



# Mercury methylation by metabolically versatile and cosmopolitan marine bacteria

Heyu Lin<sup>1</sup> · David B. Ascher<sup>2,3</sup> · Yoochan Myung<sup>2,3</sup> · Carl H. Lamborg<sup>4</sup> · Steven J. Hallam<sup>5,6</sup> · Caitlin M. Gionfriddo<sup>7,10</sup> · Kathryn E. Holt<sup>8,9</sup> · John W. Moreau<sup>1,11</sup>

Received: 3 June 2020 / Accepted: 17 December 2020  
© The Author(s) 2021. This article is published with open access

## Abstract

Microbes transform aqueous mercury (Hg) into methylmercury (MeHg), a potent neurotoxin that accumulates in terrestrial and marine food webs, with potential impacts on human health. This process requires the gene pair *hgcAB*, which encodes for proteins that actuate Hg methylation, and has been well described for anoxic environments. However, recent studies report potential MeHg formation in suboxic seawater, although the microorganisms involved remain poorly understood. In this study, we conducted large-scale multi-omic analyses to search for putative microbial Hg methylators along defined redox gradients in Saanich Inlet, British Columbia, a model natural ecosystem with previously measured Hg and MeHg concentration profiles. Analysis of gene expression profiles along the redoxcline identified several putative Hg methylating microbial groups, including Calditrichaeota, SAR324 and Marinimicrobia, with the last the most active based on *hgc* transcription levels. Marinimicrobia *hgc* genes were identified from multiple publicly available marine metagenomes, consistent with a potential key role in marine Hg methylation. Computational homology modelling predicts that Marinimicrobia HgcAB proteins contain the highly conserved amino acid sites and folding structures required for functional Hg methylation. Furthermore, a number of terminal oxidases from aerobic respiratory chains were associated with several putative novel Hg methylators. Our findings thus reveal potential novel marine Hg-methylating microorganisms with a greater oxygen tolerance and broader habitat range than previously recognized.

**Supplementary information** The online version of this article (<https://doi.org/10.1038/s41396-020-00889-4>) contains supplementary material, which is available to authorized users.

✉ John W. Moreau  
john.moreau@glasgow.ac.uk

<sup>1</sup> School of Earth Sciences, The University of Melbourne, Parkville, VIC 3010, Australia

<sup>2</sup> Structural Biology and Bioinformatics, Department of Biochemistry and Molecular Biology, Bio21 Molecular Science and Biotechnology Institute, The University of Melbourne, Parkville, VIC 3010, Australia

<sup>3</sup> Computational Biology and Clinical Informatics, Baker Heart and Diabetes Institute, PO Box 6492, Melbourne, VIC 3004, Australia

<sup>4</sup> Department of Ocean Sciences, University of California, Santa Cruz, CA 95064, USA

<sup>5</sup> Department of Microbiology and Immunology, University of British Columbia, Vancouver, BC V6T 1Z1, Canada

## Introduction

Mercury (Hg), a highly toxic metal, is widespread in the environment from primarily anthropogenic sources, leading to increased public concern over the past few decades. [e.g., 1–3]. Methylmercury (MeHg) is recognized as a

<sup>6</sup> Genome Science and Technology Program, University of British Columbia, Vancouver, BC V6T 1Z4, Canada

<sup>7</sup> Biosciences Division, Oak Ridge National Laboratory, PO Box 2008, Oak Ridge, TN 37831, USA

<sup>8</sup> Department of Infectious Diseases, Central Clinical School, Monash University, Monash, VIC 3800, Australia

<sup>9</sup> Department of Infection Biology, London School of Hygiene & Tropical Medicine, London WC1E 7HT, UK

<sup>10</sup> Present address: Smithsonian Environmental Research Center, Edgewater, MD 21037, USA

<sup>11</sup> Present address: Currently at School of Geographical & Earth Sciences, University of Glasgow, Glasgow G12 8QQ, UK

potent neurotoxin that bioaccumulates through both marine and terrestrial food webs [2, 4], with potential deleterious impacts on human health. With implementation of the Minamata Convention on Mercury [3], a better understanding is expected of potential MeHg sources, in the context of global biogeochemical cycles and factors influencing Hg speciation. For example, oxygen gradients in seawater have been observed to expand to both a wider depth range and larger geographical area in response to climate change [5], and the resulting impacts on biogeochemical cycles, e.g., carbon, nitrogen, and sulfur, have been studied [6]. Effects on Hg cycling, however, are rarely considered.

The environmental transformation of Hg(II) to MeHg is a microbially-mediated process, for which the proteins are encoded by the two-gene cluster *hgcA* and *hgcB* [7]. Possession of the *hgcAB* gene pair is a predictor for Hg methylation capability [8], and the discovery of *hgc* genes has stimulated a search for potential Hg methylating microbes in diverse environments [9]. To date, all experimentally confirmed Hg methylators are anaerobes [10] from three *Deltaproteobacteria* clades (sulfate-reducing bacteria, Fe-reducing bacteria and syntrophic bacteria), one clade belonging to fermentative Firmicutes, and an archaeal clade (*Methanomicrobia*). These microorganisms are ubiquitous in soils, sediments, seawater, freshwater, and extreme environments, as well as the digestive tracts of some animals [9]. A few other *hgc* carriers, not only from anaerobic but also microaerobic habitats, were discovered using culture-independent approaches, including Chloroflexi, Chrysiogenetes, Spirochaetes [9], *Nitrospina* [11, 12], and Verrucomicrobia [13]. Two more recent studies have expanded the Hg-methylating community [14, 15] and revealed the role played by Hg methylators in the context of global biogeochemical cycling.

Apart from Hg methylation, MeHg demethylation also plays a key role in global Hg cycling. This process is mainly catalyzed by alkylmercury lyase encoded by the *merB* gene, which has been largely found in aerobes (e.g., *Bacillus cereus*, *Staphylococcus aureus*, and *Escherichia coli*). Yet some obligate and facultative anaerobes (e.g., *Geobacter bemidjensis* and *Geobacter sulfurreducens*) also possess *merB* and can contribute to demethylation [16] and therefore total environmental MeHg levels.

A global Hg survey [17] found a prevalence for MeHg in suboxic waters, where dissolved O<sub>2</sub> and H<sub>2</sub>S concentrations are typically very low (<50 μM and <10 μM, respectively), especially in regionally widespread oxygen gradients at upper and intermediate ocean depths. Oxygen concentrations can be reduced by a combination of physical and biological forcing effects [6, 18]. These gradients can exist as permanent features in the water column, impinge on coastal margins, or manifest more

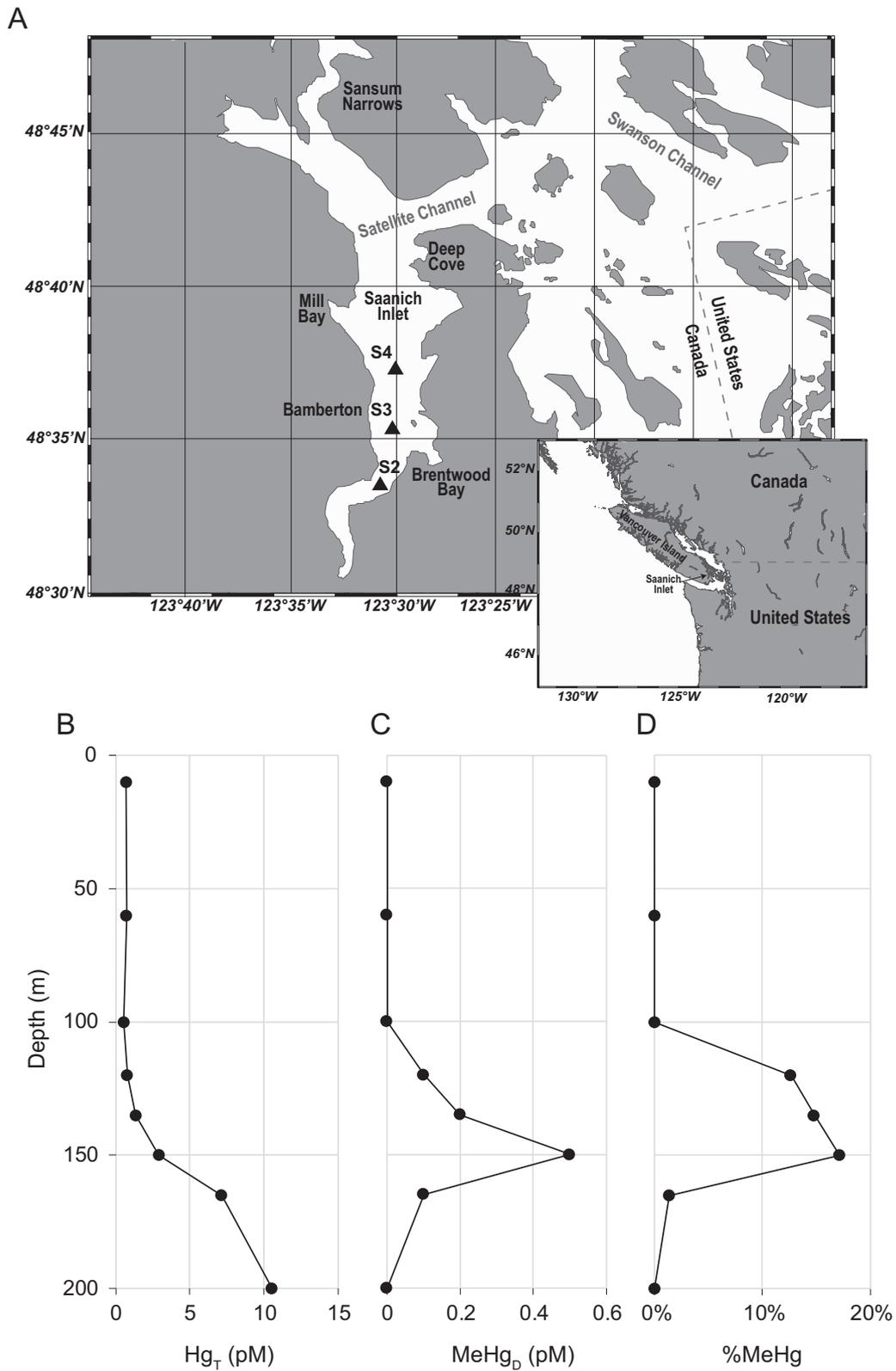
transiently (e.g., induced by phytoplankton blooms). As oxygen levels decrease, metabolic energy gets increasingly diverted to alternative electron acceptors, resulting in coupling of other biogeochemical cycles, e.g., C, N, S, Fe, and Mn [6, 19, 20]. Recent findings of microaerophilic microbial Hg methylation potential in sea ice and seawater [e.g., 11, 21–23] raise the possibility that this process contributes significantly to ocean MeHg biomagnification.

