UNSW SYDNEY

Exploring the one-carbon metabolism of a novel, dichloromethane-fermenting bacterium, 'Candidatus Formamonas warabiya'

Author:
Holland, Sophie

Publication Date:
2020

DOI:
https://doi.org/10.26190/unsworks/3948

License:
https://creativecommons.org/licenses/by-nc-nd/3.0/au/
Link to license to see what you are allowed to do with this resource.

Downloaded from http://hdl.handle.net/1959.4/67429 in https://unsworks.unsw.edu.au on 2024-04-28

Exploring the one-carbon metabolism of a novel, dichloromethane-fermenting bacterium, '*Candidatus* Formamonas warabiya'

Sophie Holland

A thesis in fulfilment of the requirement for the degree of
Doctor of Philosophy

School of Civil and Environmental Engineering
Faculty of Engineering
University of New South Wales

March 2020

Thesis Dissertation sheet

Surname/Family Name	:	**Holland**
Given Name/s	:	**Sophie Isobel**
Abbreviation for degree as give in the University calendar	:	**PhD**
Faculty	:	**Engineering**
School	:	**Civil and Environmental Engineering**
Thesis Title	:	**Exploring the one-carbon metabolism of a novel, dichloromethane-fermenting bacterium, '*Candidatus* Formamonas warabiya'**

Abstract 350 words maximum: (PLEASE TYPE)

Anaerobic microbial metabolism of dichloromethane (DCM; CH_2Cl_2), quaternary amines, and methanol has important implications for carbon and nitrogen cycling in oligotrophic environments and the atmospheric flux of climate-active trace gasses. A novel, strictly anaerobic member of the *Peptococcaceae* family, strain DCMF, is the dominant organism in a DCM-fermenting enrichment culture (DFE) and one of very few known bacteria capable of fermenting DCM to the innocuous end product acetate. Long read, whole genome sequencing provided a single, circularised 6.44 Mb chromosome for strain DCMF, which contains 5,772 predicted protein-coding genes including an abundance of MttB superfamily methyltransferases. Genomic comparison of anaerobic, DCM-degrading bacteria provided a relatively small core genome, including the Wood-Ljungdahl pathway.

Strain DCMF is the first non-obligate anaerobic DCM-degrading bacterium. Genomic, physiological and proteomic experiments confirmed that it is an anaerobic methylotroph, able to metabolise DCM, methanol, and methyl groups from quaternary amines via the Wood-Ljungdahl pathway. The quaternary amine choline was converted to glycine betaine, which was demethylated to sarcosine with a glycine betaine methyltransferase, then reductively cleaved to methylamine and acetate. Methanol (via a methanol methyltransferase) and DCM were fermented to acetate. Comparative proteomics revealed a methyltransferase system that was significantly more abundant in cells grown with DCM than glycine betaine. The novel, putative DCM methyltransferase genes are highly conserved between anaerobic DCM-degrading bacteria. Genomic and physiological evidence support placement of strain DCMF in a novel genus, for which we propose the name '*Candidatus* Formamonas warabiya'.

Cohabiting bacteria in the DFE community have persisted despite repeated attempts to isolate strain DCMF, yet strain DCMF-free enrichments demonstrated that most are unable to utilise DCM, quaternary amines, or methanol. Five MAGs were generated from the long-read sequencing data and a metaproteogenomic approach suggested that the cohabiting organisms persist in the culture via necromass fermentation, i.e. oxidation of carbohydrates, proteins, and sugars released from expired strain DCMF cells. The DFE culture is a long-term stable-state community that highlights interactions between foundation species and supporting bacteria, as well as important pathways of carbon and nitrogen cycling.

Declaration relating to disposition of project thesis/dissertation

I hereby grant to the University of New South Wales or its agents a non-exclusive licence to archive and to make available (including to members of the public) my thesis or dissertation in whole or in part in the University libraries in all forms of media, now or here after known. I acknowledge that I retain all intellectual property rights which subsist in my thesis or dissertation, such as copyright and patent rights, subject to applicable law. I also retain the right to use all or part of my thesis or dissertation in future works (such as articles or books).

…… ……..……..……

Signature Date

The University recognises that there may be exceptional circumstances requiring restrictions on copying or conditions on use. Requests for restriction for a period of up to 2 years can be made when submitting the final copies of your thesis to the UNSW Library. Requests for a longer period of restriction may be considered in exceptional circumstances and require the approval of the Dean of Graduate Research.

Originality statement

‘I hereby declare that this submission is my own work and to the best of my knowledge it contains no materials previously published or written by another person, or substantial proportions of material which have been accepted for the award of any other degree or diploma at UNSW or any other educational institution, except where due acknowledgement is made in the thesis. Any contribution made to the research by others, with whom I have worked at UNSW or elsewhere, is explicitly acknowledged in the thesis. I also declare that the intellectual content of this thesis is the product of my own work, except to the extent that assistance from others in the project's design and conception or in style, presentation and linguistic expression is acknowledged.’

Signed .....

Date .....

Copyright statement

‘I hereby grant the University of New South Wales or its agents a non-exclusive licence to archive and to make available (including to members of the public) my thesis or dissertation in whole or part in the University libraries in all forms of media, now or here after known. I acknowledge that I retain all intellectual property rights which subsist in my thesis or dissertation, such as copyright and patent rights, subject to applicable law. I also retain the right to use all or part of my thesis or dissertation in future works (such as articles or books).’

‘For any substantial portions of copyright material used in this thesis, written permission for use has been obtained, or the copyright material is removed from the final public version of the thesis.’

Signed …….....

Date …….....

Authenticity statement

‘I certify that the Library deposit digital copy is a direct equivalent of the final officially approved version of my thesis.’

Signed …….....

Date ……...…….....

Inclusion of publications statement

UNSW is supportive of candidates publishing their research results during their candidature as detailed in the UNSW Thesis Examination Procedure.

Publications can be used in their thesis in lieu of a Chapter if:

- The candidate contributed greater than 50% of the content in the publication and is the "primary author", i.e. the candidate was responsible primarily for the planning, execution and preparation of the work for publication
- The candidate has approval to include the publication in their thesis in lieu of a Chapter from their supervisor and Postgraduate Coordinator.
- The publication is not subject to any obligations or contractual agreements with a third party that would constrain its inclusion in the thesis

Please indicate whether this thesis contains published material or not:

☐ This thesis contains no publications, either published or submitted for publication

(if this box is checked, you may delete all the material on page 2)

☒ Some of the work described in this thesis has been published and it has been documented in the relevant Chapters with acknowledgement

(if this box is checked, you may delete all the material on page 2)

☐ This thesis has publications (either published or submitted for publication) incorporated into it in lieu of a chapter and the details are presented below

CANDIDATE'S DECLARATION

I declare that:

- I have complied with the UNSW Thesis Examination Procedure
- where I have used a publication in lieu of a Chapter, the listed publication(s) below meet(s) the requirements to be included in the thesis.

Candidate's Name	Signature	Date (dd/mm/yy)
Sophie Holland		

Abstract

Anaerobic microbial metabolism of dichloromethane (DCM; CH_2Cl_2), quaternary amines, and methanol has important implications for carbon and nitrogen cycling in oligotrophic environments and the atmospheric flux of climate-active trace gasses. A novel, strictly anaerobic member of the *Peptococcaceae* family, strain DCMF, is the dominant organism in a DCM-fermenting enrichment culture (DFE) and one of very few known bacteria capable of fermenting DCM to the innocuous end product acetate. Long read, whole genome sequencing provided a single, circularised 6.44 Mb chromosome for strain DCMF, which contains 5,772 predicted protein-coding genes including an abundance of MttB superfamily methyltransferases. Genomic comparison of anaerobic, DCM-degrading bacteria provided a relatively small core genome, including the Wood-Ljungdahl pathway.

Strain DCMF is the first non-obligate anaerobic DCM-degrading bacterium. Genomic, physiological and proteomic experiments confirmed that it is an anaerobic methylotroph, able to metabolise DCM, methanol, and methyl groups from quaternary amines via the Wood-Ljungdahl pathway. The quaternary amine choline was converted to glycine betaine, which was demethylated to sarcosine with a glycine betaine methyltransferase, then reductively cleaved to methylamine and acetate. Methanol (via a methanol methyltransferase) and DCM were fermented to acetate. Comparative proteomics revealed a methyltransferase system that was significantly more abundant in cells grown with DCM than glycine betaine. The novel, putative DCM methyltransferase genes are highly conserved between anaerobic DCM-degrading bacteria. Genomic and physiological evidence support placement of strain DCMF in a novel genus, for which we propose the name '*Candidatus* Formamonas warabiya'.

Cohabiting bacteria in the DFE community have persisted despite repeated attempts to isolate strain DCMF, yet strain DCMF-free enrichments demonstrated that most are unable to utilise DCM, quaternary amines, or methanol. Five MAGs were generated from the long-read sequencing data and a metaproteogenomic approach suggested that the cohabiting organisms persist in the culture via necromass fermentation, i.e. oxidation of carbohydrates, proteins, and sugars released from

expired strain DCMF cells. The DFE culture is a long-term stable-state community that highlights interactions between foundation species and supporting bacteria, as well as important pathways of carbon and nitrogen cycling.

Acknowledgements

To my supervisors, Mike Manefield and Matt Lee, endless thank yous for your patience, guidance, and support over the years. It has been a privilege to explore the world of anaerobic dechlorination under your combined expertise. Mike – thanks for supporting my science career all the way from undergrad through Honours to my PhD. Your enthusiasm for my research and staunch advocation for putting my mental health first has been invaluable. Matt – thank you for putting so much of your energy into making this project what it is, for always being available for any questions I might have, and for your kind support when my anxiety would get the better of me.

Miriam – what is there to say, other than "THIS IS FINE". I couldn't have done this without you. Thanks for being there for the good times, the bad times, and the downright crazy times. You help me put it all in perspective, you kept me laughing through all the stress, and are a fountain of wisdom both professionally and personally. Thanks dudi <3

I am grateful to the rest of the Manefield lab both past and present for the good company and fruitful discussions. Huge gratitude to Haluk, who put many hours into creating detailed lists of genes. Iris, thanks for always looking out for me, taking care of my babies when I was away, and proving it is possible to finish one PhD and want to come back for a second one. Thanks to Önder for dealing with the ordering system so we don't have to and for all the good advice over the years. To Gan, Mol, Priyanka and Jay for the company and camaraderie. Thanks to Mac and Brittney for being badass women in STEMM. Thanks to Rob, Sabrina, Emile, and the rest for the good chats and fun camping trips way back at the start.

I'm grateful to have begun exploring bioinformatics under the guidance of Dr Rich Edwards – his work on the strain DCMF genome assembly kicked this whole thesis off, and advice on metagenome assembly opened to door to Chapter 5 for me. Thank you also to Dr Ling Zhong at the BMSF for her help with proteomics and careful answering of all my questions and to Dr James McDonald for sharing his LC-MS/MS and GC-TQMS expertise with me. I'm also grateful to Dr Stéphane Vuilleumier for many illuminating conversations about the genomic comparison work.

To all my friends and family – thank you, thank you, thank you for your support, especially in these last six months, and for your patience as I fumbled my way through four years of last-minute cancelling plans and running late because I was stuck in the lab.

Jesse – you've dated the PhD, you've helped me birth the PhD, you've never questioned my mixed metaphors about what your relationship is to this damn PhD. I don't know how I would have gotten through the write-up without you. You've held space for all my big emotions on this rollercoaster and I will forever be grateful for your support and your love. Thanks for being my biggest cheerleader.

Talina, my #1 bun, you have been a source of so much strength. You've looked after me through thick and thin this whole time, always with the biggest, most heartfelt encouragement. You've shown me the magic that awaits on the other side and even from the depths of the Amazon jungle you've provided more support in your messages than most people manage to do in person. I am so endlessly grateful to have you in my life.

To housemates past and present – thank you for the laughs, the boogies, the cups of tea, and the shoulders to cry on. You've all contributed at some point and I'm grateful to have lived with friends who are family. To Dan, Si, and everyone at the Erko house, thank you for being my second home, for feeding me so many delicious meals, and for letting me hang out for days on end. I am deeply appreciative of what you've created and let me be a part of.

To my family – thank you for your support not just during this PhD, but always. Thanks Mum for all your messages, cards, phone calls, baked goods, and constant supply of cute Tazzy photos. Thanks Dad for always encouraging me to pursue my dreams, whatever they might be. And thanks Seth and Phoebe for being there for a chat or some R&R at home whenever I needed it.

And finally – thanks to the amazing teachers at Sydney Pole; to everyone who's ever uploaded a 4 hour rainforest thunderstorm recording or sick folktronica mix to YouTube; to my psych; to the party organisers who've provided spaces to dance my troubles away. You've all helped keep me sane. This thesis was only possible because of the support of everyone mentioned here. I'm lucky to have them.

Table of Contents

List of Publications

Holland SI*, Edwards RJ*, Ertan H, Wong YK, Russell TL, Deshpande NP, Manefield MJ and Lee M. 2019. Whole genome sequencing of a novel, dichloromethane-fermenting *Peptococcaceae* from an enrichment culture. *PeerJ* **7**:e7775. (*These authors contributed equally to this work.)

Elements of this publication are incorporated into Chapter 2. The author contributions to the *chapter* (NB the chapter has a wider scope than the publication) are as follows:

Author	**Contribution**
Sophie I Holland	Conception and design (40%), analysis and interpretation of data (60%), writing (90%).
Richard J Edwards	Conception and design (10%), analysis and interpretation of data (15%), writing (10%), critically reviewing the text (20%).
Haluk Ertan	Analysis and interpretation of data (15%), critically reviewing the text (10%).
Yie Kuan Wong	Contributed sections to the paper that are not incorporated into this Chapter.
Tonia L Russell	Analysis and interpretation of data (5%).
Nandan P Deshpande	Analysis and interpretation of data (5%).
Michael J Manefield	Conception and design (25%), critically reviewing the text (35%).
Matthew Lee	Conception and design (25%), critically reviewing the text (35%).

List of Figures

List of Tables

List of Abbreviations

AAI	average amino acid identity
ANI	average nucleotide identity
ATP	adenosine triphosphate
BES	2-bromoethanosulfonate
bp	basepairs
BLAST	Basic Local Alignment Search Tool
BMC	bacterial microcompartment
C1	one-carbon
Cas	CRISPR associated
CAZyme	carbohydrate active enzyme
CDS	coding sequences
CH_2=THF	5,10-methylenetetrahydrofolate
CH_3-THF	methyl-tetrahydrofolate
CI	confidence interval
CoP	corrinoid protein
CRISPR	clustered regularly interspaced short palindromic repeats
DAPI	4,6-diamidino-2-phenylindole
DCM	dichloromethane
DFE	DCM-fermenting enrichment
DGGE	denaturing gradient gel electrophoresis
DNA	deoxyribonucleic acid
DNAPL	dense non-aqueous phase liquid
EDTA	ethylenediaminetetraacetic acid
EUT	ethanolamine utilising
FASP	filter-aided sample preparation
FDR	false discovery rate
FISH	fluorescence *in situ* hybridisation
GC-FID	gas chromatography with flame ionization detector
GC-PDD	gas chromatography with pulse discharge detector
GC-TQMS	gas chromatography with triple quadrupole mass spectrometry
Gg	gigagrams

GI	genomic island
GTDB	Genome Taxonomy Database
HGT	horizontal gene transfer
IMG	Integrated Microbial Genomes
JGI	Joint Genome Institute
Kb	kilobases
LC-MS/MS	liquid chromatography with tandem mass spectrometry
LFC	$\log_2$ fold change
LFQ	label free quantitative
MAG	metagenome-assembled genome
Mb	megabases
MeCoCo	Metabolically Cohesive Consortium
MICFAM	MicroScope gene families
MOPS	3-(*N*-morpholino)propanesulfonic acid
MRM	multiple reaction monitoring
MT_1	methyltransferase 1, i.e. the methyltransferase responsible for transferring a methyl group from a substrate to a corrinoid protein
MT_2	methyltransferase 2, i.e. the methyltransferase responsible for transferring a methyl group from the corrinoid protein to the final accepting compound (e.g. tetrahydrofolate or coenzyme M)
NAD^+/NADH	nicotinamide adenine dinucleotide (oxidised/reduced)
NCBI	National Centre for Biotechnology Information
nd	not described
OSWER	Office of Solid Waste and Emergency Response
PCA	Principle Components Analysis
PCR	polymerase chain reaction
PDU	propanediol utilising
qPCR	quantitative polymerase chain reaction
rcf	relative centrifugal force
rRNA	ribosomal ribonucleic acid
SD	standard deviation
SDS	sodium dodecyl sulphate
SECIS	selenocysteine insertion sequence

SMRT	Singe Molecule Real Time™
Tg	teragrams
THF	tetrahydrofolate
tRNA	transfer ribonucleic acid
U.S.	The United States of America
V	volts
VFA	volatile fatty acid
VSLS	very short-lived substance
y	year

1. Introduction

1.1 Dichloromethane is a toxic pollutant

Dichloromethane (DCM; CH_2Cl_2) is a chlorinated one-carbon compound (Figure 1.1) that has been used extensively in industry and domestically over the past century. It belongs to a group of widely used chemicals (organochlorines) that are notoriously toxic and recalcitrant environmental pollutants. Within the chloromethanes – tetrachloromethane (CCl_4), trichloromethane ($CHCl_3$, commonly known as chloroform), DCM, and chloromethane (CH_3Cl) – all except chloromethane have been classed as priority pollutants by the U.S. Environmental Protection Agency, the Ministry of Environmental Protection of China, and the European Commission (Huang *et al.*, 2014).

DCM is a non-polar, volatile liquid that is only slightly water-soluble (13 g l^{-1}) but miscible with many organic solvents under standard conditions. While these properties have made it valuable for use in industry, they have also contributed to its recalcitrance as a pollutant. Within industry, DCM has primarily been used in paint strippers, as a process and extraction solvent, and in the manufacturing of pharmaceuticals (WHO Regional Office for Europe, 2000). It has also been used in aerosols, foam blowing of polyurethane, metal degreasing, cleaning products, as an adhesive or sealant, and to decaffeinate coffee (WHO Regional Office for Europe, 2000).

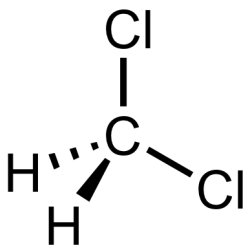


Figure 1.1 Chemical structure of dichloromethane.

1.1.1 Natural and anthropogenic sources of dichloromethane

DCM production globally, including all anthropogenic and natural sources, is estimated to exceed 900 gigagrams per year (Gg y^{-1}) (Gribble, 2010). In the post-WWII industrial boom, global DCM production increased over six-fold from 93 Gg in

1960 to 570 Gg in 1980 (Brooke and How, 1994; Hi Valley Chemical, 2020). DCM continues to be widely used, with over 118 Gg y^{-1} still being produced in the U.S. (United States Environmental Protection Agency, 2017).

