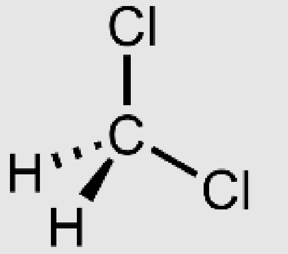
Gene Markers for Monitoring Anaerobic Dichloromethane Biodegradation: Current Progress and Future Directions

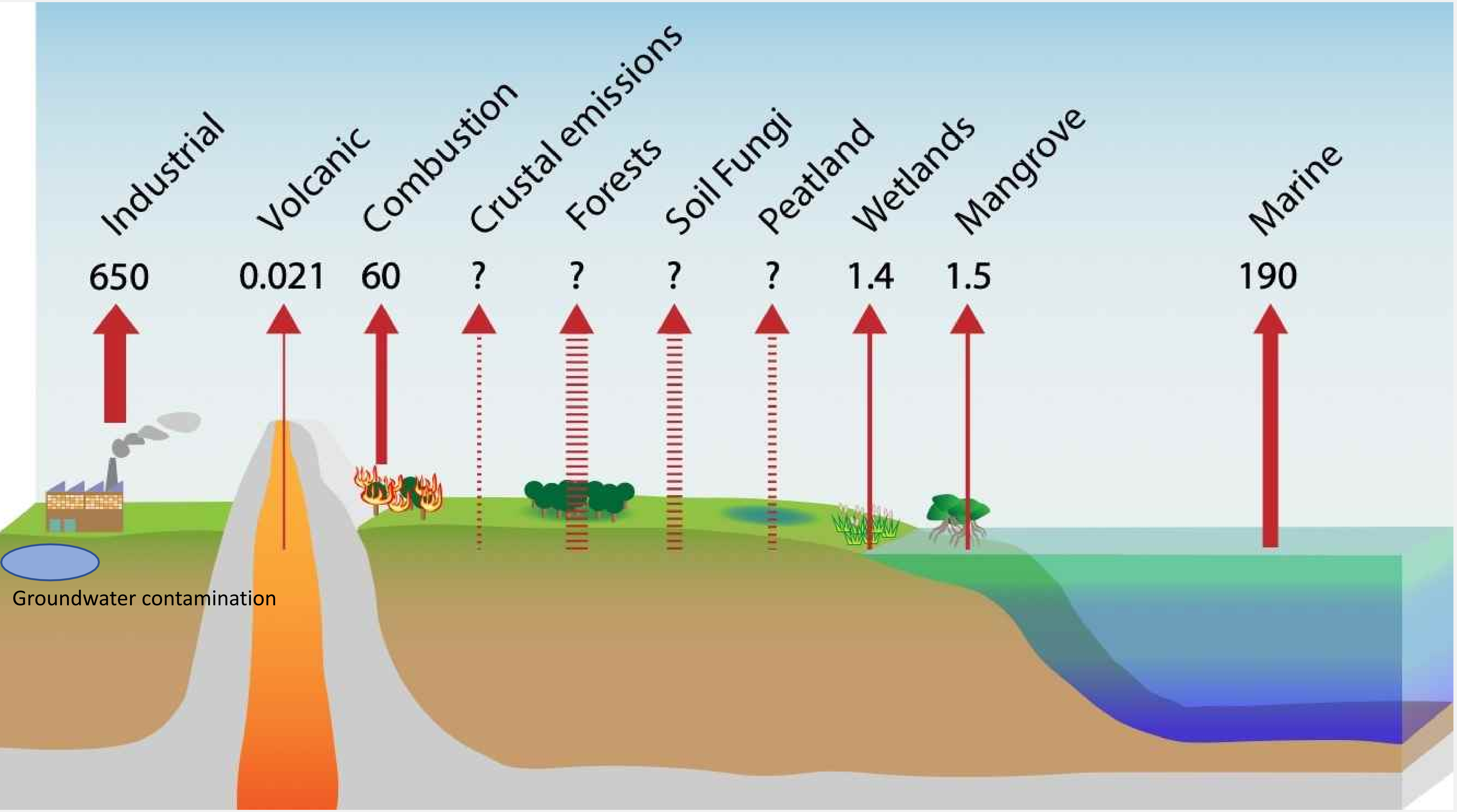
Robert W. Murdoch[1,]*
Fadime Kara Murdoch[1]
Gao Chen[2]
Frank E. Loeffler[2]

[1]Battelle Memorial Institute
[2]University of Tennessee Knoxville
[*]murdoch@battelle.org

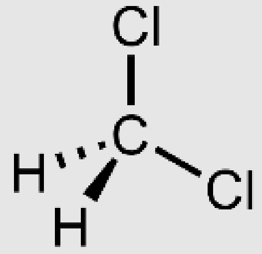
Sources of dichloromethane (DCM) atmospheric emissions



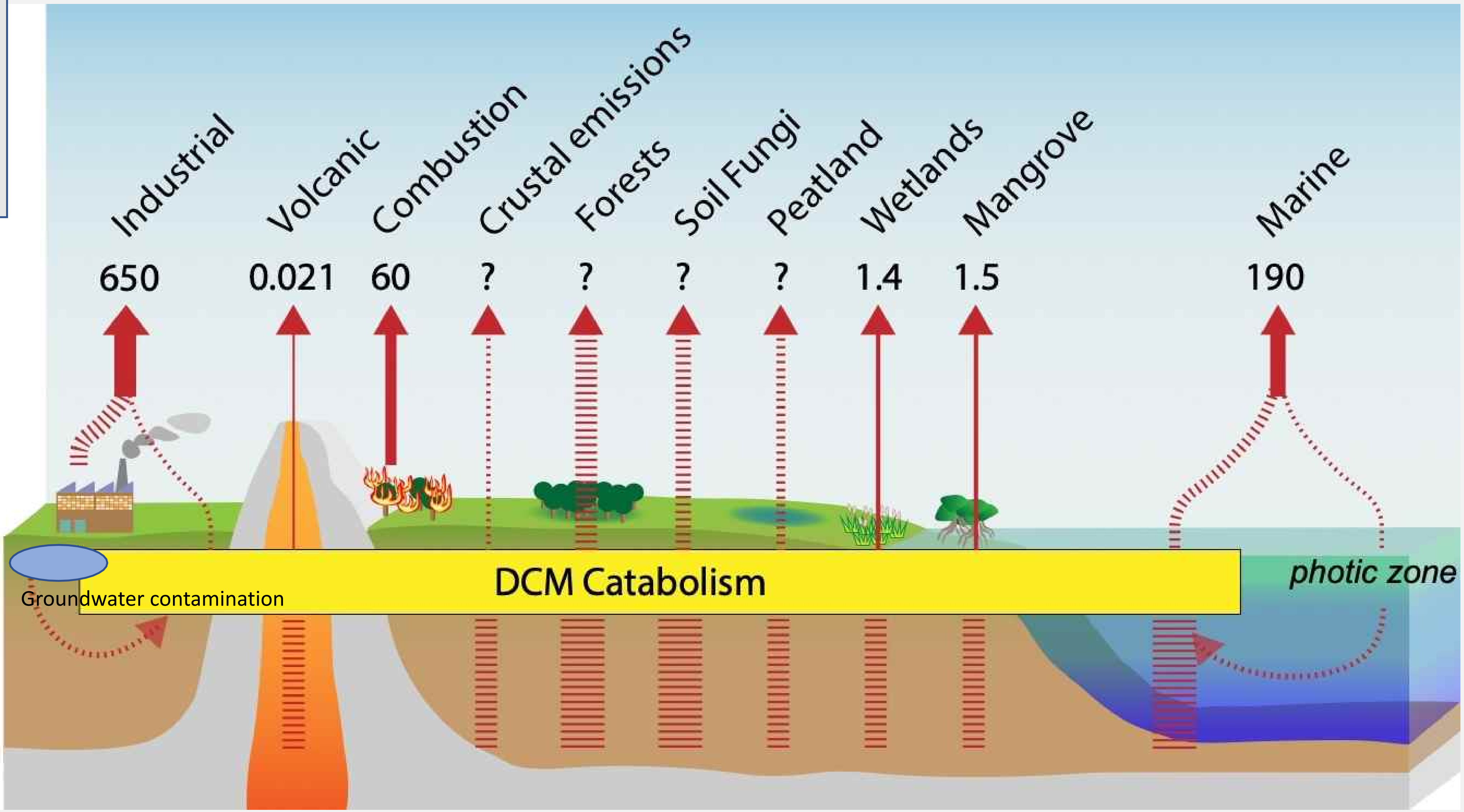
Dichloromethane,
a.k.a. methylene chloride



Sources of dichloromethane (DCM) atmospheric emissions



Dichloromethane,
a.k.a. methylene chloride

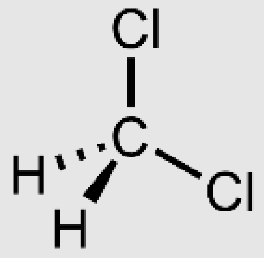


Global Change Biology

RESEARCH ARTICLE

Identification and widespread environmental distribution of a gene cassette implicated in anaerobic dichloromethane degradation

Robert W. Murdoch, Gao Chen, Fadime Kara Murdoch, E. Erin Mack, Manuel I. Villalobos Solis, Robert L. Hettich, Frank E. Löffler



Dichloromethane,
a.k.a. methylene chloride

In 2022

- Conserved genes for anaerobic DCM biodegradation identified (the "*mec* cassette")
- qPCR primers for key functional genes designed and tested

Today:

- New detections of *mec* cassettes in natural environments
- Location of *mec* cassettes in their natural hosts
  - Radically different from laboratory enrichments
  - What drives this?

Comparative Genetics of DCM-degraders Identifies a Highly-Conserved Gene Cassette

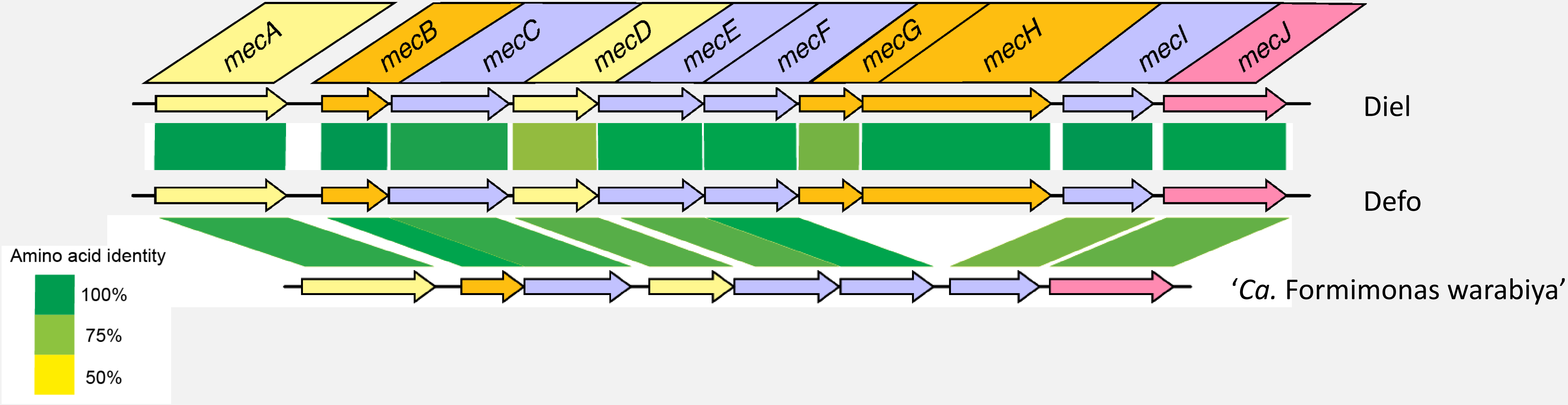
Dehalobacterium formicoaceticum
(Defo)

Vs.