In this study, we measured the concentrations of Hg (total) and MeHg in the waters of Saanich Inlet, a seasonally anoxic fjord on the coast of Vancouver Island (British Columbia, Canada), and performed targeted metagenomic and metatranscriptomic analyses of seawater sampled from varying depths. Saanich Inlet, as a model natural ecosystem for studying microbial activity along defined oxygen gradients [24], provides an ideal site to search for novel putative Hg methylators and study microbially-mediated Hg cycling in low-oxygen environments. Physico-chemical and biogeochemical parameters, including nutrient and dissolved gas concentrations, have been published from data acquired from 2006 to 2014 [25]. This study clearly illustrates the dynamic character of the oxygen gradient in Saanich Inlet waters. Computational homology modelling was performed to predict the functionality of HgcAB proteins encoded for by putative novel Hg-methylators. Finally, we scanned global metagenomic datasets for recognizable *hgcA* genes, to reassess the environmental distribution of microbial mercury methylation potential.

## Results and discussion

### Hg and MeHg concentrations along redox gradients in Saanich Inlet

Concentrations of total dissolved Hg (Hg<sub>T</sub>) and monomethylmercury (MeHg) from filtered Saanich Inlet S3 station (see Fig. 1A for a map) seawater samples, obtained and analyzed in March–April of 2010, were vertically profiled at eight different depths from surface (10 m) to bottom (200 m) waters. The concentration of Hg<sub>T</sub> at sea surface was ~0.70 pM and remained nearly constant in seawater above 120 m depth, increasing to 1.35 pM and ~10.56 pM at 135 m and 200 m depths, respectively (Fig. 1B). MeHg was below detection limit (<0.1 pM) for seawater above 100 m depth, but increased to 0.50 pM (17.2% of total Hg) at 150 m depth. However, MeHg then decreased to 0.1 pM at 165 m depth, becoming undetectable at 200 m (Fig. 1C and 1D). Similar MeHg peaks under suboxic conditions have been observed in other seawater depth profiles: the Pacific Ocean [26], Arctic Ocean [27], Southern Ocean [28], Arabian Sea [29], and Mediterranean Sea [30].



**Fig. 1 Map of Saanich Inlet and mercury profiles of station “S3”.** **A** Map of Saanich Inlet showing locations of the station “S2”, “S3” and “S4” by triangles. **B** Concentration of total dissolved Hg at station

“S3”. **C** Concentration of dissolved MeHg at station “S3”. **D** MeHg as a percentage of total dissolved Hg at station “S3”.

## MAG (metagenome-assembled genome) reconstruction and putative Hg methylator identification

A total of 2088 MAGs with completeness >70% and contamination <5% were recovered from all Saanich Inlet metagenomic datasets (generated across all sampling depths, 96% being from station S3, obtained from 2008 to 2013; [24]). From these, 56 MAGs, belonging to seven phyla: Proteobacteria, Marinimicrobia, Verrucomicrobia, Firmicutes, Calditrichaeota, Spirochaetes, and Nitrospinae (Table S1), were identified as having genes homologous with *hgcA* (Fig. S1). Among these, Marinimicrobia and Calditrichaeota have not previously been implicated in Hg methylation. No *hgcAB* genes were recovered from archaeal MAGs, suggesting that archaeal Hg methylators were rare or absent. As most known Hg-methylating microorganisms derive from a limited number of phylogenetic groups [8, 14, 15], our findings expand the database of putative microbial Hg-methylators and their microenvironments.

Fifteen *hgcA*-carrying MAGs represented members of Marinimicrobia, a widespread but uncultured marine phylum that couples C, N, and S cycling [31]. These Marinimicrobia contained the same *hgcA* gene sequence (100% nucleotide identity), and uniformly possessed *hgcB* genes downstream of *hgcA*. Marinimicrobia genome association with *hgcAB* was supported strongly by emergent self-organizing maps (ESOMs; Fig. S2A). One Marinimicrobia-associated MAG, SI037\_bin139, exhibited the highest binning quality score (Table S1), with 97.8% completeness and 0% contamination, as estimated by CheckM, having 75 contigs in total. Many Marinimicrobia MAGs also carried 16S rRNA genes, further supporting taxonomic classification (Fig. S3). Notably, Marinimicrobia-HgcA sequences fell within the *Euryarchaeota* in the HgcA phylogenetic tree (Fig. 2; see Fig. S4 for a more detailed tree with accession numbers and strain names), suggesting a different evolutionary pathway to other bacterial *hgc* sequences.

Three MAGs containing *hgcAB* genes were associated with Calditrichaeota, another cosmopolitan marine phylum [32]. HgcA sequences from Calditrichaeota formed a cluster with those of the PVC group, Planctomycetes, Acidobacteria, and Candidatus *Brocadia* spp., which also have close phylogenetic relationships to *Deltaproteobacteria*, as shown in Fig. 2. ESOMs were constructed (Fig. S2B) that supported the presence of *hgcAB* genes in Calditrichaeota MAGs.

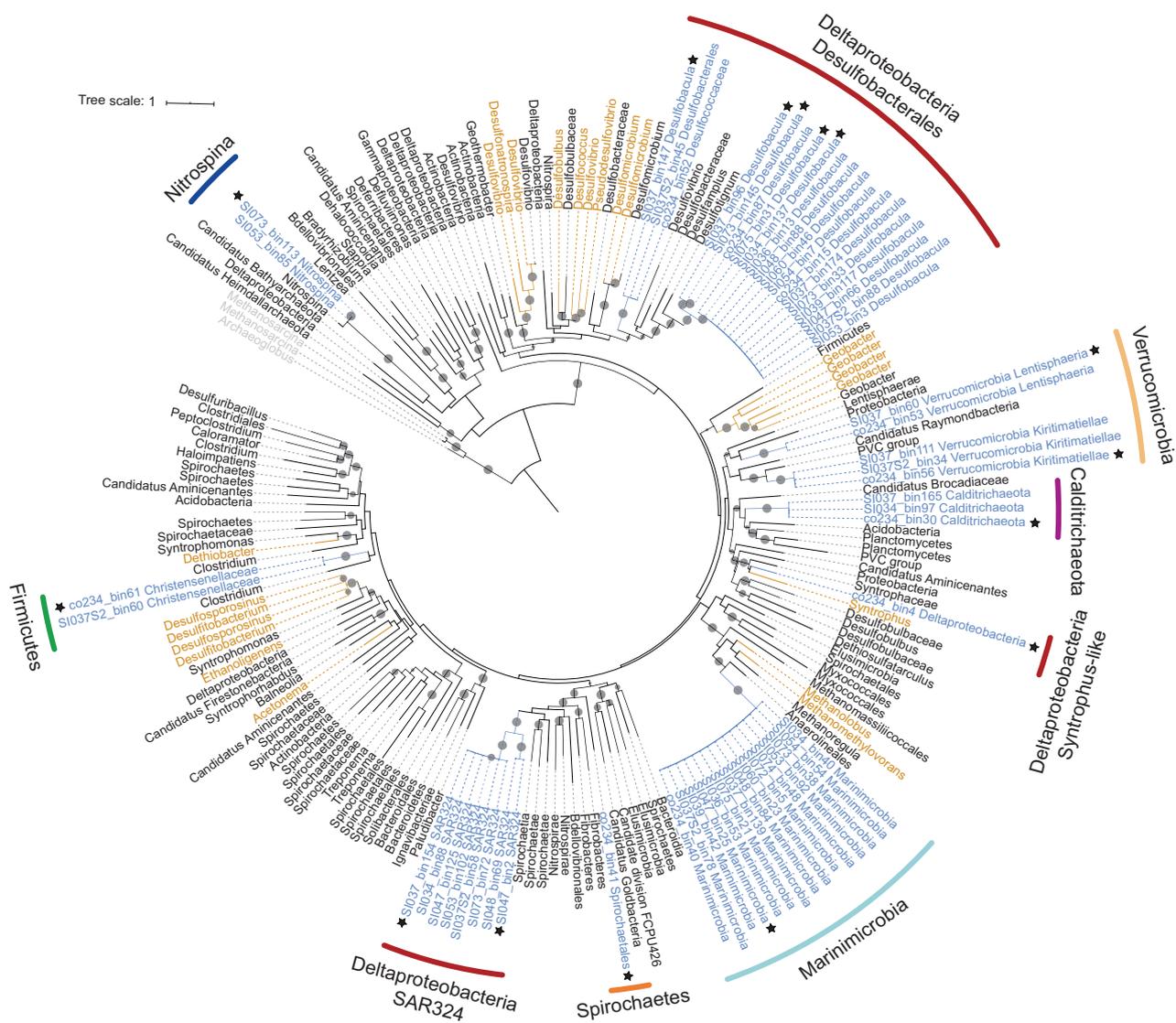
HgcA sequences were found in eight Proteobacteria MAGs, all belonging to the *Deltaproteobacteria* group SAR324. *Deltaproteobacteria* are the most diverse Hg-methylating clade currently known [9]. However, the SAR324 *hgcA* sequences clustered separately from those of three previously described *Deltaproteobacteria* methylating

clades (*Desulfovibrio*-like, *Geobacter*-like, and *Syntrophus*-like) [8, 9], forming a potentially novel clade of Hg-methylators (Fig. 2). SAR324 is affiliated with a group of *Deltaproteobacteria* abundant in the deep sea and in low-oxygen settings, with the ability to metabolize sulfur, organic carbon, and C<sub>1</sub> compounds [33]. A contiguous *hgcB* gene was also recognized downstream of *hgcA* in SAR324 MAGs SI047\_bin2 and SI048\_bin69. Six other SAR324 MAGs did not contain recognizable *hgcB* genes, but several *hgcB*-like candidates encoding for tandem [CX<sub>2</sub>CX<sub>2</sub>CX<sub>3</sub>C] motifs were recognized in different contigs in the genome. The presence of *hgc* genes in SAR324 MAGs was well supported by ESOM (Fig. S2A).

## Relative abundance of *hgcA* genes and transcripts in Saanich Inlet

The 56 *hgcA* sequences were clustered into 15 groups using a 99% sequence identity threshold. One *hgcA* sequence was selected arbitrarily from each group, with corresponding MAG (Table 1; also indicated by stars in Fig. 2), for further analysis. These sequences were used to recruit reads from Saanich Inlet metagenomic and metatranscriptomic datasets, in order to calculate relative abundance and expression. Results showed that *hgcA* genes were widely distributed in Saanich Inlet water samples, and the abundance of *hgcA* gene copies increased with depth, peaking at ~200 m depth (Fig. 3A). Since *hgcA* is usually present as a single copy per genome, this finding suggests that Hg methylators were most abundant at this depth. Notably, Marinimicrobia was the most abundant putative Hg-methylator detected throughout the water column, accounting for >2% of the microbial community for many samples from 200 m depth (Fig. 3B). *Deltaproteobacteria* with *hgcA* also showed higher abundance in 200 m depth samples, suggesting an overlapping habitat range with Marinimicrobia. Although *hgcA*-carrying *Nitrospina* was not a dominant phylum in Saanich Inlet datasets, the relative abundance of *Nitrospina*-hosted *hgcA* remained nearly constant from sea surface to bottom water, implying the adaptability of this phylum across the redoxcline.