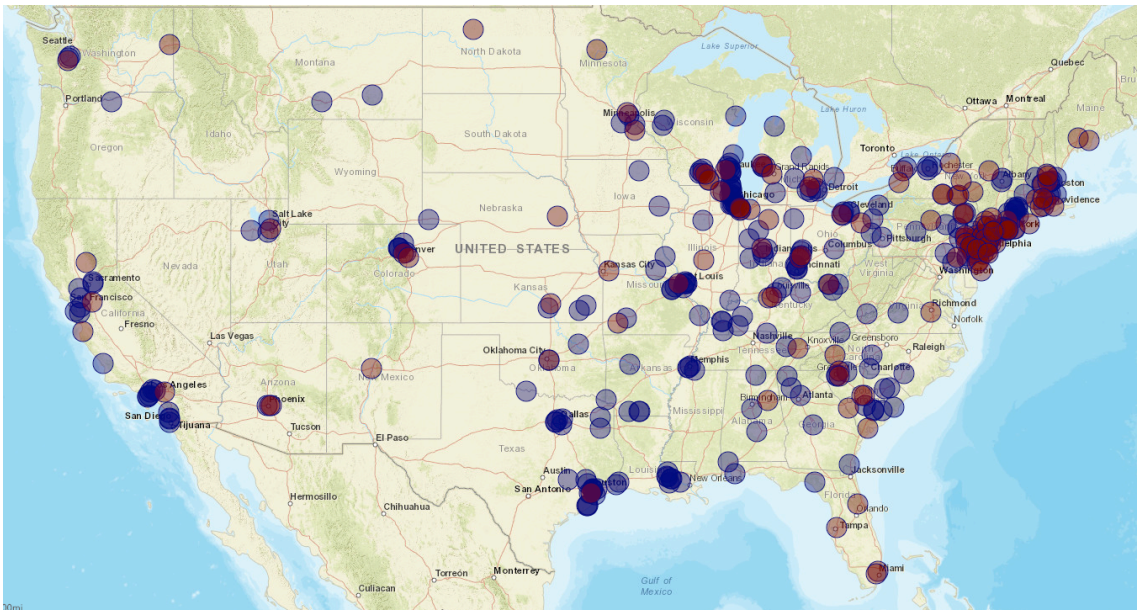


Figure 1.2 Map of the U.S. showing National Priority List sites containing DCM (red circles) and releases of DCM to the environment according to the 2016 Toxic Release Inventory (blue circles). Image was taken from TOXMAP (US National Library of Medicine, 2019) on 19/11/18[1].

The extensive use of DCM, as well as historically lax attitudes to chemical storage, handling, and disposal, have led to widespread environmental contamination. DCM is one of the most frequently encountered subsurface pollutants in industrial areas (Shestakova and Sillanpää, 2013). In 2019, it was present at 30% of the National Priority List sites within the U.S. and its territories (Figure 1.2); these sites are chosen for the concentration and toxicity of contaminants present (US National Library of Medicine, 2019). More locally, Australia reported 0.15 Gg of DCM emissions across 168 sites from 2017-2018, the vast majority of which were to the atmosphere (Australian Government - Department of the Environment and Energy, 2019). Atmospheric concentrations of DCM have increased 38-69% in the last decade alone, due to its frequent use and release into the air (Hossaini *et al.*, 2015b; Leedham Elvidge *et al.*, 2015). Although DCM contributes to only a small fraction of

[1] Unfortunately, the TOXMAP site was closed in December 2019, so a more up-to-date image could not be obtained.

the atmospheric chlorine pool (~0.7%) (Laube *et al.*, 2008), anthropogenic activity accounts for 90% of the tropospheric abundance of DCM (Montzka *et al.*, 2011).

Due to its density and polarity, DCM that is released into the environment tends to sink down through the vadose zone and the groundwater until it reaches an impermeable geological structure such as bedrock or clay (Figure 1.3). Large quantities of organochlorines (including DCM) may form a dense, non-aqueous phase liquid (DNAPL) source zone, due to their immiscibility with the surrounding groundwater. As the groundwater flows over this DNAPL zone, small quantities of the pollutant are solvated and transported in the direction of the flow, forming a dissolved plume that moves away from the initial site of contamination (Figure 1.3).

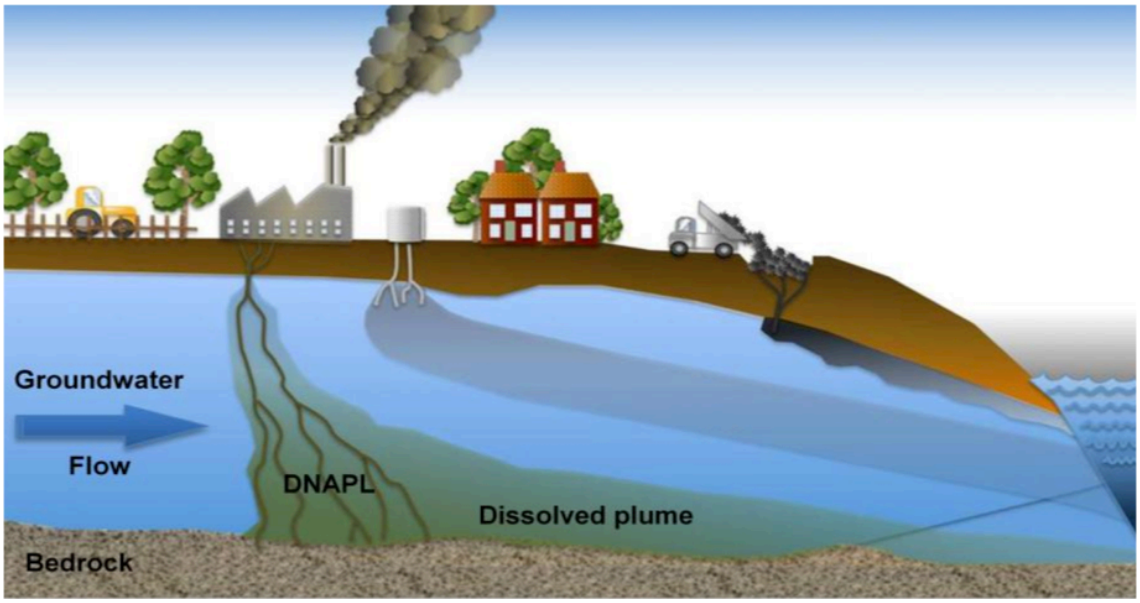


Figure 1.3 Schematic of organochlorine, i.e. DCM, release into the environment. Due to its density and polarity, DCM sinks to the bottom of the water table, forming a pool of dense, non-aqueous phase liquid (DNAPL). Small quantities of DCM are picked up as groundwater flows over the DNAPL, forming a dissolved plume of decreasing solvent concentration that spreads away from the site. Figure adapted from Koenig *et al* (2014).

As well as anthropogenic sources, approximately 30% of all DCM is of natural origin (Gribble, 2010). The bulk of this is oceanic sources and biomass burning (Gribble, 2010), e.g. the former was estimated to contribute 68 Gg y^{-1} of DCM (Kolusu *et al.*, 2017). Trace amounts of DCM are also released from volcanic activity, marine macroalgae, and wetlands (Gribble, 2010). Production from tropical terrestrial mangrove swamps was estimated at 1– 2 Gg y^{-1} (Kolusu *et al.*, 2018), while

macrophytes have been estimated to contribute 0.32 Gg y^{-1} to environmental DCM emissions (Baker *et al.*, 2001). It should be considered that environmental sources of DCM may be even larger than the reported estimates, as they do not take DCM sinks (i.e. microbial degradation) into account.

DCM is also produced from the microbial reductive dechlorination of trichloromethane ($CHCl_3$, commonly known as chloroform; Eq. 1) by *Dehalobacter* and *Desulfitobacterium* lineages (Grostern *et al.*, 2010; Ding *et al.*, 2014; Wong *et al.*, 2016). Anaerobic co-metabolism of trichloromethane by genera such as *Acetobacterium*, *Clostridium*, *Methanosarcina*, and *Pantoea* can also produce DCM (Egli *et al.*, 1988; Gälli and McCarty, 1989; Mikesell and Boyd, 1990; Bagley and Gossett, 1995; Baeseman and Novak, 2001; Shan *et al.*, 2010). Trichloromethane is even more toxic and a more common pollutant than DCM, present at ~36% of all National Priority List sites (US National Library of Medicine, 2019). Therefore, new DCM-contaminated sites continue to emerge alongside legacy sites, and its degradation is often a crucial step in the remediation of sites contaminated with higher-chlorinated methanes.

$$CHCl_3 + H_2 \rightarrow CH_2Cl_2 + H^+ + Cl^- \qquad \text{Eq. 1}$$

1.1.2 Adverse health and environmental effects of dichloromethane

DCM is a problematic pollutant because of the adverse effects it has on the natural environment. Organic solvents such as DCM are known to challenge cell membrane integrity, causing leakage or even complete lysis (Sikkema *et al.*, 1995; Rodriguez Martinez *et al.*, 2008; Sherry *et al.*, 2014; Lueders, 2017). The compound is inhibitory to microorganisms. When released into the soil, it can inhibit indigenous enzyme activity, although this inhibition has been shown to abate over time (Kanazawa and Filip, 1986). As a methane analogue, DCM can also inhibit methanogenesis via competitive binding to key enzymes. Minimum inhibitory concentrations of as low as 1.8 mg l^{-1} (21 μM) DCM have been noted in anaerobic sludge cultures (Stuckey *et al.*, 1980), although other systems had minimum inhibitory concentrations around twice as high (Stuckey *et al.*, 1980; Yu and Smith, 2000). Numerous studies have investigated the IC_{50} of DCM (the concentration at which 50% of function, i.e. gas production, is lost), with results ranging from 8 – 204

mg l^{-1} (94 – 2,400 μM) (Bauchop, 1967; Thiel, 1969; Mack, 1973; Stuckey *et al.*, 1980; Vargas and Ahlert, 1987; Byers and Sly, 1993; Sanz *et al.*, 1997; Yu and Smith, 2000). However, the diminishing inhibitory effect of continuous DCM amendment on methanogenesis has also been observed in culture (Thiel, 1969; Stuckey *et al.*, 1980; Vargas and Ahlert, 1987).

Yu *et al* (2000) proposed inhibition of methanogenesis by chloromethanes can occur via direct and indirect mechanisms. In the former, the compounds preferentially bind to intracellular reduced corrinoid or porphinoid enzymes, due to their analogous structure to methane. Due to their difference in redox potential, trichloromethane (+0.56 V) has a higher affinity for these enzymes than DCM (+0.49 V) and is thus more inhibitory (Yu and Smith, 2000). Indirect inhibition of methanogenesis is only demonstrated by trichloromethane and chloroethenes (i.e. not DCM), as they can also bind free intracellular corrinoids/porphinoids, which shifts the equilibrium of these cofactors away from protein-bound to free-form and creates intermediates that redirect electron flow away from methanogenesis to dechlorination. Direct, competitive inhibition of methanogenesis like that caused by DCM is more effective and hence chloromethanes are more inhibitory to methanogens than chloroethenes (Yu and Smith, 2000).

As well as the harmful effects that DCM can have on the environment, it is dangerous to human health and classed as a possible (group 2B) carcinogen (International Agency for Research on Cancer, 1986). Exposure to DCM may adversely affect the central nervous system and reproductive system in humans, and is also associated with kidney and liver toxicity (Agency for Toxic Substances and Disease Registry, 2000; Starr *et al.*, 2006; Evans and Caldwell, 2010; Olvera-Bello *et al.*, 2010). Sustained inhalation can be fatal (e.g. Stewart and Hake, 1976; Bonventre *et al.*, 1977; Hall and Rumack, 1990). There have also been reports of carbon monoxide poisoning, following metabolism of DCM within the body (Fagin *et al.*, 1980).

1.2 Microbial transformation of dichloromethane

For terrestrial sites contaminated with organohalides such as DCM, bioremediation is an attractive option for pollutant removal. The practice of stimulating indigenous microorganisms (biostimulation) or applying exogenous microbial cultures

(bioaugmentation) has gained increasing favour in the remediation industry over the past few decades due to its ability to be carried out *in situ*, cost-effectiveness, and efficacy.

1.2.1 Aerobic dichloromethane transformation

Under aerobic conditions, a wide range of bacteria are capable of degrading DCM, most commonly methylotrophs. Species from the genera *Albibacter* (Doronina *et al.*, 2001), *Ancylobacter* (Firsova *et al.*, 2009), *Bacillus* (Wu *et al.*, 2007), *Gottschalkia* (Firsova *et al.*, 2010), *Lysinobacillus* (Wu *et al.*, 2009), *Paracoccus* (Doronina *et al.*, 1998), *Methylophilus* (Bader and Leisinger, 1994), *Methylopila* (Brunner *et al.*, 1980), and *Methylobacterium* (La Roche and Leisinger, 1990) are all capable of using DCM as a sole source of carbon and electrons. A comprehensive review of aerobic DCM-degrading strains is available from Muller *et al* (2011).

Aerobic dehalogenation reactions are catalysed by a DCM dehydrogenase from the glutathione S-transferase family (La Roche and Leisinger, 1990). The dehydrogenase is encoded by *dcmA* and expression is regulated by *dcmR* (La Roche and Leisinger, 1990, 1991). DCM dehalogenases are split into two groups, A and B, which differ primarily in their kinetic properties. Group B enzymes, such as the one found in *Methylophilus leisingeri* DM11, dechlorinate DCM significantly faster under substrate saturation conditions (Scholtz *et al.*, 1988). Although catalytic activity of Group A DcmA enzymes is typically low, it is highly inducible and can comprise 12-20% of total cell protein (Kohler-Staub and Leisinger, 1985). The more efficient Group B enzymes represent a lower percentage (~7%) of the total protein (Leisinger and Braus-Stromeyer, 1995).

Aerobic dechlorination with DcmA requires glutathione as a cofactor and results in the formation of a *S*-chloromethyl glutathione conjugate, which likely undergoes non-enzymatic hydrolysis to form *S*-hydroxymethyl glutathione. Decomposition of this compound produces formaldehyde, a central metabolite of methylotrophic bacteria; hydrochloric acid; and regenerates free glutathione (Stucki *et al.*, 1981) (Figure 1.4). Thus, the overall reaction can be described by Equation 2.

$$CH_2Cl_2 + H_2O \rightarrow CH_2O + 2\ H^+ + 2\ Cl^- \qquad \text{Eq. 2}$$

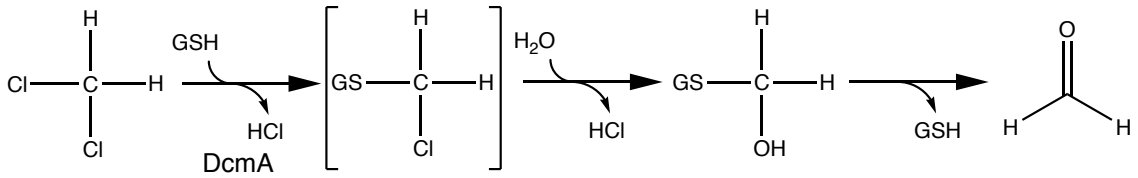


Figure 1.4 Aerobic degradation of dichloromethane is catalysed by DcmA, a glutathione S-methyltransferase.

DcmA was initially thought to be the only requirement for aerobic DCM dechlorination, but later work showed that additional genes and proteins may also be required (Kayser *et al.*, 2000; Kayser and Vuilleumier, 2001; Kayser *et al.*, 2002; Vuilleumier, 2002). For example, an efficient DNA repair system is necessary to negate the genotoxic effects of DcmA-mediated DCM transformation (Kayser *et al.*, 2000; Kayser and Vuilleumier, 2001).

1.2.2 Anaerobic dichloromethane transformation

Fewer bacteria have been observed to degrade DCM under anaerobic conditions. This is often crucial for natural attenuation, given the propensity of DCM to form DNAPL zones with low oxygen availability and redox potential. While anaerobic respiration of DCM is theoretically possible, no organisms with this metabolism have been identified. Rather, this compound can be utilised as a sole source of carbon and electrons by some strains of *Acinetobacter* and *Hyphomicrobium* under denitrifying conditions (Melendez *et al.*, 1993; Freedman *et al.*, 1997). Other consortia degrade DCM under methanogenic conditions (Freedman and Gossett, 1991; Stromeyer *et al.*, 1991).

Only two bacteria capable of transforming DCM under fermentative conditions have been characterised: *Dehalobacterium formicoaceticum* strain DMC and *'Candidatus* Dichloromethanomonas elyunquensis' strain RM (Mägli *et al.*, 1996; Kleindienst *et al.*, 2017). *D. formicoaceticum* is the only DCM-fermenting bacteria currently isolated in pure culture, and has only recently come under renewed interest after a 20 year gap in the literature (Mägli *et al.*, 1996; Chen *et al.*, 2017b). It was enriched from a mixed culture in a charcoal-packed fixed bed reactor fed with contaminated anaerobic groundwater (Stromeyer *et al.*, 1991) and isolated in the mid 1990s (Mägli *et al.*, 1996). DCM fermentation resulted in the formation of formate and acetate in a 2:1 molar ratio (Eq. 3). Studies with cell-free extracts and ^{13}C-labelled

DCM indicated that this transformation likely occurred via the Wood-Ljungdahl pathway (Mägli *et al.*, 1996, 1998; Chen *et al.*, 2020), which genome sequencing later confirmed is present in its entirety (Chen *et al.*, 2017b).

$$3\ CH_2Cl_2 + 4\ H_2O + CO_2 \rightarrow CH_3COO^- + 2\ HCOO^- + 9\ H^+ + 6\ Cl^- \qquad \text{Eq. 3}$$

The Wood-Ljungdahl pathway is typically found in anaerobic acetogens and is used to convert H_2 and CO_2 into acetate. It is suggested to be one of the first metabolic cycles to evolve on early Earth (Fuchs, 1989; Wood, 1991; Russell and Martin, 2004; Berg *et al.*, 2010; Weiss *et al.*, 2016). The exact mechanism by which the chlorine moiety of DCM is removed from the molecule remains unknown, but Mägli and colleagues showed that DCM is mostly transformed into 5,10-methylenetetrahydrofolate, which enters the pathway and can then be simultaneously oxidised to formate and reduced to acetate (Figure 1.5) (Mägli *et al.*, 1996, 1998).

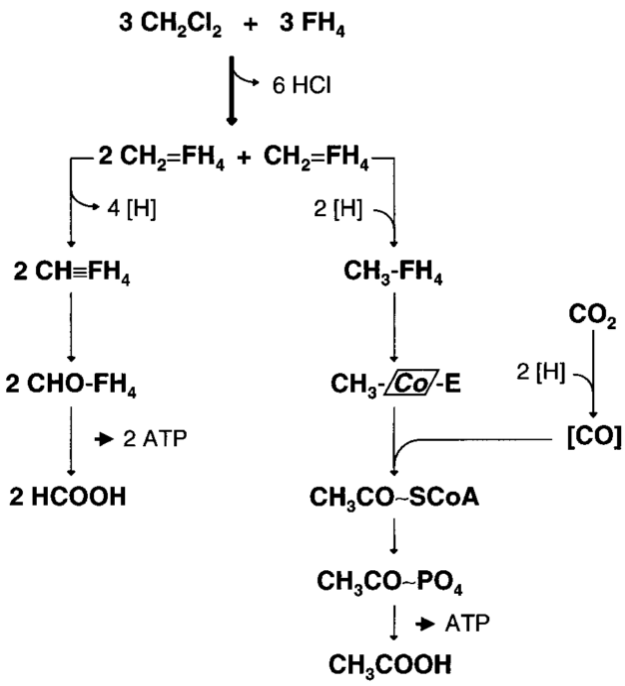


Figure 1.5 Proposed schema of the transformation of DCM to formate and acetate by *Dehalobacterium formicoaceticum* strain DMC. Taken from Mägli *et al* (1998)**.**

'*Ca*. Dichloromethanomonas elyunquensis' strain RM is a more recent discovery enriched from pristine river sands in Puerto Rico (Kleindienst *et al.*, 2017). A genome sequence is also available for this organism, although it has proven resistant to isolation and exists in a enrichment culture, RM (Kleindienst *et al.*, 2016, 2017).