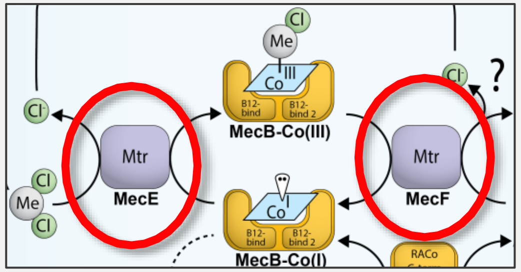
Ca. Dichloromethanomonas elyunquensis
(Diel)

The mec cassette

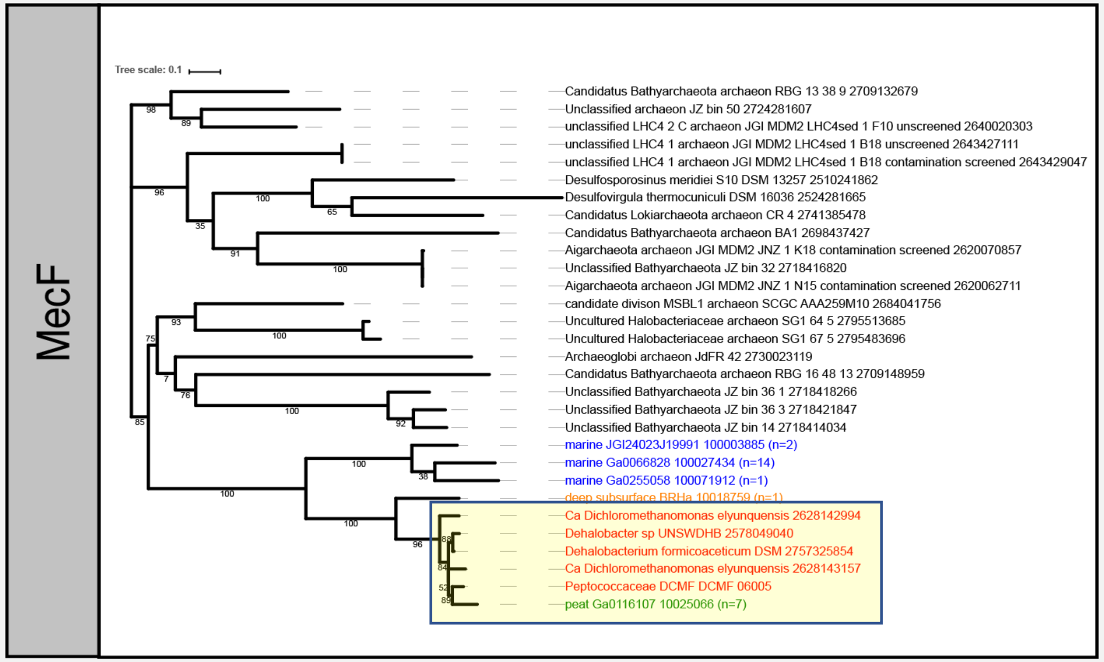
(”mec” = methylene chloride)



qPCR primer assays: The same gene cassette emerges again in new anaerobic DCM enrichments

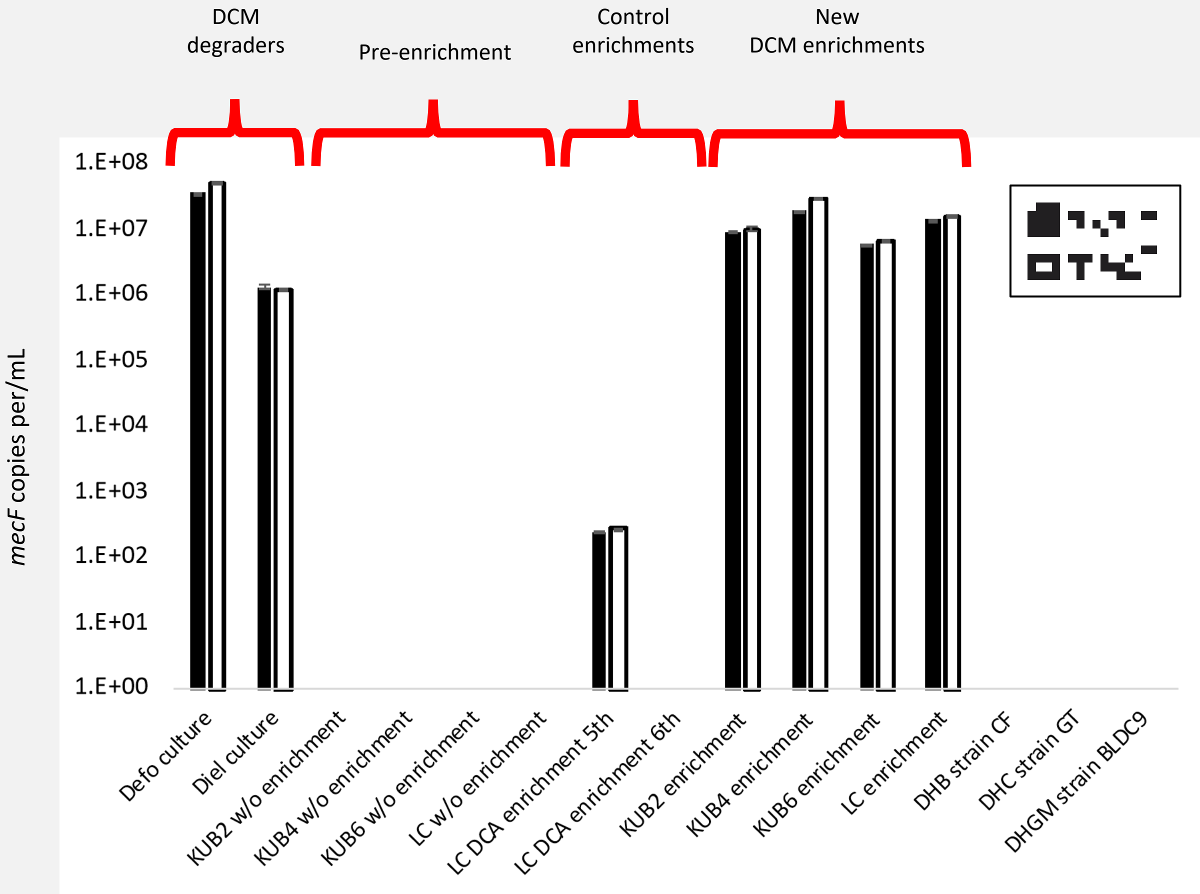


Developed qPCR assays targeting conserved core dehalogenating methyltransferases *mecE* and *mecF*



Capture diversity of genes in DCM degraders and mec genes in peatlands with single primer set for each gene

mecE and *mecF* qPCR of new DCM enrichment cultures



Mec genes emerge again in 4 new DCM enrichments

DCM contaminated site groundwater samples

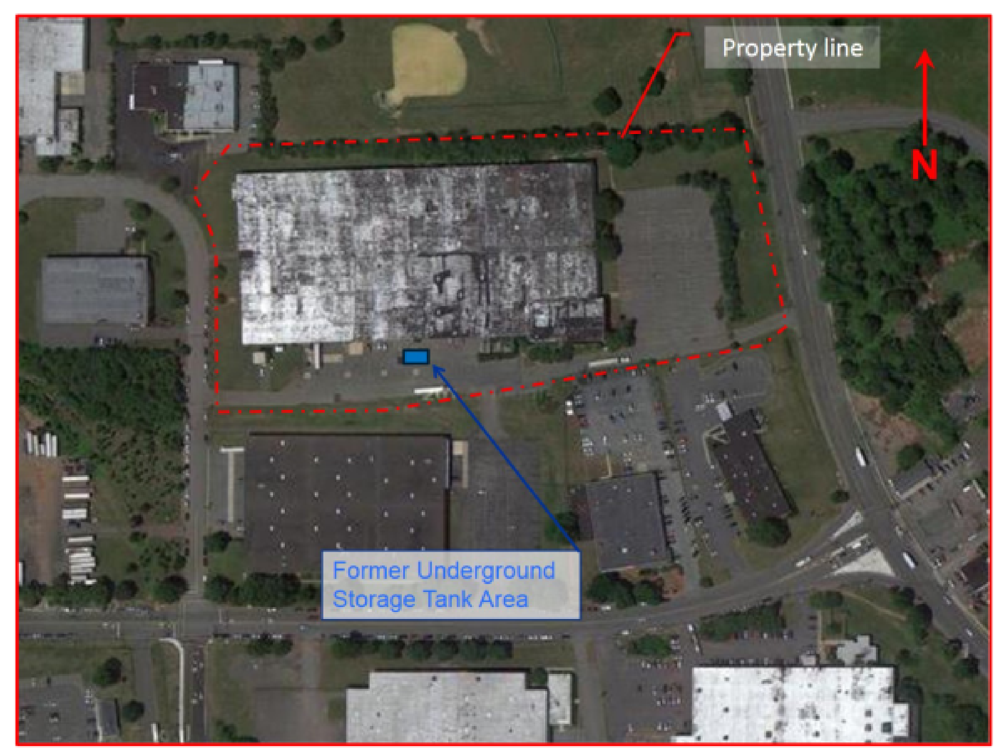
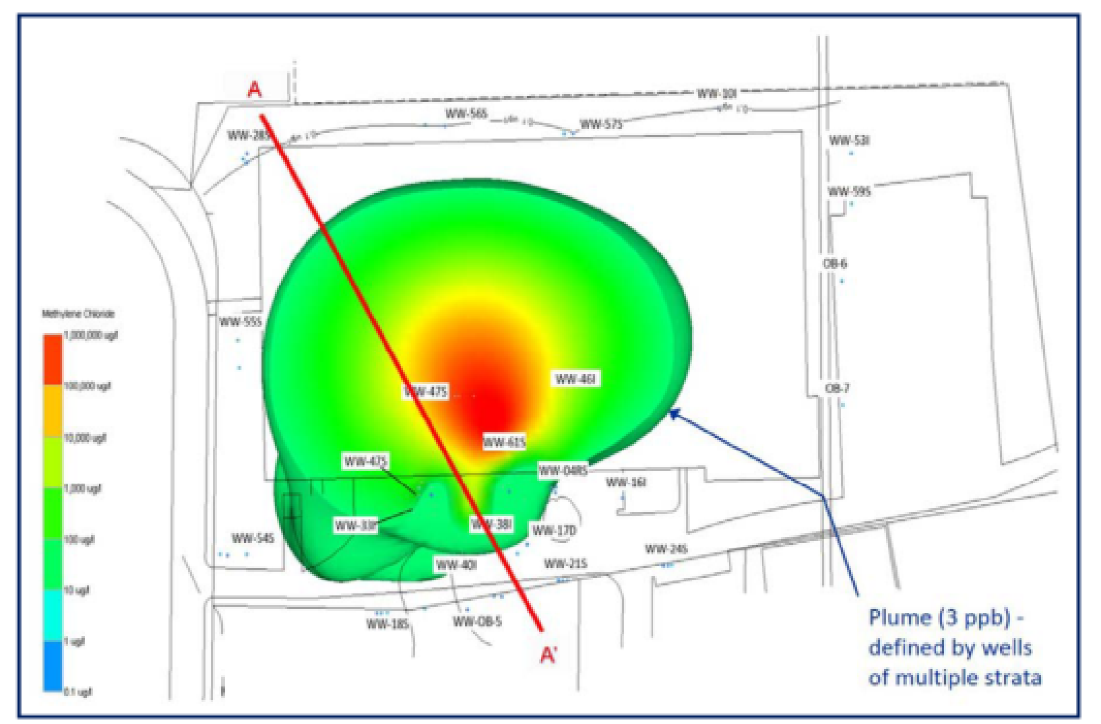


FIGURE 1. Site aerial photograph.



