Global distribution of anaerobic dichloromethane degradation potential

Short Title: Anaerobic dichloromethane biodegradation

Robert W. Murdoch[1]†, Gao Chen[1,2], Fadime Kara Murdoch[1,7]†, E. Erin Mack[5], Manuel I. Villalobos Solis[6], Robert L. Hettich[6], and Frank E. Löffler[1,2,3,4,6]*

[1]Center for Environmental Biotechnology, [2]Department of Civil and Environmental Engineering, [3]Department of Microbiology, [4]Department of Biosystmes Engineering and Soil Science, University of Tennessee, Knoxville, TN 37996, USA

[5]Corteva Environmental Remediation, Corteva Agriscience, Wilmington, DE 19805, USA

[6]Biosciences Division, Oak Ridge National Laboratory, Oak Ridge, TN 37831, USA

† Current address: Battelle Memorial Institute, Columbus, Ohio 43201

* Corresponding author: Frank E. Löffler. Center for Environmental Biotechnology, University of Tennessee, 676 Dabney Hall, 1416 Circle Drive, Knoxville, TN 37996-1605, USA

Phone: (865) 974 4933 **Email:** frank.loeffler@utk.edu

Competing Interest Statement: The authors declare no competing interest.

Classification: Biological Sciences (Major); Environmental Sciences (Minor)

Keywords: Dichloromethane fluxes, Degradation Potential, Bioremediation, Ozone Destruction

Abstract

Anthropogenic activities and natural processes release dichloromethane (DCM), a toxic chemical with substantial ozone-depleting capacity. Specialized anaerobic bacteria metabolize DCM; however, the genetic basis for this process has remained elusive. Comparative genomics of the three known anaerobic DCM-degrading bacterial species revealed a homologous gene cluster, designated the methylene chloride catabolism (*mec*) gene cassette, comprising eight to ten genes with predicted 79.6 – 99.7% amino acid identity. Functional annotation identified genes encoding a corrinoid-dependent

methyltransferase system, and shotgun proteomics applied to two DCM-catabolizing cultures revealed high expression of proteins encoded on the *mec* gene cluster during anaerobic growth with DCM. In a DCM-contaminated groundwater plume, the abundance of *mec* genes strongly correlated with DCM concentrations (R^2 = 0.71 – 0.85) indicating their value as process-specific bioremediation biomarkers. *mec* gene clusters were identified in metagenomes representing peat bogs, the deep subsurface, and marine ecosystems including oxygen minimum zones (OMZs), suggesting DCM turnover in diverse habitats. The broad distribution of anaerobic DCM catabolic potential suggests a relevant control function for emissions to the atmosphere, and a role for DCM as a microbial energy source in critical zone environments. The findings imply that the global DCM flux might be far greater than emission measurements suggest.

Importance

Dichloromethane (DCM) is an increasing threat to stratospheric ozone with both anthropogenic and natural emission sources. Anaerobic bacterial metabolism of DCM has not yet been taken into consideration as a factor in the global DCM cycle. The discovery of the *mec* gene cassette associated with anaerobic bacterial DCM metabolism and its widespread distribution in environmental systems highlight a strong attenuation potential for DCM. Knowledge of the *mec* cassette offers new opportunities to delineate DCM sources, enables more robust estimates of DCM fluxes, supports refined DCM emission modeling and simulation of the stratospheric ozone layer, reveals a novel, ubiquitous C_1 carbon metabolic system, and provides prognostic and diagnostic tools supporting bioremediation of groundwater aquifers impacted by DCM.

Introduction

Dichloromethane (DCM, methylene chloride) is a widely distributed halomethane, produced both naturally and industrially. While anthropogenic DCM has received attention due to widespread groundwater contamination and, more recently ozone destruction potential, analysis of Antarctic ice cores has demonstrated that DCM was present in the atmosphere prior to the industrial era at approximately 10% of modern levels (Trudinger et al., 2004). The natural sources of DCM are diverse, encompassing both abiotic (Isidorov et al., 1990; Kanters & Louw, 1996) and biotic (Eustáquio et al., 2008; Hoekstra et al., 1998; Wuosmaa & Hager, 1990) processes and are estimated to contribute up to one third of total emissions (Gribble, 2010). Since the 1960's, atmospheric DCM concentrations rose steadily, with a mean annual increase of approximately 8% (Hossaini et al., 2017) although the reported worldwide production and use has been steady or declining since 2010 (McCulloch, 2017). Possible explanations include undocumented production, rogue emmissions, or increased natural emissions reflecting environmental (e.g., climate) change responses.

Atmospheric measurements and corresponding efforts to extrapolate to global-scale emissions have led to the perception that marine systems and biomass combustion (e.g., wildfires) are the primary non-industrial sources of DCM (Gribble, 2010), releasing estimated amounts of 190 and 60 Gg of DCM each year, respectively. Natural and deliberate forest fires have increased in frequency and size (Haines et al., 2020), a global trend that can be expected to lead to further formation of DCM. Wetlands emit up to 2 Gg DCM per year (Cox et al., 2004; Hu et al., 2017; Kolusu et al., 2018) (Supplementary Information), and volcanic activity contributes an estimated amount of 0.021 Gg/y (Gribble, 2010). Halomethanes occur in crustal minerals, and DCM release from rocks from the near-surface and deep subsurface have been reported (Mulder et al., 2013; Svensen et al., 2009). Knowledge gaps remain and not-yet identified environmental sources of DCM are likely. The current understanding of global DCM fluxes, as opposed to emissions, is very limited (Figure 1) (McCulloch, 2017).

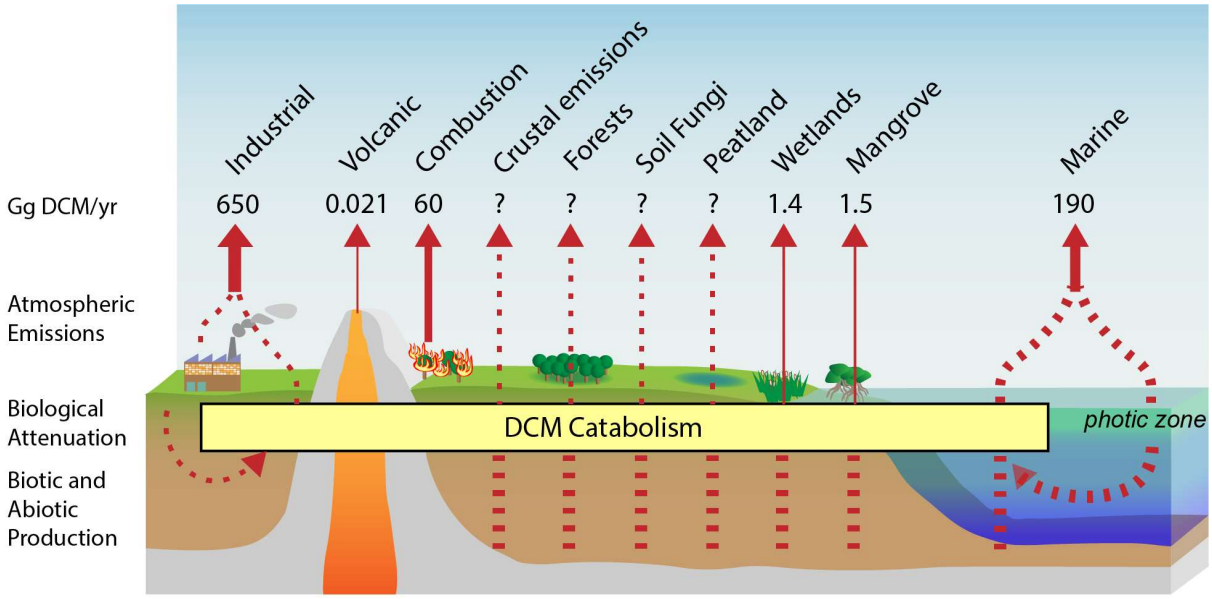

Figure 1. Major reported or potential DCM atmospheric emission sources. Width of the solid arrows is log-proportional to the magnitude of DCM emission estimates or potential. Dashed lines represent putative DCM sources, fluxes, and emissions that have not been directly investigated.

The rates of DCM production and consumption are unclear, and biological attenuation in anaerobic systems prior to release to the atmosphere has not been incorporated into existing atmospheric emission models (Hossaini et al., 2017).

Anaerobic metabolism of DCM was described in a bacterial isolate, *Dehalobacterium formicoaceticum* strain DMC (Defo) (*19*). Two additional bacterial populations responsible for DCM-metabolism in anaerobic enrichment cultures have been identified, '*Ca.* Dichloromethanomonas elyunquensis' (Diel) (Justicia-Leon et al., 2012; Kleindienst et al., 2016, 2017) and '*Ca.* Formimonas warabiya' (Dcmf) (Holland et al., 2021). All three DCM degraders are members of the family Peptococcaceae within the phylum Firmicutes. Bacterial dechlorination of DCM under oxic conditions is catalyzed by the glutathione-S-transferase (GST) DcmA (Muller et al., 2011); however, GST enzymes are generally absent in obligate anaerobes (Allocati et al., 2009). Accordingly, anaerobes metabolize DCM following a distinct strategy, in which the C_1 group is transferred to methylene-THF, a process that likely employs a corrinoid-dependent methyltransferase. The resulting methylene-THF is then channeled into the Wood-Ljungdahl pathway

(WLP). Interestingly, Diel generates hydrogen during DCM mineralization to CO_2 and chloride, which necessitates a syntrophic partnership with a hydrogen-consuming population (G. Chen, Kleindienst, et al., 2017). In contrast, '*Ca.* Formimonas warabiya' ferments DCM via intracellular syntrophy to acetate and chloride (Holland et al., 2021; Wiechmann et al., 2020) and axenic Defo cultures ferment DCM to acetate, formate and chloride (G. Chen, Murdoch, et al., 2017).

Due to the genetic intractability of anaerobic DCM-degrading bacteria, comparative genomic approaches were applied to unravel underlying conserved genes involved in DCM metabolism, which led to discovery of a novel conserved gene cluster. Proteomics applied to Diel and Defo grown with DCM supported a role of this gene cluster in anaerobic DCM metabolism. Tracing this gene cluster by analyzing public metagenome datasets and performing targeted qPCR assays revealed the prevalence of this gene cassette and the potential DCM-metabolizing phenotype in various environmental systems. The findings suggest DCM formation and consumption in diverse natural ecosystems and provide new opportunities for assessing how global changes in climate and habitat patterns impact DCM emissions and associated ozone destruction.

Methods

Comparative genomics

Initial identification of homologous genes shared between the genomes of Diel and Defo was performed by BLASTP-based reciprocal best hit (RBH) analysis within the Integrated Microbial Genomes (IMG) system (I.-M. A. Chen et al., 2019) and by using GView (Petkau et al., 2010). The two homologous gene clusters present in the Diel genome were manually delineated. A homologous gene cluster was identified in the genome of '*Ca.* Formimonas warabiya' by application of local BLASTP searches (Altschul et al., 1990). Functional annotations (COG, pfam, KEGG, TIGRFAM) for Defo and Diel genes were obtained from the IMG system, while those of '*Ca.* Formimonas warabiya' were assigned using the WebMGA server (Wu et al., 2011) for COG (Tatusov et al., 2000), pfam (El-Gebali et al., 2019), and TIGRFAM (Haft et al., 2012) annotations and GhostKOALA (Kanehisa et al., 2016) for KEGG annotations (Kanehisa et al., 2017).

Metagenome searches

All 18,314 metagenomes in the IMG database publicly available as of January 7, 2020 were subjected to BLASTP query with the Defo MecE protein sequence with a minimum bit score cutoff of 150 (approximately 40% identity). The resulting protein set was further filtered by applying a RBH criterion, retaining only proteins whose closest BLAST-P hit to the IMG genomes database was found among the MecE sequences located in the putative *mec* gene cassettes. The candidate metagenome MecE homologs were then further filtered by retaining only proteins encoded by genes co-localized with at least one other *mec* cassette protein, applying the same RBH criterion. All proteins encoded by genes located on scaffolds where the *mec* protein homologs were identified were downloaded, subjected to local BLASTP query using the ten *mec* gene cassette proteins from Defo, and plotted using GenoPlotR and custom R scripts (Guy et al., 2010). Gene copy per genome for metagenome *mec* cassettes, provided in Dataset S1, were calculated by dividing the read depth of the corresponding scaffold by an average read depth of ten single copy conserved protein-encoding genes (ribosomal proteins L11 (COG0080), L1 (COG0081), L3 (COG0087), L4 (COG0088), L2 (COG0090), L22 (COG0091), L5 (COG0094), L15 (COG0200), L10 (COG0244), and L29 (COG0255)).