Initially thought to ferment DCM to acetate in a manner similar to *D. formicoaceticum* (Kleindienst *et al.*, 2017), further investigation showed that '*Ca.* Dichloromethanomonas elyunquensis' in fact completely mineralises DCM to H_2 and CO_2 (Eq. 4) (Chen *et al.*, 2020). These products are then used by homoacetogens and methanogens present in the enrichment, explaining the observed acetate (Eq. 5) and methane (Eq. 6) formation, respectively (Chen *et al.*, 2020).

$$CH_2Cl_2 + 2\ H_2O \rightarrow CO_2 + 2\ H_2 + 2\ Cl^- + 2\ H^+ \qquad \text{Eq. 4}$$

$$2\ CO_2 + 4\ H_2 \rightarrow CH_3COO^- + H^+ + 2\ H_2O \qquad \text{Eq. 5}$$

$$CO_2 + 4\ H_2 \rightarrow CH_4 + 2\ H_2O \qquad \text{Eq. 6}$$

Unlike *D. formicoaceticum*, '*Ca.* Dichloromethanomonas elyunquensis' encodes reductive dehalogenases in its genome, as well as the full set of genes for the Wood-Ljungdahl pathway (Kleindienst *et al.*, 2016). Reductive dehalogenases are enzymes found in organohalide respiring bacteria that catalyse the transfer of electrons to an organohalide compound acting as the terminal electron acceptor of a respiratory chain (reviewed in Jugder *et al.*, 2016). Proteomics of culture RM revealed that two reductive dehalogenases were among the most abundant proteins expressed by '*Ca.* Dichloromethanomonas elyunquensis', alongside four corrinoid-dependent methyltransferases. All proteins for the Wood-Ljungdahl pathway were also expressed (Kleindienst *et al.*, 2019).

The difference in fermentation end products and presence of reductive dehalogenases strongly suggest that '*Ca.* Dichloromethanomonas elyunquensis' utilises the Wood-Ljungdahl pathway for DCM metabolism differently to *D. formicoaceticum*. This was supported by carbon-chlorine isotope work that indicated unique C—Cl bond-breaking mechanisms likely operated in the two organisms – S_N2 in *D. formicoaceticum* and S_N1 in '*Ca.* Dichloromethanomonas elyunquensis' (Chen *et al.*, 2018). In summary, *D. formicoaceticum* utilises the Wood-Ljungdahl pathway primarily in both oxidative and reductive directions, whilst '*Ca.* Dichloromethanomonas elyunquensis' utilises it exclusively in the oxidative direction (Chen *et al.*, 2020).

There are a handful of other reports of anaerobic bacteria capable of transforming DCM, including *Dehalobacterium* species (Trueba-Santiso *et al.*, 2017) and two

Dehalobacter species (Justicia-Leon *et al.*, 2012; Lee *et al.*, 2012). The former fermented DCM to formate and acetate, similar to *D. formicoaceticum*, but had the distinguishing additional capability to dehalogenate dibromomethane to near completion (Trueba-Santiso *et al.*, 2017, 2020). Both *Dehalobacter* species were identified in DCM-degrading enrichment cultures but, upon further study, disappeared in favour of other novel DCM-metabolising organisms. The *Dehalobacter* reported by Justicia-Leon *et al* (2012) existed in the enrichment RM, which became dominated by the novel DCM-degrading organism '*Ca.* Dichloromethanomonas elyunquensis' (Kleindienst *et al.*, 2017).

The other *Dehalobacter* species whose growth was linked to DCM transformation was enriched in a parent culture to this study – the methanogenic, DCM-fermenting culture DCMD (Lee *et al.*, 2012). The Archaeal population in DCMD was dominated by a hydrogenotrophic methanogen from the genus *Methanoculleus* and *Dehalobacter* sp. growth was inhibited by addition of excess hydrogen. This suggested that hydrogen was a product of DCM fermentation and utilised by *Methanoculleus* in a syntrophic interaction with the *Dehalobacter* species (Lee *et al.*, 2012). Following removal of the methanogens from culture DCMD, hydrogen levels in the culture likely rose, preventing the growth of the *Dehalobacter* sp. and enabling growth of the novel, non-hydrogenogenic, DCM-fermenting *Peptococcaceae* described in this thesis, strain DCMF (Holland *et al.*, 2019).

1.3 Microbial metabolism of quaternary amines

1.3.1 Significance and distribution of quaternary amines

There are no reports in the literature of anaerobic bacteria capable of metabolising both DCM and quaternary amines, though some methylotrophs capable of aerobic DCM degradation could also utilise methylated amines and methane for growth (Brunner *et al.*, 1980; Doronina *et al.*, 2000). Quaternary amines (also known as quaternary ammonium cations) are amines with four organic groups attached to the central nitrogen moiety. Unlike primary, secondary, and tertiary amines, they retain their positive charge regardless of the surrounding pH. Both choline (Figure 1.6 G) and glycine betaine (also known as trimethylglycine, Figure 1.6 D) are quaternary amines that play significant roles in environmental and human contexts.

Trimethylamine (Figure 1.6 A) is a tertiary amine whose metabolism is linked to that of the quaternary amines via a number of microbial processes.

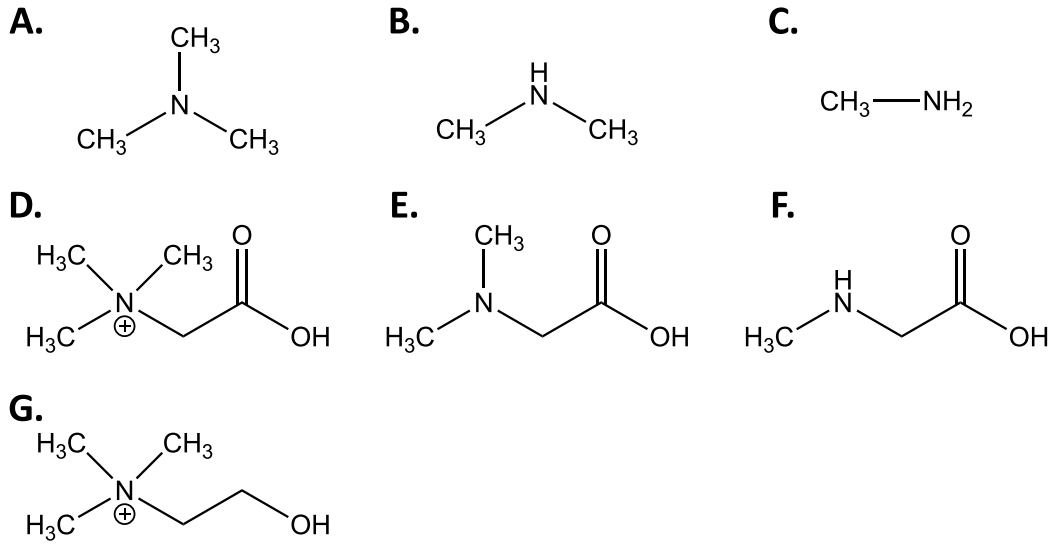


Figure 1.6 The chemical structure of the methylated amines: trimethylamine (A), dimethylamine (B), methylamine (C); the methylated glycines: glycine betaine (trimethylglycine, D), dimethylglycine (E), sarcosine (methylglycine, F); and choline (G).

Choline and glycine betaine are environmentally significant osmolytes and their metabolism is closely linked throughout nature (Neill *et al.*, 1978; King, 1984). The former is predicted to be more abundant in the environment, although predominantly as a part of larger molecules such as the eukaryotic phospholipids phosphatidylcholine and sphingomyelin. Evidence for the rapid bacterial uptake of choline (Snipes *et al.*, 1974; Salvano *et al.*, 1989; Kiene, 1998; Lucchesi *et al.*, 1998), along with bioinformatic analysis showing the near ubiquity of a biochemical pathway to oxidise choline to glycine betaine in soil and water environments, suggest that this compound is a critical precursor to glycine betaine (Wargo, 2013). While some bacteria can harness energy from choline, preference appears to be given to the conversion into glycine betaine when communities are under osmotic stress (Kiene, 1998).

Glycine betaine is a compatible solute capable of protecting proteins in saline environments. It has been shown to act as a osmoregulant in bacteria (Galinski and Trüper, 1982; Imhoff, 1986; Csonka, 1989), marine algae (Blunden *et al.*, 1982), marine invertebrates (Beers, 1967), plants (Larher *et al.*, 1982), and even some vertebrates (Yancey *et al.*, 1982). Glycine betaine is also an important source of

nitrogen, comprising up to 20% of the total nitrogen in hypersaline environments (King, 1988b). Both choline and glycine betaine have been suggested to play an accessory role in shaping microbial communities, based on their limited availability in the environment and importance for coping with osmotic stress (Wargo, 2013).

In addition to their roles in the environment, both compounds are important for human and animal health. For humans, choline is an essential nutrient that has roles in cell membrane integrity and lipid metabolism (Sheard and Zeisel, 1989; Zeisel, 2000), while glycine betaine can act as a methyl donor in the liver and as a protective osmolyte in the kidney (Craig, 2004). Both compounds are precursors of trimethylamine and trimethylamine *N*-oxide, readily converted by gut microflora (Martínez-del Campo *et al.*, 2015). In excess, the former compound can cause trimethylaminuria (also known as fish odour syndrome) (Mitchell and Smith, 2001), while the latter has recently been linked to increased risk of cardiovascular disease (Wang *et al.*, 2011). Thus, manipulation of gut microflora and promotion of microbial metabolic pathways that do not transform choline and glycine betaine into trimethylamine/trimethylamine *N*-oxide may be a novel therapeutic approach worthy of exploration.

1.3.2 Anaerobic microbial metabolism of choline and glycine betaine

Fermentation of choline to trimethylamine and acetate has been reported in sulphate-reducing bacteria (Hayward and Stadtman, 1959; Fiebig and Gottschalk, 1983; King, 1984), *Clostridia* (Bradbeer, 1965; Möller *et al.*, 1986), and *Pelobacter* species (Schink, 1985; Jameson *et al.*, 2019). Early studies reported acetaldehyde was an intermediate metabolite of choline fermentation (Hayward, 1960), but it was not until much later that the reaction mechanism was revealed as a novel glycyl radical enzyme: choline-trimethylamine lyase (Figure 1.7) (Craciun and Balskus, 2012). The conversion is catalysed by the *cut* (choline utilisation) cluster of genes within a bacterial microcompartment (Kuehl *et al.*, 2014; Jameson *et al.*, 2016b; Herring *et al.*, 2018). This gene cluster is found in a diverse yet unevenly distributed range of bacterial taxa, particularly those found in the human gut and marine ecosystems – reflecting two environments where choline is relatively abundant and

organisms are subject to significant osmotic stress (Martínez-del Campo *et al.*, 2015; Jameson *et al.*, 2016a).

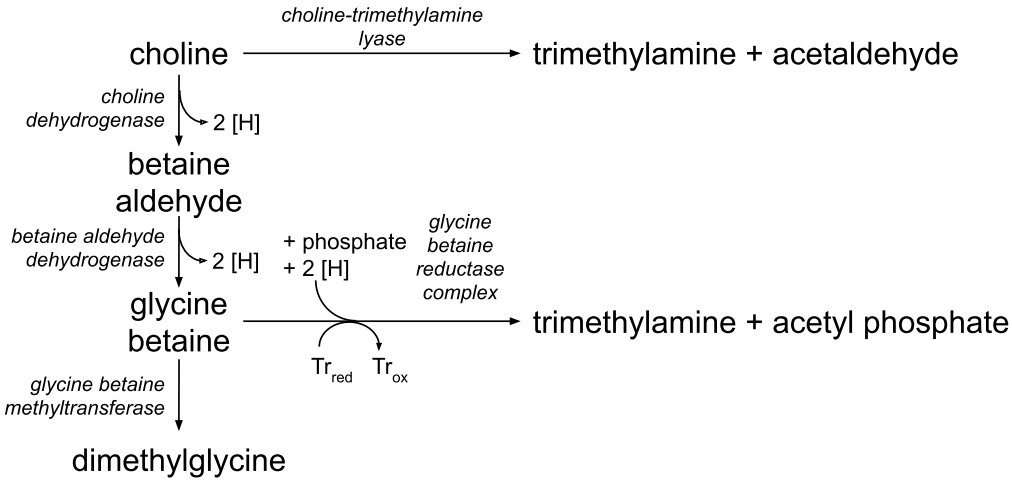


Figure 1.7 Schema of anaerobic quaternary amine metabolism. Redox cofactors are represented as electron equivalents [H] entering or leaving reactions. Tr, thioredoxin; red, reduced; ox, oxidised.

As mentioned above, there is also a near-ubiquitous pathway in soil and water bacteria to transform choline into the osmoprotectant glycine betaine (Wargo, 2013). This conversion is typically a two-step process via the intermediate betaine aldehyde (Figure 1.7). Oxidation of choline is catalysed by a membrane-bound choline dehydrogenase and the resulting betaine aldehyde is oxidised by a soluble betaine aldehyde dehydrogenase (Landfald and Strom, 1986; Choquet *et al.*, 1991; Boch *et al.*, 1994). *Alcaligenes* species appear to be an exception to this, encoding a soluble choline oxidase that performs both steps (Ohta-Fukuyama *et al.*, 1980).

Providing that glycine betaine isn't simply stored as an osmoprotectant, there are three broad mechanisms for its metabolism in anaerobic bacteria: (1) demethylation to dimethylglycine; (2) reductive cleavage to trimethylamine and acetyl phosphate, which requires an external electron donor; (3) a combination of both demethylation and reductive cleavage, in which the reducing equivalents for reductive cleavage are generated from oxidation of the released methyl group(s) (Figure 1.7). These mechanisms are limited to organisms with the Wood-Ljungdahl pathway, which is used for methyl group oxidation, i.e. acetogenic bacteria (Müller *et al.*, 1981; Eichler and Schink, 1984), sulphate-reducing bacteria (Heijthuijsen and Hansen, 1989; Ticak *et al.*, 2014), and methanogenic archaea (Watkins *et al.*, 2014).

Bacteria capable of growth via demethylation of glycine betaine include *Eubacterium limosum* (Müller *et al.*, 1981) and numerous *Acetobacterium* species (Eichler and Schink, 1984; Tanaka and Pfennig, 1988; Kotsyurbenko *et al.*, 1995). Demethylation is catalysed by a glycine betaine methyltransferase (MtgB; Figure 1.7). Ticak *et al* (2014) were the first to report that a non-pyrrolysine member of the widespread trimethylamine methyltransferase (MttB) family was in fact a glycine betaine methyltransferase. The enzyme is a corrinoid dependent methyltransferase (MT_1) that transfers the methyl group from glycine betaine onto a cognate corrinoid protein (CoP), from which a second methyltransferase (MT_2) transfers it onto tetrahydrofolate (THF) (Ticak *et al.*, 2014; Visser *et al.*, 2016; Lechtenfeld *et al.*, 2018). MtgB homologs were found in a large number of species, particularly within the Firmicutes and alpha proteobacteria, indicating that glycine betaine demethylation may be more widespread than initially observed (Ticak *et al.*, 2014). Both *E. limosum* and *Acetobacterium woodii* produce dimethylglycine (Figure 1.6 E) without further demethylating this product to sarcosine (also known as methyl glycine; Figure 1.6 F), suggesting that their enzymes are specific to glycine betaine alone (Müller *et al.*, 1981; Lechtenfeld *et al.*, 2018).

The second mechanism of anaerobic glycine betaine metabolism is reductive cleavage to trimethylamine and acetyl-phosphate (Figure 1.7). This is essentially a modified Stickland fermentation, a reaction that couples the oxidation and reduction of amino acids to organic acids (Nisman, 1954). Glycine betaine acts as the electron acceptor, and hydrogen, formate or various amino acids such serine can act as the electron donor. A range of fermentative bacteria, including numerous clostridial species (Naumann *et al.*, 1983; Möller *et al.*, 1986), *Haloanaerobacter salinarius* (Mouné *et al.*, 1999), and *Peptoclostridium acidaminophilum* (previously *Eubacterium* (Galperin *et al.*, 2016)) (Zindel *et al.*, 1988) perform reductive cleavage of glycine betaine. *Desulfuromonas acetoxidans* is uniquely able to use reducing equivalents generated by the oxidation of acetate (produced from the reductive cleavage of glycine betaine) to CO_2 in the reductive cleavage, negating the requirement for an external electron donor (Heijthuijsen and Hansen, 1989).

Thirdly, *Sporomusa* species are unique in combining the demethylation and reductive cleavage mechanisms (Möller *et al.*, 1984; Visser *et al.*, 2016). *Sporomusa*

ovata An4 demethylated a small proportion of glycine betaine to dimethylglycine and then sarcosine, producing reducing equivalents via the oxidation of the removed methyl groups to CO_2 via the Wood-Ljungdahl pathway. These were then funnelled into reductive cleavage of the majority of the glycine betaine to trimethylamine (Möller *et al.*, 1984; Visser *et al.*, 2016). Interestingly, *S. ovata* H1 was grown with dimethylglycine, trimethylamine was still amongst the end products, indicating that a portion of the dimethylglycine was methylated to form glycine betaine for reductive cleavage to trimethylamine (Möller *et al.*, 1984).

1.4 Microbial utilisation of methanol

Methanol is widely distributed in the terrestrial biosphere from the breakdown of pectin and lignin in plant cell walls (Schink and Zeikus, 1980). Significant concentrations of methanol have also been measured in the ocean (50 – 400 nM) (Galbally and Kirstine, 2002; Singh *et al.*, 2003; Williams *et al.*, 2004; Kameyama *et al.*, 2010; Beale *et al.*, 2011, 2013), much of which originates from atmospheric flux (Yang *et al.*, 2013). Globally, sources and sinks of methanol are approximately equivalent, at 340 and 270 Tg y^{-1}, respectively (these values are in balance according to the margin of error on the authors' calculations) (Heikes *et al.*, 2002). This suggests that methanol is both widely available to microorganisms and widely used by them.

Methanol can be utilised by numerous aerobic and facultatively anaerobic methylotrophic bacteria, which convert it to formaldehyde via a methanol dehydrogenase (Kolb, 2009). Anaerobically, the dominant pathway for methanol transformation is via methanol methyltransferases. Similar to glycine betaine methyltransferases, these enzymes are found in acetogenic bacteria, sulphate-reducing bacteria, and methanogenic archaea, which compete against each other for methanol in anoxic environments (Oremland and Polcin, 1982). Methanol methyltransferases are mechanistically similar to their glycine betaine counterparts, comprising an MT_1 (MtaB), MT_2 (MtaA), and CoP (MtaC), which ultimately transfer the methyl group onto THF in bacteria or coenzyme M in archaea (van der Meijden *et al.*, 1983a, 1983b, 1984a, 1984b; Stupperich and Konle, 1993; Sauer *et al.*, 1997). While methanol methyltransferase systems have been well-studied in methanogens, their bacterial counterparts are less thoroughly explored.

Acetogenic bacteria capable of growth on methanol include *Butyribacterium methylotrophicum* (Lynd and Zeikus, 1983), *Eubacterium limosum* (van der Meijden *et al.*, 1984b), *Moorella thermoacetica* (Daniel *et al.*, 1990), *Acetobacterium woodii* (Bache and Pfennig, 1981), and *Sporomusa ovata* (Möller *et al.*, 1984). Detailed studies have only recently been published for the latter two organisms (Visser *et al.*, 2016; Kremp *et al.*, 2018), characterising their use of the Wood-Ljungdahl pathway for further metabolism of the methyl-THF produced by the methanol methyltransferase.