Plume (3 ppb) - defined by wells of multiple strata

FIGURE 2. Site contaminant plume area.

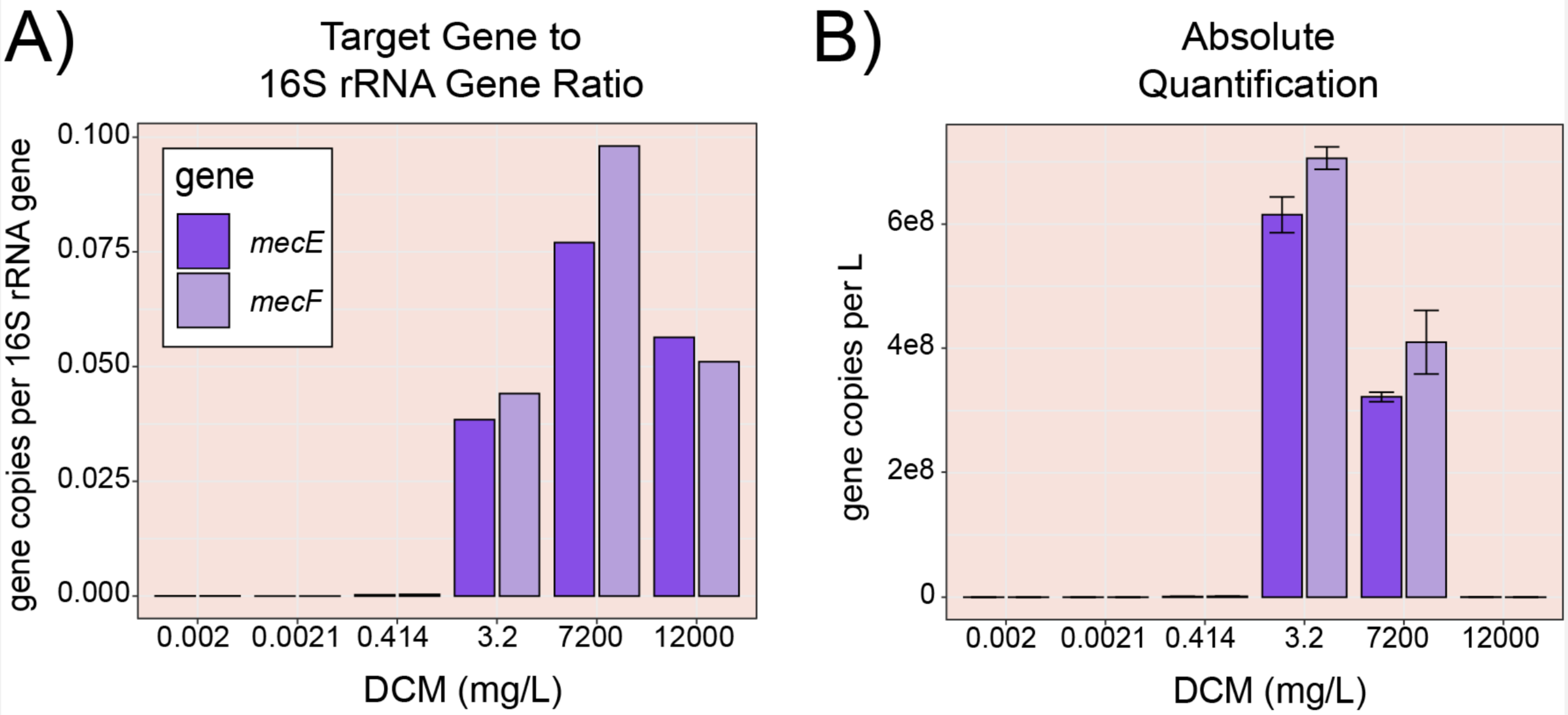
Described in:

Advanced Field Testing to Support Monitored Natural Attention of a Dichloromethane Groundwater Plume in Bedrock

Third International Symposium on Bioremediation and Sustainable Environmental Technologies (Miami, FL; May 2015). ISBN 978-0-9964071-0-6, Battelle Memorial Institute, Columbus, OH

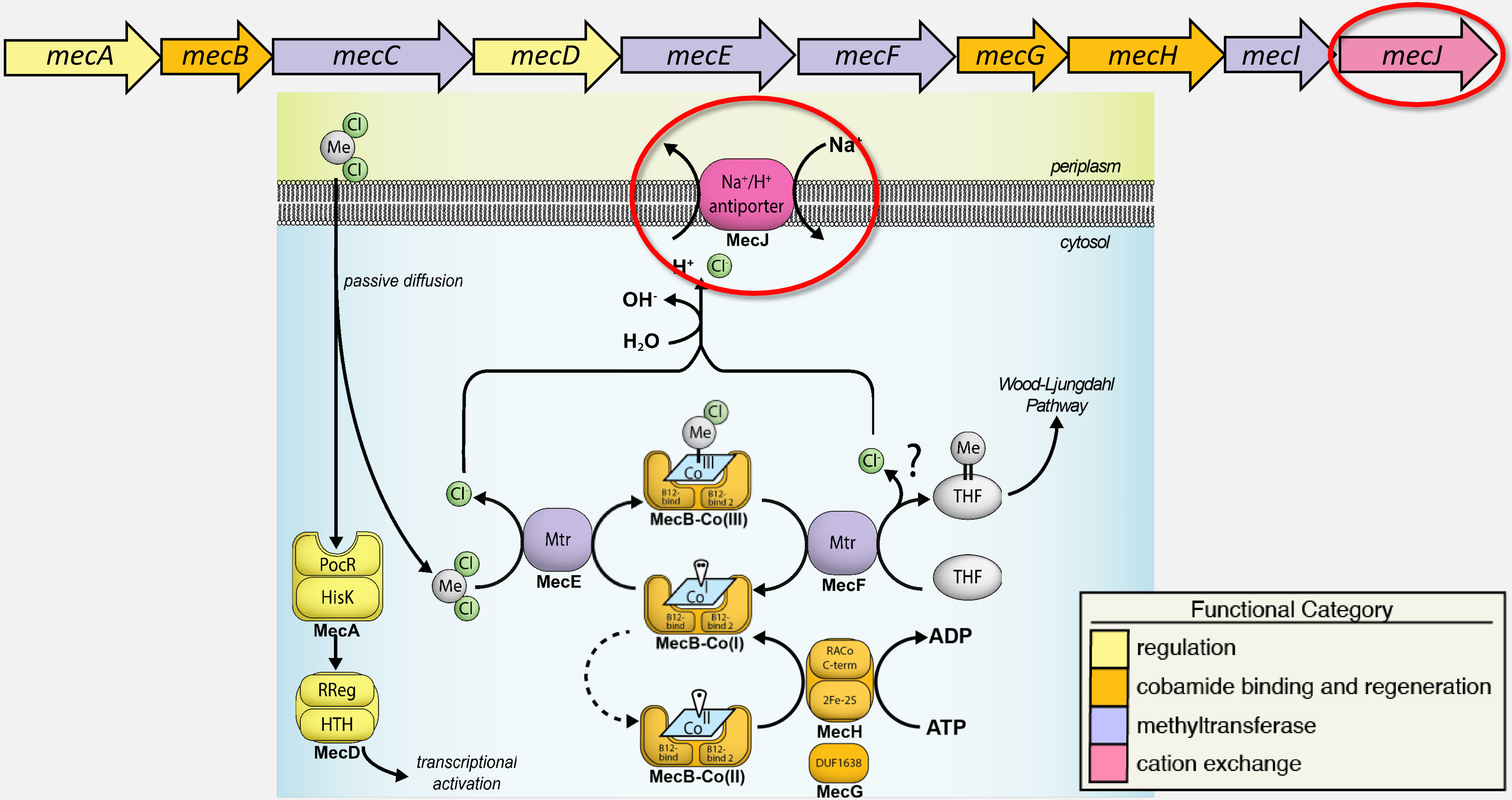
David A. Marabello (marabelloda@cdmsmith.com)
Ricky Chenenko, Stephen Mirabello (CDM Smith, Edison, NJ, USA) Tamzen Macbeth (CDM Smith, Helena, MT, USA)
Regina Lamendella and Justin Wright (Juniata College, Huntingdon, PA, USA) Tomasz Kuder (University of Oklahoma, Norman, OK, USA)

Biomarker quantification: Mec genes are frequent in a DCM contaminated groundwater plume (qPCR)



- Fresh plume samples were assayed for DCM concentration and *mecE*, *mecF*, and 16S rRNA gene abundances

Anaerobic DCM degradation: Methyltransferse-mediated dechlorination

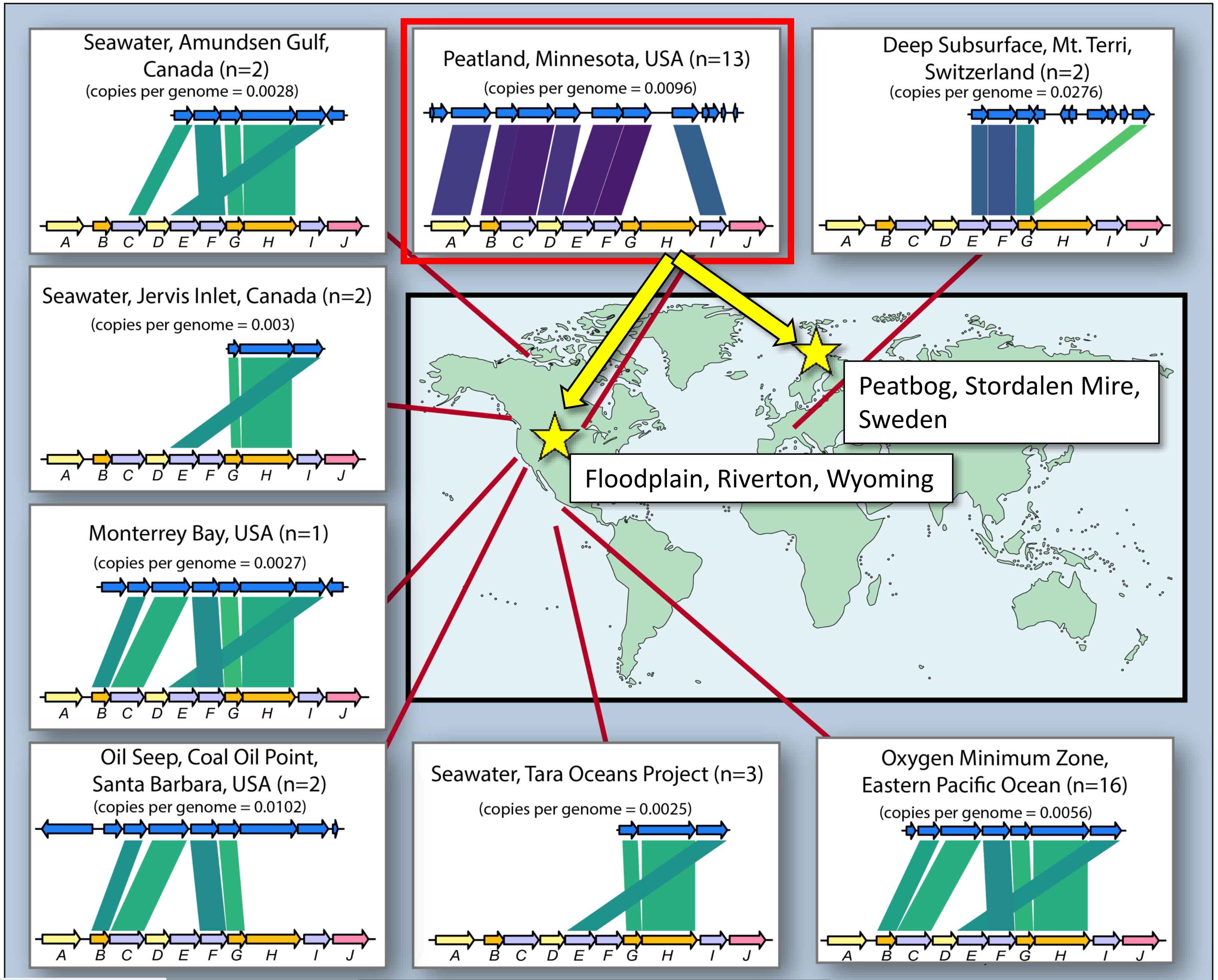


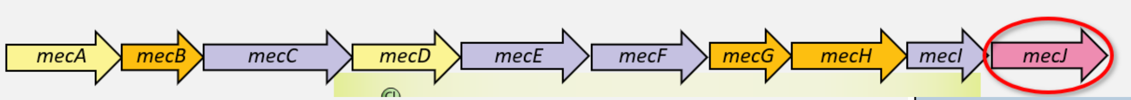
mec cassette homologs are found around the globe