Phylogenetic reconstruction

The top 20 most similar proteins to each of the proteins encoded by each of the Defo genes located on the *mec* cassette were obtained by searching the IMG genome database using BLASTP with a confidence threshold of 1e-5, except in the case of MecC, for which the threshold was 1e-2. All *mec* cassette genes located in both genomes and metagenomes were aligned and subjected to phylogenetic reconstruction alongside the top 20 most similar genes located in microbial genomes in the IMG database. Proteins encoded by genes from metagenomes were clustered at 80% similarity using CD-hit (Li & Godzik, 2006). Sequences were aligned using MAFFT G-INS-I with 1,000 maximum iterations (Katoh & Standley, 2013), trimmed using trimAl-gappyout (Capella-Gutiérrez et al., 2009) and subjected to phylogenetic reconstruction using FastTree2 maximum-likelihood estimation (Gamma-LG model) (Price et al., 2010). The resulting Newick tree files were visualized using the Interactive Tree of Life (Letunic & Bork, 2016).

148

149 *Preparation of Diel and Defo cultures for global proteomics*

150 The DCM-degrading consortium RM harboring '*Ca.* Dichloromethanomonas elyunquensis' (Diel) and the
151 axenic culture *Dehalobacterium formicoaceticum* (Defo) were grown in triplicate in 100 mL of anoxic
152 mineral basal salt medium with 0.2 mM sodium sulfide, 0.2 mM L-cysteine (Löffler et al., 2005) and 30
153 mM bicarbonate (pH 7.3) under a headspace of N_2/CO_2 (80:20, vol/vol) with 156 µmol (10 µL) of DCM as
154 the sole energy source. Cultures were initiated with a 5% (vol/vol) inoculum, incubated at 30°C in the dark
155 without agitation, and provided one additional feeding of DCM once the initial amendment was consumed.
156 Biomass for (meta)proteomic analyses was collected after 2 weeks of incubation when approximately
157 95% of the second DCM feedings were consumed. Culture suspensions were passed through Sterivex™
158 0.22 µm membrane filter units (EMD Millipore Corporation, Billerica, MA, US) to capture cells. The outlet
159 of the filters were capped, and 1.5 mL of boiling SDS lysis buffer (4% SDS in 100 mM Tris/HCl buffer, pH
160 8.0) were added to each of them. Filter unit inlets were then capped and placed in a laboratory rocker for
161 1-hour at room temperature. The SDS lysis buffer was removed by connecting 3 mL plastic syringes to
162 the inlets of the cartridges, and then holding the syringes and filter units vertically and pushing air into
163 each cartridge in order to withdraw as much lysate as possible by back pressure. In addition, filters were
164 rinsed once more with 0.5 mL of fresh SDS lysis buffer. Lysate mixtures were centrifuged at 21,000 g for
165 15 mins and the clean protein supernatant transferred to fresh Eppendorf plastic tubes. Proteins were
166 precipitated with trichloroacetic acid (TCA), denaturated in 8 M urea, reduced with dithiothreitol (DTT),
167 alkylated with iodoacetamide (IAM), and digested with sequencing grade trypsin (Promega, 1:50 trypsin-
168 to-protein [wt/wt]) (Yang et al., 2012). Protein concentrations were estimated with the BCA assay (Pierce
169 Biotechnology, Waltham, MA, US) and crude protein and peptide extracts were stored at -80°C for
170 subsequent LC-MS/MS analysis.

171

172 *Global proteomics analyses of Diel and Defo cultures*

173 Global proteomics analyses were performed with an Orbitrap Q Exactive Plus mass spectrometer
174 (Thermo Fisher Scientific, Waltham, MA, US) equipped with a nano-electrospray source (ESI) interfaced
175 with a Proxeon EASY-nLC™ 1200 system. Peptides (2 µg) from each sample were suspended in solvent

A (2% acetonitrile / 0.1% formic acid) and injected onto a C18 resin 75 µm microcapillary column (1.7 µm, 100Å, Phenomenex). Separation was accomplished at a constant flow rate of 250 nL/min with a 90-minute gradient from 2 to 30% solvent B (0.1% formic acid / 80% acetonitrile) followed by an increase to 40% solvent B within 10 minutes. Tandem mass spectrometry data (MS/MS) were collected using the Thermo Xcalibur software version 4.2.47 with similar parameters as reported before (Ganusova et al., 2021). Raw spectral files were searched against protein databases from the IMG annotated genomes of enrichment culture RM (which contains Diel) and Defo (IMG genome IDs 3300005804 and 2811995020, respectively), to which common laboratory contaminant proteins were appended. For standard database searching, the peptide MS/MS data was searched using Proteome Discoverer v2.4. The MS/MS data were searched using the SEQUEST HT algorithm (Eng et al., 1994) which was configured to derive fully tryptic peptides with the following settings: Maxium of 2 missed cleavage sites per peptide, minimum peptide length of 2, MS1 mass tolerance of 10 ppm and a MS2 tolerance of 0.02 Da. In addition, carbamidomethylations on cysteines (+57.0214 Da) and methionine oxidations (+5.9949 Da) were searched on peptides as static and dynamic modifications, respectively. Peptide spectrum match (PSM) confidence was evaluated with Percolator (Käll et al., 2007). PSMs and peptides were considered identified at a *q value* of < 0.01. Abundance values were converted to log2 values for ease of visualization. The IMG gene IDs of detected culture RM proteins were mapped to proteins contained in the IMG annotated genome of Diel (IMG genome ID 2627853586, Dataset S2).

DCM measurements

DCM was quantified by manual headspace injections (0.1 mL) into an Agilent 7890 gas chromatograph (GC) (Santa Clara, CA, USA) equipped with a DB-624 column (60 m length, 0.32 mm i.d., 1.8 mm film thickness) and a flame ionization detector (FID). To analyze DCM concentrations in groundwater, 1-mL samples were collected, immediately transferred to sealed 20-mL glass vials, and the DCM concentration determined in the headspace. Aquoeus phase concentrations were determined using a dimensionless Henry's law constant of 0.0895 (Gossett, 1987).

Environmental samples

Anaerobic digester sludge was collected from two wastewater treatment plants, one located in Knoxville (KUB) and the other in Lenoir City (LC), TN. Groundwater samples from six monitoring wells representing within plume, fringe and outside locations at a DCM-contaminated site were obtained from CDM Smith (Wright et al., 2017). The groundwater samples were shipped with an overnight carrier in a cooler with ice and analyzed immediately upon receipt. Frank Stewart (Montana State University) provided archived DNA samples from two vertical transcects from the ETNP OMZ.

DCM enrichments

Microcosms were established in 160-mL glass serum bottles containing 98 mL of anoxic mineral basal salt medium amended with 156 µmol (10 µL) of DCM. The microcosms were seeded with 2 mL of digester sludge, and additional DCM feedings occurred upon the depletion of DCM. Microcosms showing DCM degradation were sequentially transferred to fresh anoxic medium with DCM as the sole electron donor with an inoculation volume of 3 mL. After eight consective transfers, solids-free enrichment cultures were obtained that degraded DCM under anoxic conditions. DNA samples were extracted from the new DCM enrichment cultures and used to examine the presence of *mecE* and *mecF* genes by qPCR.

DNA extraction

DNA extraction from 1 mL anaerobic digester sludge was performed using the DNeasy PowerSoil DNA extraction kit (Qiagen, Valencia, CA). DNA from Defo and Diel cultures was extracted from 5 mL culture suspensions collected onto 0.22 µm Durapore membrane filters (Millipore, Cork, Ireland) and DNA using the DNeasy PowerLyzer PowerSoil DNA extraction kit (Qiagen, Valencia, CA) following the manufacturer's instructions.

Biomass from groundwater samples (950 mL) was collected on Supor® 0.2 µm membrane filters (Pall Lab., Ann Arbor, MI). Each filter was cut in half using a sterile scalpel and each piece was placed into a separate bead-beating tube for extraction with the DNeasy PowerLyzer PowerSoil DNA kit. The extracted

DNA was concentrated with the Zymo DNA Clean and Concentrator-25 Kit (Zymo Research, Irvine, CA). DNA concentrations were determined with fluorometry and DNA was stored at -80°C until qPCR analysis.

Primer design, PCR and qPCR analyses

Primer sets were developed for both *mecE* and *mecF* genes. The design was based on the target gene alleles identified in the genomes of Diel, Defo, Dcmf, and strain UNSWDHB and the most similar homologs from peat bog metagenomes. Additional primers were designed based on the most common *mecE* and *mecF* alleles identified in Eastern Pacific OMZ metagenomes (IMG genome Ga0066828, gene IDs 100177932 [*mecE*] and 100027434 [*mecF*]). The respective target gene sequences were aligned using ClustalW and primer sets were designed using the Primer 3 plug-in in Geneious R11.0.2 (Kearse et al., 2012) for PCR and SYBR qPCR assays. The primer sequences were blasted against NCBI nr database using the Primer BLAST program to verify specificity of the assays. The primers were obtained from a commercial supplier (Integrated DNA Technologies, Coralville, IA). For quantification of total bacterial 16S rRNA genes, previously reported Bac1055YF/Bac1392R (Ritalahti et al., 2006) and EUB338F/EUB518R primers were used (Lane, 1991; Muyzer et al., 1993). Primer sequences are listed in Table 1.

248 **Table 1.** Novel primers used for qPCR. The column "Target" refers to the source of the template
249 alleles used to design and validate the primers. peat; *mec* gene homologs derived from peat
250 metagenomes.

Assay	Primer	Sequence (5'…..3')	Target	Reference
qPCR	mecE 828F	ACCATATTGTCTTTTTGCCYCAG	Defo, Diel, Dcmf, *Dehalobacter* UNSWDHB, and peat	this study
qPCR	mecE 1007R	TACCGCCCAAATTTYTCTGC	Defo, Diel, Dcmf, *Dehalobacter* UNSWDHB, and peat	this study
qPCR	mecF 554F	TGCTTGACATGGCCGTAMTGGAC	Defo, Diel, Dcmf, *Dehalobacter* UNSWDHB, and peat	this study
qPCR	mecF 641R	GCAGGATADCCATATTTGTCTTT	Defo, Diel, Dcmf, *Dehalobacter* UNSWDHB, and peat	this study
qPCR	mecE 98F	ACGGCCTGACCTACAATGTC	ETNP OMZ samples	this study
qPCR	mecE 191R	GCCGTGATGTCATAGCCGTA	ETNP OMZ samples	this study
qPCR	mecF 612F	GCTCAAGGACAAGTACGGCT	ETNP OMZ samples	this study
qPCR	mecF 698R	CCGTATTGCTTCTTGCCGTG	ETNP OMZ samples	this study
qPCR	EUB338F	ACTCCTACGGGAGGCAGCAG	All samples	[85]
qPCR	EUB518R	ATTACCGCGGCTGCTGG	All samples	[86]
qPCR	Bac1055YF	ATGGYTGTCGTCAGCT	All samples	[84]
qPCR	Bac1392R	ACGGGCGGTGTGTAC	All samples	[84,85]

251
252
253 qPCR was performed in 10-µL volumes consisting of 5 µL 2X Power SYBR Green PCR Master Mix
254 (Applied Biosystems, Foster City, CA), 0.5 µL of each primer (final concentration of 300 nM) and 2 µL of
255 template DNA (undiluted, 1:10 and 1:100 dilutions). qPCR analysis was conducted using a QuantStudio
256 12K Flex Real Time qPCR System (Life Technologies, Carlsbad, CA) and the thermocycling program was
257 followed as initial step of 10 min at 95°C, 40 cycles of 15 sec at 95°C and 1 min at 60°C. Specific
258 amplification was confirmed by melt curve analysis and agarose gel electrophoresis as described (Hatt &
259 Löffler, 2012). Standard curves were established with linear GeneArt DNA fragments of target genes.
260 Standards were run as triplicate on each plate using ten-fold dilution series in the range of 10^1 to 10^8 gene

copies/µL. The amplification efficiencies, linear dynamic range, slope, Y-intercept and R^2 values are listed in Dataset S3. Amplification efficiencies (AE) were calculated using the equation $10^{(-1/slope)} - 1$.

Metatranscriptomics

Unprocessed Illumina metatranscriptome sequencing data generated from groundwater from a DCM contamination plume was provided by CDM Smith. This same site was previously the subject of 16S rRNA gene amplicon library analysis (Wright et al., 2017). Raw data were trimmed using Trimmomatic v0.35 with a 6:25 sliding window quality trim, Illumina adapter read-through contamination removal, and final minimum length of 25 bp (Bolger et al., 2014). Trimmed reads were assembled *de novo* using Trinity v2.8.5 under default parameters (Grabherr et al., 2011). Prokaryotic ribosomal RNA genes were detected in the transcript contigs using barrnap v0.9-2 (Seeman, 2018) and removed using a custom R script employing Biostrings v2.5.4 (Pagès et al., 2019). Remaining transcripts were queried against a database consisting of all Defo *mec* genes using TBLASTX. Transcripts aligning to at least *mecE* with a bit-score of >1,000 (approximately 70% full-length amino acid alignment) were included. Transcript coverage was calculated using kallisto (Bray et al., 2016) and TPM values were calculated using the Trinity utility script align_and_estimate_abundance.pl.

