysed using GraphPad Prism 9

Different authors were responsible for field sample collection (G.S., P.A.N., S.L., S.E.M., Y-J.C., A.J.K., P.L.M.C.), gas chromatography measurements (Z.F.I., G.S., T.J.), biogeochemical analysis (G.S., P.M.L., P.A.N., C.G.), thermodynamic modelling (P.M.L, C.G., F.B., P.M.L.C), culture based work (Z.F.I., T.J., G.S., T.J.W., R.C., C.G.), metagenome analysis (R.L., G.S., S.L., C.G.) and environmental driver and phylogenetic analysis (R.L., G.N., C.G.).

Timing and spatial scale

Field sample collection started on 20/03/2019 and finished on 23/07/2019, with Port Phillip Bay samples obtained on 20/03/2019, Heron Island samples obtained on 09/07/2019 and the Munida Transect samples obtained on 23/07/2019. Field samples were collected on the same day for each sampling location and were collected <1 m below the surface of the water. Ex situ microcosm activity assays were carried out immediately post field sample collection for each field site.

Growth curve and gas chromatography measurements for Sphingopyxis alaskensis, Robiginitalea biformata and Marinovum algicola were collected between 18/11/2019 and 23/12/2020; note extended timescale was due to the impact of COVID-19 related lockdowns in Melbourne, Australia. Quantitative RT-PCR analysis on S. alaskensis was carried out between 28/09/2020 and 29/10/2020.

Data exclusions

The metagenomes were filtered to remove reads corresponding to major organisms detected in the extraction control. One timepoint in the ex-situ gas chromatography measurements was removed due to machine error.

Reproducibility

All experiments and measurements undertaken were successful, and replicates were consistent within so we did not attempt to repeat the experiments beyond the technical and biological replications used within each experiment.

Randomization	Samples that were obtained in the field were returned to the laboratory, homogenized and sub-sampled for the various treatments for the ex situ microcosm and gas chromatography experiments. For axenic culture experiments, bacterial cultures were independently inoculated into triplicate flasks for growth, gas chromatography and qRT-PCR experiments, leading to three biologically distinct populations of bacterial cells per treatment condition.
Blinding	Blinding was not relevant to our study as it relates to solely microbial processes and experimental data was collected electronically and analysed digitally.

Did the study involve field work? ☒ Yes ☐ No

Field work, collection and transport

Field conditions	For the Port Phillip Bay and Heron Island samples, these were collected manually from shallow water (~1m deep) on fine weather days. Carrum Beach is situated in Port Phillip Bay, ~ 50 km from the central business district of Melbourne, Victoria, Australia. The site is accessible to general public. Port Phillip Bay is a relatively large, shallow embayment (1930km2, mostly < 8 m deep, max depth 24 m) which provides protection to its beaches from severe weather. Seawater temperature ranged from 13.6 to 20.3°C based on historical record (https://www.seatemperature.org/australia-pacific/australia/port-phillip.htm). Heron Island is a natural coral cay located approximately 80km north east of Gladstone, Queensland, Australia. It is accessible to the public by boat or helicopter. Seawater temperatures range from 21.5 to 27.3°C based on historical record (https://seatemperature.info/heron-island-water-temperature.html)/. The Munida Microbial Observatory Time Series is a surface water transect that extends 65 km off the coast of Otago, New Zealand, and includes neritic, subtropical and sub-Antarctic surface waters (https://www.otago.ac.nz/mots/about/). It is only accessible by a research vessel moored at Otago University, New Zealand. Surface water temperatures range from 9 to 13°C based on historical records (https://www.st.nmfs.noaa.gov/copepod/time-series/nz-10101/).
Location	Eight samples were collected from across the Munida Microbial Observatory Time-Series transect (Otago, New Zealand) on 23/07/2019, in calm weather, on the RV Polaris II. Eight equidistant stations were sampled travelling east, ranging from approximately 15 km to 70 km from Taiaroa Head (~45.7828° S, 170.7333° E to ~45.8°S, 171.6° E). At each station, water was collected at 1 m depth using Niskin bottles. Four samples were also collected from the temperate Port Phillip Bay at Carrum Beach (38.0765° S, 145.1206° E, Victoria, Australia) on 20/03/2019 and two were collected from the tropical Heron Island (23.4423° S, 151.9148° E, Queensland, Australia) on 9/7/2019. At both sites, near-shore surface microlayer and surface water samples were collected in the subtidal zone (water depth ca. 1 m; sampling depth of ~20 cm). At Port Phillip Bay, two samples were also collected at 7.5 km and 15 km east of the mouth of the Patterson River, labelled 'Intermediate' and 'Centre' respectively. All samples were collected at sea level.
Access & import/export	Suitable footwear, gloves and equipment, including pre-sterilised Schott bottles, were used during the sampling process to minimize anthropogenic effects for all samples collected. Water samples were collected in compliance with Australian biosecurity regulations, with those from the Munida Transect also subject to New Zealand biosecurity laws.
Disturbance	The study was minimally disruptive to the sampling locations as water was only collected from the surface and surface microlayer using Schott bottles and manual glass plate samplers, respectively. In total, <20 L of water was collected from the Munida Transect, <15 L of water was collected from Port Phillip Bay and <10 L of water was collected from Heron Island. Disturbance was negligible as small volumes of samples were collected from each sampling location in a non-invasive manner.

Reporting for specific materials, systems and methods

We require information from authors about some types of materials, experimental systems and methods used in many studies. Here, indicate whether each material, system or method listed is relevant to your study. If you are not sure if a list item applies to your research, read the appropriate section before selecting a response.

Materials & experimental systems

n/a	Involved in the study
☒	☐ Antibodies
☒	☐ Eukaryotic cell lines
☒	☐ Palaeontology and archaeology
☒	☐ Animals and other organisms
☒	☐ Clinical data
☒	☐ Dual use research of concern

Methods

n/a	Involved in the study
☒	☐ ChIP-seq
☒	☐ Flow cytometry
☒	☐ MRI-based neuroimaging