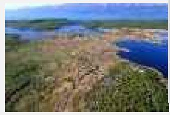
- New highly similar *mec* gene cassettes identified in Wyoming and Sweden

Key differences between DCM enriched microbes and natural systems

- No natural *mec* includes *mecJ*



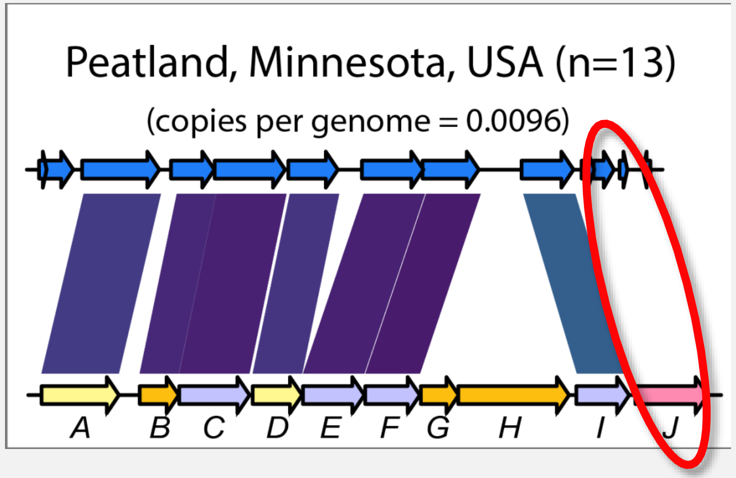




SOURCE	Phylum	Class	Order	Family	Genus	Species
DCM-enrichment	Bacillota	Clostridia	Eubacteriales	Peptococcaceae	Dehalobacterium	formicoaceticum
DCM-enrichment	Bacillota	Clostridia	Eubacteriales	Peptococcaceae	*Ca.* Dichloromethanomonas	elyunquensis
DCM-enrichment	Bacillota	Clostridia	Eubacteriales	Peptococcaceae	*Ca.* Formiminas	warabiya
Peatland	Thermodesulfobacterota	Syntrophia	Syntrophales	Smithellaceae	Smithella	None
Peatland	Bacillota	Clostridia	Thermoanaerobacterales	Thermoanaerobacteraceae	Thermacetogenium	None

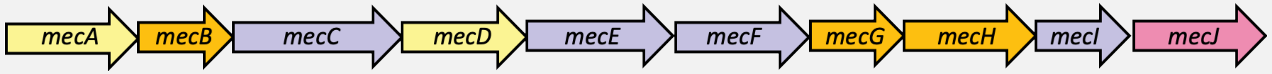
- **In nature, the mec cassette is found in entirely different bacterial taxa**
    - *Smithella* and *Thermacetogenium* are known as anaerobes living in syntrophy with methanogens

- Naturally-occurring *mec* cassettes also lack mecJ, which encodes a pH-acclimating ion channel

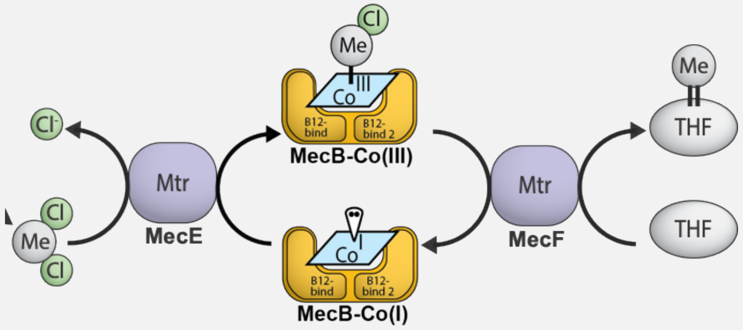


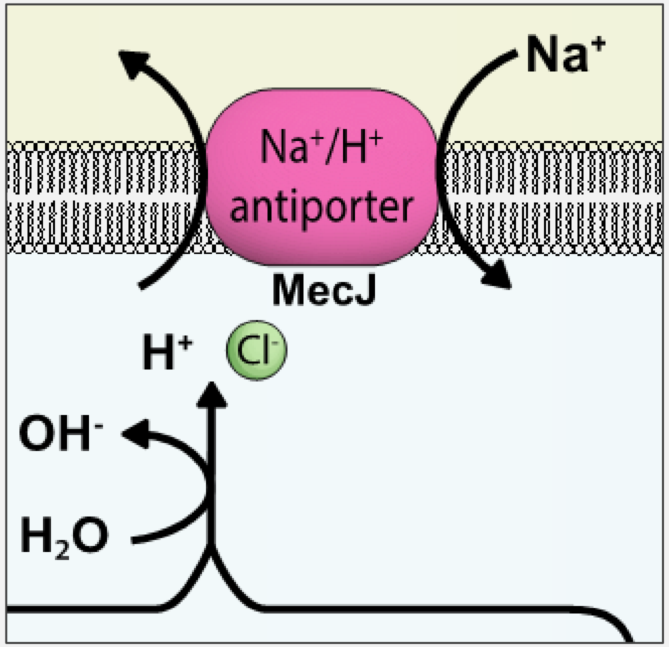
Does enrichment drive the functional cassette into a host that has better background for very high rates of dehalogenation?

Summary



- Mec cassette encodes for anaerobic DCM degradation
- Functional gene qPCR primers give reliable signals
- Present in diverse natural environments BUT:
  - Is harbored in entirely different bacterial taxa
  - Lacks *mecJ*





Implications and Future Work

- *mecE* and *mecF* qPCR assays are ready for refinement and field testing
- General lesson: Taxonomy in enrichments may not reflect natural reservoirs of functional capacity

Acknowledgements

Tamzen MacBeth, CDM Smith

Thank you!

Questions?

Appendix

Co-localization of *mecJ* homologs with dehalogenase-encoding genes

Gene ID	AA % ID	Length	E-value	Bit Score	Genome Name	adjacent dehalogenase
2628143161	96.5	397	9.90E-266	827	*Ca.* Dichloromethanomonas elyunquensis	*MEC cassette*
2628141546	77.6	398	2.50E-214	680	*Ca.* Dichloromethanomonas elyunquensis	HADase X 2
2753034167	63.5	400	2.80E-169	551	Dehalobacter sp. FTH1	RdhA
2775644313	60.5	394	1.70E-158	521	Dehalobacter sp. KB-1_124TCB1	RdhA
2578048931	57.5	395	2.80E-155	511	Dehalobacter sp. UNSWDHB	RdhA X 2
2520955686	57.5	395	2.80E-155	511	Dehalobacter sp. 11DCA	RdhA X 2
2520070062	57.5	395	2.80E-155	511	Dehalobacter sp. CF	RdhA X 2
2753034216	57.4	395	3.90E-153	505	Dehalobacter sp. FTH1	RdhA
2578049036	53.1	399	1.10E-146	487	Dehalobacter sp. UNSWDHB	*MEC cassette*
2776022298	52.9	620	7.60E-145	481	Clostridiaceae bacterium mt12	none
2775643454	52.2	397	3.70E-143	476	Dehalobacter sp. KB-1_124TCB1	RdhA

Top 10 *mecJ* homologs

The Team

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

University of Tennessee, Knoxville, USA

[1]Center for Environmental Biotechnology
[2]Department of Civil and Environmental Engineering
[3]Department of Microbiology
[4]Department of Biosystems Engineering and Soil Science

Corteva Agriscience
[5]Corteva Environmental Remediation, Corteva Agriscience

Oak Ridge National Laboratory, USA
[6]Biosciences Division
[7]Joint Institute for Biological Sciences,

[#]Current affiliation: **Battelle Memorial Institute**, Columbus, Ohio 43201

Robert W. Murdoch, PhD
Lead Bioinformaticist
Battelle CBRNE Biosciences
murdoch@battelle.org

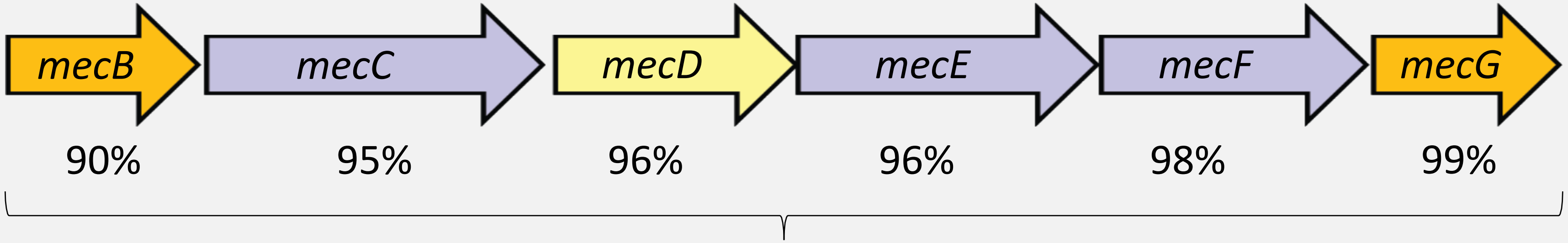
Global Change Biology

RESEARCH ARTICLE

Identification and widespread environmental distribution of a gene cassette implicated in anaerobic dichloromethane degradation

Robert W. Murdoch, Gao Chen, Fadime Kara Murdoch, E. Erin Mack, Manuel I. Villalobos Solis, Robert L. Hettich, Frank E. Löffler

Mec genes are expressed the DCM contaminated groundwater plume (metatranscriptomics)