Results

Anaerobic DCM degraders share a common gene cassette

Comparative analysis of the Defo and Diel genomes revealed that the eight most similar genes (and 10 of the top 25 most similar genes) in terms of percent predicted amino acid identity were located in genetic clusters. The genome of Defo harbors a single 10-gene cluster and Diel has two highly similar clusters A and B (Figure 2, Dataset S4).

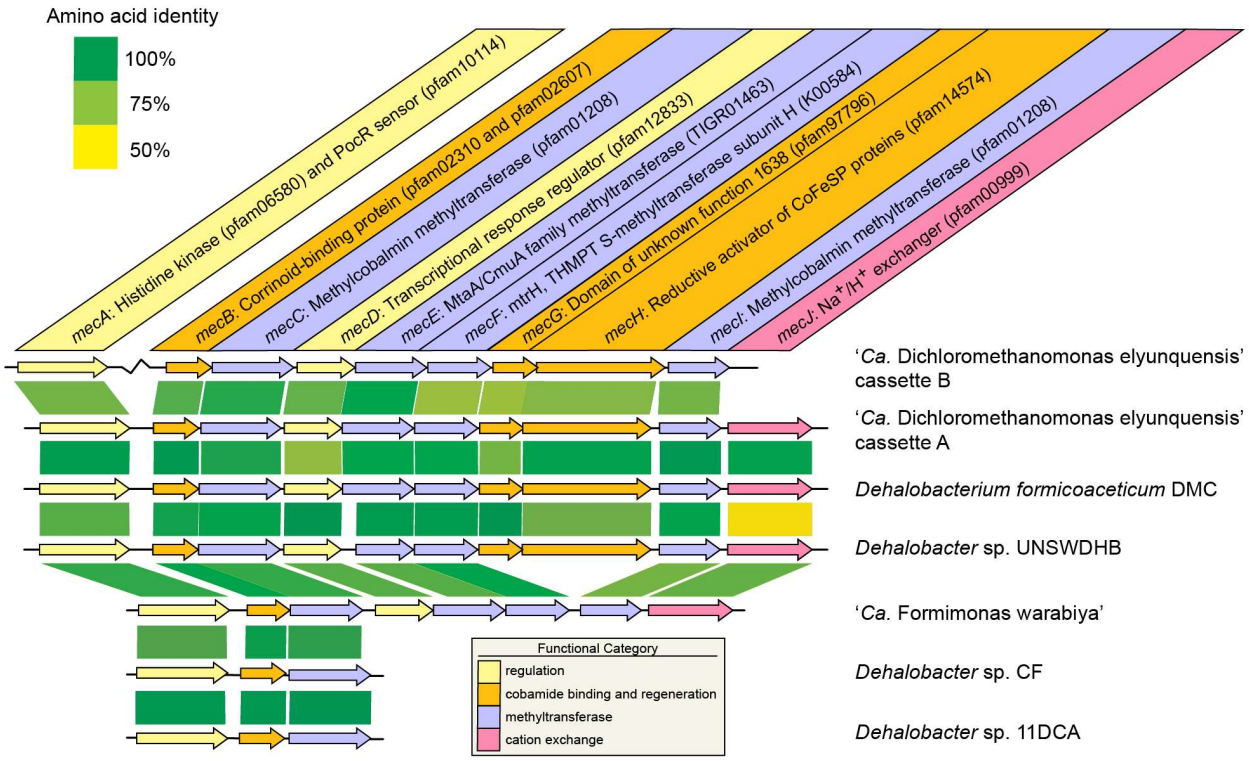

Figure 2. *mec* metabolic gene cassettes and close homologs identified in genomes. Shaded boxes represent BLASTP amino acid identity scores. The colors represent the general functional category of the encoded protein. Function was inferred from functional annotation systems, with priority given to the TIGRFAM and KEGG systems. GenoPlotR(*26*) was used to parse BLASTP results and gene coordinates to generate the figure.

The gene arrangement in these clusters is identical (i.e., syntenic) and predicted amino acid identities between clusters range from 79.6% to 99.7% (Figure 2). This novel, conserved 10-gene *mec* cassette harbors *mecA* thorough *mecJ* implicated in **me**thylene **c**hloride catabolism. A homologous gene cassette was also found in the newly described DCM degrader '*Ca.* Formimonas warabiya' (Holland et al., 2019), although lacking *mecG* and *mecH.* Aside from *mecJ*, the only close homologs in any prokaryotic genome to any of the *mec*-encoded proteins are found in the genomes of three chloroform degrading bacteria; a homologous 10-gene cassette in *Dehalobacter* sp. strain UNSWDHB and partial gene cassettes in *Dehalobacter* sp. strain CF and *Dehalobacter* sp. strain 1,1-DCA (Figure 2). Functional annotation of the *mec* cassette gene products revealed a histidine kinase sensory protein and an associated regulatory

protein (MecAD), an MtaA/CmuA methyltransferase (MecE), an MtrH methyltransferase (MecE), two methyltransferases of indeterminate function (MecCI), a corrinoid-binding protein (MecB), a cation transporter/antiporter (MecJ), a reductive activator of corrinoid proteins (MecH), and a protein with a conserved domain of unknown function (MecG) (Table 2 and Supplementary Text).

Table 2. Consensus functional annotations of the putative *mec* cassette gene product proteins. COG, Clusters of Orthologous Genes; TIGRFAM, The Instutute for Genomic Research's Database of Protein Families; pfam, Protein Familes; KEGG, Kyoto Encyclopedia of Genes and Genomes.

Protein	Annotation	Description
MecA	pfam06580	histidine kinase
	pfam10114	sensory domain found in PocR
MecB	COG5012	methanogenic corrinoid protein MtbC1
	pfam02310	B_{12}-binding
	pfam02607	B_{12}-binding 2
	TIGR02370	methyltransferase cognate corrinoid proteins
	K00548	metH, 5-methyl-THF-homocysteine methyltransferase
MecC	COG0407	uroporphyrinogen III decarboxylase
	pfam01208	uroporphyrinogen decarboxylase
	K01599	uroporphyrinogen decarboxylase
MecD	pfam00072	translational response regulator receiver domain
	pfam12833	helix-turn-helix domain
MecE	COG0407	uroporphyrinogen III decarboxylase
	pfam01208	uroporphyrinogen decarboxylase
	TIGR01463	methyltransferase, MtaA/CmuA family
MecF	COG1962	tetrahydromethanopterin S-methyltransferase, subunit H
	pfam02007	MtrH, tetrahydromethanopterin S-methyltransferase subunit H
	TIGR01114	N5-methyltetrahydromethanopterin:coenzyme M methyltransferase subunit H
	K00584	MtrH, tetrahydromethanopterin S-methyltransferase subunit H
MecG	pfam07796	domain of unknown function 1638
MecH	COG3894	uncharacterized 2Fe-2 and 4Fe-4S clusters-containing protein
	pfam14574	C-terminal of reductive activator of CoFeSP (RACo)
	pfam00111	2Fe-2S iron-sulfur cluster binding domain
MecI	pfam01208	uroporphyrinogen decarboxylase
MecJ	COG0475	Kef-type K+ transport system, membrane component KefB
	pfam00999	sodium/hydrogen exchanger family
	K03455	monovalent cation:H+ antiporter-2, CPA2 family
	K03499	KtrA, trk system potassium uptake protein

Mec proteins are expressed during growth on DCM

When grown with DCM, a total of 1,781 proteins were detected in the axenic Defo culture (Dataset S5) and 1,743 proteins were detected in the metaproteome of mixed culture RM, 797 of which were assigned to Diel (Dataset S2). The majority of proteins encoded by the *mec* gene cassettes were detected in the proteomes of both DCM degraders (Figure 3, Table S1).

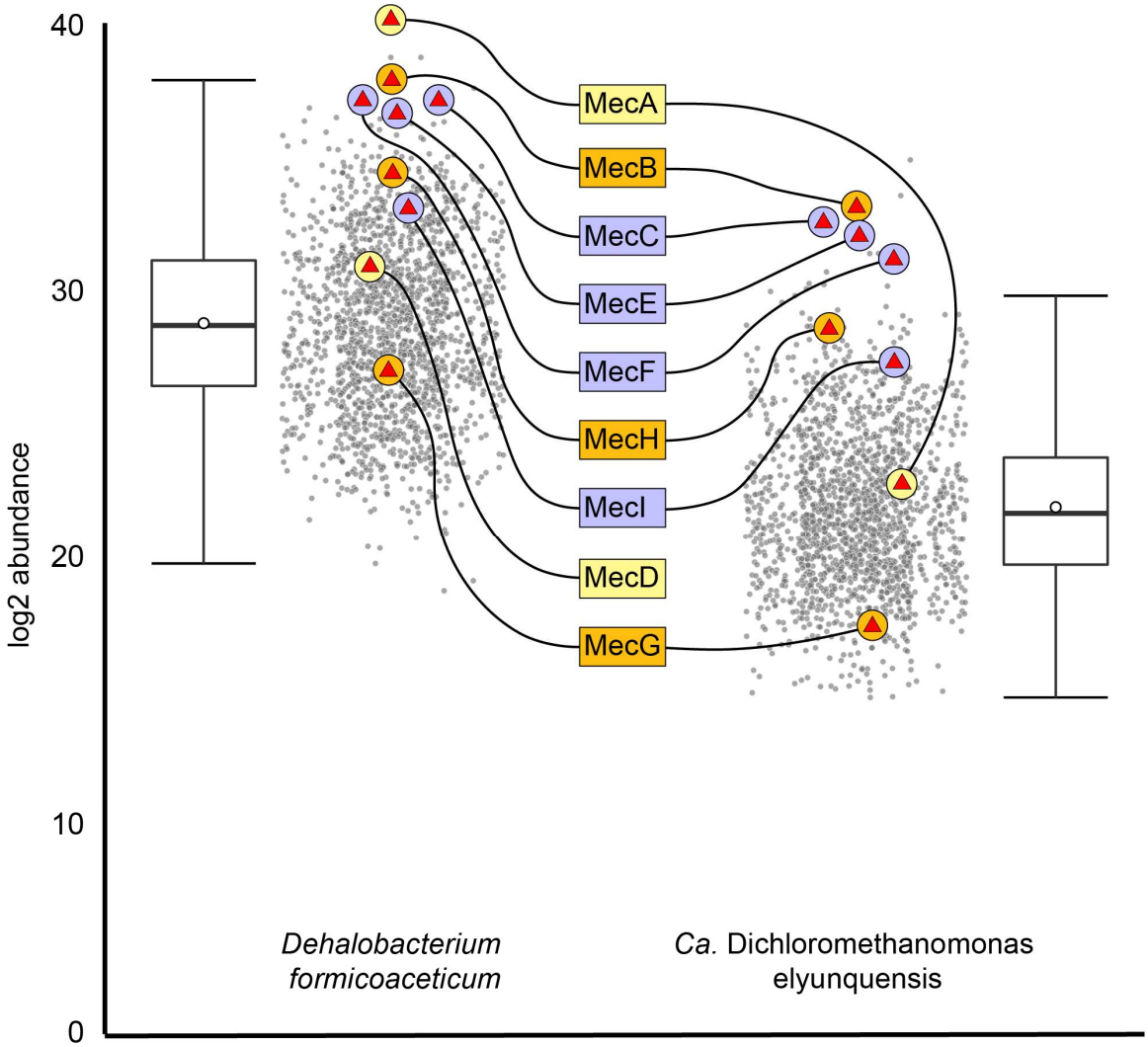

Figure 3. Mec protein abundance in the proteomes of Defo and Diel when grown anaerobically with DCM as the sole carbon and energy source. Box and whisker plots show median (central horizontal line), upper and lower quartile range (box), highest and lowest values excluding outliers (upper and lower whiskers) and mean value (open circle). Protein products of

the *mec* cassette genes are labelled, with general functional category indicated by color (orange, corrinoid-related; purple, methyltransferase; yellow, regulatory).