Reads per kilobase per million mapped reads (RPKM) values were calculated for *hgcA* transcripts in each sample to assess *hgc* gene expression in a subset of 36 samples from site S3 (Fig. 3C). Consistent with relative DNA abundance, *hgcA* transcripts also increased with depth to a maximum at 200 m (mean RPKM = 5.1); the most abundant *hgcA* transcripts were associated with Marinimicrobia (Fig. 3D). The abundance of *hgcA* transcript was normalized to housekeeping genes *rpoB* and *gyrB* to study the *hgcA* expression profile (Figure S5). Intriguingly, both the ratio of *hgcA/rpoB* RPKM and *hgcA/gyrB* RPKM increased with depth, but were similar within uncertainties for bottom



**Fig. 2** Maximum-likelihood phylogenetic tree of *HgcA* amino acid sequences (1000 ultrafast bootstrap replicates; values >90% are shown by black dots at the nodes). *HgcA* sequences recovered in this study are highlighted in blue. *HgcA* sequences retrieved from public databases are shown in black. Experimentally confirmed *HgcA* from previous studies are shown in brown. *HgcA* paralogues from

non-methylating bacteria were used as outgroups and are shown in gray. The 15 representative *HgcA* sequences used in this study are indicated by stars. Taxonomic classifications of the *hgcA*-carrying genomes are labeled in the outer circle by different colors. Scale bar indicates substitutions per site.

waters (from 0 to ~0.6 and ~1.0, respectively). Considering the substantial increase of total *hgcA* transcripts with depth (Fig. 4C) this observation suggests a relatively constant rate of *hgcA* expression by individual microorganisms. Therefore, the increase in *hgcA* transcripts is likely explained by increasing abundance of Hg methylators rather than by an increase in cell-specific *hgcA* expression. This result is consistent with previous studies that showed *hgcA* expression is constitutive [34, 35]. Furthermore, no *hgcA* transcripts were recovered from 10 m depth samples, although a number of *hgcA* DNA sequences were recovered. In addition, no *hgcA* transcript associated with Verrucomicrobia

(*Lentisphaeria* and *Kiritimatiellae*) was detected for any dataset, which may reflect the relatively lower abundance of Verrucomicrobia-*hgcA* genes, as well as a low transcription level. Predicted amino acid sequences from Saanich Inlet metaproteomic datasets were also scanned for expression of putative *hgcA* genes (Fig. S6; also indicated by stars in Fig. 3D). Excepting from phyla Calditrichaeota, Verrucomicrobia, Spirochaetes, and Firmicutes, representative *HgcA* sequences could be detected in matched metaproteomic datasets. Notably, of many *hgcA* DNA sequences from the upper depths of Saanich Inlet, only a small number were actively expressed (Fig. 3C and 3D), suggesting an

**Table 1** Summary of 15 representative *hgcA*-carrying MAGs<sup>a</sup>.

MAGs	%Compl. <sup>b</sup>	%Conta. <sup>b</sup>	Size (Mbp)	# Contigs	%GC	16S	Taxonomy
SI034_bin137	<b>97.58</b>	<b>2.992</b>	4.94	304	40		<i>Desulfobacula</i> sp.
SI037_bin96	71.02	<b>3.87</b>	2.76	633	39.5		<i>Desulfobacula</i> sp.
SI073_bin145	87.48	5.322	4.26	657	39.2		<i>Desulfobacula</i> sp.
SI075_bin31	88.61	<b>2.293</b>	4.89	394	40.1		<i>Desulfobacula</i> sp.
co234_bin4	81.15	5.545	5.91	1000	50.1		Order <i>Desulfatiglandales</i>
SI037_bin147	89.67	<b>1.935</b>	3.28	289	38.6		Order <i>Desulfobacterales</i>
SI037_bin154	89.02	<b>3.361</b>	7.85	512	45.5		SAR324 group
SI047_bin2	74.15	<b>2.941</b>	3.63	142	36.2		SAR324 group
SI047_bin25	<b>96.7</b>	<b>0</b>	3.17	92	37.4	+	Phylum Marinimicrobia
co234_bin30	<b>96.08</b>	<b>1.648</b>	4.54	293	42.9		Phylum Calditerrhaeota
SI037_bin60	83.67	<b>3.914</b>	8.81	2073	63.9		Class <i>Lentisphaeria</i>
co234_bin56	<b>95.6</b>	<b>2.027</b>	4.85	162	60.6	+	Class <i>Kiritimatiellae</i>
co234_bin61	<b>98.25</b>	<b>0</b>	2.67	97	47.7		Order <i>Christensenellales</i>
co234_bin41	72.27	<b>1.253</b>	4.68	1025	48.9		Order <i>Spirochaetales</i>
SI073_bin113	74.89	<b>1.709</b>	2.92	321	43.1		<i>Nitrospina</i> sp.

<sup>a</sup>Representative MAGs selected by 99% *hgcA* identity threshold. See Table S1 for all 56 *hgcA*-carrying MAGs.

<sup>b</sup>Completeness above 95% and contamination below 5% are shown in bold; quality of MAGs was determined by CheckM with the “lineage\_wf” pipeline.

influence of oxygen level or other environmental factors. We therefore caution that abundance of *hgcA* genes, as measured by metagenomic or amplicon sequencing, cannot effectively predict the actual extent of MeHg production.

### Abundance of *merB* genes and transcripts in Saanich Inlet

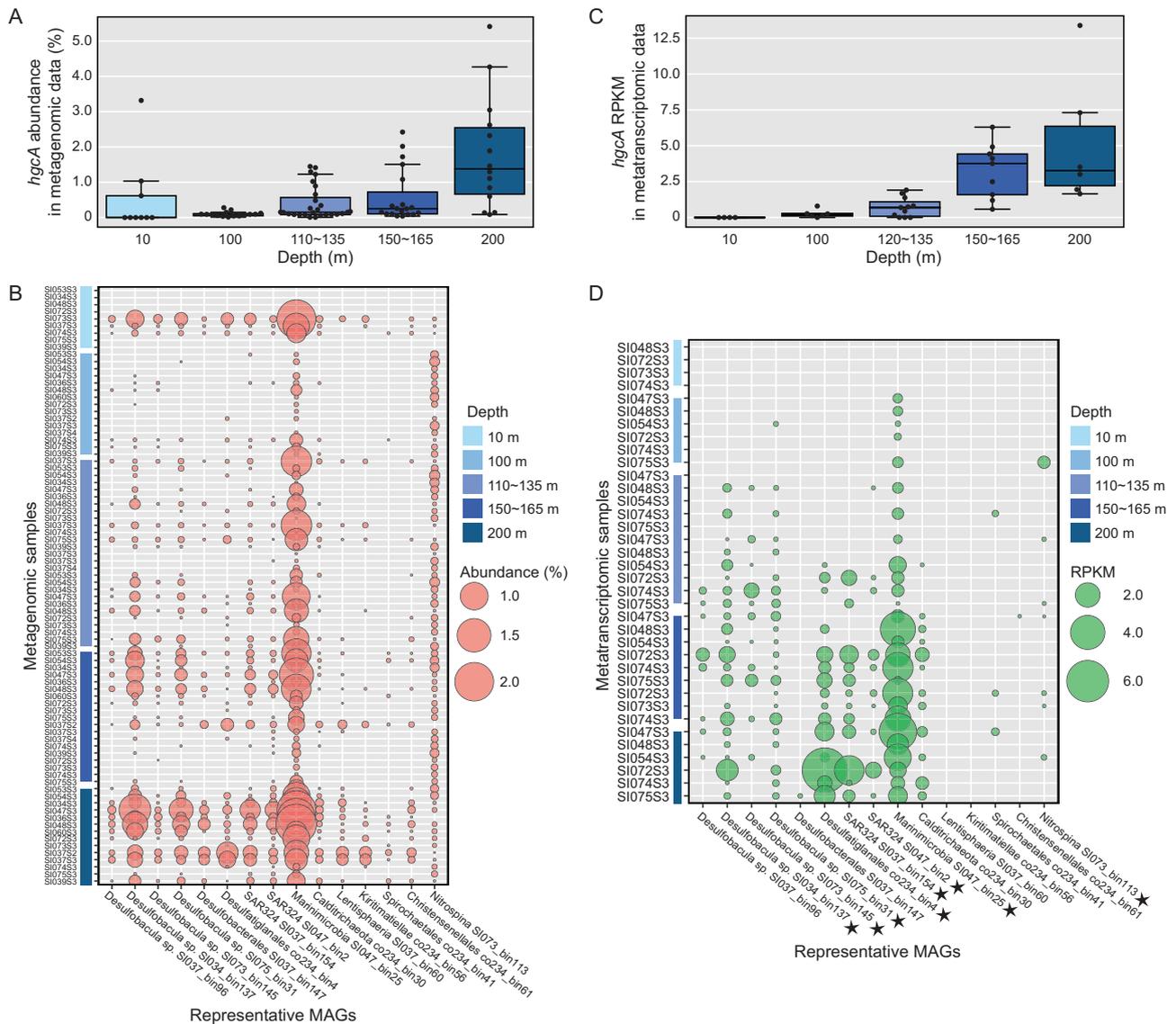
As MeHg concentrations in the Saanich Inlet water column may reflect net Hg methylation and demethylation, the distribution with depth of *merB*, a gene encoding for an organomercury lyase, was also assessed. Results showed the relative abundance of *merB* genes peaked at 200 m depth, as ~0.6% of total microbial community genes (Fig. 4A). RPKM values of *merB* transcripts in each metatranscriptomic sample were calculated and compared with *hgcA* transcripts. Consistent with metagenomic results, *merB* transcripts peaked at 200 m depth (mean RPKM = 5.2), whereas no *merB* transcripts were detected at 10 m depth (Fig. 4B). Average RPKM values of *merB* were higher than those of *hgcA* at depths <120 m, but were exceeded by *hgcA* RPKM values at 120–200 m depth. Interestingly, the average RPKM of *hgcA* was again exceeded by that of *merB* at 200 m depth, although transcription of both genes increased to maxima at that same depth (Fig. 4C). Contrasting the relative abundance of potential methylators observed in the benthic zone, the observed increased abundance of *merB* genes and transcripts may provide a viable hypothesis for why MeHg levels decreased over certain depth ranges (Fig. 1D).