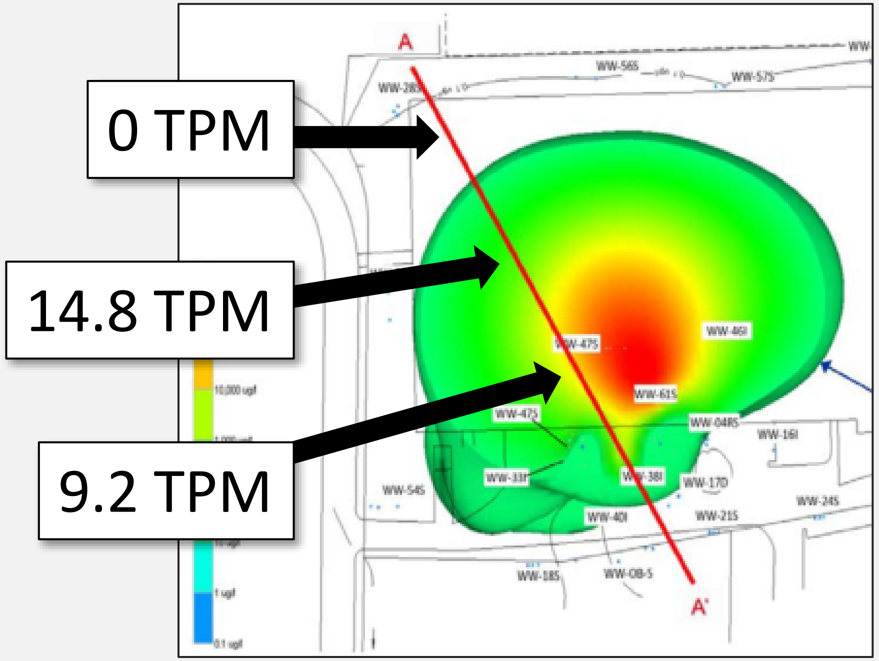
Amino acid identity to *Dehalobacterium formicoaceticum*

mecB	*mecC*	*mecD*	*mecE*	*mecF*	*mecG*
90%	95%	96%	96%	98%	99%

Single mRNA molecule

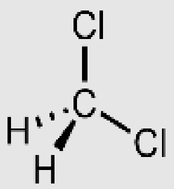
mec gene transcripts were detected in core and edge samples from the DCM contamination plume, but not outside

TPM = 'transcripts per million transcripts'



Dichloromethane as a Contaminant

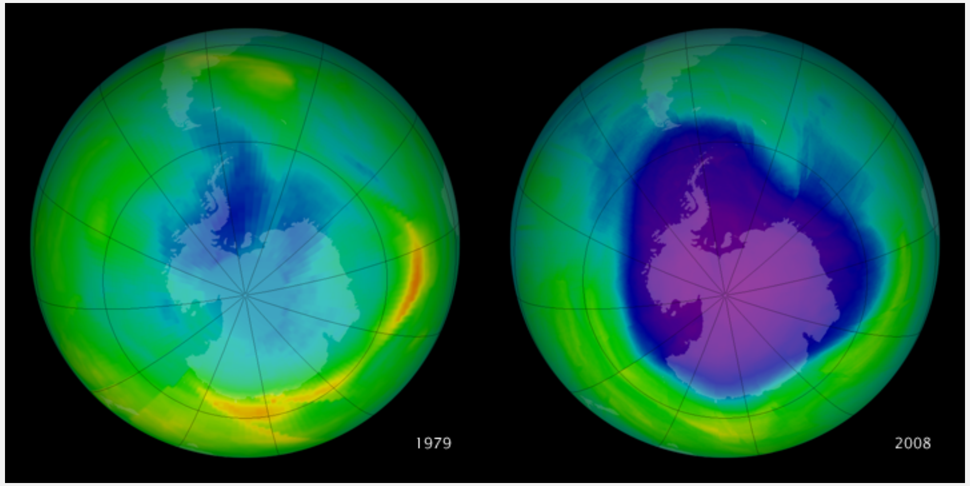
- Probable human carcinogen
- Neurotoxin
- Ozone-depleting agent



Dichloromethane,
a.k.a. methylene chloride

Dichloromethane as a Natural Chemical

- Halomethanes form naturally
- Widespread, ancient energy source?



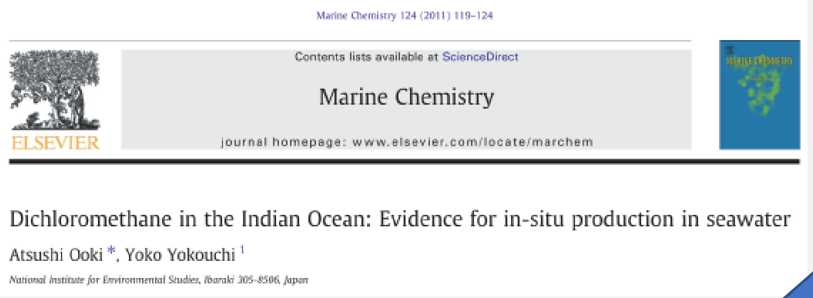
Dehalobacterium formicoaceticum
Contaminated groundwater

Ca. Dichloromethanomonas elyunquensis
Pristine stream sediment

$$CH_2Cl_2 \xrightarrow{?} \text{Me-THF} \longrightarrow \begin{matrix}\text{WL-pathway} \\ \text{(1-carbon metabolism)}\end{matrix}$$

How dehalogenation and methyl-capture is catalyzed is unclear

Evidence of marine production and/or persistence well below the photic zone



Marine Chemistry 124 (2011) 119–124

Contents lists available at ScienceDirect

Marine Chemistry

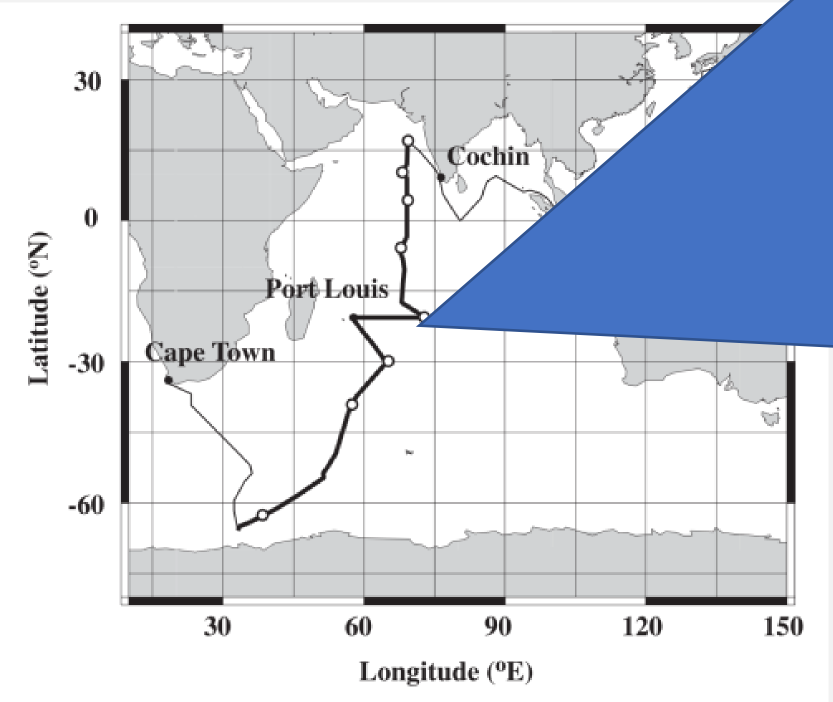
journal homepage: www.elsevier.com/locate/marchem

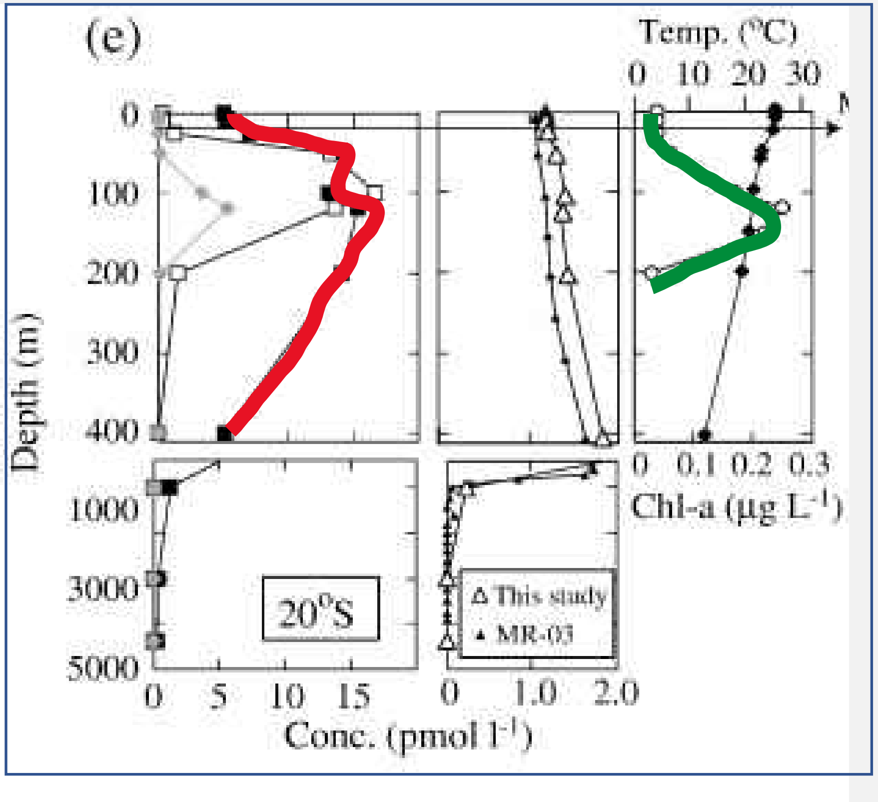
ELSEVIER

Dichloromethane in the Indian Ocean: Evidence for in-situ production in seawater

Atsushi Ooki *, Yoko Yokouchi [1]

National Institute for Environmental Studies, Ibaraki 305-8506, Japan



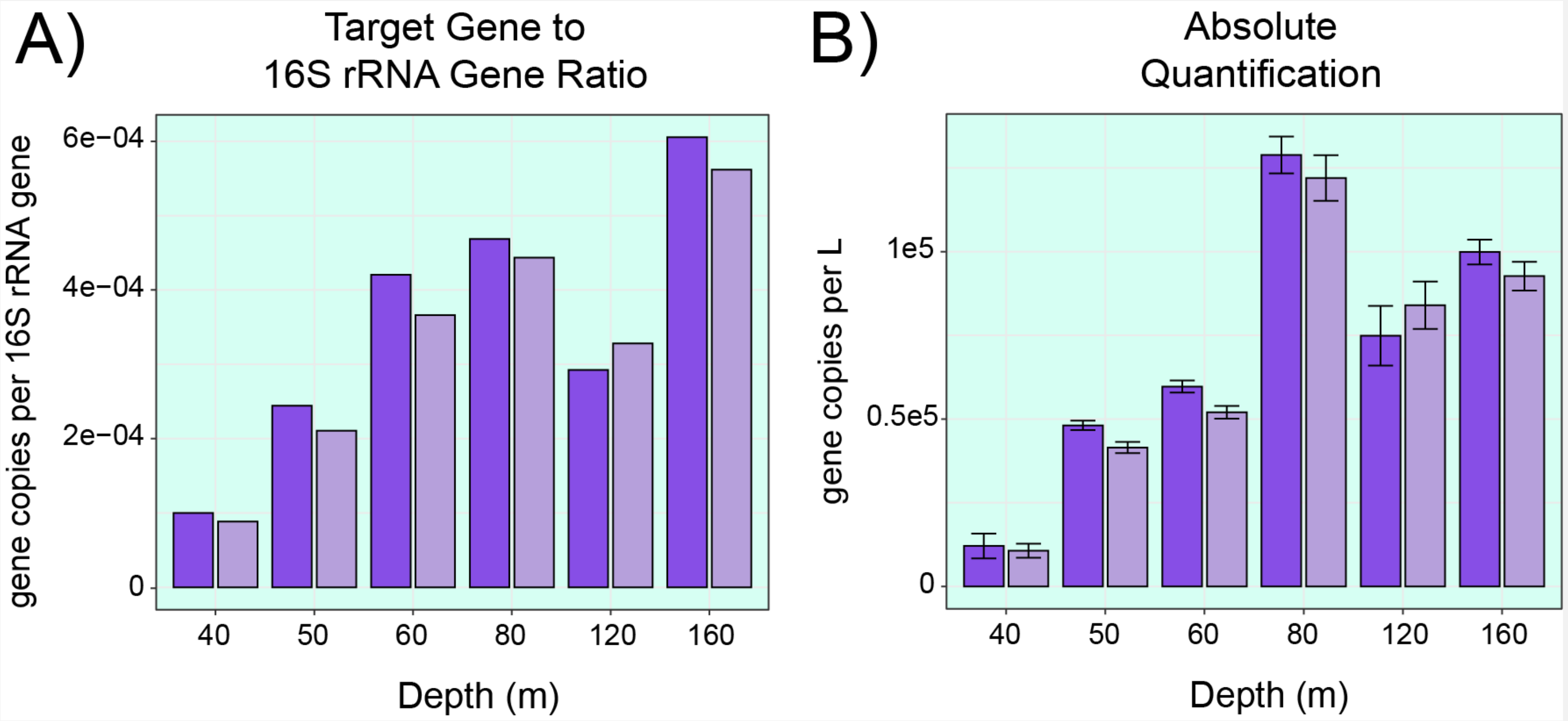


DCM

✓ Chlorophyll-a

Mec genes may be more frequent in the Eastern Pacific Oxygen Minimum Zone than indicated by metagenomics