In Diel, all proteins of *mec* cassette B were detected except for MecD and MecJ while in Defo, all but MecJ were detected. The corrinoid-binding protein MecB was the 2nd and 3rd most abundant protein in Defo and Diel proteomes, respectively. The three methyltransferases MecC, MecE, and MecF were in the top 1% most abundant proteins in both proteomes. The fourth methytransferase MecI and the corrinoid protein reductive activator MecH were all in the upper quartile of detected proteins. The sensor histidine kinase MecA was 1st and 38th most abundant in Defo and Diel proteomes, respectively. In neither case was MecJ detected, although its predicted eleven transmembrane alpha helices suggest strong association with the cytoplasmic membrane, likely hindering detection in the proteomics measurements (Vit & Petrak, 2017).

DCM enriches for bacteria harboring mec genes

Targeted qPCR assays for the *mecE* and *mecF* genes did not yield quantifiable signals with template DNA extracted from anaerobic digestor sludge. Following enrichment with DCM using the same anaerobic sludge as inoculum, 1 x 10^6 to 5 x 10^7 gene copies/mL of both *mecE* and *mecF* were measured in transfer cultures, consistent with the observed consumption of DCM (Figure S1).

Groundwater samples from a DCM plume provided a unique opportunity to explore *mec* gene abundance and expression in response to varying DCM concentrations. *mecE*- and *mecF*-targeted qPCR assays yielded signals in a DCM dose-dependent manner, with ratios of *mec* gene copy number versus total bacterial 16S rRNA gene copy numbers ranging between 4% and 10% in groundwater from wells with 3.2 mg/L DCM or higher (Figure 4B).

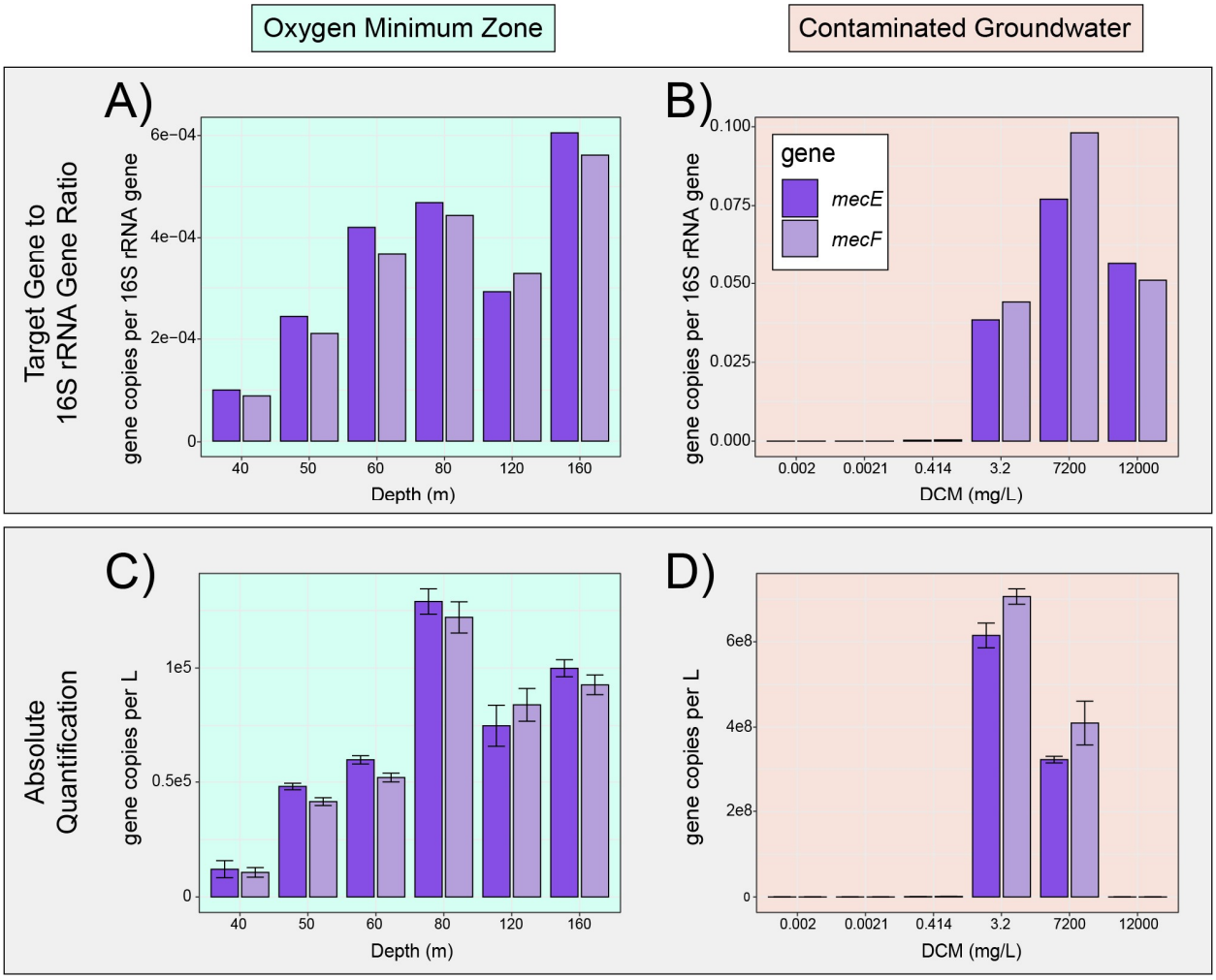

Figure 4. Quantitative PCR of *mecE* and *mecF* in the Eastern Pacific oxygen minimum zone (OMZ) and from groundwater samples from a DCM plume. A) and B) represent relative quantification results expressed as gene copies per 16S rRNA gene copy number for OMZ and DCM groundwater plume samples, respectively. Panels C) and D) represent absolute quantification (i.e., qPCR) results for OMZ and DCM groundwater plume

At lower DCM concentrations of 1.47 mg/L, the relative abundance of *mec* genes dropped to 0.029%, and in wells at the fringe of the plume with DCM in the low µg/L range, *mec* gene to 16S rRNA gene ratios dropped to 0.0004 – 0.0026% (Figure 4B). Absolute *mec* gene copy numbers followed this trend, except for samples collected from the well with the highest DCM concentration of 12 g/L, where lower abundances of bacterial and archaeal 16S rRNA genes were observed (i.e., 1.44 x 10^6 versus 3.57 x 10^9

– 1.60 x 10^{10} 16S rRNA genes per L at locations with lower DCM concentrations) (Figure 4D). Absolute target gene copy numbers and ratios of target genes to total bacterial 16S rRNA copy numbers both covaried with measured DCM concentrations, with no detections outside the plume and *mec* gene to 16S rRNA gene ratios of up to 10% within the plume, indicating that roughly 1 in 10 bacterial genomes in the plume harbored a *mec* cassette.

Analysis of metatranscriptomes, previously obtained for groundwater microbiomes collected from the same DCM-contaminated site, revealed expression of *mec* cassette genes. *mecE* and *mecF* were consistently identified on the same transcript, with relative expression reaching its highest value of 44.6 transcripts per million transcripts (TPM) in wells located in the plume fringes. *mec* cassette transcripts were also detected in groundwater samples in the core plume at relative expression levels up to 21.2 TPM (Table S2).

Environmental distribution of the mec cassette

The search of over 18,000 metagenomes identified similar gene cassettes in 41 metagenomes from peatland, the deep subsurface, and marine systems (Figure 5, Figure S2). The SI provides gene and genome IDs and detailed BLAST-P results (Table S3, Dataset S6, Dataset S7).

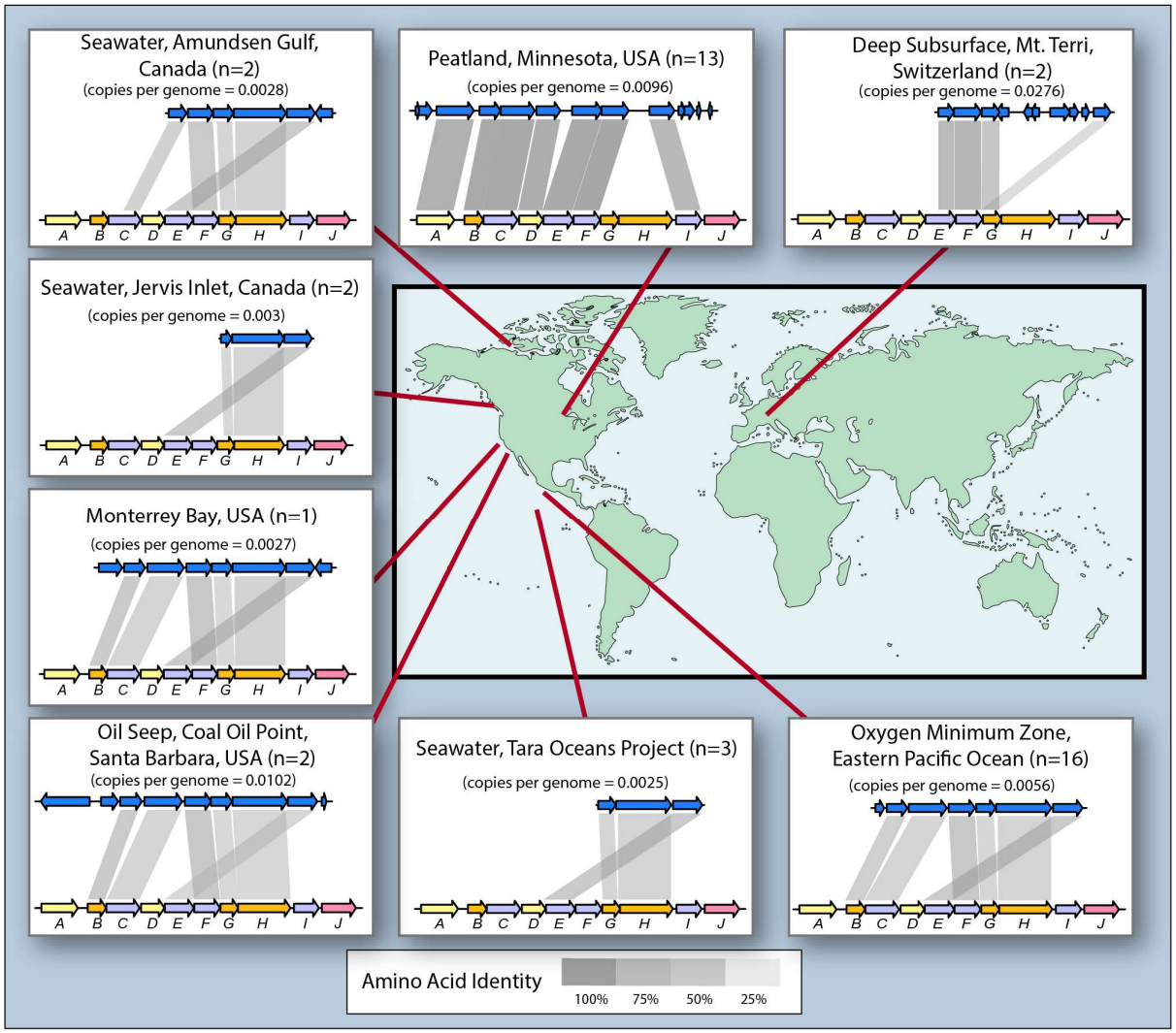

Figure 5. **Distribution of 41 gene cassettes with homology and synteny to the Defo *mec* cassette.** Identified in the IMG metagenome database, details provided in Dataset S7. Representative cassettes from the eight environments, in which the cassettes were identified, are shown. The red lines point to the approximate locations from where the metagenomes were derived. In each gene cassette illustration, the lower, multicolored gene cassette represents the *D. formicoaceticum mec* cassette, while the blue, upper illustration represents the gene cassette found in the indicated environmental metagenome. "n" indicates the number of individual samples, in which a syntenic cassette was identified, while "copies per genome" indicates the estimate of cassette copy per total genomes present in the metagenome, as determined by comparing cassette read depth to the average read depth of ten single-copy protein-encoding

385 genes for the cassette shown. Shaded boxes show the BLAST-P-derived amino acid identity
386 allowing comparisons of gene-product identities.
387
388 *mec* cassettes were identified in 13 of 117 metagenomes from ombrotrophic peat bogs in the Marcell
389 experimental forest in Minnesota, USA. The *mec* cassettes were predominantly (i.e., 10 out of 13 gene
390 cassettes) identified in metagenomes from samples collected between 1 – 1.5 m depth, where the pH
391 was approximately 4.5 and anoxic conditions prevailed (Chris Schadt, personal communication). Two
392 cassettes organized as *mecABCDEFI* displayed predicted amino acid identity scores to the Defo *mec*
393 cassette genes above 80% (Figure 2) and close phylogenetic affiliation (Figure S3). This *mecABCDEFI*
394 cassette was most prevalent, present in nearly 1% of all genome copies (i.e., approximately 1 out of 100
395 prokaryotic cells in the community harbor a *mec* cassette) (Dataset S7).
396
397 A total of 23 *mec* gene cassettes were identified in metagenomes from samples collected beneath the
398 photic zone of the oceanic water column. The majority of cassettes was identified in metagenomes from
399 the Eastern Pacific oxygen minimum zone (OMZ) at depths of 150 m to 400 m, wherein oxygen was
400 below the limit of detection (Thamdrup et al., 2019). Additional cassettes detected in metagenomes
401 derived from three coastal sites, Monterrey Bay, CA, Jervis Inlet in British Columbia, Canada, and
402 Amundsen Gulf, in Arctic northern Canada, and from an open ocean sample from the Eastern Tropical
403 North Pacific (ETNP). The marine *mec* cassettes were syntenic, with the order *mecBCFGHE* (Figure 5).
404 *mecE* and *mecF* were the most similar to the corresponding Defo homologs, with average amino acid
405 identities of 49.7 - 51.6%. All of the marine *mec* genes were more closely related to one another and to
406 the *mec* cassette genes of the characterized DCM degraders than to any other gene in assembled
407 metagenomes or genomes available in NCBI or the IMG database. The marine *mec* gene clusters
408 formed distinct, deeply branching clades (Figure S3). The highest marine *mec* cassette occurrence was
409 observed in a sample from Coal Oil Point, CA, a natural marine petroleum seep area (present in 1.02% of
410 total genome copies) (Dataset S7). qPCR applied to water column samples from two ETNP OMZ
411 locations detected *mecE* and *mecF* at a higher frequency (9 out of 11) than they were found in the ETNP

OMZ metagenome assemblies (16 out of 90), with target gene-to-bacterial 16S rRNA gene ratios of 0.01 – 0.06% (Figure 4, Dataset S8).