### Phylogenetic analysis of *hgcA*-carrying Marinimicrobia

All 15 *hgcA*-carrying Marinimicrobia and 409 non-*hgcA*-carrying Marinimicrobia spp., derived from this study and public databases (NCBI and IMG), were used to build a phylogenetic tree based on concatenated alignment of housekeeping genes (Fig. 5A, Fig. S7). All 15 *hgcA*-carrying Marinimicrobia clustered into a monophyletic clade and carried identical *hgcA* gene sequences, consistent with a single horizontal transfer of *hgcA* into a sublineage of Marinimicrobia. Furthermore, while all Marinimicrobia-associated *hgcA* sequences were identical, four different 16S rRNA genes were represented by these MAGs, with minimum identity of 96.0% (Fig. S3). The four representative 16S sequences from *hgcA*-carrying Marinimicrobia were further searched in the SILVA database [36], and all of them have best hits associated with Marinimicrobia phyla, with identities ranging from 91.17 to 95.71%. This finding suggests that *hgcA*-carrying Marinimicrobia represent a novel candidate species or genus, according to Yarza et al. [37].

### Marinimicrobia is a widespread potential Hg methylator

The phylum Marinimicrobia, a yet uncultivated but apparently cosmopolitan marine group, is thought to play an important role in global biogeochemical cycling, especially across redox gradients [31, 38]. However, the

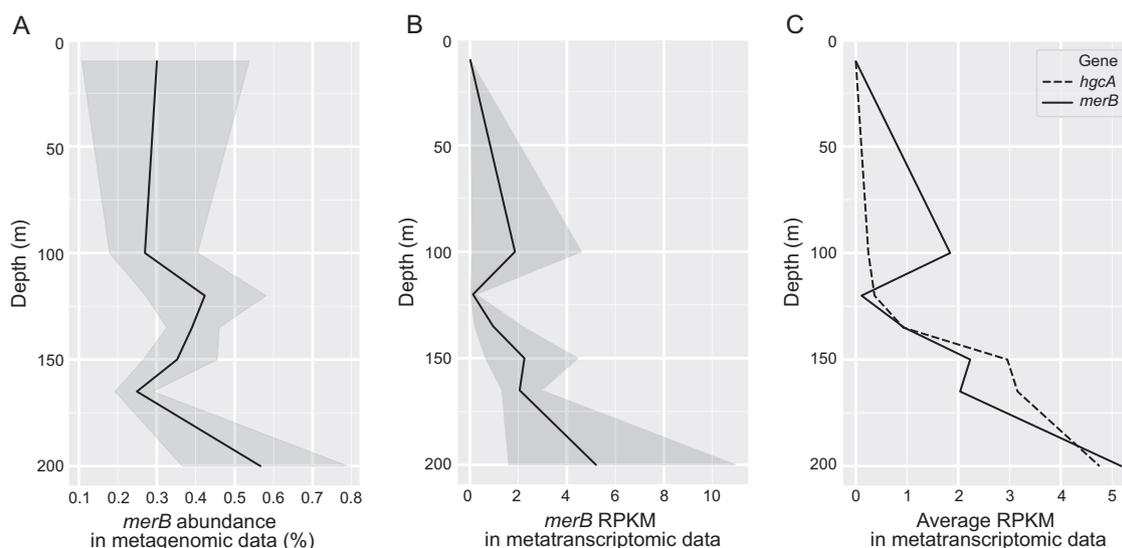


**Fig. 3** Relative abundance of representative *hgcA* genes and transcripts in different samples from Saanich Inlet. Gene abundance is normalized by gene length and genome equivalent for each sample, and transcript abundances are represented as RPKM values. **A** Relative abundance of the sum of representative *hgcA* genes for different depths. **B** Relative abundance of different representative *hgcA* genes from different metagenomic datasets. Larger circles indicate a

higher percentage of the whole microbial community. Different shades of blue indicate different depths from which the samples were taken. **C** RPKM values of the sum of representative *hgcA* transcripts in different depths. **D** RPKM values of different representative *hgcA* transcripts in different metatranscriptomic samples. *HgcA* sequences from metaproteomic samples are indicated with stars.

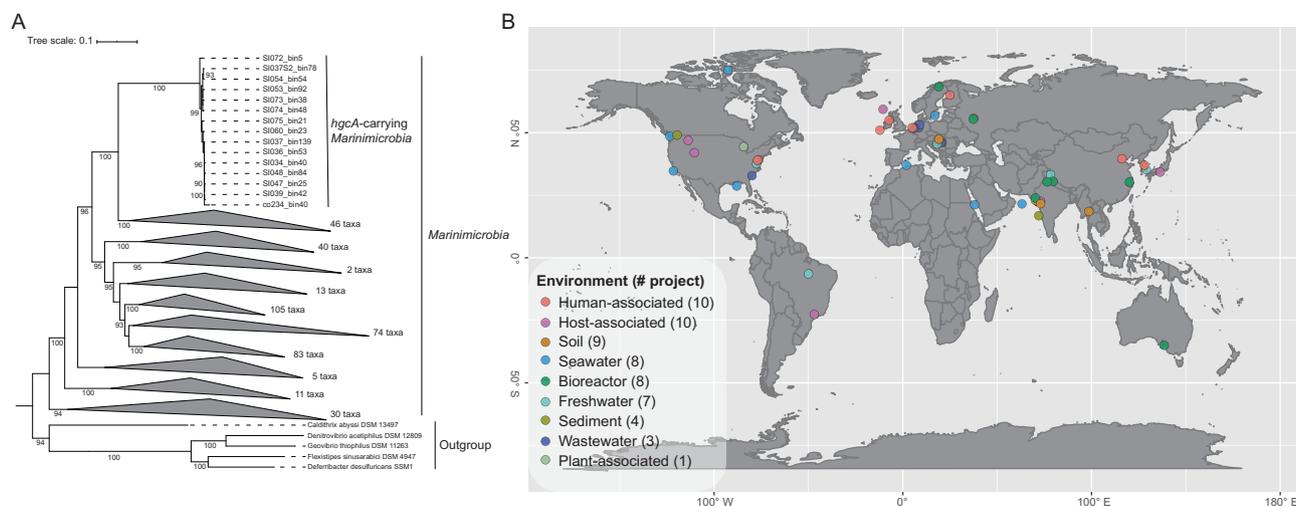
Hg-methylation potential of Marinimicrobia has not previously been recognized. A few HgcA sequences from Saanich Inlet were discovered by Podar et al. [9] from metagenomic analyses; these previously unidentified sequences are shown here to co-locate phylogenetically with the Marinimicrobia-HgcA identified in this study at whole genome resolution. In order to explore the global distribution of *hgcA*-carrying Marinimicrobia, a wider search was performed against NCBI SRA metagenomic datasets. In total, 58 BioProjects from a range of

environments, including seawater, soil, sediment, freshwater, industrial wastewater, and plant rhizomes, contained Marinimicrobia-*hgcA* associated reads (Fig. 5B, Table S2), suggesting a wide ecological distribution. Interestingly, our findings of *hgcA*-carrying Marinimicrobia from metagenomic datasets from Gulf of Mexico (PRJNA288120) and Canadian Arctic (PRJNA266338) waters contaminated by oil spills [39, 40] may provide one explanation for observed in situ MeHg production [41].



**Fig. 4** DNA and RNA abundance of *merB* in Saanich Inlet datasets changing with depth. **A** *merB* gene abundance in Saanich Inlet metagenomic datasets. Percentage abundance was normalized by gene length and genome equivalent for each sample. Line plot depicts the average abundance of *merB* with depth; shaded area depicts 95% confidence interval. **B** *merB* transcript abundance in Saanich Inlet

metatranscriptomic datasets, as represented by RPKM values. Line plot depicts mean RPKM value of *merB* with depth; shaded area depicts 95% confidence interval. **C** Comparison between RPKM values for *hgcA* (orange) and *merB* (green) in Saanich Inlet metatranscriptomic datasets.



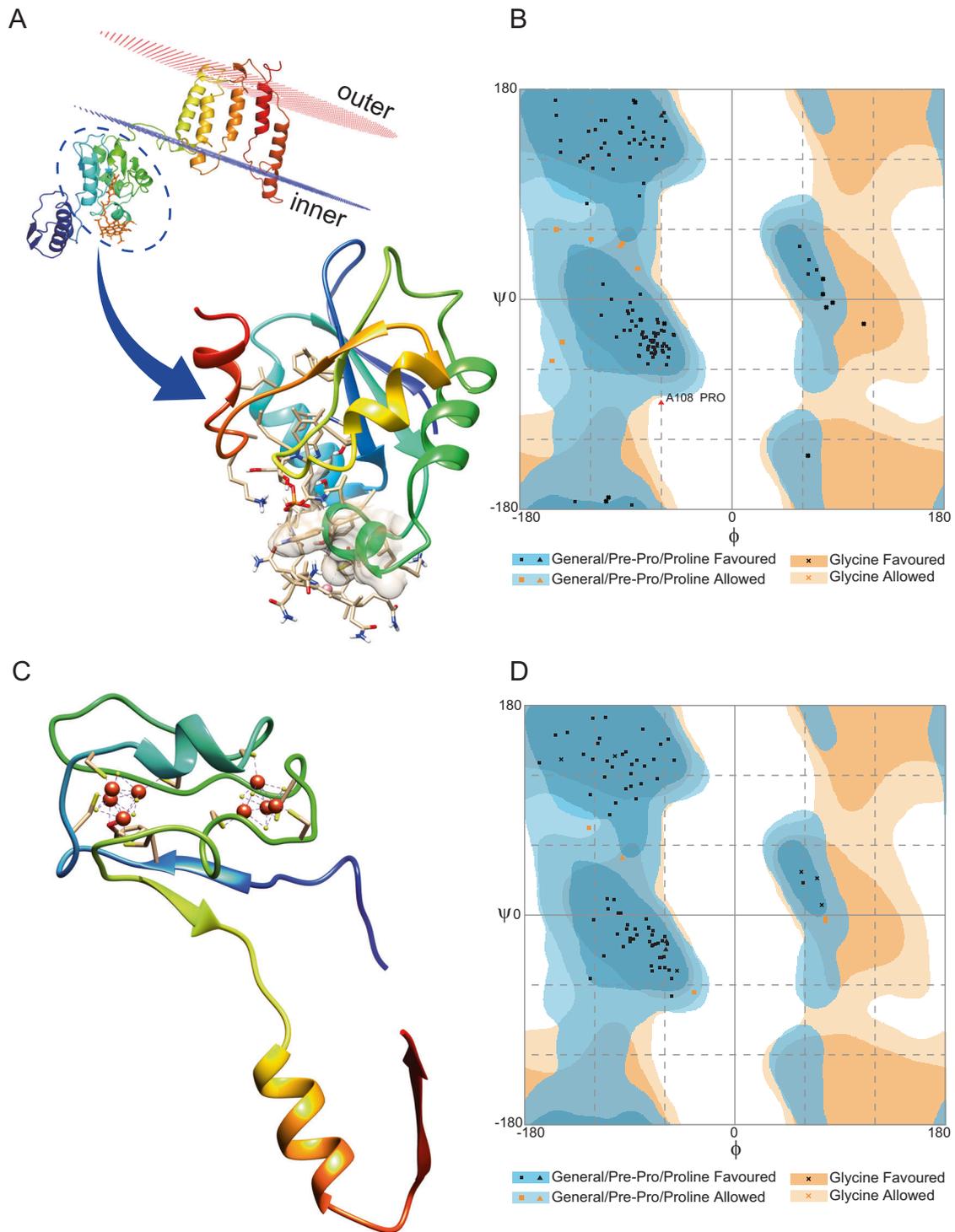
**Fig. 5** Phylogeny and global distribution of phylum *Marinimicrobia*. **A** Maximum-likelihood phylogenetic tree of phylum *Marinimicrobia* based on concatenated housekeeping genes (1000 ultrafast bootstrap replicates; values >90% are shown at the nodes). Scale bar indicates substitutions per site. A total of 424 *Marinimicrobia* genomes were used for the tree; *Marinimicrobia* lineages

without *hgcA* genes were collapsed to simplify presentation (see Fig. S6 for a more detailed tree). **B** Distribution of *hgcA*-carrying *Marinimicrobia* in various environments globally, shown as different colors; total numbers of BioProjects from each environment are shown in brackets.