Amongst sulphate-reducing bacteria, species of *Desulfosporosinus* (Klemps *et al.*, 1985), *Desulfobacterium* (Szewzyk and Pfennig, 1987; Schnell *et al.*, 1989), *Desulfotomaculum* (Fardeau *et al.*, 1995; Liu *et al.*, 1997; Tebo and Obraztsova, 1998; Goorissen *et al.*, 2003; Balk *et al.*, 2007), and *Desulfovibrio* (Nanninga and Gottschal, 1987; Qatibi *et al.*, 1991) have all been shown to utilise methanol as an electron donor for sulphate reduction. However one unique sulphate-reducing bacterium, *Desulfotomaculum kuznetsovii*, in fact encodes enzymes for both methanol dehydrogenases and methanol methyltransferases (Sousa *et al.*, 2018).

1.5 Links between volatile organic compound cycling and the climate

Although seemingly disparate substrates, the microbial metabolism of choline, glycine betaine, DCM, and methanol are all closely linked to carbon and nitrogen cycling, with potential impacts for global climate. The latter two compounds have a direct effect on atmospheric chemistry, while the former two have an indirect effect via their close links to methylated amines and methane. DCM, methanol, methane, and methylated amines are all volatile organic compounds with atmospheric impacts.

In anaerobic ecosystems, both choline and glycine betaine are readily transformed into trimethylamine, which is then utilised almost exclusively by methanogens (King, 1984, 1988a). Trimethylamine is thought to be responsible for the bulk of methane production in intertidal mudflats and saltmarshes (Oremland *et al.*, 1982; King *et al.*, 1983). Within these environments, sulphate-reducing bacteria typically outcompete acetoclastic and hydrogenotrophic methanogens for these substrates

because of the higher energy yield from sulphate reduction compared to methanogenesis (Fenchel and Blackburn, 1979; Morris and Whiting, 1986). Production of trimethylamine therefore has a significant impact on methane flux in such environments because it is a non-competitive substrate. More recently, three strains of *Methanococcoides* have also been reported that can use glycine betaine directly as a substrate for methanogenesis (Watkins *et al.*, 2014). Methylated amine particles in the atmosphere have also been implicated in global climate via their enhancement of aerosol particle nucleation (Almeida *et al.*, 2013; Yao *et al.*, 2018), which influences cloud seeding (Almeida *et al.*, 2013; Intergovernmental Panel on Climate Change, 2013). Thus, microbial metabolism of methylated amines and glycines both directly and indirectly affects the cycling of climate-active trace gasses.

DCM has recently been recognised as a potent greenhouse gas with an increasing negative impact on stratospheric ozone levels (Hossaini *et al.*, 2017). It is a so-called very short-lived substance (VSLS), having a lifetime of around five months (Montzka *et al.*, 2011). Over 80% of chlorinated VSLSs such as DCM are predicted to reach the troposphere (Carpenter and Reimann, 2015), which can affect atmospheric chemistry on a regional and global scale. Halogenated VSLSs in particular have a disproportionately large effect on radiative forcing and climate because their breakdown into reactive chlorine species leads to ozone depletion at lower, climate-sensitive altitudes (Hossaini *et al.*, 2015a).

VSLSs were overlooked in the Montreal Protocol, as they were thought to play only a minor role in ozone depletion due to their relatively short atmospheric lifetime. However, industrial activity, particularly from developing countries, has caused tropospheric measurements of DCM to increase over the last decades (Carpenter and Reimann, 2015; Hossaini *et al.*, 2015a, 2015b; Leedham Elvidge *et al.*, 2015), leading researchers to reconsider the effect that this chemical may have on the ozone layer if environmental releases continue at this rate. If left unchecked, increasing DCM emissions may delay the return of Antarctic ozone to pre-1980 levels by up to a decade (Hossaini *et al.*, 2017). The relative importance of DCM and other VSLSs also continues to increase as the observed and projected decreases in longer-lived anthropogenic chlorocarbons (e.g. those banned by the Montreal Protocol) reduce their effect on the atmosphere.

Finally, methanol is a ubiquitous background volatile organic carbon compound and the second-most abundant organic gas in the atmosphere after methane (Heikes *et al.*, 2002). Methanol is a significant sink for hydroxy radicals in the atmosphere (Singh *et al.*, 1995), a reaction that produces formaldehyde (Millet *et al.*, 2006) and carbon monoxide (Duncan *et al.*, 2007). Oxygenated hydrocarbon species such as methanol can also influence atmospheric ozone formation through reactions with nitrous oxides (Finlayson-Pitts and Pitts, 2000). Thus, microbial metabolism of methanol, DCM, and quaternary amines reported in this thesis has wider implications for the climate.

1.6 Research aims and chapter summary

Despite the prevalence of DCM-contaminated sites worldwide, many of which are anaerobic, relatively little is known about anaerobic microbial metabolism of DCM. The overarching aim of this work was therefore to characterise a novel, DCM-fermenting organism (strain DCMF) enriched from a contaminated aquifer near Botany Bay, Sydney, Australia (Holland *et al.*, 2019). Strain DCMF exists within a DCM-fermenting enrichment (DFE) culture, where it is thought to be the only species capable of directly metabolising DCM. DFE is a non-methanogenic culture that was enriched from the previously reported DCM-fermenting culture DCMD (Lee *et al.*, 2012; Holland *et al.*, 2019). Investigation of the role of strain DCMF, potential mechanisms of DCM and wider substrate metabolism, and the role of the non-dechlorinating DFE cohabitants is organised into four chapters, the aims of which are summarised below:

1. Whole genome sequencing of the novel, DCM-fermenting bacterium (strain DCMF) and genomic comparison with the two other anaerobic DCM-metabolising bacteria;
2. Characterisation of DCM, quaternary amine and methanol metabolism in strain DCMF;
3. Comparative proteomic analysis of strain DCMF cells grown on DCM, glycine betaine, choline and methanol, with particular focus on highly abundant proteins in DCM-grown cells that may be implicated in DCM dechlorination;
4. Exploration of the role of the wider bacterial community in the DCM-fermenting enrichment culture (DFE).

Although anaerobic DCM-degrading bacteria are known to utilise the Wood-Ljungdahl pathway for DCM transformation, the enzyme(s) catalysing the initial dechlorination step remain unknown (Mägli *et al.*, 1998; Kleindienst *et al.*, 2019; Chen *et al.*, 2020). In order to investigate possible pathways for DCM transformation by strain DCMF, we determined that a detailed and accurate genome annotation was necessary. Chapter 2 reports the assembly of a complete, circular genome for strain DCMF was achieved via long-read PacBio sequencing technology. Extensive manual curation of the genome was carried out to ascertain the presence of potential metabolic pathways. The genome of strain DCMF was then compared to the

available genomes of two other anaerobic DCM-degrading bacteria (*D. formicoaceticum* and '*Ca.* Dichloromethanomonas elyunquensis'), to assess areas of commonality and difference between them. The core genome of the three bacteria was relatively small, with the Wood-Ljungdahl pathway and an F_1F_0-type ATP synthase being the key features shared between all three. Strain DCMF appears far closer to *D. formicoaceticum* in terms of shared genomic traits.

Stable isotope work using ^{13}C-labelled DCM and bicarbonate in Chapter 3 confirmed that DCM is metabolised via the Wood-Ljungdahl pathway in strain DCMF. This chapter then investigated the ability of strain DCMF to metabolise substrates other than DCM, a capability that was suggested by the presence of a wide range of MttB superfamily methyltransferases and glycine/betaine/sarcosine reductases in the genome. It reports the ability of strain DCMF to utilise the quaternary amines choline and glycine betaine for growth, as well as the one-carbon compound methanol. Genome-based metabolic models for growth on these compounds are suggested.

The revelation that strain DCMF could also metabolise quaternary amines and methanol provided the opportunity to perform comparative proteomics analysis, which is reported in Chapter 4. This work revealed a putative DCM methyltransferase gene cluster that was significantly more abundant in cells grown with DCM than glycine betaine and highly conserved amongst anaerobic DCM-degrading bacteria.

Finally, the role of the wider bacterial community in the DFE culture is investigated in Chapter 5. Illumina 16S rRNA amplicon sequencing showed how the relative abundance of the cohabitants shifts over batch cultivation cycles and metaproteogenomics revealed the presence of genes and proteins for metabolism of carbohydrates, sugars, and amino acids. We therefore propose that the cohabitants persist in the DFE culture via necromass recycling, i.e. metabolism of components of expired strain DCMF cells.

This work paves the way for identification of a DCM-dechlorinating enzyme active under anoxic conditions, which would be of interest to the bioremediation industry for assessing and monitoring bioremediation potential and/or efficacy at contaminated sites. The establishment of a new taxonomic group involved in DCM,

quaternary amine and methanol transformation also has implications for the flux of climate-active trace gasses from coastal subsurface environments. Exploration of the role of the non-dechlorinating community in an anaerobic DCM-metabolising culture also demonstrates the importance of a keystone species (strain DCMF) as the sole, primary substrate degrader, on which all other organisms are dependent (as they persist via oxidation of necromass components). This has implications for understanding community dynamics and carbon and nitrogen cycling in contaminated groundwater sites and beyond.

2 Whole genome sequencing of a novel *Peptococcaceae* bacterium and genomic comparison of anaerobic DCM-degrading bacteria

2.1 Introduction

Bacteria capable of dechlorinating the toxic environmental contaminant dichloromethane (DCM; CH_2Cl_2) are of great interest for potential bioremediation applications. Anerobic DCM transformation is of particular importance for remediating contaminated sites, as DCM is denser than water and thus migrates downwards in groundwater to anoxic zones. To date, only two bacteria capable of anaerobically metabolising DCM have been characterised and genome sequenced – *Dehalobacterium formicoaceticum* strain DMC (Mägli *et al.*, 1996, 1998; Chen *et al.*, 2017b) and '*Candidatus* Dichloromethanomonas elyunquensis' strain RM (Kleindienst *et al.*, 2016, 2017). Although the former ferments DCM to formate and acetate, while the latter completely mineralises it to H_2 and CO_2, both organisms utilise variations of the Wood-Ljungdahl pathway for the transformation (Mägli *et al.*, 1996, 1998; Chen *et al.*, 2020). However, the enzyme(s) responsible for the initial dechlorination step are unknown.

This chapter reports the whole genome sequencing and assembly of a novel, DCM-fermenting bacterium, strain DCMF. The bacterium is the dominant organism in a non-methanogenic, DCM-fermenting enrichment culture (DFE). In order to investigate possible pathways for DCM transformation within strain DCMF, it was determined that a detailed and accurate genome annotation was necessary. As it can be difficult to assemble a high-quality genome from a mixed culture and no closely related reference genome was available, a thorough genome sequencing and assembly strategy was sought to overcome these challenges. The genome was then compared with strain DMC and strain RM, with the aim of identifying areas of similarity that might provide clues to the DCM dechlorination mechanism employed by these bacteria. We also sought any evidence of horizontal gene transfer between the three organisms. In light of the findings, suggestions are made regarding the evolution and ecological niche of each anaerobic DCM-degrading bacterium.

2.2 Materials and Methods

2.2.1 Culture medium

DFE cultures were grown in anaerobic minimal mineral salts medium that comprised (g l^{-1}): $CaCl_2.2H_2O$ (0.1), KCl (0.1), $MgCl_2.6H_2O$ (0.1), $NaHCO_3$ (2.5), NH_4Cl (1.5), NaH_2PO_4 (0.6), 1 ml of trace element solution A (1000×), 1 ml of trace element solution B (1000×), 1 ml of vitamin solution (1000×), 10 ml of 5 g l^{-1} fermented yeast extract (100×), and resazurin 0.25 mg l^{-1}. Trace element solutions A and B were prepared as described previously (Wolin *et al.*, 1963), as was the vitamin solution (Adrian *et al.*, 1998). Medium was sparged with N_2 during preparation and the pH was adjusted to 6.8 – 7.0 by a final purge with N_2/CO_2 (4:1). Aliquots were dispensed into glass serum bottles that were crimp sealed with Teflon faced rubber septa (13 mm diameter, Wheaton) before the medium was chemically reduced with sodium sulphide (0.2 mM). DCM (1 mM) was supplied as the sole electron source via a glass syringe. All cultures were incubated statically at 30˚C in the dark.

2.2.2 Preparation of spent media as a co-factor solution

A stock fermented yeast extract solution was prepared by inoculating anoxic yeast extract (5 g l^{-1}) in defined minimal mineral salts medium (described above, excluding DCM) with the DFE culture. The culture was incubated for one week at 30˚C before being filter sterilised. The filtered, spent media was re-inoculated with DFE and incubated for a further week, to ensure that growth was no longer possible on the fermented yeast extract (i.e. that it had been energetically exhausted). The spent media was then filter-sterilised again before use.

2.2.3 DNA extraction

Genomic DNA was extracted as previously described (Urakawa *et al.*, 2010). Briefly, cells were lysed with lysis buffer and bead-beating, before DNA was extracted with phenol-chloroform-isoamyl, precipitated using isopropanol, and resuspended in molecular grade water. The nucleic acid concentration was quantified using a Qubit instrument and assay as per the manufacturer's instructions (Life Technologies).

2.2.4 Illumina genome sequencing

DNA was prepared with the Nextera XT library prep kit (Illumina). Sequencing was carried out on an Illumina MiSeq with a v2 500-cycle kit (2 × 250 bp run) at the

Ramaciotti Centre for Genomics (UNSW Sydney, Australia). Three MS110-2 libraries were used for the run. Library size ranged from 200 - 3000 bp, with an average of 955 bp. Raw reads were trimmed and filtered with SolexaQA (DynamicTrim.pl and LengthSort.pl) (Cox *et al.*, 2010) and then submitted to the NCBI Sequence Read Archive with the identifier SRR5179547.

2.2.5 Pacific Biosciences SMRT sequencing

A MagAttract HMW DNA kit (Qiagen) was used to extract high-molecular weight genomic DNA, followed by purification using AMPure PB beads (Beckman Coulter). DNA concentration and purity were checked by Qubit and NanoDrop instruments, respectively. A 0.75% Pippin Pulse gel (Sage Science) was performed by the Ramaciotti Centre for Genomics (UNSW Sydney, Australia) to further verify integrity. A SMRTbell library was prepared with the PacBio 20 kb template protocol excluding shearing (Pacific BioSciences). Additional damage repair was carried out following minimum 4 kb size selection using Sage Science BluePippin.

Whole genome sequencing was performed on the PacBio RS II (Pacific Biosciences), employing P6 C4 chemistry with 240 min movie lengths. DNA was initially sequenced using two Single Molecule Real Time™ (SMRT) cells. A third SMRT™ cell was added to compensate for low quality data from the first two, due to degraded DNA yield from the sample. The SMRTbell library for this cell was prepared with the PacBio 10 kb template protocol, without size selection, and a lower input (3,624 ng) of DNA was used. In total, the three SMRT cells yielded 463,878 subreads from 169,180 ZMW, with a combined length of 1,712,588,985 bp. Reads were submitted to the NCBI Sequence Read Archive with the identifier SRR5179548.

2.2.6 Genome assembly and annotation

An overview of the genome sequencing, assembly, and annotation pipeline is provided in Figure 2.1. PacBio subreads were assembled using HGAP3 (Chin *et al.*, 2013) as implemented in SMRT Portal. In-house software, SMRTSCAPE (SMRT Subread Coverage & Assembly Parameter Estimator; http://rest.slimsuite.unsw.edu.au/smrtscape) was used to predict optimal HGAP settings for several different assemblies with different predicted genome size and minimum correction depths (Table S1). The assembly with the greatest depth of coverage used for seed read error correction that still yielded a full-length (6.44 Mb)

intact chromosome was selected for the draft genome. This corresponded to: minimum read length 4,010 bp; minimum seed read length 8,003 bp; minimum read quality 0.86; minimum 10× correction coverage. The genome was corrected with Quiver (Chin *et al.*, 2013) using all subreads and circularised by identifying and trimming overlapping ends, then annotated in-house using Prokka (Seemann, 2014).

Based on draft annotation, the genome was re-circularised to have its break-point in the intergenic region between the 3' of two hypothetical genes. To ensure that the ends were jointed correctly, the re-circularised genome was subjected to a second round of Quiver correction to make sure the manually joined region was of high quality, and every base was covered by long reads spanning at least 5 kb 5' and 3' (Figure S1 B). Filtered Illumina reads were mapped onto the Quiver-corrected genome using BWA-MEM v0.7.9a (Li, 2013) and possible errors were identified with Pilon (Walker *et al.*, 2014). Manual curation was then performed to check any discrepancies between the PacBio and Illumina data and correct small indels. Raw PacBio reads were mapped onto the completed genome with BLASR (Chaisson and Tesler, 2012). The corrected genome was re-annotated with Prokka and uploaded to the Integrated Microbial Genomes and Microbiomes (IMG/M) system of the Joint Genome Institute (JGI) for independent annotation (Chen *et al.*, 2019).

Twenty-nine pairs of fragmented genes and four truncated genes were subject to additional manual curation and correction where a pyrrolysine or selenocysteine residue had been erroneously translated as a stop codon (Table S2). The IMG annotation was publicly updated to reflect these manual annotations, and this annotation was used for all genomic analyses. The genome has subsequently been re-annotated by NCBI.

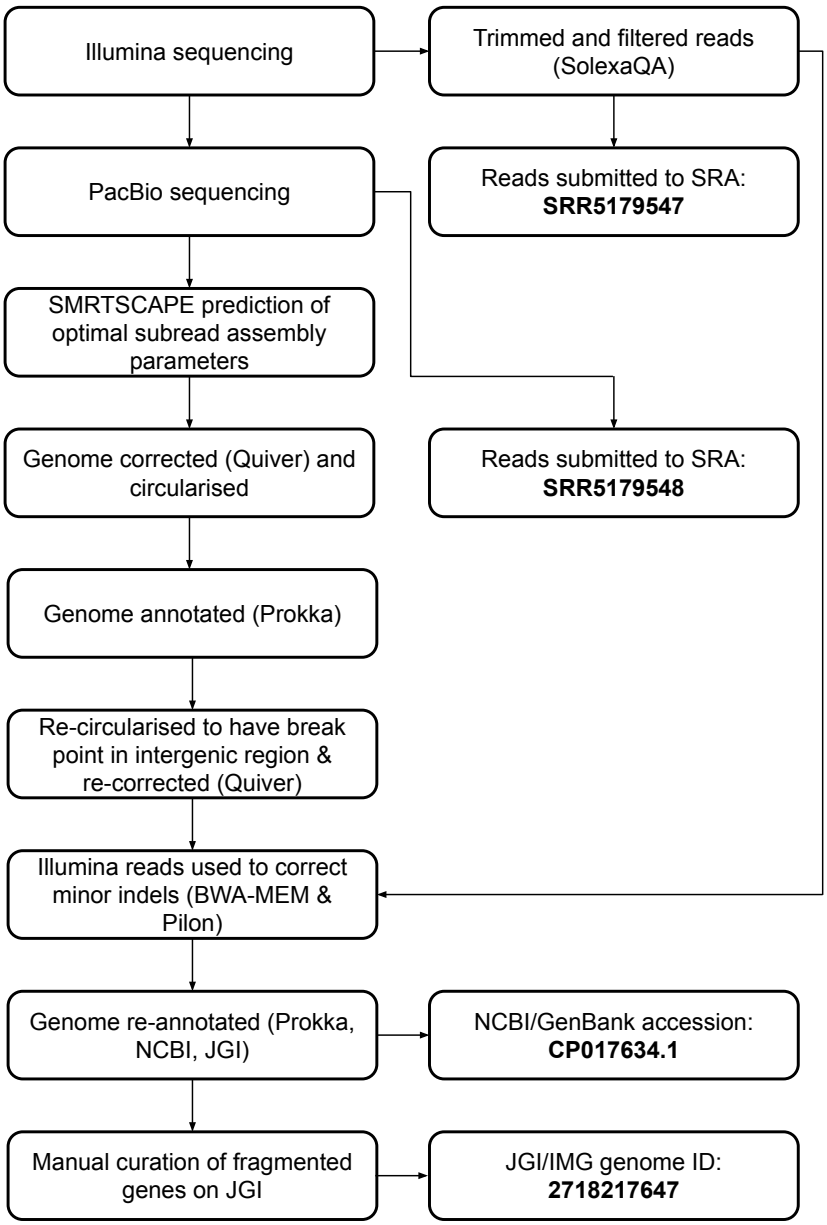


Figure 2.1 Research pipeline for strain DCMF genome sequencing, assembly, and annotation.