Evidence was obtained for the presence of *mecEFG* in anoxic porewater from a hydrogen-amended borehole in Opalinus Clay rock situated 300 m beneath Mt. Terri, Switzerland (Bagnoud et al., 2016). The amino acid identities for the putative methyltransferases MecE and MecF were 64.0 and 66.7%, respectively. Accordingly, phylogenetic analysis revealed close relationships with *mecE* and *mecF* of the known DCM degraders (Figure S3). The deep subsurface *mec* cassettes were present in 2.8% of total estimated genome copies.

Genomes of chloroform-respiring organisms harbor mec gene orthologs

The *mec* gene orthologs comprise cohesive, deeply branching clades (Figure S3). Aside from *mecJ,* no close homologs (i.e., proteins with >35% amino acid identity) to the *mec* genes are found in any publicly available bacterial or archaeal genomes, with three notable exceptions. Homologs of *mecA*, *mecB*, and *mecC* are found on genomes of the chloroform (CF) respirers *Dehalobacter* sp. strain CF and *Dehalobacter* sp. strain 11DCA, both of which were reported to generate DCM as an end product during growth with CF as electron acceptor (Grostern et al., 2010). In both of these genomes, the *mecABC* homologs are found immediately adjacent to the CF reductive dehalogenase and anchor protein encoding genes *cfrAB*, suggesting functional association between the two gene clusters. In the third case, a complete 10-gene *mec* cassette is located on the genome of *Dehalobacter* sp. strain UNSWDHB (Figure 2), also a CF-respiring organism that lacks the ability to utilize DCM (Wong et al., 2016). Close inspection of the *mec* cassette of strain UNSWDHB reveals that *mecE*, implicated in the initial chloromethyltransfer reaction, is truncated at the 5' end, which is projected to lead to a ~70 amino acid shorter protein (Figure S4), consistent with a loss of function (Wong et al., 2016).

Discussion

Identification DCM biomarker genes

The 10 gene *mec* cassette was initially identified by comparative genomic analyses between the anaerobic DCM degrading bacteria *Dehalobacterium formicoaceticum* (Defo) and '*Ca.*' Dichloromethanomonas elyunquensis (Diel). A third highly similar and syntenic gene cassette was also identified in the newer DCM degrader '*Ca.*' Formamonas warabiya (Dcmf). This high degree of gene identity and gene order among DCM degraders and the near total absence of any closely-related gene cassette in any other bacterial genome placed strong suspicion on this gene cassette as being involved in anaerobic DCM catabolism. To confirm the association of these genes with anaerobic growth on DCM, additional enrichments were generated from municipal wastewater in Eastern Tennessee, which led to an increase in the abundance of *mec* genes from undetectable to levels similar to those seen in Defo and Diel cultures (1e6 to 1e8 gene copies per mL). Mec proteins were also among the most abundant proteins in the proteomes of DCM-grown Defo and Diel. Furthermore, it was observed that *mec* genes are enriched in the bacterial community at a DCM-contaminated groundwater site and were expressed in proportion to DCM concentration. The multiple lines of evidence espouse a strong and exclusive coordination between anaerobic DCM metabolism and *mec* cassette gene, transcript and protein abundance.

Functional annotations of the mec cassette are consistent with a DCM dehalogenating methyltransferase system

The DCM-degrading bacteria are strict anaerobes and no genes consistent with aerobic DCM catabolism by a GST (e.g., *dcmA*) were identified in the genomes of strains Defo, Diel, or Dcmf. Previous biochemical studies on Defo demonstrated that DCM is channeled into the WLP via methylene-THF catalyzed by a corrinoid methyltransferase (G. Chen et al., 2018, 2020; A Mägli et al., 1998; Andreas Mägli et al., 1996). Guided by functional annotation (Table 2) and comparison to analogous metabolic systems, the proteins encoded by the *mec* cassette genes may comprise a DCM catabolic pathway that is compatible with available biochemical evidence (Figure 6).

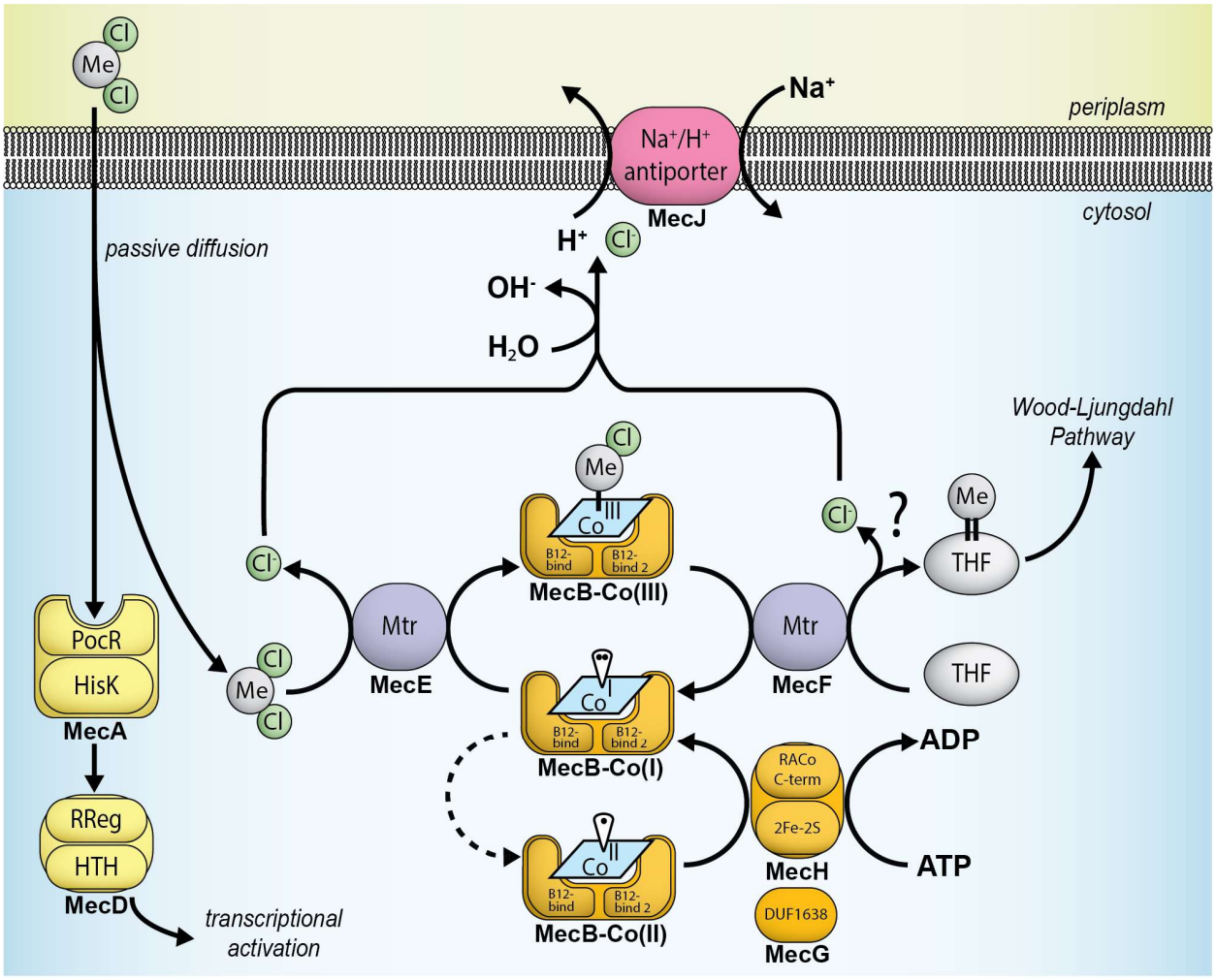

Figure 6. Proposed pathway for anaerobic DCM catabolism catalyzed by proteins encoded by the *mec* cassette identified in this study. Proteins with regulatory (yellow), methyltransferase (purple), B_{12}-related (orange), transport (pink) functions are depicted. PocR, sensory domain found in PocR; HisK, histidine kinase; RReg, translational response regulator receiver domain; HTH, helix-turn-helix domain; Mtr, methyltransferase; B12-bind, B_{12}-binding domain; RACo, reductive activator of CoFeSP; THF, tetrahydrofolate. Mechanistic understanding about the release of the second chlorine substituent has not been obtained, as indicated by the question mark. Experimental CSIA data suggest that different mechanisms operate in different strains (*40*).

The core system involved in C_1 group transfer is likely composed of two methyltransferases, MecE and MecF, and a corrinoid-binding protein, MecB. MecE is a member of the MtaA/CmuA family (TIGR01463), an ortholog family defined by methylamine, methanol and chloromethane methyltransferases. The relationship with such methyltransferases, especially the chloromethane dehalogenase CmuA of the aerobic bacterium *Hyphomicrobium* sp. strain MC1 (Borodina et al., 2004), renders MecE a candidate for catalyzing the initial chloromethyltransfer reaction to remove one chlorine substituent from DCM. Chloromethane is dechlorinated by a corrinoid-dependent methyltransferase (CmuA), suggesting MecE could perform an analogous single dechlorination of DCM during C_1 group transfer to the corrinoid-binding protein MecB. Such a reaction would yield a hypothetical chloromethyl corrinoid protein (Figure 6). Methyltransferase MecF is a member of the MtrH enzyme family responsible for transfer of a corrinoid-bound methyl group to the organic methyl carrier tetrahydromethanopterin (THMPT), a cofactor for C_1 transfer analogous to tetrahydrofolate (THF). MecF is the best candidate for the final methyltransfer reaction to THF generating methylene-THF, for which biochemical evidence exists (A Mägli et al., 1998). This methyltransfer reaction could lead to C-Cl bond cleavage and release of the second chlorine substituent. Alternatively, another methyltransferase, such as MecC or MecI, or another enzyme not necessarily conserved among all anaerobic DCM degraders, is involved in the release of the chlorine substitutent. Previously, C and Cl stable isotope measurements provided evidence for distinct dechlorination mechanisms in anaerobic DCM degraders, and ^{13}C-tracer experiments demonstrated different end products, even though both Diel and Defo employ the WLP for DCM catabolism (G. Chen et al., 2018, 2020).

Features of the remaining *mec* cassette gene products suggest roles in support of the dehalogenating C_1 group transfer system and gene expression regulation. MecGH share domain organization with proteins implicated in activating the corrinoid-binding protein involved in methyltransferase-mediated methionine synthesis in a variety of bacteria (Price et al., 2018). Thus, MecGH may function as a reductive activator for MecB. MecJ is predicted to be a monovalent cation:proton antiporter, possibly supporting DCM catabolism by maintaining acid-base homeostasis (E Pinner, 1994; Roosild et al., 2010), counteracting acidification of the cytosol by hydrochloric acid (i.e., proton) generation during dechlorination of DCM

(Ferguson et al., 2000). Co-localization of dehalogenase genes (e.g., reductive dehalogenases, haloacid dehalogenases) and *mecJ* homologs is common among related organisms (Table S4), and the gene cluster *cmuABC* for aerobic chloromethane utilization is 6 kb upstream and in the same orientation of a gene (IMG ID 650984269) encoding a putative chloride carrier/channel (ClC) protein, a protein family also implicated in acid-base homeostasis (Iyer et al., 2002). The histidine kinase MecA likely represents a DCM sensor kinase that regulates the DNA-binding protein MecD following detection of DCM via the PocR domain (Anantharaman & Aravind, 2005). The high abundance of MecA sensor kinase in Defo and Diel proteomes (1st and 38th most abundant protein, respectively) may suggest an additional or alternative role. In the marine *mec* cassettes, the regulatory genes *mecA* and *mecD* were not detected; however, in two cases, homologs of *dcmR*, which encodes the DCM sensor and regulatory protein utilized by aerobic DCM degraders (La Roche & Leisinger, 1991), were identified directly adjacent to the *mec* cassette (Figure 7) and could fulfill regulatory functions.