- New PCR primers for Eastern Tropical North Pacific OMZ
- DNA samples from OMZ vertical transects
- *mecE and mecF* detected at all depths throughout the OMZ



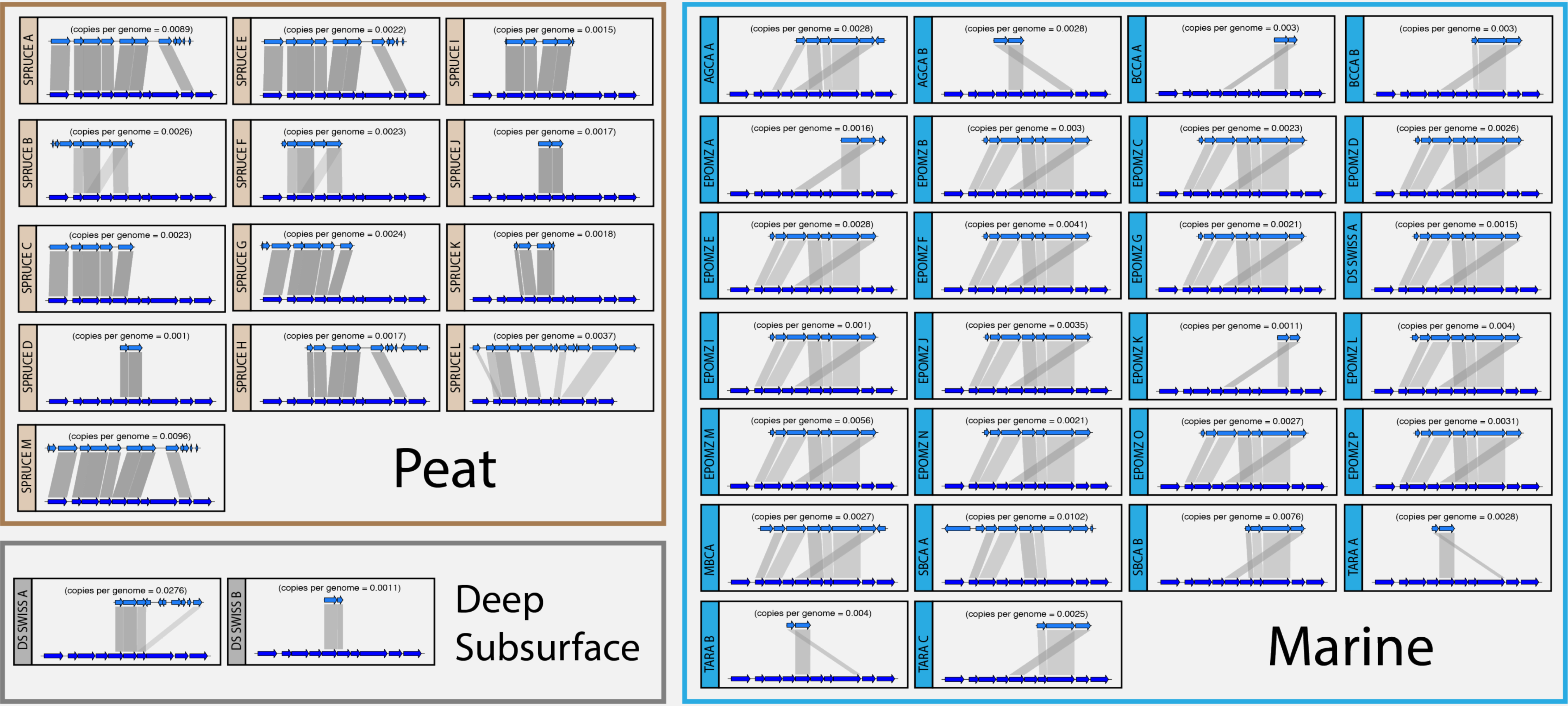
Comparative Genetics of DCM-degraders:
Many of the most similar genes are co-localized

Dehalobacterium formicoaceticum vs. *Ca*. Dichloromethanomonas elyunquensis 25 most similar genes

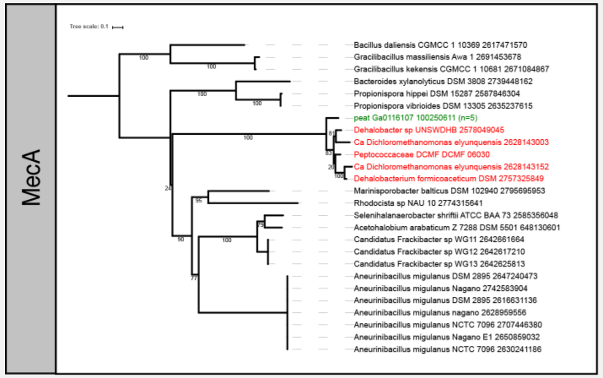
D. formicoaceticum Gene	D. elyunquensis Gene	Percent Ident	Bit Score	E-value
2812836294 uroporphyrinogen decarboxylase	2628143160 Uroporphyrinogen decarboxylase (URO-D)	99		

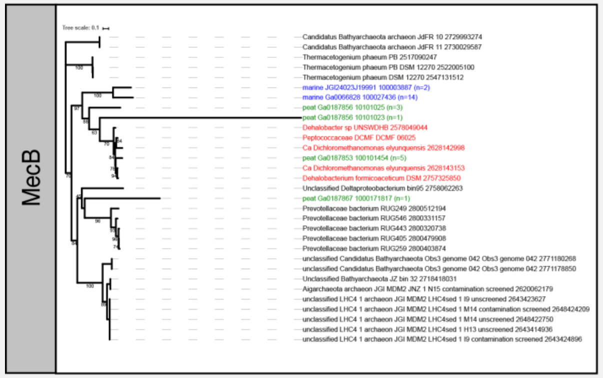
[illegible]	[illegible]	[illegible]	[illegible]	.7	582.4	9.60E-168
[illegible]	[illegible]	[illegible]	methyltransfe[illegible] rinoid proteins	98.1	407.9	2.40E-115
[illegible]	[illegible]	[illegible]	Kef-t[illegible] membrane component KefB	96.5	723.4	4.80E-210
2812836293	uncharacterized 2Fe-2S/4Fe-4S cluster protein (DUF4445 family)	2628143159	Uncharacterized 2Fe-2 and 4Fe-4S clusters-containing protein, contains DUF4445 domain	96.2	1147.5	0
2812836290	MtaA/CmuA family methyltransferase	2628142995	Uroporphyrinogen decarboxylase (URO-D)	94	654.1	3.00E-189
2812836291	tetrahydromethanopterin S-methyltransferase subunit H	2628143157	tetrahydromethanopterin S-methyltransferase, subunit H	93	556.6	5.80E-160
2812836286	histidine kinase	[illegible]	Histidine kinase	[illegible]	[illegible]	[illegible]
2812836288	uroporphyrinogen decarboxylase	[illegible]	uroporphyrinogen decarboxylase	[illegible]	[illegible]	[illegible]
2812836311	type IV secretory pathway VirB4 component	2628143553	Type IV secretory pathway, VirB4 component	92	1426.4	0
2812837477	type IV secretory pathway VirB4 component	2628143553	Type IV secretory pathway, VirB4 component	90.9	1406.7	0
2812836847	O-acetyl-ADP-ribose deacetylase (regulator of RNase III)	2628143219	O-acetyl-ADP-ribose deacetylase (regulator of RNase III), contains Macro domain	90.5	414.8	3.10E-117
2812837991	stage V sporulation protein S	2628142327	stage V sporulation protein S	89.5	151.4	1.60E-38
2812836317	hypothetical protein	2628143548	hypothetical protein	89.3	149.1	7.00E-38
2812836316	hypothetical protein	2628143549	hypothetical protein	89.2	497.3	3.90E-142
2812839492	SSU ribosomal protein S10P	2628143433	SSU ribosomal protein S10P	88	179.9	5.00E-47
2812836891	SSU ribosomal protein S21P	2628141537	SSU ribosomal protein S21P	87.7	100.1	2.90E-23
2812836319	type IV secretion system protein VirD4	2628143546	type IV secretion system protein VirD4	87.2	1031.6	0
2812838819	prevent-host-death family protein	2628142161	prevent-host-death family protein	86.7	157.9	1.80E-40
2812839512	RNA polymerase sigma-30 (SigH) subunit	2628143719	RNA polymerase, sigma 30 subunit, SigH	86	357.8	2.80E-100
2812839500	LSU ribosomal protein L12P	2628143441	LSU ribosomal protein L12P	85.8	199.1	9.90E-53
2812836289	AraC family two component transcriptional regulator	2628143155	two component transcriptional regulator, AraC family	84.9	458	2.50E-130
2812837467	type IV secretion system protein VirD4	2628143546	type IV secretion system protein VirD4	84.8	1010	3.80E-296
2812837438	chromosome partitioning protein	2628143527	chromosome partitioning protein	83.9	448.7	1.60E-127
2812836351	chromosome partitioning protein	2628143527	chromosome partitioning protein	83.6	446.8	5.90E-127
2812836292	uncharacterized protein DUF1638	2628143158	Protein of unknown function (DUF1638)	79.6	316.6	7.00E-88

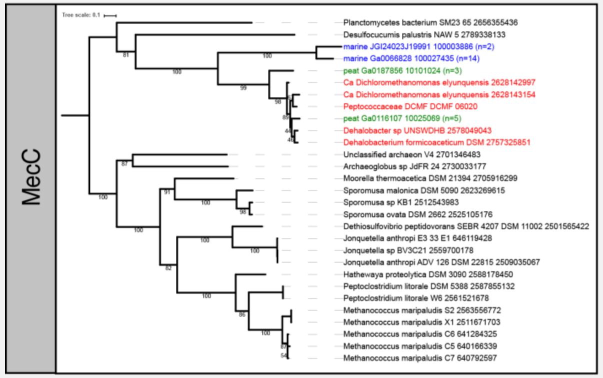
Detection of mec cassette homologs in public metagenomes