2.2.7 16S rRNA gene identification and phylogeny

The strain DCMF 16S rRNA gene consensus sequence was searched against the NCBI prokaryotic 16S rRNA BLAST database as well as the 16S rRNA gene sequences of the two other known anaerobic DCM-degrading bacteria (absent from that database), strain DMC (NCBI locus tags CEQ75_RS05455, CEQ75_RS05490, CEQ75_RS13675, CEQ75_RS13970, CEQ75_RS17045) and strain RM (KU341776.1). The closest phylogenetic relatives and an outgroup, *Moorella perchloratireducens* strain An10 (NR_125518.1), were aligned with MAFFT program v.7 (Kuraku *et al.*,

2013) and a neighbour-joining tree constructed with 1000 bootstraps resampling a 200PAM/k = 2 scoring matrix using 1,365 nucleotides. This was performed using Archaeopteryx (Han and Zmasek, 2009), as well as manual curation. In addition, strain DCMF 16S rRNA gene sequences were mapped to taxa using the SILVA Alignment, Classification and Tree Service (Pruesse *et al.*, 2012) with default values.

2.2.8 Strain DCMF genomic analysis

CheckM (Parks *et al.*, 2015) was used to assess the completeness and contamination in the strain DCMF genome. Whole genome taxonomic analysis was carried out with the GTDB-Tk (Genome Taxonomy Database toolkit) (Chaumeil *et al.*, 2019). SPADE (Mori *et al.*, 2019) was used to analyse repeat regions in the genome, using default parameters.

The 81 full-length predicted trimethylamine methyltransferase protein sequences were aligned with MAFFT v7.310 (Katoh *et al.*, 2002) and a Maximum-Likelihood tree (1000 bootstraps) inferred by IQTree v1.6.1 using ModelFinder (Nguyen *et al.*, 2015; Kalyaanamoorthy *et al.*, 2017). Global pairwise percentage identities were calculated using GABLAM v2.28.2 (Davey *et al.*, 2006) from an all-by-all BLAST 2.5.0+ blastp search (Camacho *et al.*, 2009).

Putative selenocysteine-containing proteins were verified via multiple lines of evidence. The presence of a selenocysteine insertion sequence (SECIS) was confirmed in either the IMG annotation or via bSECISearch (Zhang and Gladyshev, 2005). Glycine/betaine/sarcosine reductase genes were checked for the presence of the conserved cysteine(s) present either before (CxxU in GrdA) (Kreimer and Andreesen, 1995) or after (UxxCxxC in GrdBFH) the selenocysteine residue (Wagner *et al.*, 1999).

In order to determine the substrate specificity of predicted glycine/betaine/sarcosine reductases, the amino acid sequences encoding the two subunits of B component (GrdB/F/H and GrdE/G/I) were aligned with those from B components of known substrate specificity from *Clostridium sticklandii*, *Peptoclostridium acidaminophilum*, *Peptoclostridium litorale*, and *Sporomusa ovata* An4 with MUSCLE in UGENE v1.32 (Okonechnikov *et al.*, 2012). An unrooted Maximum Likelihood tree (1000 bootstraps) was inferred by IQ-Tree v1.6.1 using ModelFinder (Nguyen *et al.*, 2015; Kalyaanamoorthy *et al.*, 2017) and visualized in

iTOL v4.4.2 (Letunic and Bork, 2019), with the clustering of the proteins used to infer the substrate specificity of those from strain DCMF.

2.2.9 Genomic comparison of anaerobic dichloromethane-degrading bacteria

The genomes of strain DCMF (IMG 2718217647), strain DMC (GenBank CP022121.1) and strain RM (GenBank LNDB00000000.1) were uploaded to the MicroScope platform (Médigue *et al.*, 2019) for independent annotation for the purpose of comparative genomic analysis. The following analyses were all performed automatically by the platform. Protein localisation was predicted with PSORTb v3.0.2 (Yu *et al.*, 2010). Genome completeness and contamination were assessed with CheckM (Parks *et al.*, 2015). Putative CRISPR (clustered regularly interspaced short palindromic repeats) arrays were identified with CRISPRCasFinder (Couvin *et al.*, 2018) and Cas (CRISPR-associated) genes were identified with MacSyFinder (Abby *et al.*, 2014).

Genome-wide synteny statistics were calculated with the MicroScope 'PkGDB Synteny Statistics' tool. Orthologous genes were defined by MicroScope as those gene couples satisfying either: bi-directional best hit (BBH) status, or a blastp alignment threshold of minimum 35% sequence identity along 80% of the length of the smaller protein. Clusters of adjacent orthologous genes were defined as syntons, and all possible chromosomal rearrangements (e.g. inversions, indels) were permitted within a synton. The gap parameter (i.e. the maximum number of consecutive genes not involved in a synton) was set to five genes.

Functional classification of genes was implemented by eggNOG-mapper v1.0.3 (Huerta-Cepas *et al.*, 2017) using the eggNOG database v4.5.1 (Huerta-Cepas *et al.*, 2016). In order to normalise results to the differing genome sizes, each class was reported as a percentage of the individual organism's total genome. The average genome percentage of each eggNOG class was also calculated (n = 3), and values more than one standard deviation from the class mean were assigned as significantly different.

The core and pan genomes were analysed using the 'Pan/Core-Genome' tool, based on MICFAMS (MicroScope gene families) computed with an algorithm implemented in SiLiX software (Miele *et al.*, 2011). This clusters homologous genes utilising the

"The friends of my friends are my friends" principle – if two homologous genes are clustered and one of them is already clustered with another gene, then all three will be clustered into the same MICFAM. A MICFAM is considered part of the core genome if it contains as least one gene from every compared genome.

Genes of interest from each genome were listed by keyword searching for "methyltransferase" or "transposase", respectively, in any of the fields: Gene annotations, COG, EGGNOG, FigFam results, TIGRFams, or InterPro. Results were manually curated; for methyltransferases, only those putatively involved in methylamine or unknown substrate metabolism were included in the final list (e.g. RNA and DNA methyltransferases were excluded).

For analyses outside of the MicroScope platform, the GenoScope annotation of each genome was still used for consistency. The average nucleotide identity (ANI) tool from the Kostas lab (Rodriguez-R and Konstantinidis, 2014) was used to calculate ANI values between strain DCMF and strain DMC. CompareM (https://github.com/dparks1134/CompareM) was used to calculate the two-way average amino acid identity (AAI) between the anaerobic DCM degraders and other related bacteria in the family *Peptococcaceae*. PHASTER (Arndt *et al.*, 2016) and Prophage Hunter (Song *et al.*, 2019) were used to identify putative prophage regions in the genomes. Genomic islands were predicted by IslandViewer 4 (Bertelli *et al.*, 2017).

2.3 Results

2.3.1 Genome assembly and annotation

Attempts were initially made to sequence the dominant, DCM-degrading organism using Illumina short read technology, which yielded 5,040,903 filtered read pairs for a total of 1,827,383,271 bp. However, the presence of the additional organisms in the DFE culture and lack of a reference genome hindered this approach. A PacBio long read strategy was subsequently used to assemble a full-length gap-free circular genome for strain DCMF. Trimmed and filtered Illumina reads (average 242× coverage) were used for final, minor error correction. The final genome assembly had an average of 132× PacBio coverage (min >50×) and no regions of unusual read depth (Figure S1 A). The genome was circularised at overlapping ends and every base was covered by long reads spanning at least 5 kb 5' and 3' (Figure S1 B). In addition to these assessments, CheckM evaluated the genome as 98.98% complete with a contamination rate of 2.73%.

The strain DCMF genome is 6,441,270 bp long and has a G+C content of 46.44% (IMG/JGI genome ID 2718217647; GenBank accession CP017634.1). IMG annotation initially revealed 5,801 predicted protein-coding genes. Manual curation of the 29 pairs of genes fragmented by the presence of the amino acids pyrrolysine and selenocysteine (encoded by in-frame UAG and UGA stop codons, respectively; Table S2) brought this total down to 5,772 protein coding genes.

2.3.2 16S rRNA gene and whole genome phylogeny

The strain DCMF genome contains four full-length 16S rRNA genes (IMG locus tags Ga0180325_11664, 11677, 113771, 114507), which share 99.87% identity when aligned. Based on the consensus 16S rRNA gene sequence, the closest relative to strain DCMF was *Dehalobacterium formicoaceticum* strain DMC (93.62% identity), suggesting that strain DCMF may be the first cultured representative of a novel genus. This was closely followed by *Dehalobacter restrictus* strain PER-K23 (88.89% identity), *Desulfosporosinus acidiphilus* strain SJ4 (88.81% identity) and '*Ca.* Dichloromethanomonas elyunquensis' strain RM (88.36%; Figure 2.2). The lowest taxonomic rank of SILVA classification was the family *Peptococcaceae*. Strain DCMF shared an even higher (94.58%) percentage nucleotide sequence identity with a

recently discovered anaerobic, glycine betaine-degrading organism, '*Candidatus* Betaina sedimentti' (Jones *et al.*, 2019). However, as the only 16S rRNA gene sequence available for this organism was a 240-bp fragment from Illumina amplicon sequencing (accession number MK313791), rather than a full-length gene, this organism was not included in the 16S rRNA phylogenetic tree (Figure 2.2).

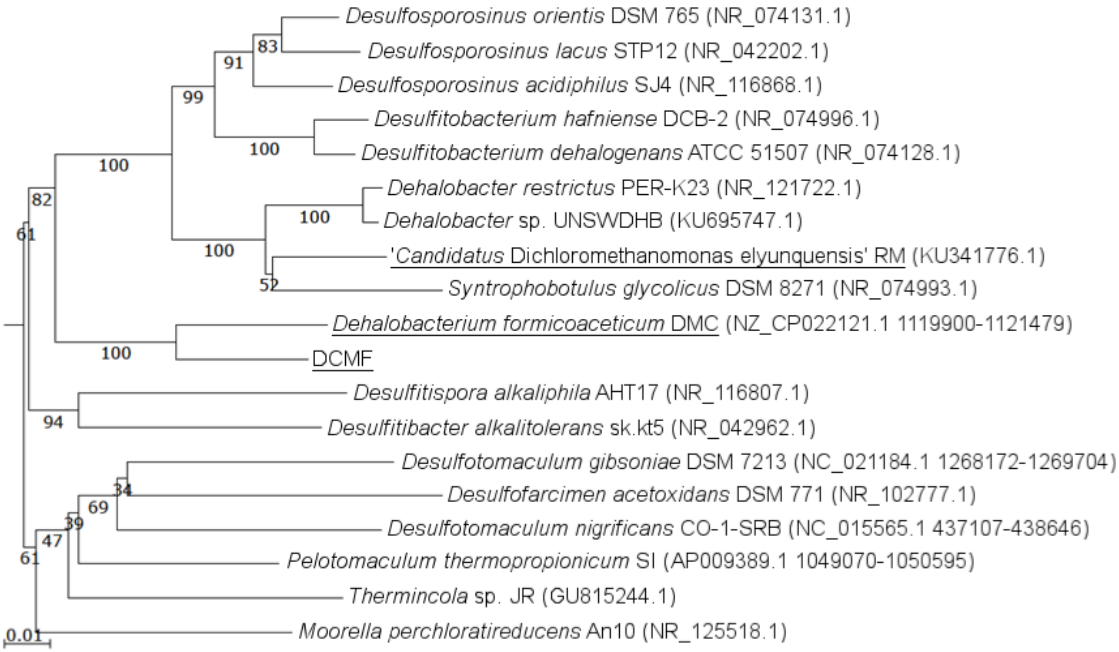


Figure 2.2 16S rRNA gene phylogenetic tree of strain DCMF with closely related bacteria (94-87% identity). The three known DCM-fermenting bacteria are underlined. GenBank accession numbers are provided in parentheses. Numbers indicate percentage of branch support from 1000 bootstraps. The scale bar indicates an evolutionary distance of 0.01 amino acid substitutions per site. Sequences were aligned in MAFFT program v.7 and a neighbour-joining tree (1000 bootstraps) resampling a 200PAM/k =2 scoring matrix was inferred using Archaeopteryx, with manual curation.

Whole genome taxonomic analysis of strain DCMF with the GTDB-Tk also identified its closest relative as strain DMC. However, it placed the organism within the novel family taxon *Dehalobacteriaceae* (order *Dehalobacteriales*, class *Dehalobacteriia*, phylum *Firmicutes*). This is a result of the GTDB re-classifying a wide range of bacterial taxa based on its analysis pipeline, including splitting the traditional class of *Clostridia* (which includes the family *Peptococcaceae*) into a variety of more specific, monophyletic classes (Parks *et al.*, 2018). In essence, the GTDB-Tk result confirms the taxonomic placement indicated by 16S rRNA gene analysis, but with

taxonomic names specific to its platform. A comparison of NCBI and GTDB taxonomy for the family *Peptococcaceae* can be found at: https://gtdb.ecogenomic.org/searches?q=%25peptococcaceae%25&s=al.

2.3.3 Genomic features of strain DCMF

Several pertinent metabolic pathways were identified in the strain DCMF genome, including a full set of genes for the Wood-Ljungdahl pathway (Dataset S1). No reductive dehalogenases were identified in the genome by any of the three independent annotation pipelines. The genome also contains an abundance of methylamine methyltransferase genes (96 in total), including 82 copies of methyltransferases in the MttB superfamily (Dataset S1). The MttB superfamily (InterPro entry IPR038601) contains trimethylamine and quaternary amine methyltransferases, as well as many of uncharacterised substrate specificity. There is a high diversity amongst the predicted MttB superfamily proteins, with an average amino acid sequence identity of only 30.3%. A number of genes annotated as di- and monomethylamine methyltransferases are also encoded in the genome.

The diversity of MttB superfamily protein sequences and scarcity of characterised representatives makes it difficult to predict the substrate range of all putative MttB superfamily proteins encoded in the strain DCMF genome. However, the top five highest percentage identity homologs to glycine betaine methyltransferase (MtgB) genes from *Acetobacterium woodii*, *Desulfitobacterium hafniense* Y51, *Sporomusa ovata* An4 and *S. ovata* H1 and the proline betaine methyltransferase (MtpB) from *Eubacterium limosum* ATCC 8486 formed a distinct clade in the phylogenetic tree (Figure 2.3).

Associated with the presence of these methyltransferase genes are all five genes necessary to synthesise and utilise pyrrolysine (Dataset S1), a non-canonical amino acid residue present in 23 of the 96 total methylamine methyltransferases in the genome. In a maximum likelihood phylogenetic tree constructed from the 81 full-length predicted MttB superfamily proteins (one truncated copy was omitted), the pyrrolysine-containing copies tend to cluster together at the bottom of the tree (Figure 2.3). The pyrrolysine gene cluster in strain DCMF includes the dedicated tRNA (*pylT*), tRNA synthetase (*pylSc and pylSn*), and associated biosynthetic enzymes (*pylBCD*).

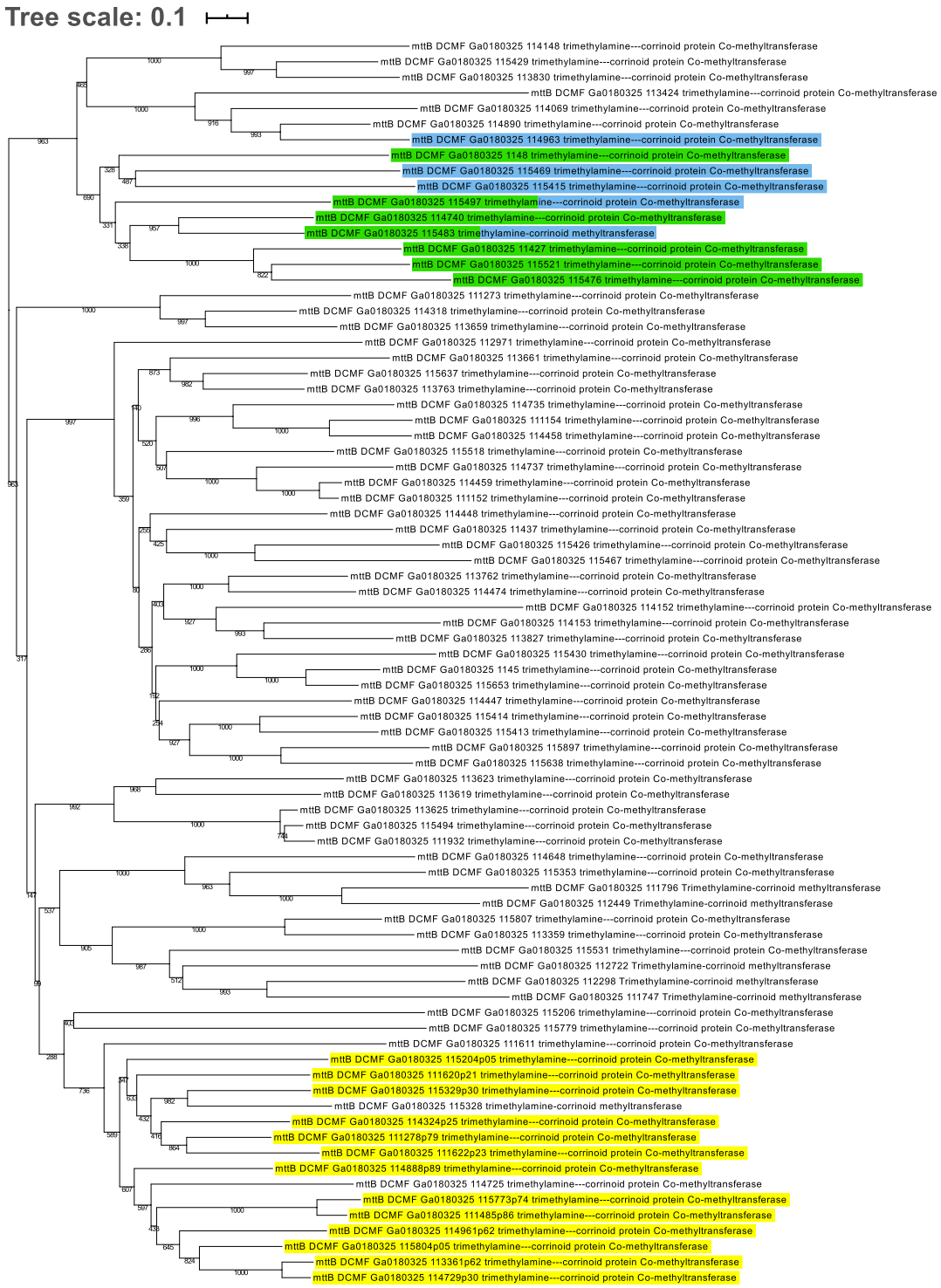


Figure 2.3 Phylogenetic tree of all predicted MttB superfamily methyltransferases in strain DCMF. Amino acid sequences were aligned with MAFFT and a Maximum Likelihood tree computed in IQ-Tree. The top five highest percentage amino acid identity homologs to known glycine betaine methyltransferase (MtgB) proteins from *Acetobacterium woodii*, *Desulfitobacterium hafniense* Y51, *Sporomusa ovata* An4 and *S. ovata* H1 are highlighted in green, and those to the proline betaine methyltransferase (MtpB) from *Eubacterium limosum* ATCC

8486 in blue. The 23 pyrrolysine-containing MttB proteins are highlighted in yellow and cluster together at the bottom of the tree.