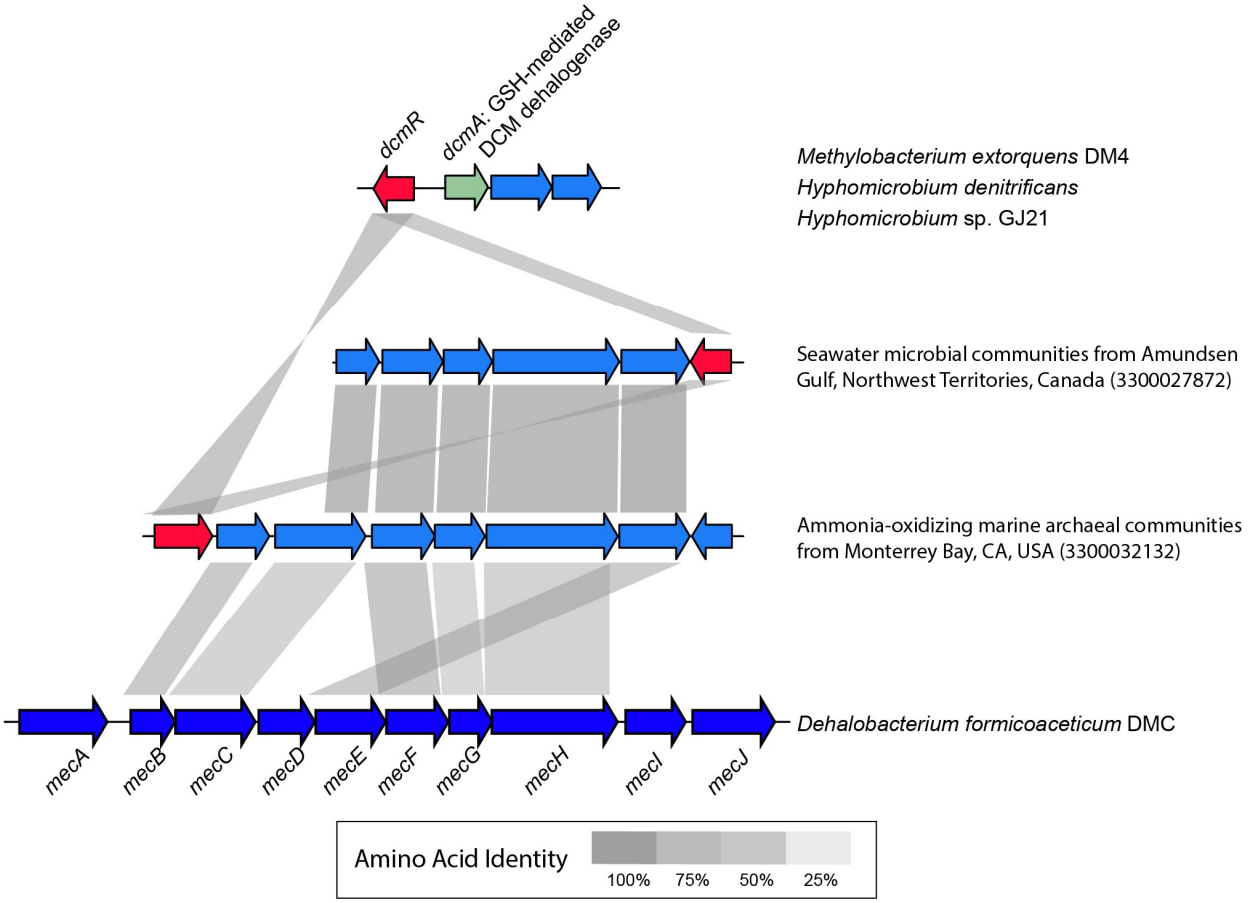

Figure 7. Two mec cassettes identified in metagenomes (from Monterrey Bay and Amundsen Gulf, CA) are flanked by genes, colored red, encoding proteins annotated as DcmR sensory domain and regulator of DCM dehalogenase DcmA (KEGG K17071). The products of these genes share 70.6% amino acid identity and are 44.6% - 46.4% identical to DcmR found in *Methylbacterium extorquens* strain DM4. The indicated gene cassette (*dcmRA*) encodes glutathione-mediated DCM dehalogenation.

The proposed involvement of additional methyltransferases (i.e., MecC and MecI) is not unprecedented. For example, the chloromethane utilization gene cluster in *Hyphomicrobium* sp. strain MC1 also contains a third putative methyltransferase encoding gene, *cmuC*, which is required for growth with chloromethane but whose function has yet to be revealed (Vannelli et al., 1999). Methanol-grown *Desulfotmaculum kuznetsovii*, which employs a similar methanol corrinoid methyltransferase system, also expresses three

methyltransferases all located in the same gene cluster (Sousa et al., 2018). Additional details of the functional annotations of the *mec* cassette genes are provided in Supplementary Information.

Highly similar mec cassettes are broadly distributed in the environment. A search of public metagenomes led to identification of *mec* cassettes in disparate ecosystems, including peatland, marine oxygen minimum zones, and the deep subsurface. However, there is good reason to suspect each of these environments as being hotspots for DCM flux. Peat bogs have been demonstrated to be rich in chlorinated organic material (Biester et al., 2004) and have been identified as a source of halomethanes (Dimmer et al., 2001). The occurrence of halomethanes in subsurface rock formations has been demonstrated (Mulder et al., 2013), although information about quantities and fluxes is lacking. Marine systems are net producers of DCM and considered major emitters to the atmosphere (Gribble, 2010). In the water column, DCM concentrations peak along with chlorophyll concentrations (Ooki & Yokouchi, 2011), consistent with production by phytoplankton and the abiotic chlorination of planktonic iodo- and bromomethanes (Ooki & Yokouchi, 2011). DCM is detected beneath the photic zone (Ooki & Yokouchi, 2011), suggesting that mixing events carry DCM into deeper waters, or other sources of DCM exist in the deep ocean, possibly hydrothermal vents (Eklund et al., 1988; Isidorov et al., 1990; Jordan et al., 2000) or settling dead biomass (Wever & Barnett, 2017). Importantly, analysis of ETNP OMZ samples using targeted qPCR assays led to much high detection rate (9 of 11) than was obtained by metagenome analyses (16 of 90). This increase in frequency of detection is likely due to the higher sensitivity of qPCR compared to shotgun metagenomics (Suttner et al., 2020) and implies that the *mec* cassette is more broadly distributed in marine systems than the metagenomics survey results suggest.

Implications of widespread anaerobic DCM degradation potential

DCM predates the anthropocene and the mechanisms underlying natural releases are far from fully characterized. Based on the available information, DCM is an energy source readily available to microorganisms in various environments, and the anaerobic microbial consumption of DCM is likely a major attenuation factor, eliminating DCM in anoxic systems prior to atmospheric release, and thus a relevant process for reducing DCM emissions. Environmental change, including global warming, has

high potential to alter the flux of DCM with unpredictable consequences for the integrity of the ozone layer. Whether low oxygen marine systems and peat bogs, for example, are net producers or net sinks of DCM is currently unclear. OMZs are expanding at accelerating rates worldwide (Stramma et al., 2008), and uncertainty exists over the impact of environmental change on net DCM emissions. Likewise, climate change induced melting of permafrost in the northern hemisphere will create more active peat bogs, and warming of peat bogs is expected to mobilize recalcitrant carbon and stimulate the breakdown of accumulated organic materials (Gill et al., 2017), which will likely increase halomethane, including DCM formation. As is the case with low oxygen marine systems, the degree to which DCM production is counterbalanced by consumption in peat bogs is unknown. A widely distributed phenotype for anaerobic DCM catabolism is likely to affect DCM pools, fluxes and thus the global DCM budget, which until present considers atmospheric releases and abiotic stratospheric breakdown, but not microbial attenuation (i.e., sinks). The increased knowledge of microbial DCM catabolism offers opportunities to include relevant sink and attenuation terms and generate refined flux models with more predictive power.

Acknowledgments

We thank F. Stewart and A. Bertagnolli for valuable discussions and providing Eastern Tropical North Pacific Oxygen Minimum Zone DNA samples for qPCR analysis. T. Macbeth, K. Sorenson, and R. Chenenko from CDM Smith graciously provided groundwater samples and metatranscriptomic data from a DCM-contaminated site. The authors acknowledge N. Ivanova and R. Seshadri from the Joint Genome Institute for providing invaluable insight into use of the IMG database.

Competing Interests: Authors declare that they have no competing interests.

Data and Code Availability: All genomes, genes, and protein sequences studied are available in the IMG database (I.-M. A. Chen et al., 2019) under the specified ID numbers. Code used for data processing and figure creation is available via RPubs links located in the methods section. All spectral data collected from the Defo axenic cultures used in this study have been deposited in the MASSIVE and ProteomeXchange repositories with identifiers MSV000087235 and PXD025479, respectively. Spectral

data collected from the Diel axenic cultures used in this study have been deposited in the MASSIVE and ProteomeXchange repositories with identifiers MSV000086520 and PXD022742, respectively. Scripts used to perform gene copy per genome copy calculations, pairwise gene cassette alignments, and to construct gene cassette synteny plots can be found at https://rpubs.com/rmurdoch/mec_cassette_abundance_and_synteny. Detailed description of gene phylogeny pipelines can be found at https://rpubs.com/rmurdoch/mec_cassette_trees. Full metatranscriptome analysis pipeline description and scripts used to operate programs can be found at https://rpubs.com/rmurdoch/mec_transcriptomes.

References

Allocati, N., Federici, L., Masulli, M., & Di Ilio, C. (2009). Glutathione transferases in bacteria. In *FEBS Journal* (Vol. 276, Issue 1, pp. 58–75). John Wiley & Sons, Ltd. https://doi.org/10.1111/j.1742-4658.2008.06743.x

Altschul, S. F., Gish, W., Miller, W., Myers, E. W., & Lipman, D. J. (1990). Basic local alignment search tool. *Journal of Molecular Biology*, *215*(3), 403–410. https://doi.org/10.1016/S0022-2836(05)80360-2

Anantharaman, V., & Aravind, L. (2005). MEDS and PocR are novel domains with a predicted role in sensing simple hydrocarbon derivatives in prokaryotic signal transduction systems. *Bioinformatics*, *21*(12), 2805–2811. https://doi.org/10.1093/bioinformatics/bti418

Bagnoud, A., Chourey, K., Hettich, R. L., de Bruijn, I., Andersson, A. F., Leupin, O. X., Schwyn, B., & Bernier-Latmani, R. (2016). Reconstructing a hydrogen-driven microbial metabolic network in Opalinus Clay rock. *Nature Communications*, *7*(1), 12770. https://doi.org/10.1038/ncomms12770

Biester, H., Keppler, F., Putschew, A., Martinez-Cortizas, A., & Petri, M. (2004). Halogen retention, organohalogens, and the role of organic matter decomposition on halogen enrichment in two Chilean peat bogs. *Environmental Science and Technology*, *38*(7), 1984–1991. https://doi.org/10.1021/es0348492

Bolger, A. M., Lohse, M., & Usadel, B. (2014). Trimmomatic: a flexible trimmer for Illumina sequence data. *Bioinformatics*, *30*(15), 2114–2120. https://doi.org/10.1093/bioinformatics/btu170

Borodina, E., McDonald, I. R., & Murrell, J. C. (2004). Chloromethane-dependent expression of the cmu gene cluster of Hyphomicrobium chloromethanicum. *Applied and Environmental Microbiology*, *70*(7), 4177–4186. https://doi.org/10.1128/AEM.70.7.4177-4186.2004

Bray, N. L., Pimentel, H., Melsted, P., & Pachter, L. (2016). Near-optimal probabilistic RNA-seq quantification. *Nature Biotechnology*, *34*(5), 525–527. https://doi.org/10.1038/nbt.3519

Capella-Gutiérrez, S., Silla-Martínez, J. M., & Gabaldón, T. (2009). trimAl: a tool for automated alignment trimming in large-scale phylogenetic analyses. *Bioinformatics (Oxford, England)*, *25*(15), 1972–1973. https://doi.org/10.1093/bioinformatics/btp348