## Homology models of *Marinimicrobia*-HgcA and -HgcB proteins

As no cultivated representative of *Marinimicrobia* is currently available to assay for functional Hg methylation, we constructed homology models to test the hypothesis that HgcA and HgcB proteins encoded for by *Marinimicrobia*

possess the three-dimensional structural functionality required for methylating Hg. Analysis of the putative *Marinimicrobia*-HgcA sequence revealed a comparable domain structure (Fig. 6A) to the previously characterized HgcA from the functionally validated Hg-methylating strain *Desulfovibrio desulfuricans* ND132 [7], including the presence of a globular domain at the N-terminus and five



**Fig. 6 Three-dimensional homology models of Marinimicrobia-HgcA and -HgcB proteins.** **A** Model of full-length Marinimicrobia-HgcA, shown relative to the cell membrane. An enlarged view of the functional domain bound to cobalamin is shown, with the ionic interaction between cysteine and cobalt represented by a dotted line. **B** Ramachandran plot of Marinimicrobia-HgcA model showing that 93.2%, 5.9%, and 0.8% of residues lie within favoured, allowed, and outlier regions, respectively.

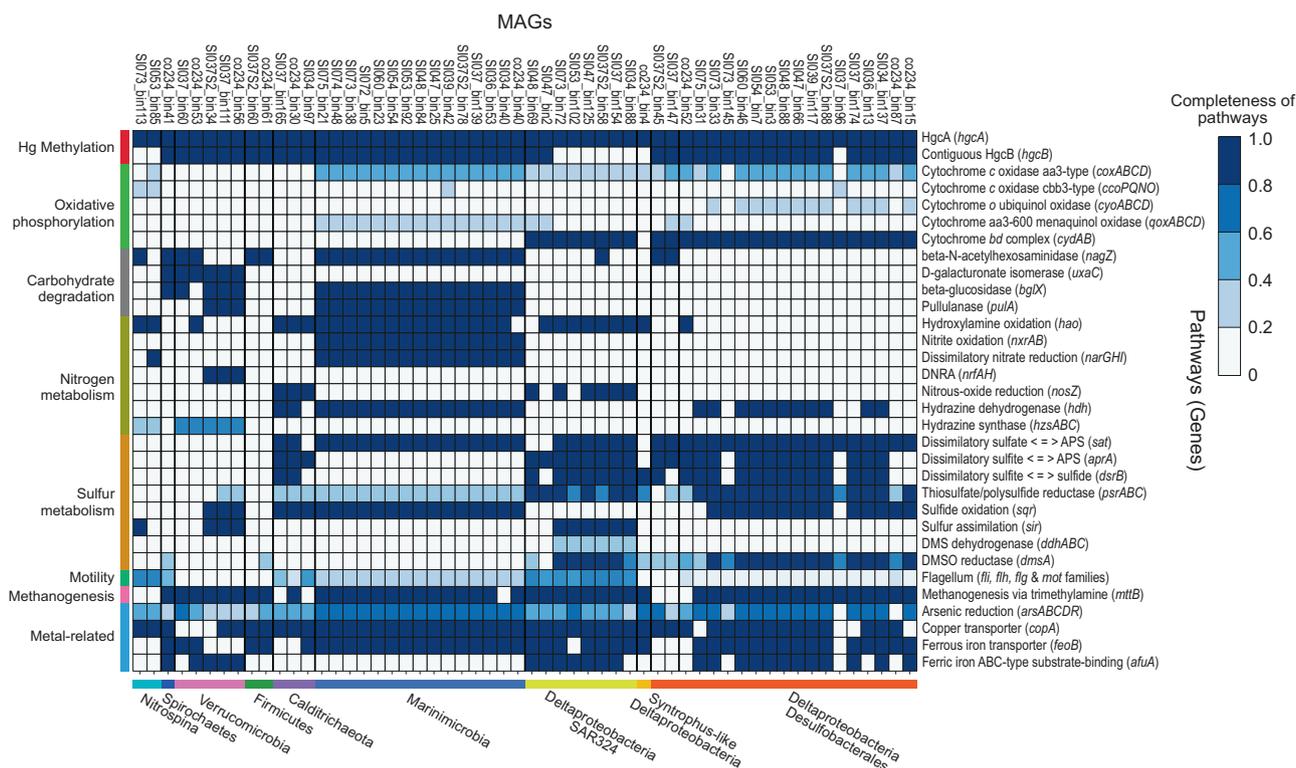
The backbone Phi and Psi ( $\phi$  and  $\psi$ ) dihedral angles are shown, along with the energetically favoured regions. **C** Model of full-length Marinimicrobia-HgcB in complex with two [4Fe4S] clusters. Interactions between the protein and iron are shown by dotted lines. **D** Ramachandran plot of Marinimicrobia-HgcB model is shown, revealing 94.1%, 5.9%, and 0% of residues lay within favoured, allowed, and outlier regions, respectively.

transmembrane spanning helices at the C-terminus. The cap-helix structure and coordinated Cys residues required for Hg(II) methylation by ND132 were conserved in the globular domain of our putative *Marinimicrobia-HgcA* (Fig. S8A and Fig. S8B), consistent with a conserved catalytic mechanism. The structure of ND132-HgcA was solved bound to the co-factor cobalamin required for catalytic methylation activity. The cobalamin binding pocket for ND132 was comparable to that of the *Marinimicrobia-HgcA* homology model (726.5 Å<sup>3</sup> and 678.4 Å<sup>3</sup>, respectively), with 69% conservation of residues interacting with cobalamin. Modelling of cobalamin in the binding site of the *Marinimicrobia-HgcA* homology model revealed a similar H-bonding network for the two structures (Table S3). Modelling of the *Marinimicrobia-HgcB* protein (Fig. 6C) demonstrated binding with 2[4Fe-4S] clusters at the N-terminal, similar to the structure of ND132-HgcB (Fig. S8C and Fig. S8D). The conserved two ferredoxin motifs (CX2CX2CX3C), and the C-terminal Cys tail in ND132-HgcB required for potentially transferring electrons and binding the Hg(II) substrate [7, 42], were also observed for *Marinimicrobia-HgcB*. The structural qualities of *Marinimicrobia-HgcA* and -HgcB were further compared using Ramachandran plots (Fig. 6B and 6D), with most residues located in the favored or allowed regions in both

structures. Overall, this analysis strongly supports the conservation of Hg(II) methylation activity for the putative *Marinimicrobia-HgcA* and -HgcB proteins. In addition, similar HgcAB homology models constructed for novel putative Hg methylators *Calditrichaeta* and SAR324 also support functionality (Fig. S9). The same methodology has been adapted to infer functionality for HgcAB proteins from *Nitrospina*, another uncultivated potential Hg methylator [11]. This approach represents a promising method for predicting the functionality of proteins encoded for by uncultivated microorganisms.

### Genome composition and adaptation strategies of *hgcA*-carrying MAGs

Metabolic pathways were inferred for MAGs by reference to Kyoto Encyclopedia of Genes and Genomes (KEGG) annotations (Fig. 7). Hg methylators living in oxygen gradients, such as those observed in Saanich Inlet [25], may require the capability to tolerate intermittent lower levels of oxygen exposure. Genes encoding for multiple terminal oxidases were detected in the *hgcA*-carrying MAGs. Specifically, all *hgcA*-carrying *Marinimicrobia* and most *Deltaproteobacteria* contained at least one gene encoding for cytochrome *c* oxidase *aa3*-type (*coxABC*), a canonical



**Fig. 7 KEGG pathways of the *hgcA*-carrying MAGs.** Taxonomic classifications of MAGs are represented at bottom of heatmap by different colors. Categories of pathways are represented at left side of

heatmap by different colors. Color of each cell refers to completeness of enzymes involved in each pathway. Corresponding genes involved in each pathway are shown in parentheses.

heme-copper coupling O<sub>2</sub> reduction to transmembrane proton pumping [43]. Other terminal oxidase systems were also recognized; for example, genes encoding for cytochrome *c* *ccb*<sub>3</sub>-type (*ccoPQNO*) oxidase were found in several members of *Deltaproteobacteria*, *Marinimicrobia*, and *Nitrospina*; this oxidase has been shown to exhibit high affinity for oxygen, enabling bacteria to respire O<sub>2</sub> under both oxic and suboxic conditions [44, 45]. All *hgcA*-carrying SAR324, and the majority of other *hgcA*-carrying *Deltaproteobacteria*, contained the complete operon encoding for cytochrome *bd* complex (*cydAB*); this type of oxidase has also been demonstrated to support adaptation to low-oxygen environments, because of its relatively high affinity for O<sub>2</sub> [46]. In addition, cytochrome *o* oxidase (*cyoABCD*), carried by many *Deltaproteobacteria*, is an enzyme proposed to function under O<sub>2</sub>-rich conditions [47]. All *Marinimicrobia* and some *Deltaproteobacteria* also contained genes encoding for cytochrome *aa*<sub>3</sub>-600 oxidase (*qoxABCD*), a heme-copper oxygen reductase that uses ubiquinol or menaquinol in place of cytochrome *c* in the canonical respiration pathway [48]. As shown here, many Hg methylator genomes contained more than one type of terminal oxidase. Such a combination of multiple terminal oxidases has been observed for other microorganisms, facilitating electron transfer to O<sub>2</sub> under variable redox potentials. In contrast, no genes involved in O<sub>2</sub> respiration were observed in *hgcA*-carrying *Calditrichaeota*, *Verrucomicrobia*, *Spirochaetes*, and *Firmicutes* genomes identified from databases. Although only partial operons for some terminal respiratory enzyme complexes were found in some MAGs, possibly resulting from incomplete genome recovery, the complete cytochrome *bd* complex was found encoded by most *Deltaproteobacteria*, suggesting a capability for O<sub>2</sub> tolerance.