The presence of all genes required for *de novo* corrinoid biosynthesis (Dataset S1) is pertinent both to certain Wood-Ljungdahl pathway proteins and the MttB superfamily methyltransferases, which require a corrinoid cofactor to function (Burke and Krzycki, 1997; Ferguson *et al.*, 2000). However the genes for methionine synthesis (*metH* and *metE*), required to form S-adenosylmethionine, which is in turn used as a methyl donor during corrin ring formation (Deeg *et al.*, 1977), were not identified in the genome. Strain DCMF may be using a novel route for *de novo* biosynthesis of this amino acid.

Additionally, five clusters of glycine/sarcosine/betaine reductase complex genes were found (Dataset S1). Four of these clusters include the thioredoxin reductase (*trxB*) and thioredoxin I (*trxA*) necessary for electron transfer to the substrate reductase, and the genome also encodes transporter genes necessary to import these compounds into the cell (Dataset S1). Components A (*grdA*) and B (*grdBE/FG/HI*) of the glycine/sarcosine/betaine reductase complex contain an integral selenocysteine residue, as does the formate dehydrogenase (Ga0180325_112876, 112877, 112878s80) encoded in the strain DCMF genome. Accordingly, the genome harbours the full complement of genes necessary for biosynthesis and incorporation of the unusual amino acid selenocysteine (*selABD, serS*; Dataset S1). All predicted selenocysteine-containing proteins also contain the SECIS downstream of the UGA stop codon, necessary for translating it as a selenocysteine residue instead.

2.3.4 Genomic comparison of anaerobic DCM-degrading bacteria

A genomic comparison was carried out between the three anaerobic DCM-degrading bacteria for which genome sequences are available: strain DCMF, *D. formicoaceticum* strain DMC and '*Ca.* Dichloromethanomonas elyunquensis' strain RM. Strain DCMF has the largest genome of the three, at 6.44 Mb, followed by strain DMC at 3.77 Mb and then strain RM at only 2.08 Mb (Table 2.1). Strain DMC (43.5%) and strain RM (43.2%) share a similar G+C content, while strain DCMF is slightly higher (46.4%; Table 2.1).

Table 2.1 Genome characteristics for the three anaerobic DCM-degrading bacteria.

	Strain DCMF	Strain DMC	Strain RM
GenBank Accession	CP017634.1	CP022121.1	LNDB00000000.1
IMG Taxon ID	2718217647	2811995020	nd[a]
Genome size (bp)	6,441,270	3,766,545	2,076,422
G+C content (%)	46.4	43.2	43.5
Contigs	1	1	53
Total genes	5,885	nd	2,395
Protein coding genes	5,772	3,935	2,323
Total rRNA genes	12	17	1
16S rRNA genes	4	5	1
Total tRNA genes	59	55	45
Completeness (%)[b]	98.98	95.63	92.47
Contamination (%)[b]	2.73	2.55	4.44
Reference	This study and (Holland *et al.*, 2019)	(Chen *et al.*, 2017b)	(Kleindienst *et al.*, 2016)

[a]not described
[b]As determined by CheckM.

Strain DCMF had 77.19% average nucleotide identity (ANI) to its closest relative, strain DMC. Given that ANI offers robust resolution primarily above 80% values (Rodriguez-R and Konstantinidis, 2014), average amino acid (AAI) analysis was instead carried out to evaluate genomic distance between the anaerobic DCM degraders and other members of the family *Peptococcaceae* (Table 2.2). Strain DCMF and strain DMC were confirmed as each other's closest relative (two-way AAI value 66.54%). However contrary to the 16S rRNA gene phylogenetic analysis, AAI placed strain RM further away from both strain DCMF (53.10%) and strain DMC (53.09%). Instead, *Thermincola potens* JR was the next closest relative to strain DCMF (55.21% AAI), while the taxon with the highest AAI to strain RM was *Dehalobacter* sp. CF (70.04%; Table 2.2), which was expected based on the analysis reported by Kleindienst *et al* (2017).

Table 2.2 Average amino acid (AAI) identity table of strain DCMF, strain DMC, strain RM and other related bacteria from the *Peptococcaceae* family. Bacteria are listed in order of highest to lowest AAI to strain DCMF and only the species with the highest AAI value was taken from each genus. Cells are colour coded from lowest (red) to highest (green) AAI.

	Strain DCMF	*Dehalobacterium formicoaceticum* DMC	*Thermincola potens* JR	*Desulfosporosinus orientis* DSM 765	*Desulfotomaculum nigrificans* CO-1-SRB	*Pelotomaculum thermopropionicum* SI	*Desulfofarcimen acetoxidans* DSM 771	*Desulfitobacterium hafniense* DSB-2	*Dehalobacter* sp. CF	*Candidatus* Dichloromethanomonas elyunquensis' RM	*Desulfitibacterium metallireducens* DSM 15288	*Syntrophobotulus glycolicus* DSM 8271
Strain DCMF	100											
Dehalobacterium formicoaceticum DMC	66.54	100										
Thermincola potens JR	55.21	54.69	100									
Desulfosporosinus orientis DSM 765	54.05	53.81	53.56	100								
Desulfotomaculm nigrificans CO-1-SRB	54.04	54.84	57.78	53.39	100							
Pelotomaculum thermopropionicum SI	54.02	53.30	58.34	52.03	61.06	100						
Desulfofarcimen acetoxidans DSM 771	53.99	53.29	56.33	53.08	59.45	60.12	100					
Desulfitobacterium hafniense DCB-2	53.41	54.54	53.19	63.60	53.21	52.09	52.36	100				
Dehalobacter sp. CF	53.28	56.05	52.50	58.23	52.01	51.46	52.58	58.52	100			
'Candidatus Dichloromethanomonas elyunquensis' RM	53.10	53.09	53.24	59.49	52.46	52.36	52.23	59.17	70.04	100		
Desulfitibacterium metallireducens DSM 15288	53.10	53.41	53.39	65.14	53.17	52.85	52.64	68.06	58.53	59.52	100	
Syntrophpbotulus glycolicus DSM 8271	52.50	53.68	52.52	57.85	52.07	51.66	52.09	58.28	62.79	63.77	58.10	100

The overall synteny of genes in each genomes was also compared. Strain DCMF and strain DMC share 614 syntons (clusters of adjacent or near-adjacent genes, as defined in Section 2.2.9) with each other, over double what each organism shares with strain RM (292 and 261 syntons for strain DCMF and strain DMC, respectively; Figure 2.4). However, when considered as a proportion of genome size, strain RM contains a higher proportion of genes in synteny with the other two organisms (39.32% and 43.64%), than they do with it (17.11 and 24.56%). This suggests that strain RM has a more streamlined genome optimised for DCM dechlorination only, while strains DCMF and DMC encode additional metabolic pathways to this.

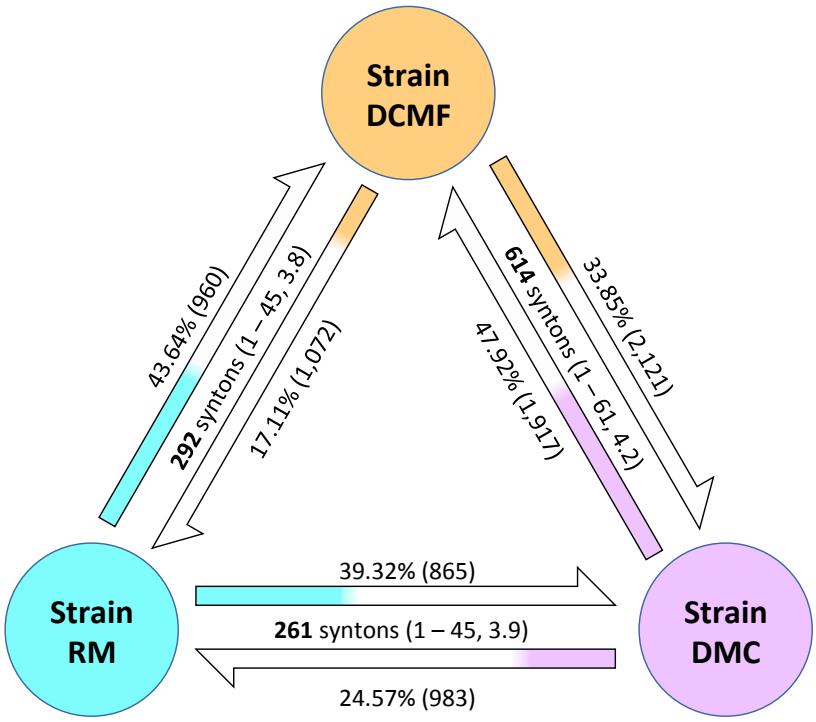


Figure 2.4 Synteny analysis comparing the strain DCMF, strain DMC, and strain RM genomes. Between each pair of arrows is written the number of syntons (minimum – maximum genes per synton, average number of genes per synton). Each directional arrow is also labelled with the percentage of CDS (absolute number of CDS) in the genome that are in syntons with the organism that the arrow is pointing to.

The eggNOG functional annotation of each genome was compared to analyse whether any particular groups of genes were responsible for the variation in genome size between the three organisms. Strain DCMF had a significantly higher proportion of genes from orthologous groups involved in metabolism, including those in the classes for energy production and conversion; amino acid transport and metabolism; coenzyme transport and metabolism; inorganic ion transport and metabolism; and secondary metabolite biosynthesis, transport and catabolism. Strain RM, on the other hand, had a higher abundance of orthologous groups for cellular processes and signalling, being over-represented in the classes of cell wall/membrane/envelope biogenesis; cell motility; posttranslational modification, protein turnover, chaperones; signal transduction mechanisms; and intracellular trafficking, secretion, and vesicular transport (Figure 2.5). Each of the three organisms had a similar proportion of genes that were either classified into orthologous groups of unknown function (23 – 26%), or unclassified by eggNOG (14 – 18%).

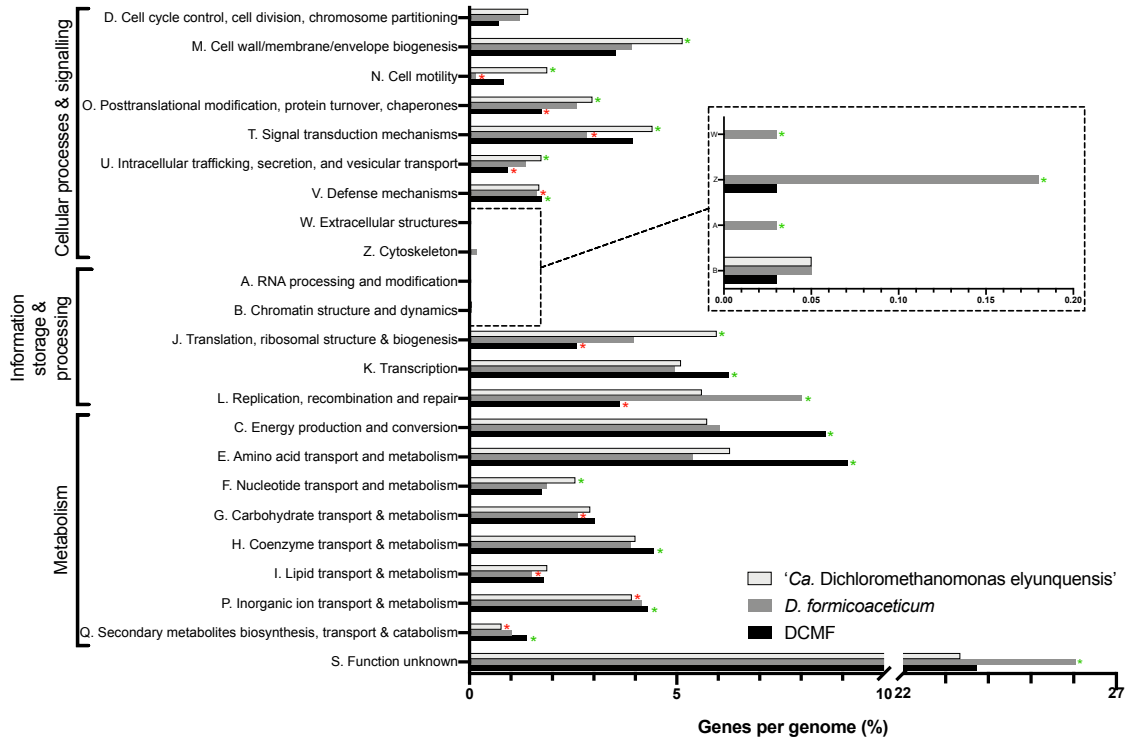


Figure 2.5 Abundance of genes classified into each eggNOG group, as a percentage of each organism's genome. Genes were classified with eggNOG v4.5.1 and eggNOG-mapper v1.0.3 and are reported as the percentage of each organism's total gene count in order to normalise between the differently sized genomes. Values >1 SD (*) or <1 SD (*) from the group mean (*n* = 3) are marked.

2.3.5 The core and pan genome of anaerobic dichloromethane-degrading bacteria

The core and pan genome of the three DCM-dechlorinating bacteria was analysed based on MicroScope gene families (MICFAMs, explained in Section 2.2.9). As the anaerobic DCM-degrading bacteria likely represent three different genera, the analysis was run with permissive alignment parameters of minimum 50% amino acid identity and 80% amino acid alignment coverage. This resulted in a core genome of 611 MICFAM (containing 2,491 genes) (Figure 2.6), as opposed to only 47 MICFAM (containing 146 genes) when the amino acid identity threshold was 80% (data not shown).

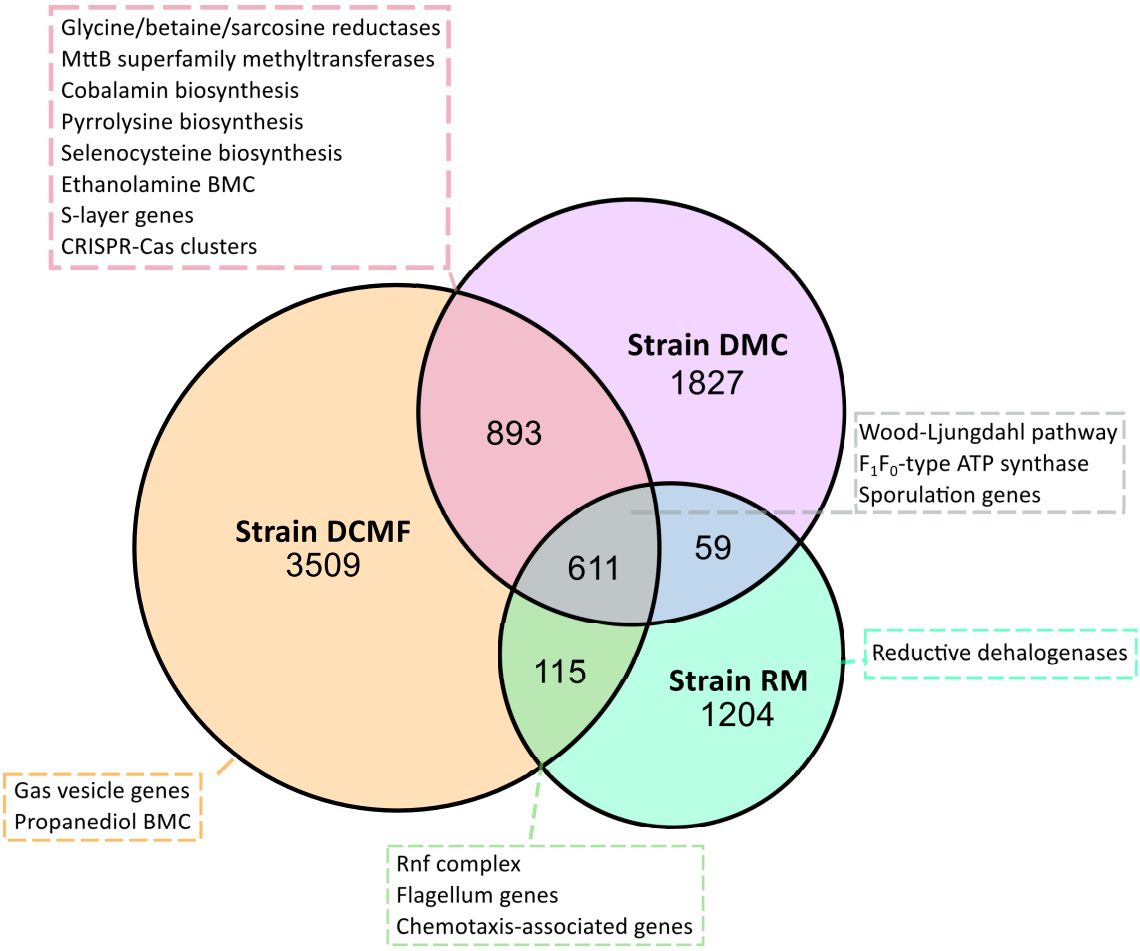


Figure 2.6 Proportional Venn diagram of the core, variable and pan genome of strain DCMF, strain DMC, and strain RM. Values are the number of MICFAM (MicroScope gene families) in each segment. Annotated unique and common genes/pathways of interest in the genomes are also shown, based on manual curation and data from Chen *et al* (2017b) and Kleindienst *et al* (2016, 2017, 2019).

In addition to the MICFAM analysis, pathways and gene clusters of interest (e.g. those hypothesised to be involved in DCM transformation or metabolism of other substrates) in the three genomes were manually inspected and curated to reveal those in common between the anaerobic DCM-degraders. There are a number of key features shared between all three species, including genes for the Wood-Ljungdahl pathway, an F_1F_0-type ATP synthase, and sporulation (Figure 2.6; Dataset S1). Strain DCMF and strain DMC again have more in common with each other than with the more distantly related strain RM. The former two organisms have the potential for metabolism of substrates other than DCM, as they encode genes for MttB family methyltransferases (involved in methylated amine, glycine betaine, and proline betaine metabolism), glycine/betaine/sarcosine reductases, and an ethanolamine-

utilising (EUT) bacterial microcompartment (BMC) (Dataset S1). The strain DCMF genome also encodes a propanediol-utilising (PDU) BMC (Figure 2.6; Dataset S1).

There is evidence for motility in strain DCMF and strain RM (genes for a flagellar and chemotaxis), and for an S-layer enveloping strain DCMF and strain DMC. Strain DCMF also contains genes for gas vesicle formation (Figure 2.6; Dataset S1).

2.3.6 Evidence of mobile genetic elements amongst the anaerobic dichloromethane-degrading bacteria

Strain DCMF was the only organism that definitively contained prophage in its genome, as predicted by both PHASTER and Prophage Hunter. Twenty-six of the putative proteins in this region (spanning nucleotides 957,144 – 1,004,373) had hits against the virus and prophage database, according to PHASTER, with a further 20 proteins identified only as hypothetical (Table 2.3). All other results for strain DCMF, strain DMC and strain RM were conflicting between PHASTER and Prophage Hunter, making it difficult to conclusively state the presence or absence of prophage in the genomes. Both genomes contain a number of ambiguous or incomplete prophage regions (Figure 2.7). In addition to this, three active prophage were predicted with relatively high scores by Prophage Hunter in the strain DMC genome, despite not being identified at all by PHASTER (Table 2.3). The putative prophage in both strain DCMF and strain DMC genomes may represent novel viral taxa, as they shared remarkably low identity (<1% query coverage in the majority of cases) with other bacteriophage, based on the blastn search results offered by the two programs. Thus, "closest hit" data is not reported, as it is likely irrelevant. Prophage induction in strain DCMF was not attempted.