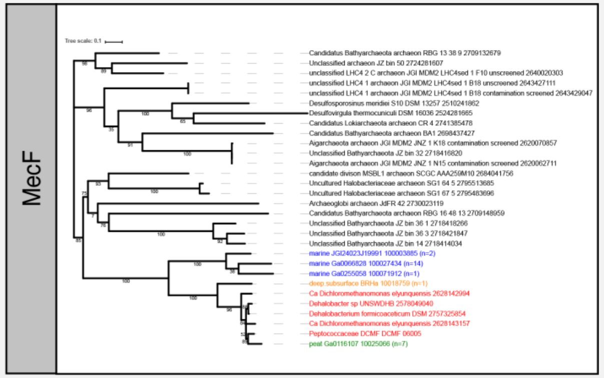


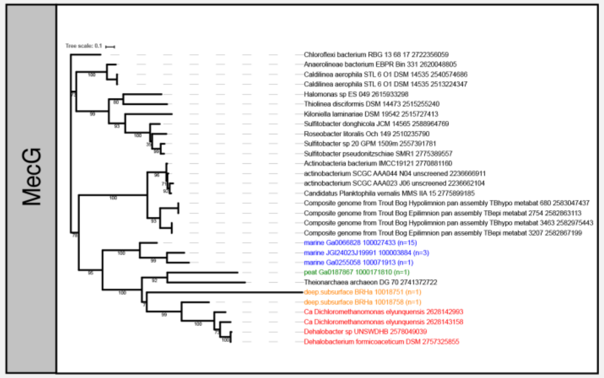
Phylogeny of mec cassette genes

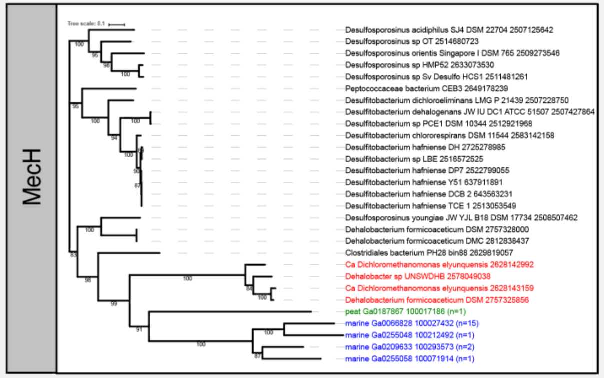


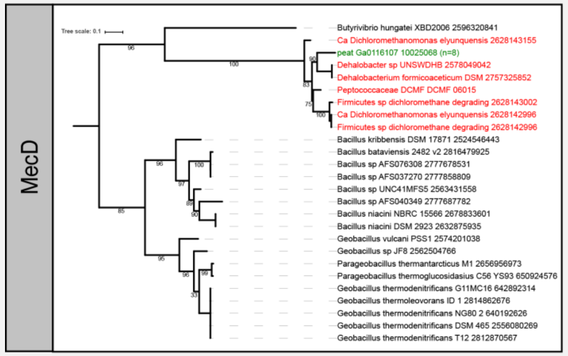


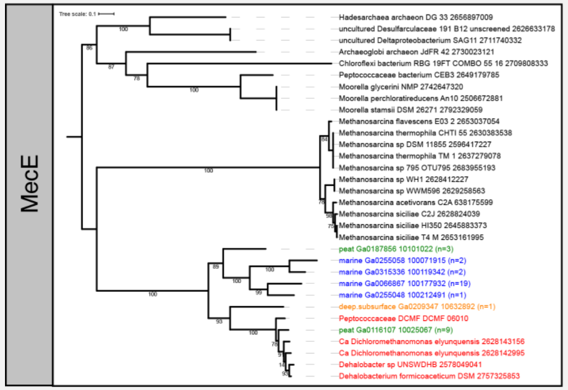


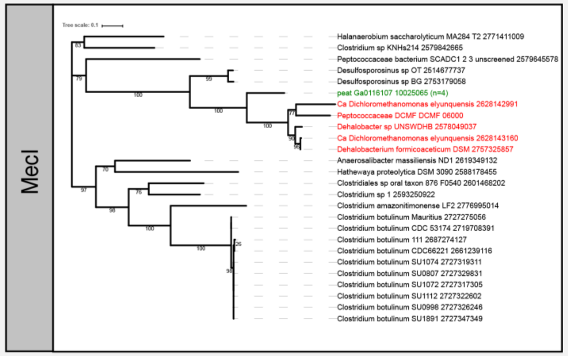


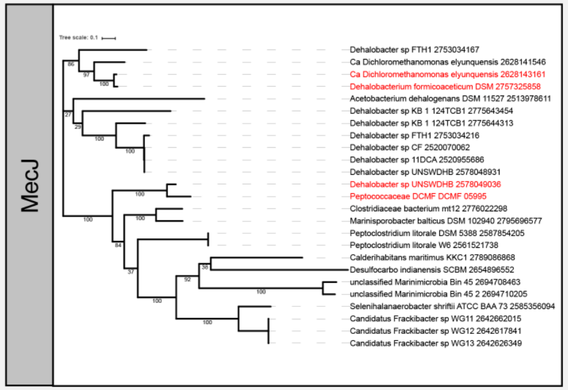






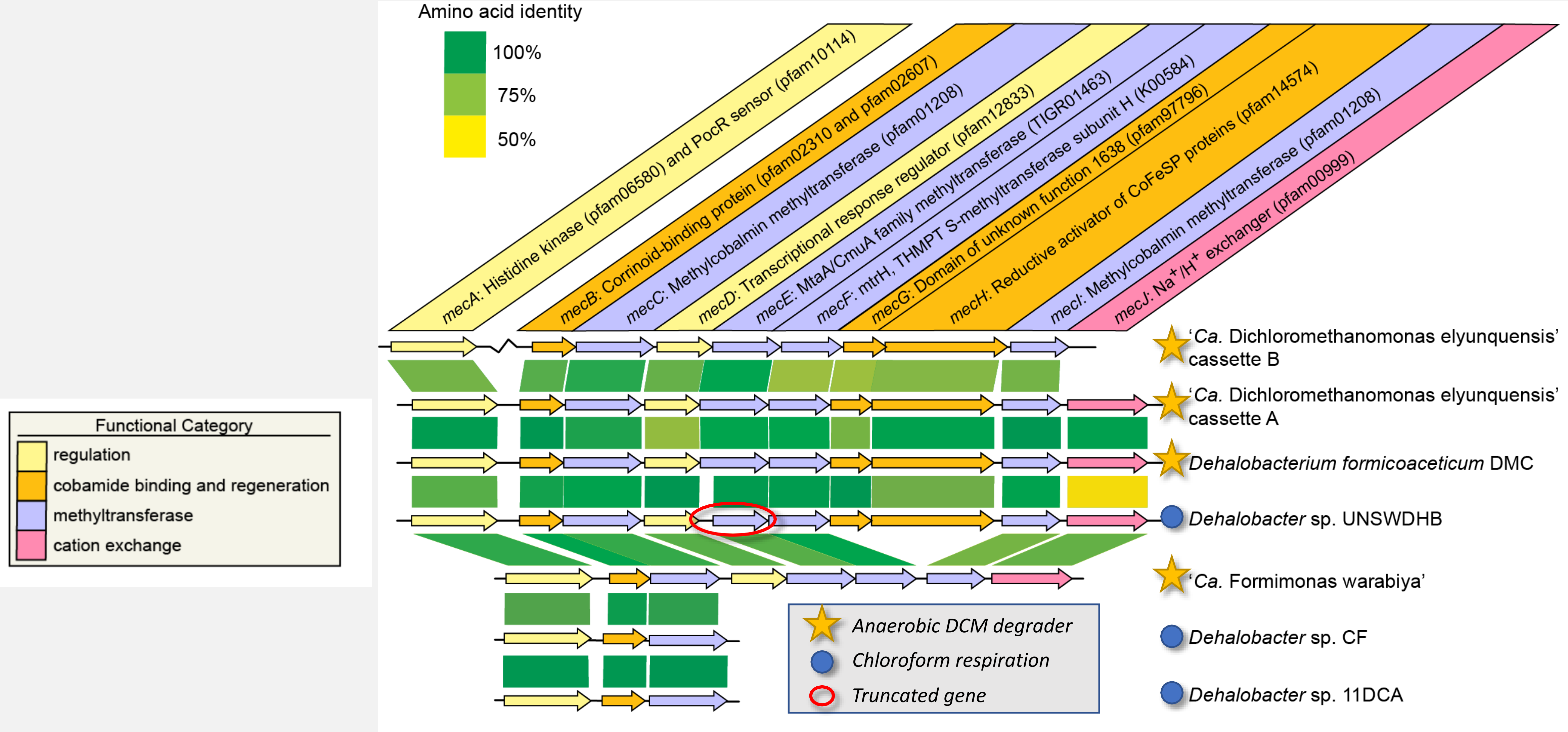






- All are “evolutionarily isolated’, except for *mecJ*
- DCM degraders and peat cassette genes form tight families