Genes involved in carbohydrate degradation pathways were found in many *hgcA*-carrying MAGs, especially in *Marinimicrobia*, *Verrucomicrobia*, and *Spirochaetes*. These MAGs also represent versatile metabolic capabilities in nitrogen and sulfur cycling. *Marinimicrobia* MAGs contained complete operons for utilizing nitrogen species in electron transfers, e.g., hydroxylamine oxidoreductase (*hao*), nitrite oxidase (*nxrAB*), nitrate reductase (*narGHI*), and hydrazine dehydrogenase (*hdh*). *Deltaproteobacteria* MAGs presented the potential for some degree of sulfur utilization, for which SAR324 were the only group to possess the gene *ddlA* encoding for a subunit of dimethyl sulfide dehydrogenase. Notably, most *hgcA*-carrying MAGs encoded for sulfide-quinone oxidoreductase (*sqr*), known to oxidize hydrogen sulfide to elemental sulfur for respiration and sulfide detoxification [49]. Nearly all MAGs harbored *mttB* genes encoding for trimethylamine methyltransferase, suggesting a capability for anaerobic one-carbon metabolism. Most MAGs also encoded genes for arsenic reduction

(*arsABCDR*), and Cu (*copA*) and Fe (*feoB*) transportation. Furthermore, genes encoding for flagellar proteins were identified in most putative Hg-methylators, enabling these bacteria to reposition optimally within redox gradients to suitable O<sub>2</sub> and nutrient levels.

Our findings point toward a previously unrecognized marine microbial Hg methylator, and are consistent with previously observed MeHg profiles in marine redoxclines. However, we acknowledge that cultivation experiments with *hgcAB*-bearing *Marinimicrobia* isolates are ultimately needed both to elucidate their lifestyle and confirm functionality for Hg methylation.

## Materials and methods

### Mercury analysis

Seawater samples for mercury analysis were collected from Saanich Inlet S3 station (48°35.500 N, 123°30.300 W, Fig. 1A) in April 2010. Total and methylmercury (monomethylmercury) determinations were made using USEPA standard methods 1631-E and 1630, respectively [50, 51]. In brief, total Hg was determined on filtered water samples following wet chemical oxidation by BrCl, followed by reduction by NH<sub>2</sub>OH and SnCl<sub>2</sub>, rendering all Hg species as volatile Hg(0). Hg was purged from solution by N<sub>2</sub> and concentrated on a gold-coated sand cartridge, which was then heated, releasing Hg for re-concentration on a second gold cartridge for final quantification by cold-vapor atomic fluorescence spectrometry (CVAFS; Tekran 2600). Methylated Hg was first separated from the seawater matrix by KCl/H<sub>2</sub>SO<sub>4</sub>/CuSO<sub>4</sub> extraction and steam distillation. Removal from seawater allowed for derivatization of MeHg into methylethyl-Hg through the use of sodium tetrathylborate, which is volatile and can be purged from solution and pre-concentrated on Tenax. The MeHg derivative was then separated from other Hg forms on a packed gas chromatography column of 15% OV-3 on Chromasorb-W at 110 °C, and then rendered into Hg<sup>0</sup> through pyrolysis for quantification by CVAFS. Detection limits for total and methylated Hg (3σ of reagent blanks) were 0.5 and 0.1 pM, respectively.

### Multi-omics data description

The multi-omic (metagenomic, metatranscriptomic, and metaproteomic) time-series samples from Saanich Inlet were taken from 2008 to 2013. A total of 84 genome-resolved metagenomic samples (Table S4) were employed in this study, including 78 samples from Saanich Inlet midpoint station “S3” (48°35.500 N, 123°30.300 W), 6 samples from the inlet mouth station “S4” (48°38.310 N,

123°30.007 W) and the inlet end station “S2” (48°33.148 N, 123°31.969 W; Fig. 1A). A set of shotgun metatranscriptomic raw data corresponding to 36 metagenomic samples (Table S4) was used to investigate the transcriptional activity of genes of interest. Protein sequences predicted from the Saanich Inlet metaproteomic dataset [52] were also used to confirm the expression of target proteins in the environment. Sample collection, DNA/ RNA/ protein extraction, and sequencing methods of the datasets were described previously [24]. Briefly, seawater samples spanning six major depths (10, 100, 120, 135, 150, and 200 m) were collected and filtered onto 0.22 µm Sterivex (Millipore) filters. 1.8 ml of RNAlater (Ambion) was added to metatranscriptomic sample filters, and 1.8 ml of sucrose lysis buffer was added to metaproteomic sample filters. Filters were stored at −80 °C until processing. For metagenomic samples, Sterivex filters were thawed on ice and incubated at 37 °C for 1 h with lysozyme (Sigma). Proteinase K (Sigma) and 20% SDS were added subsequently and incubated at 55 °C for 2 h with rotation. Filters were rinsed with sucrose lysis buffer after lysate was removed. Combined lysate was extracted with phenol-chloroform followed by chloroform. The aqueous layer was washed with TE buffer (pH 8.0) for three times and concentrated to 150–400 µl. The metagenomic samples were sequenced at the DOE Joint Genome Institute (JGI) and sequenced on the Illumina HiSeq 2000 platform. For metatranscriptomic samples, total RNA was extracted using the mirVana miRNA Isolation kit (Ambion) modified for Sterivex filters. RNAlater was removed from thawed filters by extrusion and rinsed with Ringer’s solution (Sigma). After incubation at 25 °C for 20 min with rotation, Ringer’s solution was removed by extrusion. Lysozyme was added, following by incubating at 37 °C for 30 min with rotation. Lysate was removed into 15 ml tube and subjected to organic extraction. TURBO DNA-free kit (ThermoFisher) was used to remove DNA and the RNeasy MinElute Cleanup kit (Qiagen) was used for purification. Metatranscriptomic shotgun libraries were generated at the JGI on the HiSeq 2000 platform. For metaproteomic samples, Bugbuster (Novagen) was added to thawed filters and incubated at 25 °C for 20–30 min with rotation. Filters were rinsed with 1 ml lysis buffer after removing lysate. Combined lysate was then subject to buffer exchange using Amicon Ultra 10K (Millipore) with 100 mM NH<sub>4</sub>HCO<sub>3</sub> for three times. Urea was added to a final concentration of 8 M and dithiothreitol added to a final concentration of 5 mM. Samples were incubated at 60 °C for 30 min and diluted ten times with 100 mM NH<sub>4</sub>HCO<sub>3</sub>, following by digesting at 37 °C for 6 h with trypsin. C18 solid phase extraction and strong cation exchange were carried out subsequently. Tandem mass spectrometry (MS/MS) was used to sequence the

protein samples, and MS-GFDB [53] was used to identify peptides based on the matched Saanich Inlet metagenomic sequences.

## Metagenomic assembly and binning

The raw Illumina reads from metagenomic samples (2 × 150 bp paired-end, median 13.75 Gbp per sample) were first filtered and trimmed by Trimmomatic v3.6 [54]. Samples from the same station and time point were co-assembled with MEGAHIT v1.2.8. MAGs were recovered using MetaBAT v2.14 [55] and MaxBin v2.2.7 [56], which were then merged and refined with MetaWRAP v1.2 [57]. In order to further improve the quality of the MAGs, metagenomic reads were remapped to each MAG and then reassembled by SPAdes [58] in careful mode. The qualities of derived MAGs were examined using CheckM [59]. 16S rRNA genes were predicted by RNAmmer [60] and then used to assign taxonomic classifications of the MAGs. For those MAGs that lacked 16S rRNA genes, GTDB-Tk v0.3.2 [61] was used to predict taxonomies according to genome contents. Genome annotation was conducted with Prokka v1.14 [62]. KEGG pathways of MAGs of interest were predicted with BlastKOALA (<http://www.kegg.jp/blastkoala/>), and completeness of the pathways was estimated by KEGG-Decoder [63] according to the KEGG Pathway Database (<https://www.genome.jp/kegg/pathway.html>).

## Relative abundance calculation

CD-HIT v4.6 [64] was used to choose representative *hgcA* sequences from all the MAGs under a 99% sequence identity threshold. Paired-end reads were remapped to these *hgcA* sequences with BWA-MEM algorithm [65]. BBMap (<http://sourceforge.net/projects/bbmap/>) was used to calculate the average coverage. MicrobeCensus [66] was used to estimate the genome equivalent of every sample. The relative abundance of every representative *hgcA* in every sample was calculated by the formula “*hgcA* coverage / genome equivalent”; relative abundance of *hgcA* genes was assumed to represent that of corresponding MAGs, since all reported methylators only carry one copy of *hgcA* gene in the genome. To evaluate and compare expression levels of genes of interest, RPKM values were calculated for each gene, normalized for both gene length and metatranscriptomic sequencing depth. RPKM values for housekeeping genes *rpoB* and *gyrB* were also calculated to normalize the expression of *hgcA*, with the ratio of *hgcA*/*rpoB* RPKM and *hgcA*/*gyrB* RPKM representing the expression of *hgcA* in individual microorganism regardless of the abundance variation of the total *hgcA*-carrying microorganisms.

## Search for *hgcAB* genes in MAGs

An in-house database of experimentally validated functional HgcAB sequences (see Table S5) was used to search for HgcAB encoding genes in MAGs by BLASTp, and BLAST results were further confirmed by examining the existence of conserved motifs ([N(V/I)WCA(A/G)(A/G)K] in HgcA and [CX2CX2CX3C] in HgcB), respectively. All the derived MAGs with *hgcAB* were then proceeded by the script *esomWrapper.pl* [67] to determine tetra-nucleotide frequencies signatures, followed by generating ESOM with Databionic emergent self-organizing map tools (<http://data.bionic-esom.sourceforge.net/>). The presence of *hgcAB* fragments in the MAGs were confirmed by checking manually in the ESOM.

## Phylogenetic analysis

HgcA sequences from the two most recent papers [14, 15] were retrieved and combined with the HgcA sequences predicted from genome taxonomy database (GTDB, release 89, <https://gtdb.ecogenomic.org/>) to establish an HgcA database (see Table S6 for accession numbers), after the redundant sequences were eliminated by CD-HIT [64] at 50% cutoff. These sequences were aligned with the putative HgcA sequences from Saanich Inlet by MAFFT v7 with the high-sensitivity (L-INS-i) algorithm [68]. The alignment was trimmed by trimal v1.2 [69], followed by the maximum likelihood (ML) tree was reconstructed by IQ-TREE v1.6 [70] under the LG + F + R9 protein substitution model chosen according to BIC. Publicly available *Marinimicrobia* 16S rRNA genes (see Table S7) were retrieved from the SILVA database (release 138) [36] and aligned with the 16S rRNA genes of *hgcA*-carrying *Marinimicrobia* to build an ML tree using IQ-TREE v1.6. The ML tree of all the *Marinimicrobia*-affiliated MAGs found in this study, as well as in public databases (see Table S8), was built by PhyloPhlAn2 [71] using 400 universal proteins without duplication.