Table 2.3 Summary of prophage identified in the strain DCMF, strain DMC and strain RM genomes by two bioinformatics methods. Columns on the left show results from PHASTER (Arndt *et al.*, 2016) and those on the right from Prophage Hunter (Song *et al.*, 2019). Only prophage regions reported as intact (PHASTER classification) or active (Prophage Hunter classification) are shown.

	PHASTER			Prophage Hunter		
	Region	**Proteins[a]**	**Score[b]**	**Region**	**Proteins[c]**	**Score[d]**
Strain DCMF	957144 – 1004357	49 (26)	130	957144 – 1004373	70	0.85
Strain DMC	–	–	–	460639 – 473846	13	0.87
				589065 – 618657	28	0.90
				185348 – 1925422	53	0.91
Strain RM	–	–	–	–	–	–

[a] Number of PHASTER-predicted proteins in the region (number of these with hits in the virus and phage database).

[b] PHASTER scores putative prophage regions as incomplete (<70), questionable (70 – 90), or intact (>90). See http://phaster.ca for details on scoring criteria.

[c] Number of Prophage Hunter-predicted proteins in the region

[d] Prophage Hunter scores putative prophage regions as ambiguous (0.5 – 0.8) or active (0.8 – 1.0).

There was similar diversity amongst the CRISPR-Cas loci in the three bacteria. The strain DCMF genome contains a single Type I-B system with 117 repeats. Strain DMC is more unusual, containing four CRISPR-Cas loci: a Type I-B system (59 CRISPR repeats, followed by a further 73 identical repeats), a Type II-A system (33 repeats), Type III-D (95 repeats) and a Type III-B system (54 repeats). In stark contrast to this, the strain RM does not contain any CRISPR-Cas loci (Figure 2.7).

In keeping with the elevated number of CRISPR-Cas loci in strain DMC, it also contained the highest proportion (3.2% of CDS in the genome; 122 genes) of predicted transposases. Notably, a number of these surround the Type II-A CRISPR-Cas loci described above (Figure 2.7). Strain DCMF contained the second highest

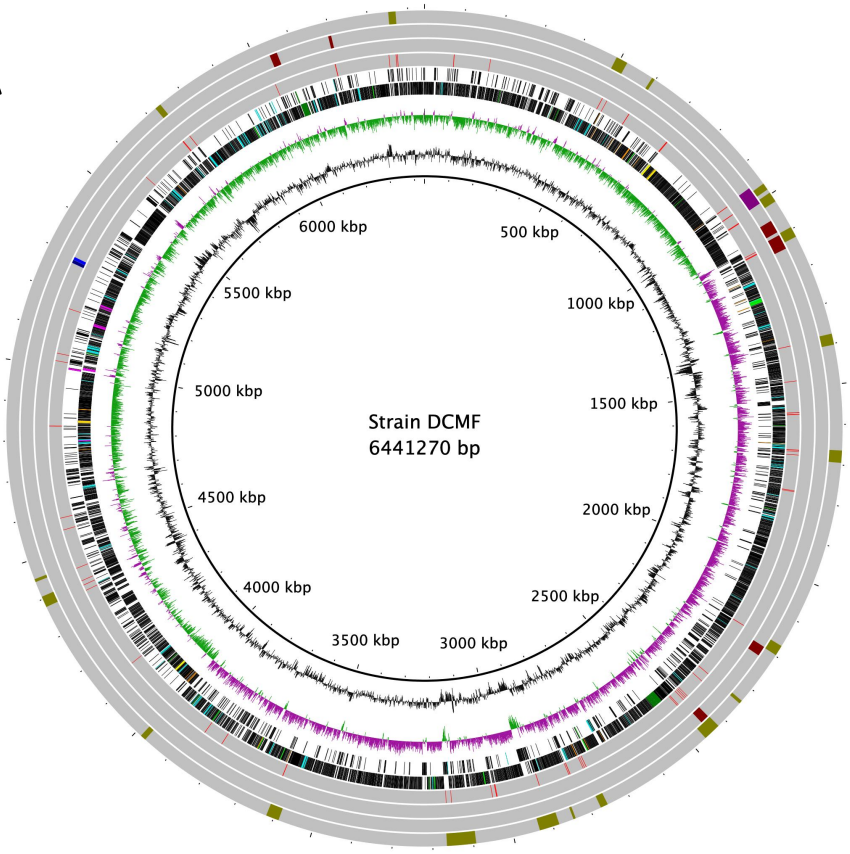
proportion of predicted transposases in its genome (1.2%, 77 genes), followed by strain RM (1.0%, 22 genes).

Both the strain DCMF and strain DMC genomes contained numerous (16 – 20) genomic islands (GIs), while strain RM contained less (Figure 2.7). IslandViewer4 initially predicted eight GIs in the latter's genome, however after discounting those that passed across contig boundaries, only four were considered for the analysis (Table 2.4). Although a predicted GI spanning many of the smaller contigs towards the end of the strain RM genome was discounted, the shortest 27 contigs (which represent only 2% of the total genome) did encode eight of the predicted 22 transposons. Indeed, the GIs in all three organisms tended to be associated with transposases and putative prophage regions, confirming the GI prediction from IslandViewer4 (Figure 2.7).

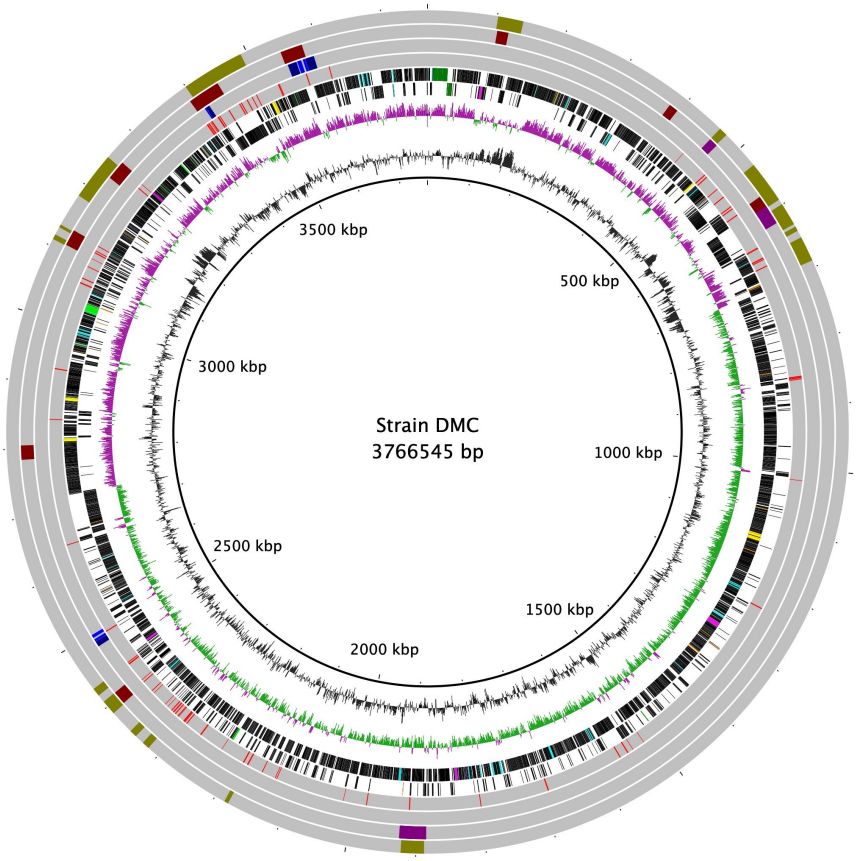
Table 2.4 Genomic islands (GIs) present in the strain DCMF, strain DMC, and strain RM genomes. Genomic islands were predicted by IslandViewer4.

Organism	# GIs	Min. length (bp)	Max. length (bp)	Total CDS in GIs	Min. CDS in GI	Max. CDS in GI
Strain DCMF	20	5,923	72,170	528	7	72
Strain DMC	16	4,225	90,984	533	7	94
Strain RM	4	9,488	34,163	81	15	32

A


Strain DCMF
6441270 bp
500 kbp
1000 kbp
1500 kbp
2000 kbp
2500 kbp
3000 kbp
3500 kbp
4000 kbp
4500 kbp
5000 kbp
5500 kbp
6000 kbp

B


Strain DMC
3766545 bp
500 kbp
1000 kbp
1500 kbp
2000 kbp
2500 kbp
3000 kbp
3500 kbp

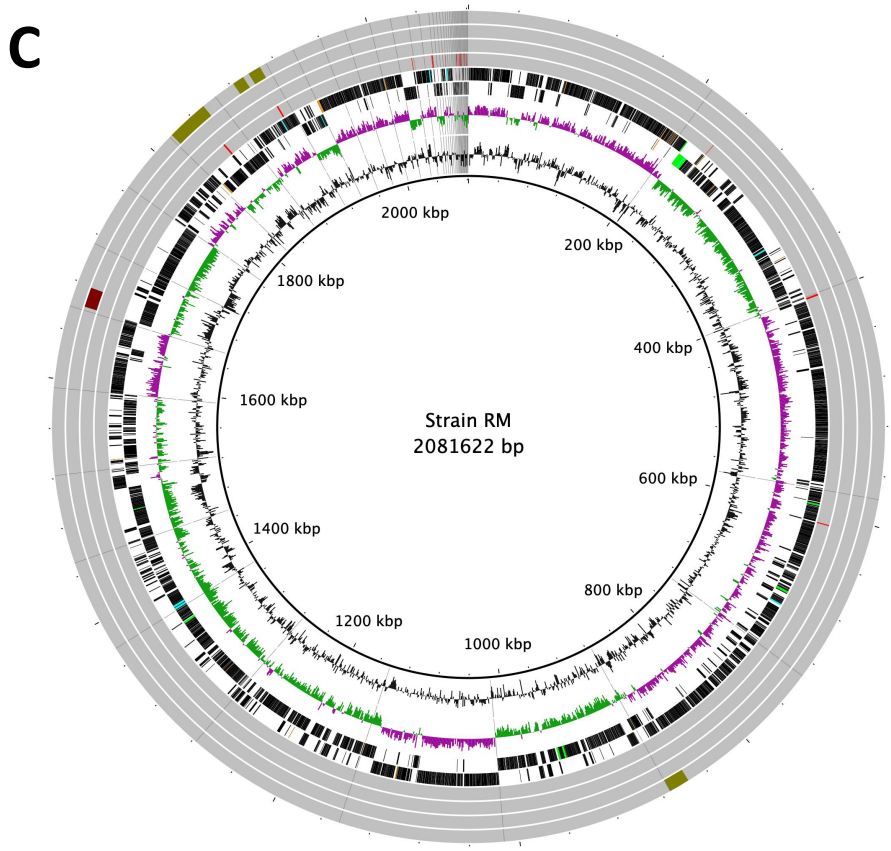


Figure 2.7 The genomes of (A) strain DCMF, (B) *Dehalobacterium formicoaceticum* strain DMC, and (C) '*Candidatus* Dichloromethanomonas elyunquensis' strain RM annotated with genes for putative metabolic pathways and elements associated with horizontal gene transfer. From inside to outside, rings are as follows: 1. Length scale; 2. GC content; 3. GC skew; 4. Forward strand; 5. Reverse strand; 6. Transposases (red); 7. CRISPR-Cas loci (Cas genes in navy, CRISPR repeats in blue); 8. Prophage regions (intact in purple, ambiguous in maroon); 9. Genomic islands (olive). Genes on the forward and reverse strands (rings 4 and 5) are mostly in black, with the following groups of interest highlighted: rRNA (yellow), tRNA (orange), Wood-Ljungdahl pathway genes (light green), methyltransferases (light blue), glycine/betaine/sarcosine reductase clusters (pink), and ethanolamine- and propanediol-utilising bacterial microcompartment genes (dark green). Contig boundaries are shown as grey lines (strain RM only).

2.4 Discussion

2.4.1 Optimisation for a high-quality genome assembly from a mixed culture

Early enrichments of the DFE culture, reported in Holland *et al* (2019), were analysed via denaturing gradient gel electrophoresis and showed a dominant band corresponding to strain DCMF. Based on the 16S rRNA gene sequence retrieved from this analysis, strain DCMF appeared to be an organism with comparatively few cultured relatives. Thus, whole genome sequencing was carried out in order to learn more about its role and function in the DFE community. The other organisms in the enrichment culture and lack of a reference genome hindered attempts to assemble the novel genome from short read sequences only, making the long-read capability of PacBio sequencing indispensable for this effort. Although long reads are prone to a higher proportion of sequencing errors than short reads, a series of checks were put in place to ensure that a high quality, uncontaminated genome assembly was obtained.

The use of SMRTSCAPE to predict the optimal HGAP settings allowed rapid comparison of various assembly parameters. By increasing the minimum correction coverage from 6× to 10×, the total size of the assembly (including contaminant organism DNA) decreased from ~16 Mb to ~8.8 Mb, while the size of the strain DCMF genome remained relatively stable at around 6.4 Mb. Increasing the minimum correction coverage one step further to 11× resulted in a significant reduction of the strain DCMF genome to 1.9 Mb, indicating that much of the assembly was likely being lost to overzealous correction (Table S1).

2.4.2 Phylogeny of the novel dichloromethane-degrading bacterium

Based on all available analyses, strain DCMF appears to be a novel taxon. The typical 16S rRNA gene identity threshold for bacteria in the same genus is 94.5% (Yarza *et al.*, 2014). Strain DCMF is on the borderline of this, sharing 94% identity with its closest relative, strain DMC (Figure 2.2). Furthermore, strain DCMF shares 66.54% AAI with strain DMC (Table 2.2), which is at the lower end of the 65-72% AAI genus boundary proposed by Konstantinidis and Tiedje (2007), but within the more recent

boundary of 55-60% proposed by Rodriguez-R and Konstantinidis (2014). GTDB-Tk analysis suggested that strain DCMF be placed in the family *Dehalobacteriaceae*, again with strain DMC as its closest relative, but did not class it within the same genus, while the lowest available taxonomic classification the SILVA database (based on the 16S rRNA gene) was the family *Peptococcaceae*. As a whole, this information suggests that strain DCMF does not fall into any pre-existing genus. At the very least, it certainly represents a novel species, as the genomic data falls below the threshold values of 98.7% 16S rRNA gene identity, 94% ANI, and 85-90% AAI (Rodriguez-R and Konstantinidis, 2014) with all closely related bacteria. Therefore, we propose that strain DCMF represents a novel taxon within family *Peptococcaceae* (NCBI taxonomy) or *Dehalobacteriaceae* (GTDB taxonomy). Taxonomic classification is revisited in Chapter 3, with phenotypic results considered alongside the genomic analysis.

2.4.3 An overview of genomes encoding anaerobic dichloromethane degradation: is size indicative of greater metabolic potential?

The large size of the strain DCMF genome distinguishes it from the two other anaerobic DCM-dechlorinating bacteria, strain DMC and strain RM (Table 2.1). When assembling a genome *de novo* from a mixed culture, sequences from co-habiting organisms can be erroneously incorporated into the assembly. This likelihood was mitigated by our assembly strategy of increasing stringency. The consistent sequencing coverage across the final genome (Figure S1) strongly indicates that there was no such mis-assembly. The CheckM contaminant rate of 2% further confirms that the large strain DCMF genome is not inflated due to contamination. Analysis of repeated sequence motifs with SPADE showed that they comprise just 21,395 bp (0.03%) of the total strain DCMF genome, which also rules this out as a source of the large genome size. The manually curated IMG annotation predicted 5,772 protein coding genes, giving a gene density of approximately 0.9 genes per kilobase, which is consistent with normal bacterial gene density (Koonin and Wolf, 2008).

The analysis of the abundance of genes classed into the various eggNOG groups showed that strain DCMF was primarily enriched for genes involved in metabolism,

particularly those involved in 'energy production and conservation' and 'amino acid transport and metabolism' (Figure 2.5). This could further explain its larger genome compared to the other two anaerobic DCM-degrading bacteria, particularly as strain DCMF encodes a large array of MttB superfamily methyltransferases and genes involved in the metabolism of quaternary amines. Both in terms of synteny (Figure 2.4) and the core/pan genome (Figure 2.6), strain DCMF appears to encompass many of the genetic elements from the other two bacteria plus additional, unique traits. The core/pan genome analysis showed that it has the highest proportion of species-specific genes (62.83%), followed by strain RM (57.35%) and then strain DMC (51%; Figure 2.6). A similar pattern was seen in the synteny analysis carried out, where both strain DMC and strain RM contain a far higher proportion of syntons with strain DCMF, than it does with either organism (Figure 2.4).

2.4.4 Central carbon and energy metabolism

The strain DCMF genome suggests that it dechlorinates DCM via incorporation into the Wood-Ljungdahl pathway as has been reported for strains DMC and RM (Mägli *et al.*, 1998; Kleindienst *et al.*, 2019). All genes for this pathway were identified within the genome, as well as those linking acetyl coenzyme A (CoA) to pyruvate and central carbon metabolism within the cell (Dataset S1). DCM catabolism via the Wood-Ljungdahl pathway would likely result in the production of acetate, and possibly formate, although it is also possible for the pathway to be used in complete mineralization of DCM into H_2 and CO_2 (Chen *et al.*, 2020). The strain DCMF genome encodes a cytoplasmic formate dehydrogenase (Ga0180325_112876, 112877, 112878s80) that could theoretically oxidise the formate to CO_2, in line with initial reports that formate did not accumulate during DCM degradation (Wong, 2015). Mägli *et al* (1996) found that, despite the presence of formate dehydrogenase activity in cell extracts, strain DMC could not further metabolise the formate it produced, and instead accumulated it with acetate in a 2:1 molar ratio. Conversely, formate doesn't accumulate in the RM consortium that strain RM is present in and formate dehydrogenase was identified in the proteome of DCM-fed cultures, indicating that it likely transforms all formate into CO_2 (Kleindienst *et al.*, 2017, 2019).

All three DCM-metabolising bacteria encode an F_1F_0-type ATPase in their genome (Figure 2.6), suggesting that they could employ a chemiosmotic mechanism for energy conservation alongside substrate-level phosphorylation of DCM (Dataset S1). Evidence for ATPase activity has been detected in cell-free extracts of strain DMC and via proteomics in strain RM (Chen *et al.*, 2017b; Kleindienst *et al.*, 2019; Mägli *et al.*, 1998). Strain DCMF and strain RM may generate a proton- or sodium-motive force for this ATPase via the Rnf complex, an ion-motive ferredoxin-NAD oxidoreductase encoded in both genomes (Figure 2.6; Dataset S1). The Rnf complex is of particular importance given the absence of any recognizable electron-bifurcating hydrogenases in the genome. Typically, the complex pumps ions out of the cell, catalyzing electron transfer from reduced ferredoxin to NAD^+ (Biegel *et al.*, 2009; Biegel and Muller, 2010), while the ATPase uses the flow of ions back into the cytoplasm to convert ADP to ATP. However, these two transmembrane protein complexes can also act in reverse in order to balance the pool of reduced electron carriers within the cell (e.g. Lechtenfeld *et al.*, 2018).