Chen, G., Fisch, A. R., Gibson, C. M., Mack, • E Erin, Seger, • Edward S, Campagna, S. R., Frank, •, & Löffler, E. (2020). Mineralization versus fermentation: evidence for two distinct anaerobic bacterial degradation pathways for dichloromethane. *The ISME Journal*, *14*, 959–970. https://doi.org/10.1038/s41396-019-0579-5

Chen, G., Kleindienst, S., Griffiths, D. R., Mack, E. E., Seger, E. S., & Löffler, F. E. (2017). Mutualistic interaction between dichloromethane- and chloromethane-degrading bacteria in an anaerobic mixed culture. *Environmental Microbiology*, *19*(11), 4784–4796. https://doi.org/10.1111/1462-2920.13945

Chen, G., Murdoch, R. W., Mack, E. E., Seger, E. S., & Löffler, F. E. (2017). Complete genome sequence of Dehalobacterium formicoaceticum strain DMC, a strictly anaerobic dichloromethane-degrading

631 bacterium. *Genome Announcements*, *5*(37). https://doi.org/10.1128/genomeA.00897-17
632 Chen, G., Shouakar-Stash, O., Phillips, E., Justicia-Leon, S. D., Gilevska, T., Sherwood Lollar, B., Mack,
633 E. E., Seger, E. S., & Löffler, F. E. (2018). Dual carbon–chlorine isotope analysis indicates distinct
634 anaerobic dichloromethane degradation pathways in two members of Peptococcaceae.
635 *Environmental Science & Technology*, *52*(15), 8607–8616. https://doi.org/10.1021/acs.est.8b01583
636 Chen, I.-M. A., Chu, K., Palaniappan, K., Pillay, M., Ratner, A., Huang, J., Huntemann, M., Varghese, N.,
637 White, J. R., Seshadri, R., Smirnova, T., Kirton, E., Jungbluth, S. P., Woyke, T., Eloe-Fadrosh, E. A.,
638 Ivanova, N. N., & Kyrpides, N. C. (2019). IMG/M v.5.0: an integrated data management and
639 comparative analysis system for microbial genomes and microbiomes. *Nucleic Acids Research*,
640 *47*(D1), D666–D677. https://doi.org/10.1093/nar/gky901
641 Cox, M. L., Fraser, P. J., Sturrock, G. A., Siems, S. T., & Porter, L. W. (2004). Terrestrial sources and
642 sinks of halomethanes near Cape Grim, Tasmania. *Atmospheric Environment*, *38*(23), 3839–3852.
643 https://doi.org/10.1016/J.ATMOSENV.2004.03.050
644 Dimmer, C. H., Simmonds, P. G., Nickless, G., & Bassford, M. R. (2001). Biogenic fluxes of
645 halomethanes from Irish peatland ecosystems. *Atmospheric Environment*, *35*(2), 321–330.
646 https://doi.org/10.1016/S1352-2310(00)00151-5
647 E Pinner, E. P. and S. S. (1994). Kinetic properties of NhaB, a Na+/H+ antiporter from Escherichia coli.
648 *The Journal of Biological Chemistry*, *269*, 26274–26279.
649 http://www.jbc.org/content/268/8/5382?ijkey=264978fabe3ca92c7251ed45ef58e1776e4abe69&keyt
650 ype2=tf_ipsecsha
651 Eklund, G., Pedersen, J. R., & Stromberg, B. (1988). Methane, hydrogen chloride and oxygen form a
652 wide range of chlorinated organic species in the temperature range 400°C-950°C. *Chemosphere*,
653 *17*(3), 575–586.
654 El-Gebali, S., Mistry, J., Bateman, A., Eddy, S. R., Luciani, A., Potter, S. C., Qureshi, M., Richardson, L.
655 J., Salazar, G. A., Smart, A., Sonnhammer, E. L. L., Hirsh, L., Paladin, L., Piovesan, D., Tosatto, S.
656 C. E., & Finn, R. D. (2019). The Pfam protein families database in 2019. *Nucleic Acids Research*,
657 *47*, D427–D432. https://doi.org/10.1093/nar/gky995
658 Eng, J. K., McCormack, A. L., & Yates, J. R. (1994). An approach to correlate tandem mass spectral data
659 of peptides with amino acid sequences in a protein database. *Journal of the American Society for*
660 *Mass Spectrometry*, *5*(11), 976–989. https://doi.org/10.1016/1044-0305(94)80016-2
661 Eustáquio, A. S., Pojer, F., Noel, J. P., & Moore, B. S. (2008). Discovery and characterization of a marine
662 bacterial SAM-dependent chlorinase. *Nature Chemical Biology*, *4*(1), 69–74.
663 https://doi.org/10.1038/nchembio.2007.56
664 Ferguson, G. P., Booth, I. R., Evans, G. J., & Vuilleumier, S. (2000). Growth inhibition of Escherichia coli
665 by dichloromethane in cells expressing dichloromethane dehalogenase/glutathione S-transferase.
666 *Microbiology*, *146*(11), 2967–2975. https://doi.org/10.1099/00221287-146-11-2967
667 Ganusova, E. E., Vo, L. T., Abraham, P. E., O'Neal Yoder, L., Hettich, R. L., & Alexandre, G. (2021). The
668 Azospirillum brasilense core chemotaxis proteins CheA1 and CheA4 link chemotaxis signaling with
669 nitrogen metabolism. *MSystems*, *6*(1). https://doi.org/10.1128/msystems.01354-20
670 Gill, A. L., Giasson, M.-A., Yu, R., & Finzi, A. C. (2017). Deep peat warming increases surface methane
671 and carbon dioxide emissions in a black spruce-dominated ombrotrophic bog. *Global Change*
672 *Biology*, *23*(12), 5398–5411. https://doi.org/10.1111/gcb.13806
673 Gossett, J. M. (1987). Measurement of Henry's Law constants for C1 and C2 chlorinated hydrocarbons.
674 *Environmental Science and Technology*, *21*(2), 202–208. https://doi.org/10.1021/es00156a012
675 Grabherr, M. G., Haas, B. J., Yassour, M., Levin, J. Z., Thompson, D. A., Amit, I., Adiconis, X., Fan, L.,
676 Raychowdhury, R., Zeng, Q., Chen, Z., Mauceli, E., Hacohen, N., Gnirke, A., Rhind, N., Di Palma,
677 F., Birren, B. W., Nusbaum, C., Lindblad-Toh, K., … Regev, A. (2011). Full-length transcriptome
678 assembly from RNA-Seq data without a reference genome. *Nature Biotechnology*, *29*(7), 644–652.
679 https://doi.org/10.1038/nbt.1883
680 Gribble, G. W. (2010). *Naturally Occurring Organohalogen Compounds - A Comprehensive Update* (Vol.
681 91). Springer Vienna. https://doi.org/10.1007/978-3-211-99323-1
682 Grostern, A., Duhamel, M., Dworatzek, S., & Edwards, E. A. (2010). Chloroform respiration to
683 dichloromethane by a Dehalobacter population. *Environmental Microbiology*, *12*(4), 1053–1060.
684 https://doi.org/10.1111/j.1462-2920.2009.02150.x
685 Guy, L., Roat Kultima, J., & Andersson, S. G. E. (2010). genoPlotR: comparative gene and genome
686 visualization in R. *Bioinformatics*, *26*(18), 2334–2335. https://doi.org/10.1093/bioinformatics/btq413

687 Haft, D. H., Selengut, J. D., Richter, R. A., Harkins, D., Basu, M. K., & Beck, E. (2012). TIGRFAMs and
688 Genome Properties in 2013. *Nucleic Acids Research*, *41*(D1), D387–D395.
689 https://doi.org/10.1093/nar/gks1234
690 Haines, A., Ebi, K. L., Ph, D., Li, S., & Ph, D. (2020). Wildfires, global climate change, and human health.
691 *The New England Journal of Medicine*, *383*, 2173–2181.
692 Hatt, J. K., & Löffler, F. E. (2012). Quantitative real-time PCR (qPCR) detection chemistries affect
693 enumeration of the Dehalococcoides 16S rRNA gene in groundwater. *Journal of Microbiological*
694 *Methods*, *88*(2), 263–270. https://doi.org/10.1016/j.mimet.2011.12.005
695 Hoekstra, E. J., M Verhagen, F. J., Field, J. A., B De Leer, E. W., & Th Brinkman, U. A. (1998). Natural
696 production of chloroform by fungi. *Phytochemistry*, *49*(1), 91–97.
697 Holland, S. I., Edwards, R. J., Ertan, H., Wong, Y. K., Russell, T. L., Deshpande, N. P., Manefield, M., &
698 Lee, M. J. (2019). Whole genome sequencing of a novel, dichloromethane-fermenting
699 Peptococcaceae from an enrichment culture. *PeerJ*, *7*, e7775.
700 https://doi.org/10.7287/peerj.preprints.27718v1
701 Holland, S. I., Ertan, H., Montgomery, K., Manefield, M. J., & Lee, M. (2021). Novel dichloromethane-
702 fermenting bacteria in the Peptococcaceae family. *The ISME Journal*, *15*, 1709–1721.
703 https://doi.org/10.1038/s41396-020-00881-y
704 Hossaini, R., Chipperfield, M. P., Montzka, S. A., Leeson, A. A., Dhomse, S. S., & Pyle, J. A. (2017). The
705 increasing threat to stratospheric ozone from dichloromethane. *Nature Communications*, *8*(1),
706 15962. https://doi.org/10.1038/ncomms15962
707 Hu, S., Niu, Z., Chen, Y., Li, L., & Zhang, H. (2017). Global wetlands: Potential distribution, wetland loss,
708 and status. *Science of The Total Environment*, *586*, 319–327.
709 https://doi.org/10.1016/J.SCITOTENV.2017.02.001
710 Isidorov, V. A., Zenkevich, I. G., & Ioffe, B. V. (1990). Volatile organic compounds in solfataric gases.
711 *Journal of Atmospheric Chemistry*, *10*, 329.
712 https://link.springer.com/content/pdf/10.1007%2FBF00053867.pdf
713 Iyer, R., Iverson, T. M., Accardi, A., & Miller, C. (2002). A biological role for prokaryotic ClC chloride
714 channels. *Nature*, *419*(6908), 715–718. https://doi.org/10.1038/nature01000
715 Jordan, A., Harnisch, J., Borchers, R., Le Guern, F., & Shinohar, H. (2000). Volcanogenic halocarbons.
716 *Environmental Science & Technology*, *34*(6), 1122–1124. https://doi.org/10.1021/ES990838Q
717 Justicia-Leon, S. D., Ritalahti, K. M., Mack, E. E., & Löffler, F. E. (2012). Dichloromethane fermentation by
718 a Dehalobacter sp. in an enrichment culture derived from pristine river sediment. *Applied and*
719 *Environmental Microbiology*, *78*(4), 1288–1291. https://doi.org/10.1128/AEM.07325-11
720 Käll, L., Canterbury, J. D., Weston, J., Noble, W. S., & MacCoss, M. J. (2007). Semi-supervised learning
721 for peptide identification from shotgun proteomics datasets. *Nature Methods*, *4*(11), 923–925.
722 https://doi.org/10.1038/nmeth1113
723 Kanehisa, M., Furumichi, M., Tanabe, M., Sato, Y., & Morishima, K. (2017). KEGG: new perspectives on
724 genomes, pathways, diseases and drugs. *Nucleic Acids Research*, *45*(D1), D353–D361.
725 https://doi.org/10.1093/nar/gkw1092
726 Kanehisa, M., Sato, Y., & Morishima, K. (2016). BlastKOALA and GhostKOALA: KEGG tools for
727 functional characterization of genome and metagenome sequences. *Journal of Molecular Biology*,
728 *428*(4), 726–731. https://doi.org/10.1016/j.jmb.2015.11.006
729 Kanters, J., & Louw, R. (1996). Thermal and catalysed halogenation in combustion reactions.
730 *Chemosphere*, *32*(1), 89–97. https://doi.org/10.1016/0045-6535(95)00230-8
731 Katoh, K., & Standley, D. M. (2013). MAFFT multiple sequence alignment software version 7:
732 improvements in performance and usability. *Molecular Biology and Evolution*, *30*(4), 772–780.
733 https://doi.org/10.1093/molbev/mst010
734 Kearse, M., Moir, R., Wilson, A., Stones-Havas, S., Cheung, M., Sturrock, S., Buxton, S., Cooper, A.,
735 Markowitz, S., Duran, C., Thierer, T., Ashton, B., Meintjes, P., Drummond, A., & Valencia, A. (2012).
736 Geneious Basic: An integrated and extendable desktop software platform for the organization and
737 analysis of sequence data. *Bioinformatics (Oxford, England)*, *28*(12), 1647–1649.
738 https://doi.org/10.1093/bioinformatics/bts199
739 Kleindienst, S., Higgins, S. A., Tsementzi, D., Chen, G., Konstantinidis, K. T., Mack, E. E., & Löffler, F. E.
740 (2017). ‘ Candidatus Dichloromethanomonas elyunquensis’ gen. nov., sp. nov., a dichloromethane-
741 degrading anaerobe of the Peptococcaceae family. *Systematic and Applied Microbiology*, *40*(3),
742 150–159. https://doi.org/10.1016/j.syapm.2016.12.001