## Searching SRA database

To investigate the distribution of *Marinimicrobia-hgcA*, the recovered *Marinimicrobia-hgcA* gene nucleotide sequence was first employed as a query to be searched in the NCBI SRA database using the SRA search tool [72], which is able to map ~1% reads of more than 100,000 public whole shotgun metagenomic samples to the query sequences using Bowtie2 [73]. Samples that contained reads associated with the *Marinimicrobia-hgcA* gene were then checked by downloading the corresponding read set from the SRA database and mapping these reads to the *hgcA* nucleotide sequence.

## Protein homology modeling

The three-dimensional structures of the putative HgcA and HgcB sequences found in this study were built and refined using Schrödinger suit v2012 (LLC, New York, USA). The transmembrane regions were predicted by CCTOP [74]. HgcAB sequences were searched against the protein data bank (PDB) using BLASTp with a word size of three. The structure of the corrinoid iron-sulfur protein (CFeSP) AcsC from *Carboxydotherrmus hydrogenoformans* (PDB ID: 2H9A; 1.9 Å; 28% identity and 55% similarity to *Marinimicrobia-HgcA*) and the CFeSP AcsC from *Moorella thermoacetica* (PDB ID: 4DJD; 2.38 Å; 31% identity and 59% similarity to *Marinimicrobia-HgcA*) were identified as appropriate templates for homology modeling of the HgcA N-terminus globular structured domain, according to their similarities and the resolution at which they were solved. The structure of the transmembrane bacteriorhodopsin Bop (PDB ID:2ZFE; 2.5 Å) was used to model the transmembrane region of HgcA, due to their comparable secondary structures. The homology model of HgcB proteins was built using the experimental structures of iron hydrogenase HydA from *D. desulfuricans* (PDB ID: 1HFE; 1.6 Å; 36% identity and 53% similarity to *Marinimicrobia-HgcB*) and the flavin-based caffeyl-CoA reductase CarE (PDB ID: 6FAH\_A; 3.133 Å; 39% identity and 45% similarity to *Marinimicrobia-HgcB*) in a similar procedure. The stereochemical quality and accuracy of reconstructed homology models were assessed and compared to each other by generating a Ramachandran plot using the Rampage server [75] and a best model was selected. Hydrogen bonding between HgcA protein and the ligand cobalamin were predicted by Arpeggio [76], PLIP [77], Chimera [78], and NGL Viewer [79], and consistent hydrogen bonding among the four programs was considered to be robust. Homology models for HgcAB from the Hg-methylating model strain *D. desulfuricans* ND132 were also built by the methods described above.

## Data availability

Data from this project have been deposited at DDBJ/ENA/GenBank under Project ID PRJNA630981. The 2088 MAGs generated by this study are available with accession numbers JABGOO000000000 to JABJQV000000000.

## Code availability

Customized scripts for conducting metagenomic and metatranscriptomic analyses are available from GitHub (<https://github.com/SilentGene/MultiOmicAnalysis>).

**Acknowledgements** HL was supported by a postgraduate fellowship from The University of Melbourne Environmental Microbiology

Research Initiative. HL and JWM designed the project and wrote the paper in consultation with all co-authors. DBA supervised the protein homology modelling, supported by an Investigator Grant from the National Health and Medical Research Council (NHMRC) of Australia [GNT1174405], and by the Victorian Government OIS Program. CHL thanks Woods Hole Oceanographic Institution for support and Tracy Mincer for help and inspiration. KEH co-advised meta-omics and phylogenetic analyses, supported by a Senior Medical Research Fellowship from the Viertel Foundation of Australia. HL, DBH, YM, RW, KEH and JWM gratefully acknowledge the use of data generated under the auspices of the US Department of Energy (DOE) Joint Genome Institute and Office of Science User Facility, supported by the Office of Science of the U.S. Department of Energy under Contract DE-AC02-05CH11231, the G. Unger Vetlesen and Ambrose Monell Foundations, and the Natural Sciences and Engineering Research Council of Canada through grants awarded to SJH.

## Compliance with ethical standards

**Conflict of interest** The authors declare that they have no conflict of interest.

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

**Open Access** This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons license, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons license and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this license, visit <http://creativecommons.org/licenses/by/4.0/>.

## References

- Fitzgerald WF, Clarkson TW. Mercury and monomethylmercury: present and future concerns. *Environ Health Perspect.* 1991;96:159–66.
- Selin NE. Global biogeochemical cycling of mercury: a review. *Annu Rev Environ Resour.* 2009;34:43–63.
- Hsu-Kim H, Eckley CS, Selin NE. Modern science of a legacy problem: mercury biogeochemical research after the Minamata Convention. *Environ Sci-Process Impacts.* 2018;20:582–3.
- Lee C-S, Fisher NS. Bioaccumulation of methylmercury in a marine copepod. *Environ Toxicol Chem.* 2017;36:1287–93.
- Stramma L, Johnson GC, Sprintall J, Mohrholz V. Expanding oxygen-minimum zones in the tropical oceans. *Science.* 2008;320:655–8.
- Wright JJ, Konwar KM, Hallam SJ. Microbial ecology of expanding oxygen minimum zones. *Nat Rev Microbiol.* 2012;10:381–94.
- Parks JM, Johs A, Podar M, Bridou R, Hurt RA Jr., Smith SD, et al. The genetic basis for bacterial mercury methylation. *Science.* 2013;339:1332–5.
- Gilmour CC, Podar M, Bullock AL, Graham AM, Brown SD, Somenahally AC, et al. Mercury methylation by novel microorganisms from new environments. *Environ Sci Technol.* 2013;47:11810–20.
- Podar M, Gilmour CC, Brandt CC, Soren A, Brown SD, Crable BR, et al. Global prevalence and distribution of genes and microorganisms involved in mercury methylation. *Sci Adv.* 2015;1:e1500675.
- Grégoire DS, Poulain AJ, Ivanova EP. Shining light on recent advances in microbial mercury cycling. *Facets.* 2018;3:858–79.
- Gionfriddo CM, Tate MT, Wick RR, Schultz MB, Zemla A, Thelen MP, et al. Microbial mercury methylation in Antarctic sea ice. *Nat Microbiol.* 2016;1:16127.
- Tada, Y, Marumoto, K, Takeuchi, A. Nitrospina-like bacteria are potential mercury methylators in the mesopelagic zone in the East China Sea. *Front Microbiol.* 2020;11:1369.
- Jones, DS, Walker, GM, Johnson, NW, Mitchell, CPJ, Coleman Wasik, JK, Bailey, JV. Molecular evidence for novel mercury methylating microorganisms in sulfate-impacted lakes. *ISME J.* 2019;13:1659–75.
- Gionfriddo, CM, Wymore, AM, Jones, DS, Wilpiszkeski, RL, Lynes, MM, Christensen, GA et al. An improved hgcAB primer set and direct high-throughput sequencing expand Hg-methylator diversity in nature. *bioRxiv.* 2020:2020.2003.2010.983866.
- McDaniel, EA, Peterson, BD, Stevens, SLR, Tran, PQ, Anantharaman, K, McMahon, KD. Expanded phylogenetic diversity and metabolic flexibility of mercury-methylating microorganisms. *mSystems.* 2020;5:20.
- Lu X, Liu Y, Johs A, Zhao L, Wang T, Yang Z, et al. Anaerobic mercury methylation and demethylation by *Geobacter bemedjiensis* Bem. *Environ Sci Technol.* 2016;50:4366–73.
- Lamborg CH, Hammerschmidt CR, Bowman KL, Swarr GJ, Munson KM, Ohnemus DC, et al. A global ocean inventory of anthropogenic mercury based on water column measurements. *Nature.* 2014;512:65–68.
- Paulmier A, Ruiz-Pino D. Oxygen minimum zones (OMZs) in the modern ocean. *Prog Oceanogr.* 2009;80:113–28.
- Moore JK, Doney SC. Iron availability limits the ocean nitrogen inventory stabilizing feedbacks between marine denitrification and nitrogen fixation. *Glob Biogeochem Cycle.* 2007;21:2.
- Bertagnoli, AD, Stewart, FJ. Microbial niches in marine oxygen minimum zones. *Nat Rev Microbiol.* 2018;16:723–9.
- Villiar, E, Cabrol, L, Heimbürger-Boavida, LE. Widespread microbial mercury methylation genes in the global ocean. *Environ Microbiol Rep.* 2020;12:277–87.
- Capo, E, Bravo, AG, Soerensen, AL, Bertilsson, S, Pinhassi, J, Feng, C et al. Marine snow as a habitat for microbial mercury methylators in the Baltic Sea. *bioRxiv.* 2020:2020.2003.2004.975987.
- Bowman KL, Collins RE, Agather AM, Lamborg CH, Hammerschmidt CR, Kaul D, et al. Distribution of mercury-cycling genes in the Arctic and equatorial Pacific Oceans and their relationship to mercury speciation. *Limnol Oceanogr.* 2020;65:S310–S320.
- Hawley AK, Torres-Beltran M, Zaikova E, Walsh DA, Mueller A, Scofield M, et al. A compendium of multi-omic sequence information from the Saanich Inlet water column. *Sci Data.* 2017;4:170160.
- Torres-Beltrán M, Hawley AK, Capelle D, Zaikova E, Walsh DA, Mueller A, et al. A compendium of geochemical information from the Saanich Inlet water column. *Sci Data.* 2017;4:170159.
- Hammerschmidt CR, Bowman KL. Vertical methylmercury distribution in the subtropical North Pacific Ocean. *Mar Chem.* 2012;132-133:77–82.
- Wang F, Macdonald RW, Armstrong DA, Stern GA. Total and methylated mercury in the Beaufort Sea: the role of local and recent organic remineralization. *Environ Sci Technol.* 2012;46:11821–8.
- Cossa D, Heimbürger L-E, Lannuzel D, Rintoul SR, Butler ECV, Bowie AR, et al. Mercury in the Southern Ocean. *Geochim Cosmochim Acta.* 2011;75:4037–52.
- Chakraborty P, Mason RP, Jayachandran S, Vudamala K, Armoury K, Sarkar A, et al. Effects of bottom water oxygen