Comparative Genetics of DCM-degraders



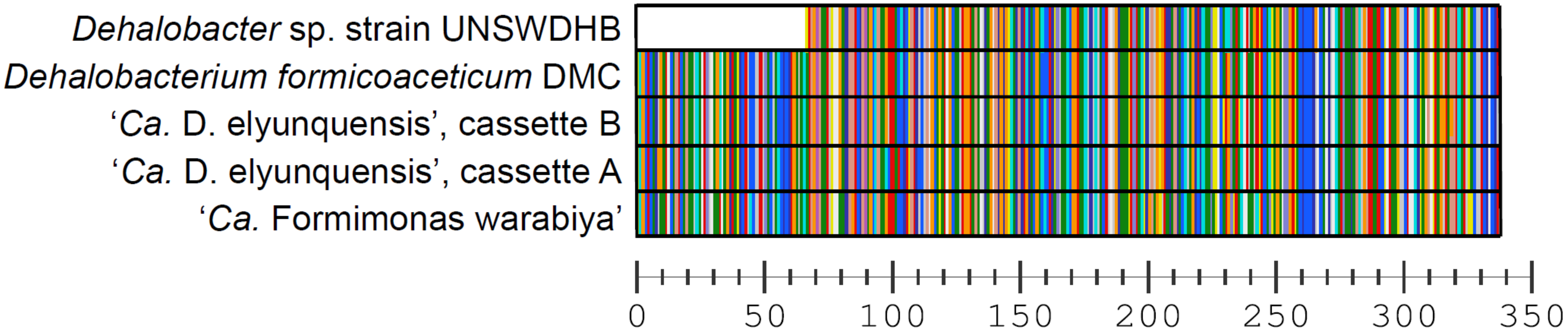
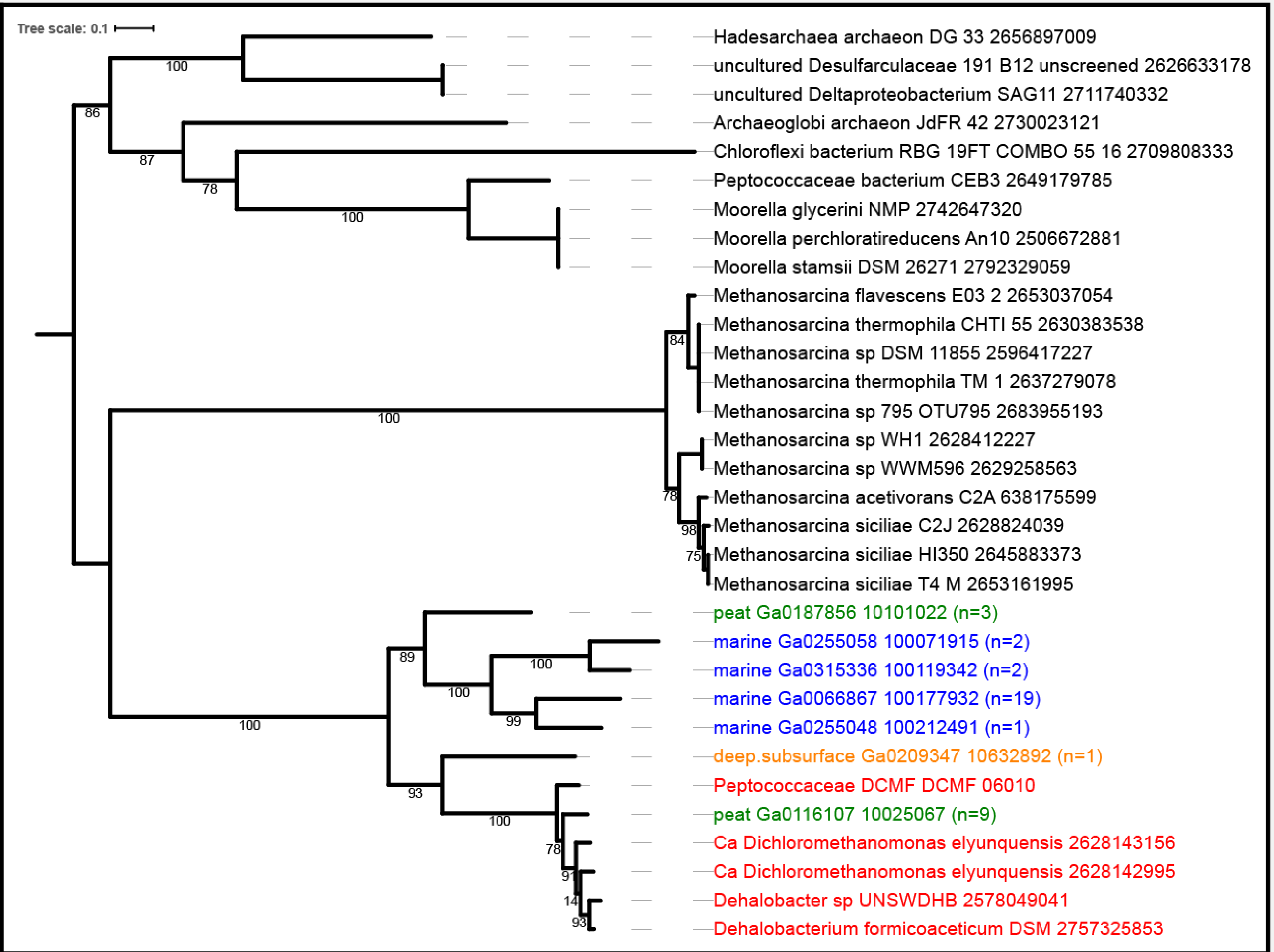
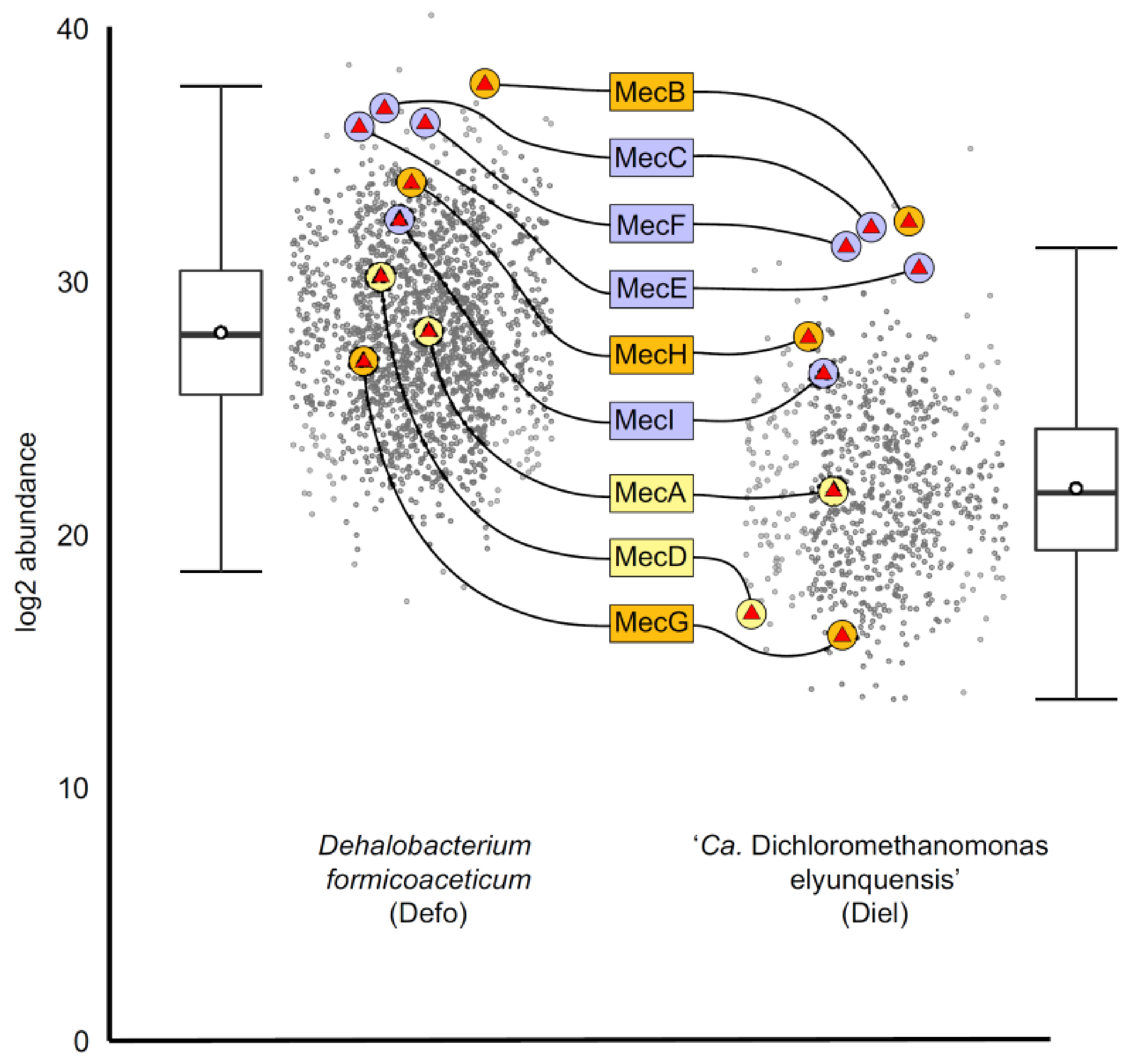


Figure S7. Simplified representation of the alignment of all genome-derived MecE homologs demonstrating the predicted N-terminal truncation in the genome of *Dehalobacter* sp. strain UNSWDHB. Sequences were aligned using MAFFT-G-ING-I with 1,000 maximum iterations. The simplified visualization was produced using Pixel (93). Each amino acid is colored according to the Jmol "shapely" color scheme (94).

MecE

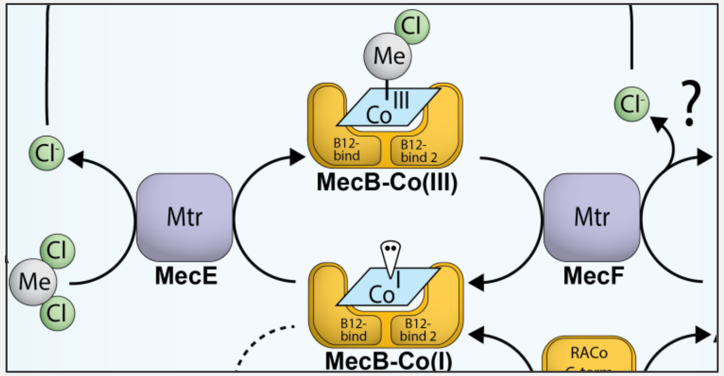

Tree scale: 0.1
Hadesarchaea archaeon DG 33 2656897009
uncultured Desulfarculaceae 191 B12 unscreened 2626633178
uncultured Deltaproteobacterium SAG11 2711740332
Archaeoglobi archaeon JdFR 42 2730023121
Chloroflexi bacterium RBG 19FT COMBO 55 16 2709808333
Peptococcaceae bacterium CEB3 2649179785
Moorella glycerini NMP 2742647320
Moorella perchloratireducens An10 2506672881
Moorella stamsii DSM 26271 2792329059
Methanosarcina flavescens E03 2 2653037054
Methanosarcina thermophila CHTI 55 2630383538
Methanosarcina sp DSM 11855 2596417227
Methanosarcina thermophila TM 1 2637279078
Methanosarcina sp 795 OTU795 2683955193
Methanosarcina sp WH1 2628412227
Methanosarcina sp WWM596 2629258563
Methanosarcina acetivorans C2A 638175599
Methanosarcina siciliae C2J 2628824039
Methanosarcina siciliae HI350 2645883373
Methanosarcina siciliae T4 M 2653161995
peat Ga0187856 10101022 (n=3)
marine Ga0255058 100071915 (n=2)
marine Ga0315336 100119342 (n=2)
marine Ga0066867 100177932 (n=19)
marine Ga0255048 100212491 (n=1)
deep.subsurface Ga0209347 10632892 (n=1)
Peptococcaceae DCMF DCMF 06010
peat Ga0116107 10025067 (n=9)
Ca Dichloromethanomonas elyunquensis 2628143156
Ca Dichloromethanomonas elyunquensis 2628142995
Dehalobacter sp UNSWDHB 2578049041
Dehalobacterium formicoaceticum DSM 2757325853
100
86
87
78
100
84
100
78
98
75
89
100
100
99
100
93
100
78
91
14
93

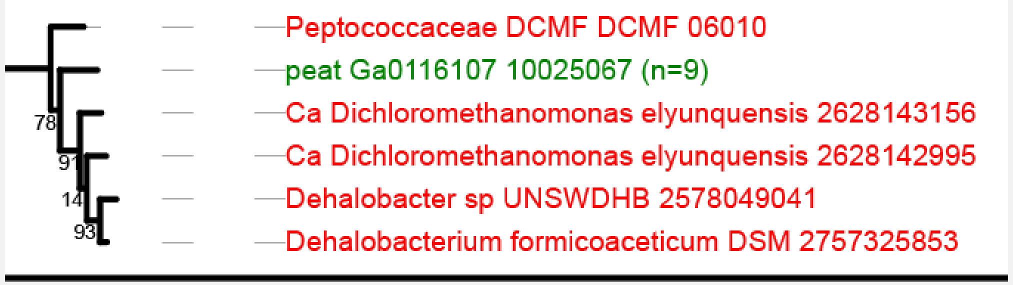


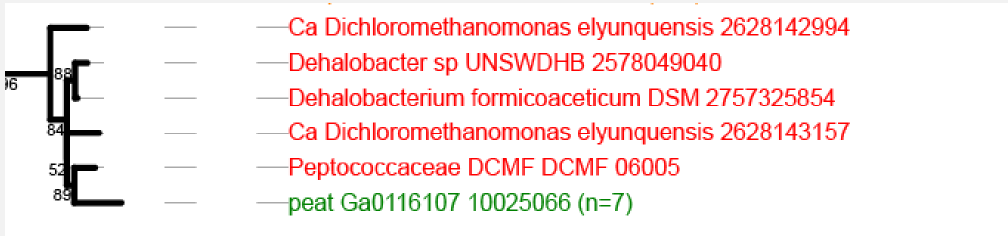
Mec proteins are common in proteomes during growth on DCM

Most proteins encoded by the mec gene cassettes were detected in both proteomes.

- The three methyltransferases MecC, MecE, and MecF were in the top 1% most abundant proteins in both proteomes.
- The corrinoid-binding protein MecB was the 4th and 3rd most abundant protein in Defo and Diel proteomes.




Peptococcaceae DCMF DCMF 06010
peat Ga0116107 10025067 (n=9)
Ca Dichloromethanomonas elyunquensis 2628143156
Ca Dichloromethanomonas elyunquensis 2628142995
Dehalobacter sp UNSWDHB 2578049041
Dehalobacterium formicoaceticum DSM 2757325853
78
91
14
93


Ca Dichloromethanomonas elyunquensis 2628142994
Dehalobacter sp UNSWDHB 2578049040
Dehalobacterium formicoaceticum DSM 2757325854
Ca Dichloromethanomonas elyunquensis 2628143157
Peptococcaceae DCMF DCMF 06005
peat Ga0116107 10025066 (n=7)
6
88
84
52
89