743 Kleindienst, S., Higgins, S. A., Tsementzi, D., Konstantinidis, K. T., Mack, E. E., & Löffler, F. E. (2016).
744 Draft genome sequence of a strictly anaerobic dichloromethane-degrading bacterium. *Genome*
745 *Announcements*, *4*(2), e00037-16. https://doi.org/10.1128/GENOMEA.00037-16
746 Kolusu, S. R., Schlünzen, K. H., Grawe, D., & Seifert, R. (2018). Determination of chloromethane and
747 dichloromethane in a tropical terrestrial mangrove forest in Brazil by measurements and modelling.
748 *Atmospheric Environment*, *173*, 185–197. https://doi.org/10.1016/J.ATMOSENV.2017.10.057
749 La Roche, S. D., & Leisinger, T. (1991). Identification of dcmR, the regulatory gene governing expression
750 of dichloromethane dehalogenase in Methylobacterium sp. strain DM4. *Journal of Bacteriology*,
751 *173*(21), 6714–6721. http://jb.asm.org/
752 Lane, D. J. (1991). 16S/23S rRNA sequencing. In Stackebrandt & M. Goodfellow (Eds.), *Nucleic acid*
753 *techniques in bacterial systematics* (pp. 115–175). John Wiley & Sons, Ltd.
754 Letunic, I., & Bork, P. (2016). Interactive tree of life (iTOL) v3: an online tool for the display and annotation
755 of phylogenetic and other trees. *Nucleic Acids Research*, *44*(W1), W242–W245.
756 https://doi.org/10.1093/nar/gkw290
757 Li, W., & Godzik, A. (2006). Cd-hit: a fast program for clustering and comparing large sets of protein or
758 nucleotide sequences. *Bioinformatics*, *22*(13), 1658–1659.
759 https://doi.org/10.1093/bioinformatics/btl158
760 Löffler, F. E., Sanford, R. A., & Ritalahti, K. M. (2005). Enrichment, cultivation, and detection of reductively
761 dechlorinating bacteria. *Methods in Enzymology*, *397*, 77–111. https://doi.org/10.1016/S0076-
762 6879(05)97005-5
763 Mägli, A, Messmer, M., & Leisinger, T. (1998). Metabolism of dichloromethane by the strict anaerobe
764 Dehalobacterium formicoaceticum. *Applied and Environmental Microbiology*, *64*(2), 646–650.
765 http://www.ncbi.nlm.nih.gov/pubmed/16349505
766 Mägli, Andreas, Wendt, M., & Leisinger, T. (1996). Isolation and characterization of Dehalobacterium
767 formicoaceticum gen. nov. sp. nov., a strictly anaerobic bacterium utilizing dichloromethane as
768 source of carbon and energy. *Archives of Microbiology*, *166*(2), 101–108.
769 https://doi.org/10.1007/s002030050362
770 McCulloch, A. (2017). *Dichloromethane in the environment: A note prepared for the European*
771 *Chlorinated Solvents Association (ECSA) and the Halogenated Solvents Industry Alliance (HSIA)*.
772 http://www.chlorinated-solvents.eu/images/Documents/Newsroom/Dichloromethane paper.pdf
773 Mulder, I., Huber, S. G., Krause, T., Zetzsch, C., Kotte, K., Dultz, S., & Schöler, H. F. (2013). A new purge
774 and trap headspace technique to analyze low volatile compounds from fluid inclusions of rocks and
775 minerals. *Chemical Geology*, *358*, 148–155. https://doi.org/10.1016/J.CHEMGEO.2013.09.003
776 Muller, E. E. L., Bringel, F., & Vuilleumier, S. (2011). Dichloromethane-degrading bacteria in the genomic
777 age. *Research in Microbiology*, *162*(9), 869–876. https://doi.org/10.1016/J.RESMIC.2011.01.008
778 Muyzer, G., De Waal, E. C., & Uitterlinden, A. G. (1993). Profiling of complex microbial populations by
779 denaturing gradient gel electrophoresis analysis of polymerase chain reaction-amplified genes
780 coding for 16S rRNA. *Applied and Environmental Microbiology*, *59*(3), 695–700.
781 https://doi.org/10.1128/aem.59.3.695-700.1993
782 Ooki, A., & Yokouchi, Y. (2011). Dichloromethane in the Indian Ocean: Evidence for in-situ production in
783 seawater. *Marine Chemistry*, *124*(1–4), 119–124. https://doi.org/10.1016/J.MARCHEM.2011.01.001
784 Pagès, H., Aboyou, P., Gentleman, R., & DebRoy, S. (2019). *Biostrings: Efficient manipulation of*
785 *biological strings* (R package version 2.54.0).
786 Petkau, A., Stuart-Edwards, M., Stothard, P., & Van Domselaar, G. (2010). Interactive microbial genome
787 visualization with GView. *Bioinformatics*, *26*(24), 3125–3126.
788 https://doi.org/10.1093/bioinformatics/btq588
789 Price, M. N., Dehal, P. S., & Arkin, A. P. (2010). FastTree 2 – Approximately maximum-likelihood trees for
790 large alignments. *PLOS ONE*, *5*(3), e9490. https://doi.org/10.1371/journal.pone.0009490
791 Price, M. N., Zane, G. M., Kuehl, J. V, Melnyk, R. A., Wall, J. D., Deutschbauer, A. M., & Arkin, A. P.
792 (2018). Filling gaps in bacterial amino acid biosynthesis pathways with high-throughput genetics.
793 *PLOS Genetics*, *14*(1), e1007147. https://doi.org/10.1371/journal.pgen.1007147
794 Ritalahti, K. M., Amos, B. K., Sung, Y., Wu, Q., Koenigsberg, S. S., & Löffler, F. E. (2006). Quantitative
795 PCR targeting 16S rRNA and reductive dehalogenase genes simultaneously monitors multiple
796 Dehalococcoides strains. *Applied and Environmental Microbiology*, *72*(4), 2765–2774.
797 https://doi.org/10.1128/AEM.72.4.2765-2774.2006
798 Roosild, T. P., Castronovo, S., Healy, J., Miller, S., Pliotas, C., Rasmussen, T., Bartlett, W., Conway, S.

J., & Booth, I. R. (2010). Mechanism of ligand-gated potassium efflux in bacterial pathogens. *Proceedings of the National Academy of Sciences of the United States of America*, *107*(46), 19784–19789. https://doi.org/10.1073/pnas.1012716107

Seeman, T. (2018). *tseemann/barrnap: Bacterial ribosomal RNA predictor* (0.9-2). https://github.com/tseemann/barrnap

Sousa, D. Z., Visser, M., van Gelder, A. H., Boeren, S., Pieterse, M. M., Pinkse, M. W. H., Verhaert, P. D. E. M., Vogt, C., Franke, S., Kümmel, S., & Stams, A. J. M. (2018). The deep-subsurface sulfate reducer Desulfotomaculum kuznetsovii employs two methanol-degrading pathways. *Nature Communications*, *9*(1), 239. https://doi.org/10.1038/s41467-017-02518-9

Stramma, L., Johnson, G. C., Sprintall, J., & Mohrholz, V. (2008). Expanding oxygen-minimum zones in the tropical oceans. *Science*, *320*(5876), 655–658. https://doi.org/10.1126/science.1153847

Suttner, B. J., Johnston, E. R., Orellana, L. H., Rodriguez-R, L. M., Hatt, J. K., Carychao, D., Carter, M. Q., & Cooley, M. B. (2020). Potential and limitations of metagenomics as a public health risk assessment tool in a study of natural creek sediments influenced by agricultural and livestock runoff. *Appl. Environ. Microbiol.*, *86*(6), e02525-19. https://doi.org/10.1128/AEM.02525-19

Svensen, H., Planke, S., Polozov, A. G., Schmidbauer, N., Corfu, F., Podladchikov, Y. Y., & Jamtveit, B. (2009). Siberian gas venting and the end-Permian environmental crisis. *Earth and Planetary Science Letters*, *277*(3–4), 490–500. https://doi.org/10.1016/J.EPSL.2008.11.015

Tatusov, R. L., Galperin, M. Y., Natale, D. A., & Koonin, E. V. (2000). The COG database: a tool for genome-scale analysis of protein functions and evolution. *Nucleic Acids Research*, *28*(1), 33–36. https://www.ncbi.nlm.nih.gov/pmc/articles/PMC102395/

Thamdrup, B., Steinsdóttir, H. G. R., Bertagnolli, A. D., Padilla, C. C., Patin, N. V., Garcia-Robledo, E., Bristow, L. A., & Stewart, F. J. (2019). Anaerobic methane oxidation is an important sink for methane in the ocean's largest oxygen minimum zone. *Limnology and Oceanography*, *64*(6), 2569–2585. https://doi.org/10.1002/lno.11235

Trudinger, C. M., Etheridge, D. M., Sturrock, G. A., Fraser, P. J., Krummel, P. B., McCulloch, A., & Etheridge, M. (2004). Atmospheric histories of halocarbons from analysis of Antarctic firn air: Methyl bromide, methyl chloride, chloroform, and dichloromethane. *J. Geophys. Res*, *109*, 22310. https://doi.org/10.1029/2004JD004932

Vannelli, T., Messmer, M., Studer, A., Vuilleumier, S., & Leisinger, T. (1999). A corrinoid-dependent catabolic pathway for growth of a Methylobacterium strain with chloromethane. *Proceedings of the National Academy of Sciences of the United States of America*, *96*(8), 4615–4620. https://doi.org/10.1073/pnas.96.8.4615

Vit, O., & Petrak, J. (2017). Integral membrane proteins in proteomics. How to break open the black box? *Journal of Proteomics*, *153*, 8–20. https://doi.org/10.1016/J.JPROT.2016.08.006

Wever, R., & Barnett, P. (2017). Vanadium chloroperoxidases: The missing link in the formation of chlorinated compounds and chloroform in the terrestrial environment? *Chemistry - An Asian Journal*, *12*(16), 1997–2007. https://doi.org/10.1002/asia.201700420

Wiechmann, A., Ciurus, S., Oswald, F., Seiler, V. N., & Müller, V. (2020). It does not always take two to tango: "Syntrophy" via hydrogen cycling in one bacterial cell. *ISME Journal*, *14*(6), 1561–1570. https://doi.org/10.1038/s41396-020-0627-1

Wong, Y. K., Holland, S. I., Ertan, H., Manefield, M., & Lee, M. (2016). Isolation and characterization of *Dehalobacter sp.* strain UNSWDHB capable of chloroform and chlorinated ethane respiration. *Environmental Microbiology*, *18*(9), 3092–3105. https://doi.org/10.1111/1462-2920.13287

Wright, J., Kirchner, V., Bernard, W., Ulrich, N., McLimans, C., Campa, M. F., Hazen, T., Macbeth, T., Marabello, D., McDermott, J., Mackelprang, R., Roth, K., & Lamendella, R. (2017). Bacterial community dynamics in dichloromethane-contaminated groundwater undergoing natural attenuation. *Frontiers in Microbiology*, *8*, 2300. https://doi.org/10.3389/fmicb.2017.02300

Wu, S., Zhu, Z., Fu, L., Niu, B., & Li, W. (2011). WebMGA: a customizable web server for fast metagenomic sequence analysis. *BMC Genomics*, *12*(1), 444. https://doi.org/10.1186/1471-2164-12-444

Wuosmaa, A. M., & Hager, L. P. (1990). Methyl chloride transferase: A carbocation route for biosynthesis of halometabolites. *Science*, *249*(4965), 160–162. https://www-jstor-org.proxy.lib.utk.edu/stable/pdf/2874545.pdf?refreqid=excelsior%3Ae86ae7d382fcc008630b1329bbb843ee

Yang, S., Giannone, R. J., Dice, L., Yang, Z. K., Engle, N. L., Tschaplinski, T. J., Hettich, R. L., & Brown,

855 S. D. (2012). Clostridium thermocellum ATCC27405 transcriptomic, metabolomic and proteomic
856 profiles after ethanol stress. *BMC Genomics*, *13*(1), 336. https://doi.org/10.1186/1471-2164-13-336
857

858