- concentrations on mercury distribution and speciation in sediments below the oxygen minimum zone of the Arabian Sea. *Mar Chem.* 2016;186:24–32.
30. Cossa D, Averty B, Pirrone N. The origin of methylmercury in open Mediterranean waters. *Limnol Oceanogr.* 2009;54:837–44.
  31. Hawley AK, Nobu MK, Wright JJ, Durno WE, Morgan-Lang C, Sage B, et al. Diverse Marinimicrobia bacteria may mediate coupled biogeochemical cycles along eco-thermodynamic gradients. *Nat Commun.* 2017;8:1507.
  32. Marshall IPG, Starnawski P, Cupit C, Fernandez Caceres E, Ettema TJG, Schramm A, et al. The novel bacterial phylum Calditrichaeota is diverse, widespread and abundant in marine sediments and has the capacity to degrade detrital proteins. *Environ Microbiol Rep.* 2017;9:397–403.
  33. Sheik CS, Jain S, Dick GJ. Metabolic flexibility of enigmatic SAR324 revealed through metagenomics and metatranscriptomics. *Environ Microbiol.* 2014;16:304–17.
  34. Goni-Urriza M, Corsellis Y, Lanceluor L, Tessier E, Gury J, Monperrus M, et al. Relationships between bacterial energetic metabolism, mercury methylation potential, and *hgcA/hgcB* gene expression in *Desulfovibrio dechloroacetivorans* BerOc1. *Environ Sci Pollut Res Int.* 2015;22:13764–71.
  35. Gilmour CC, Elias DA, Kucken AM, Brown SD, Palumbo AV, Schadt CW, et al. The sulfate-reducing bacterium *Desulfovibrio desulfuricans* ND132 as a model for understanding bacterial mercury methylation. *Appl Environ Microbiol.* 2011;77:3938–51.
  36. Quast C, Pruesse E, Yilmaz P, Gerken J, Schweer T, Yarza P, et al. The SILVA ribosomal RNA gene database project: improved data processing and web-based tools. *Nucleic Acids Res.* 2012;41:D590–D596.
  37. Yarza P, Yilmaz P, Pruesse E, Glockner FO, Ludwig W, Schleifer KH, et al. Uniting the classification of cultured and uncultured bacteria and archaea using 16S rRNA gene sequences. *Nat Rev Microbiol.* 2014;12:635–45.
  38. Bertagnolli AD, Padilla CC, Glass JB, Thamdrup B, Stewart FJ. Metabolic potential and in situ activity of marine Marinimicrobia bacteria in an anoxic water column. *Environ Microbiol.* 2017;19:4392–416.
  39. Yergeau E, Maynard C, Sanschagrín S, Champagne J, Juck D, Lee K, et al. Microbial community composition, functions, and activities in the Gulf of Mexico 1 year after the deepwater horizon accident. *Appl Environ Microbiol.* 2015;81:5855–66.
  40. Yergeau E, Michel C, Tremblay J, Niemi A, King TL, Wyglinski J, et al. Metagenomic survey of the taxonomic and functional microbial communities of seawater and sea ice from the Canadian Arctic. *Sci Rep.* 2017;7:42242.
  41. Liu B, Schaidler LA, Mason RP, Shine JP, Rabalais NN, Senn DB. Controls on methylmercury accumulation in northern Gulf of Mexico sediments. *Estuar Coast Shelf S.* 2015;159:50–59.
  42. Rush, KW. Expression and characterization of *HgcA* and *HgcB*, two proteins involved in methylmercury biosynthesis. Ph.D. thesis. Ann Arbor, MI: University of Michigan;2018.
  43. Sousa FL, Alves RJ, Ribeiro MA, Pereira-Leal JB, Teixeira M, Pereira MM. The superfamily of heme-copper oxygen reductases: types and evolutionary considerations. *Biochim Biophys Acta.* 2012;1817:629–37.
  44. Visser JM, Jong GAH, Vries S, Robertson LA, Kuenen JG. *ccb3*-type cytochrome oxidase in the obligately chemolithoautotrophic *Thiobacillus* sp. W5. *FEMS Microbiol Lett.* 2006;147:127–32.
  45. Colburn-Clifford J, Allen C. A *ccb(3)*-type cytochrome C oxidase contributes to *Ralstonia solanacearum* R3bv2 growth in micro-aerobic environments and to bacterial wilt disease development in tomato. *Mol Plant Microbe Interact.* 2010;23:1042–52.
  46. Borisov VB, Gennis RB, Hemp J, Verkhovsky MI. The cytochrome bd respiratory oxygen reductases. *Biochim Biophys Acta.* 2011;1807:1398–413.
  47. Cotter PA, Chepuri V, Gennis RB, Gunsalus RP. Cytochrome o (cyoABCDE) and d (cydAB) oxidase gene expression in *Escherichia coli* is regulated by oxygen, pH, and the *fnr* gene product. *J Bacteriol.* 1990;172:6333–8.
  48. Xu J, Ding Z, Liu B, Yi SM, Li J, Zhang Z, et al. Structure of the cytochrome aa<sub>3</sub>-600 heme-copper menaquinol oxidase bound to inhibitor HQNO shows TM0 is part of the quinol binding site. *Proc Natl Acad Sci USA* 2020;117:872–6.
  49. Marcia M, Ermiler U, Peng G, Michel H. The structure of Aquifex aeolicus sulfide:quinone oxidoreductase, a basis to understand sulfide detoxification and respiration. *Proc Natl Acad Sci USA* 2009;106:9625–30.
  50. Telliard WA. Method 1630: Methyl mercury in water by distillation, aqueous ethylation, purge and trap, and cold vapor atomic fluorescence spectrometry. US Environmental Protection Agency-Office of Water. 1998.
  51. Telliard, WA, Gomez-Taylor M. Method 1631, Revision E: Mercury in Water by Oxidation, Purge and Trap, and Cold Vapor Atomic Fluorescence Spectrometry. United States Environmental Protection Agency, Office of Water, 4303. EPA-821-R-02-019, 2002.
  52. Hawley AK, Brewer HM, Norbeck AD, Pasa-Tolic L, Hallam SJ. Metaproteomics reveals differential modes of metabolic coupling among ubiquitous oxygen minimum zone microbes. *Proc Natl Acad Sci USA* 2014;111:11395–11400.
  53. Kim S, Mischerikow N, Bandeira N, Navarro JD, Wich L, Mohammed S, et al. The generating function of CID, ETD, and CID/ETD pairs of tandem mass spectra: applications to database search. *Mol Cell Proteom.* 2010;9:2840–52.
  54. Bolger AM, Usadel B, Lohse M. Trimmomatic: a flexible trimmer for Illumina sequence data. *Bioinformatics.* 2014;30:2114–20.
  55. Kang DD, Froula J, Egan R, Wang Z. MetaBAT, an efficient tool for accurately reconstructing single genomes from complex microbial communities. *PeerJ.* 2015;3:e1165.
  56. Wu Y-W, Tang Y-H, Tringe SG, Simmons BA, Singer SW. MaxBin: an automated binning method to recover individual genomes from metagenomes using an expectation-maximization algorithm. *Microbiome.* 2014;2:26.
  57. Uritskiy GV, DiRuggiero J, Taylor J. MetaWRAP—a flexible pipeline for genome-resolved metagenomic data analysis. *Microbiome.* 2018;6:158.
  58. Bankevich A, Nurk S, Antipov D, Gurevich AA, Dvorkin M, Kulikov AS, et al. SPAdes: a new genome assembly algorithm and its applications to single-cell sequencing. *J Comput Biol.* 2012;19:455–77.
  59. Parks DH, Imelfort M, Skennerton CT, Hugenholtz P, Tyson GW. CheckM: assessing the quality of microbial genomes recovered from isolates, single cells, and metagenomes. *Genome Res.* 2015;25:1043–55.
  60. Lagesen K, Hallin P, Rødland EA, Staerfeldt H-H, Rognes T, Ussery DW. RNAmmer: consistent and rapid annotation of ribosomal RNA genes. *Nucleic Acids Res.* 2007;35:3100–8.
  61. Parks DH, Chuvochina M, Waite DW, Rinke C, Skarshewski A, Chaumeil P-A, et al. A standardized bacterial taxonomy based on genome phylogeny substantially revises the tree of life. *Nat Biotechnol.* 2018;36:996.
  62. Seemann T. Prokka: rapid prokaryotic genome annotation. *Bioinformatics.* 2014;30:2068–9.
  63. Graham ED, Heidelberg JF, Tully BJ. Potential for primary productivity in a globally-distributed bacterial phototroph. *ISME J.* 2018;12:1861–6.
  64. Godzik A, Li W. Cd-hit: a fast program for clustering and comparing large sets of protein or nucleotide sequences. *Bioinformatics.* 2006;22:1658–9.
  65. Li, H. Aligning sequence reads, clone sequences and assembly contigs with BWA-MEM. arXiv preprint arXiv:1303.3997. 2013.

66. Nayfach S, Pollard KS. Average genome size estimation improves comparative metagenomics and sheds light on the functional ecology of the human microbiome. *Genome Biol.* 2015;16:51.
67. Dick GJ, Andersson AF, Baker BJ, Simmons SL, Thomas BC, Yelton AP, et al. Community-wide analysis of microbial genome sequence signatures. *Genome Biol.* 2009;10:R85.
68. Standley DM, Katoh K. MAFFT multiple sequence alignment software version 7: improvements in performance and usability. *Mol Biol Evol.* 2013;30:772–80.
69. Silla-Martínez JM, Capella-Gutiérrez S, Gabaldón T. trimAl: a tool for automated alignment trimming in large-scale phylogenetic analyses. *Bioinformatics.* 2009;25:1972–3.
70. Schmidt HA, Minh BQ, von Haeseler A, Nguyen L-T. IQ-TREE: a fast and effective stochastic algorithm for estimating maximum-likelihood phylogenies. *Mol Biol Evol.* 2014;32:268–74.
71. Segata N, Börnigen D, Morgan XC, Huttenhower C. PhyloPhlAn is a new method for improved phylogenetic and taxonomic placement of microbes. *Nat Commun.* 2013;4:2304–2304.
72. Levi, K, Rynge, M, Abeysinghe, E, Edwards, RA. Searching the sequence read archive using Jetstream and Wrangler. *Proc Pract Exp Adv Res Comput.* 2018:1–7.
73. Langmead B, Salzberg SL. Fast gapped-read alignment with Bowtie 2. *Nat Methods.* 2012;9:357–9.
74. Dobson L, Reményi I, Tusnády GE. CCTOP: a Consensus Constrained TOPology prediction web server. *Nucleic Acids Res.* 2015;43:W408–W412.
75. Lovell SC, Davis IW, Arendall WB III, de Bakker PIW, Word JM, Prisant MG, et al. Structure validation by C $\alpha$  geometry:  $\phi$ ,  $\psi$  and C $\beta$  deviation. *Proteins.* 2003;50:437–50.
76. Jubb HC, Higuero AP, Ochoa-Montaña B, Pitt WR, Ascher DB, Blundell TL. Arpeggio: a web server for calculating and visualising interatomic interactions in protein structures. *J Mol Biol.* 2017;429:365–71.
77. Salentin S, Schreiber S, Haupt VJ, Adasme MF, Schroeder M. PLIP: fully automated protein–ligand interaction profiler. *Nucleic Acids Res.* 2015;43:W443–W447.
78. Pettersen EF, Goddard TD, Huang CC, Couch GS, Greenblatt DM, Meng EC, et al. UCSF Chimera—A visualization system for exploratory research and analysis. *J Comput Chem.* 2004;25:1605–12.
79. Rose AS, Hildebrand PW. NGL Viewer: a web application for molecular visualization. *Nucleic Acids Res.* 2015;43:W576–W579.