2.4.5 The abundance of methyltransferases in strain DCMF may indicate a key role in dichloromethane and wider metabolism

The protein responsible for the dechlorination of DCM remains elusive within anaerobic DCM-degrading bacteria. However, there is growing evidence that a novel, corrinoid-dependent methyltransferase is responsible for transforming DCM into 5,10-methylene-THF, which then enters the Wood-Ljungdahl pathway (Mägli *et al.*, 1998; Chen *et al.*, 2018; Kleindienst *et al.*, 2019). Strain DCMF encodes an abundance of predicted methyltransferase proteins (96) in its genome, hinting at a key role in metabolism (Dataset S1). The majority are for a methyltransferase system comprised of a methyltransferase 1 (MT_1), which transfers a methyl group from the substrate onto a cognate corrinoid protein (CoP), from which the methyltransferase 2 (MT_2) transfers the methyl group to the final receiving compound.

The majority (82) of the methyltransferases in strain DCMF were members of the MttB superfamily (i.e., an MT_1). This group of proteins is named after its founding member, a trimethylamine:corrinoid methyltransferase (*mttB*) discovered in

methanogenic Archaea and notable for containing the non-canonical amino acid pyrrolysine (Ferguson and Krzycki, 1997; Paul *et al.*, 2000; Krzycki, 2004). The MttB superfamily is widespread among Bacteria and Archaea, although most genes do not encode the pyrrolysine residue (Srinivasan, 2002; Ticak *et al.*, 2014). Some non-pyrrolysine members of the family have since been demonstrated to act on glycine betaine (Ticak *et al.*, 2014) and proline betaine (Picking *et al.*, 2019), providing growing support that the pyrrolysine-free majority of this family catalyse demethylation of quaternary amines, or perhaps an even wider array of substrates.

The large number (82) of MttB superfamily methyltransferases in the strain DCMF genome is unusual and implies a certain level of functional redundancy amongst them. It is more than double the number of MttB family genes in *Eubacterium limosum* SA11 (39), which was previously reported to be the highest of any bacterial genome (Kelly *et al.*, 2016). There is relatively high sequence diversity between the MttB family predicted protein sequences in strain DCMF (average pairwise amino acid sequence identity of 30.3%). This is congruent with their potential ability to demethylate a wider array of substrates and could also be due to diversification to accommodate cobalamin cofactors with various upper and lower ligands (Visser *et al.*, 2016). It has previously been shown that the chloromethane dehalogenase CmuAB is functionally similar to the methylamine methyltransferase MtaA (Studer *et al.*, 2001). Moreover, four corrinoid-dependent methyltransferases were highly expressed in the proteome of DCM-mineralising strain RM (Kleindienst *et al.*, 2019). The array of MttB superfamily genes in strain DCMF, along with its complete corrinoid biosynthetic pathway therefore raises the question of whether some of these proteins may be linked to DCM metabolism.

The high abundance of methyltransferases in the proteome of DCM-grown strain RM cells suggests that anaerobic DCM transformation may be mechanistically more similar to the aerobic dechlorination of chloromethane than DCM. Distinct from the aerobic glutathione *S*-transferase enzymes involved in aerobic DCM dechlorination, the chloromethane dehalogenase CmuAB is a two-step methyltransferase system (Vannelli *et al.*, 1999). CmuA is a bifunctional methyltransferase and corrinoid-binding protein that transfers the methyl group of chloromethane onto itself, from which CmuB then transfers it to THF, generating methyl-THF (Vannelli *et al.*, 1999;

Studer *et al.*, 2001). Given that the four abundant corrinoid-dependent methyltransferases from '*Ca.* Dichloromethanomonas elyunquensis' had relatively high identity homologs in *D. formicoaceticum*, yet lacked similarity to any previously characterised methyltransferases (Kleindienst *et al.*, 2019), perhaps a similar system to the chloromethane methyltransferases functions within anaerobic DCM-dehalogenating bacteria.

It is worth noting that strain RM is unique among the three anaerobic DCM-degrading species in encoding reductive dehalogenase genes in its genome (Figure 2.6). A proteomic study showed that two of the three dehalogenases were expressed by the organism during growth on DCM (Kleindienst *et al.*, 2019). Furthermore, it was recently reported to completely mineralise DCM to H_2 and CO_2, rather than producing acetate and/or formate as an end product (Chen *et al.*, 2020). These findings, coupled with a recent dual carbon-chlorine isotopic analysis of DCM dechlorination in strain DMC and strain RM, support distinct DCM dechlorination mechanisms operating in the DCM-degrading bacteria (Chen *et al.*, 2018; Kleindienst *et al.*, 2019). Based on its lack of reductive dehalogenases and greater genomic similarity to strain DMC than strain RM (Figure 2.4, Figure 2.6), strain DCMF is predicted to transform DCM via a mechanism more similar to the former than the latter organism.

2.4.6 Strain DCMF and strain DMC may not be obligate dichloromethane fermenters

Strains DCMF and DMC both encode MttB superfamily genes (82 and 23, respectively), as well as other methyltransferases putatively involved in methylated amine, quaternary amine, DCM, or other unknown substrate metabolism (Dataset S1). Conversely, strain RM lacks any genes from the MttB superfamily, although four corrinoid-dependent methyltransferases were among the most abundant proteins in DCM-grown cells (Kleindienst *et al.*, 2019). Based on the following arguments, we hypothesise that these genes are involved in the metabolism of methylated amines, glycine betaine, or sarcosine.

Notably, 23 of the 96 methylamine methyltransferase genes in strain DCMF contain the pyrrolysine residue, identifiable as an in-frame UAG (amber) stop codon. To date, only verified trimethylamine methyltransferases have possessed this residue,

suggesting strain DCMF may be capable of trimethylamine metabolism. The *pylTSBCD* gene cluster to synthesise and incorporate pyrrolysine is limited to a small number of bacterial genera. These include *Desulfotomaculum, Desulfitobacterium*, and *Thermincola* (Gaston *et al.*, 2011) – all members of the *Peptococcaceae* family and close relatives of strain DCMF based on 16S rRNA phylogeny and AAI analysis. Strain DMC also encodes the *pyl* genes and some methyltransferases with pyrrolysine residues, but strain RM does not (Figure 2.6). A model for growth on methylated amines has been shown in *Sporomusa* sp. strain An4, in which trimethylamine is demethylated in a stepwise manner to ammonium (Visser *et al.*, 2016). The methyl groups are transferred on to THF, forming CH_3-THF, which is then funnelled into the Wood-Ljungdahl pathway to produce CO_2 and reducing equivalents (Visser *et al.*, 2016).

Further, Ticak *et al* (2014) demonstrated that a non-pyrrolysine MttB homolog in *Desulfitobacterium hafniense* Y51 was in fact a glycine betaine methyltransferase, *mtgB*. Glycine betaine methyltransferase systems (including the MT_1, MT_2, and CoP outlined above) have also been found in *S. ovata* An4 (Visser *et al.*, 2016), *S. ovata* DSM 2662, and *Acetobacterium woodii* (Lechtenfeld *et al.*, 2018). In *S. ovata* An4, the same genes were proposed to further demethylate the resulting dimethylglycine to sarcosine (Visser *et al.*, 2016). Non-pyrrolysine MttB superfamily members have also been shown to demethylate proline betaine to *N*-methyl proline in *Eubacterium limosum* ATCC8486 (Picking *et al.*, 2019) Thus, the presence of non-pyrrolysine MttB superfamily genes in strains DCMF and DMC suggests they may be able to demethylate one or more quaternary amines. This discovery drove the testing of glycine betaine as a substrate, reported in Chapter 3.

There is further evidence for the potential metabolism of glycine betaine in the strain DCMF and strain DMC genomes due to the presence of glycine/betaine/sarcosine reductase genes (Figure 2.6; Dataset S1). The reductase complex consists of three components: the selenocysteine-containing component A (GrdA); a two- or three-subunit substrate-specific component B, one of which contains a selenocysteine residue (GrdBE for glycine, GrdFG for sarcosine, GrdHI for betaine); and component C, which is post-translationally combined into a single protein (GrdCD) (Andreesen, 2004). Reductive cleavage of glycine betaine or

sarcosine results in trimethylamine or methylamine, respectively, plus acetyl phosphate. Some clusters of these genes in the strain DCMF and strain DMC genomes also harbour the thioredoxin I and thioredoxin reductase necessary for electron transfer to the glycine betaine/sarcosine reductase, and both organisms contain the genes for synthesising and incorporating selenocysteine into the selenoproteins (Figure 2.6; Dataset S1).

Of the five B components in strain DCMF, Ga0180325_114802 (GrdG) and Ga0180325_114803s (GrdF) are likely specific to sarcosine, as they clustered with the sarcosine reductase genes from *Sporomusa* sp. An4, while Ga0180325_11115251 (GrdI) and Ga0180325_115252s54 (GrdH) clustered with the glycine betaine reductases (Figure 2.8). One of the remaining B components is likely a pseudogene (Ga0180325_11855s), as it lacks the required UxxCxxC motif after the selenocysteine residue to protect against accidental oxidation (Parther, 2003). The substrate-specificity of the remaining two B components (Ga0180325_114453 and Ga0180325_114454s56; Ga0180325_114684 and Ga0180325_114685s86) is unclear, as they did not cluster with any of the annotated reductase genes (Figure 2.8).

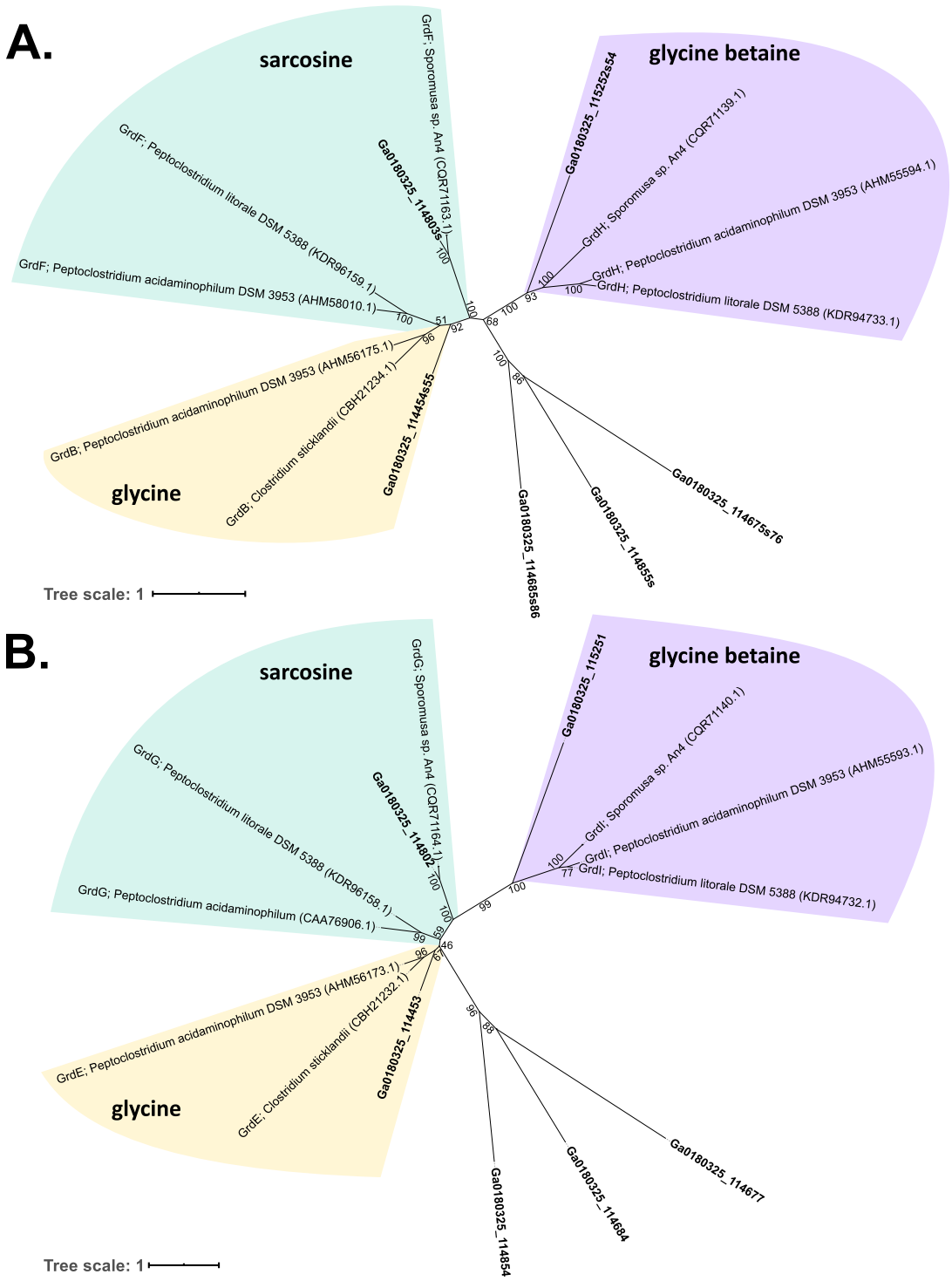


Figure 2.8 Unrooted Maximum Likelihood trees of predicted glycine/betaine/sarcosine reductase complex B proteins from strain DCMF with those of known function from other bacteria. The substrate specificity of **A.** predicted GrdB/F/H (selenocysteine-containing) and **B.** GrdE/G/I proteins from strain DCMF (bold text) was inferred via the construction of Maximum Likelihood trees with proteins of known substrate specificity (GrdBE proteins are specific to glycine; GrdFG to sarcosine, and GrdHI to glycine betaine). Numbers indicate percentage of branch support from 1000 bootstraps.

Finally, the genomes of strain DCMF and strain DMC also appear to contain genes for the formation of bacterial microcompartments (BMCs; Figure 2.6; Dataset S1). These are self-assembling proteinaceous structures within the cytoplasm that contain an enzymatic core, typically functioning to protect the cell from toxic or volatile intermediates resulting from certain metabolic processes (reviewed in Kerfeld *et al.*, 2018). Based on the classification scheme proposed by Axen *et al* (2014), the ethanolamine BMCs in both strains appear to be EUT2D-type loci, while the propanediol locus in strain DCMF is subtype PDU1. Although all three BMC gene clusters appear to be complete, they differ slightly from the typical EUT2D and PDU1 types described previously (Axen *et al.*, 2014). EUT2D-type loci typically encode an MttB superfamily methyltransferase (Axen *et al.*, 2014), which is absent from the gene cluster in strain DMCF and strain DMC, although clearly abundant throughout the rest of the genomes.

It is worth considering whether BMCs may have some function in DCM dechlorination instead of, or as well as, in the catabolism of their stated compounds. BMCs are known to protect the intracellular environment from toxic metabolic intermediates and could perhaps act similarly for DCM. DCM is a toxic chemical that may affect cell membrane integrity, as has been reported for a variety of other lipophilic hydrocarbons (Sikkema *et al.*, 1995). It can also adversely affect various corrinoid- and porphinoid-dependent enzymes by competitively binding to them (Yu and Smith, 2000). Given the importance of a corrinoid cofactor to the acetyl-CoA synthase / corrinoid iron-sulphur protein reaction in the Wood-Ljungdahl pathway, it may be necessary for the cells to keep cytoplasmic concentrations of DCM relatively low or risk a similar inhibitory effect. There is also precedent for cobalamin recycling within BMCs, which would be required by the corrinoid-dependent methyltransferases implicated in DCM transformation (Mägli *et al.*, 1998; Kleindienst *et al.*, 2019). The PDU BMC is able to internally recycle the adenosylcobalamin cofactor required by its signature enzyme (Johnson *et al.*, 2001; Bobik, 2007). In summary, the encapsulation of DCM in a BMC until it has been dechlorinated could ameliorate toxic effects.

2.4.7 Evidence for mobile genetic elements in dichloromethane-dechlorinating bacteria

Horizontal gene transfer appears to be important for DCM-degradation in oxic environments (Vuilleumier *et al.*, 2009; Muller *et al.*, 2011) and is frequently involved in the spread of reductive dehalogenase genes amongst anerobic organohalide-respiring bacteria (reviewed in Liang *et al.*, 2012). In accordance with this, Kleindienst *et al* (2019) suggested that horizontal gene transfer may have also been involved in the acquisition of the reductive dehalogenase genes found in the strain RM genome. The incidence of horizontal gene transfer within the three anaerobic DCM-degrading bacterial genomes was thus investigated to determine whether it may have been involved in the spread of any other gene clusters. However, little association was found between mobile genetic elements (transposases, prophage regions, genomic islands) and genes for the Wood-Ljungdahl pathway, methyltransferases, glycine/betaine/sarcosine reductase complexes, or BMCs (Figure 2.7).

Nonetheless, the analysis revealed that strain DMC has likely been subject to higher levels of invasion from bacteriophage than the other two anaerobic DCM-degrading bacteria. As could be expected, there was a loose association between regions denser with transposons and those predicted to be prophage regions and/or genomic islands (Figure 2.7). This was particularly pronounced in strain DMC, which had the highest number of transposases encoded in its genome (112). Furthermore, strain DMC had a higher number and variety of CRISPR-Cas loci in strain DMC, compared to the other two species. In fact, the type II-A CRISPR-Cas locus in strain DMC (3404717 – 3412916 bp) was flanked by transposase genes (Figure 2.7) suggesting that it may have been entirely acquired via horizontal gene transfer, as has been observed previously (Godde and Bickerton, 2006; Tyson and Banfield, 2008; Heidelberg *et al.*, 2009; Horvath *et al.*, 2009; Portillo and Gonzalez, 2009). Overall, it seems likely that strain DMC has been subject to greater evolutionary pressure from invading DNA, although we were unable to perceive any effect of this on the putative metabolic pathways encoded in its genome.

2.4.8 The evolution and unique environmental niches of anaerobic dichloromethane-degrading bacteria

The three anaerobic DCM-degrading bacteria were each enriched or isolated from unique environments: strain DCMF from contaminated groundwater near marine coastal waters, strain DMC from enrichments originally from an anaerobic fixed-bed charcoal reactor remediating polluted groundwater (Stromeyer *et al.*, 1991; Mägli *et al.*, 1995, 1996), and strain RM from pristine freshwater river sediment (Kleindienst *et al.*, 2017). These discrete contexts may help to understand some of the genetic differences between the organisms.

For example, strain DCMF is unique in harbouring the genes for gas vesicle production (Figure 2.6; Dataset S1). Gas vesicles occur almost exclusively in aquatic bacteria, which use them to move through the water column to depths of optimal light, oxygen or salinity (reviewed in Walsby, 1994). Their putative presence in strain DCMF suggests that the organism may have primarily evolved in the nearby marine environment (the sampling site sits adjacent to Botany Bay, Sydney, Australia). Further evidence for the possible marine evolution of strain DCMF comes from its putative ability to consume the common osmoprotectant glycine betaine. Biomass turnover in marine environments could be reasonably expected to provide a constant source of this compound.

The fact that strain RM is the only one of the three organisms to have been enriched/isolated from a pristine environment seems at odds with the fact that it appears to be an obligate DCM-degrading bacterium, both physiologically and genomically. It does not have any of the MttB superfamily methyltransferases, glycine/betaine/sarcosine reductases, or the BMC genes present in strain DCMF and strain DMC (Figure 2.6). However, there are a growing number of reports of organohalide-respiring bacteria being isolated from pristine environments, which is hypothesised to be a result of naturally produced organohalides (Atashgahi *et al.*, 2017). For example, oceanic sources are estimated to produce 68 Gg of DCM per year (Kolusu *et al.*, 2017). Sources of DCM in pristine ocean environments may include marine macrophytes (Baker *et al.*, 2001) and phytoplankton (Ooki and Yokouchi, 2011); coastal mangrove forests, particularly in tropical latitudes (Kolusu *et al.*, 2017, 2018); and flux from the atmosphere (Moore, 2004), all of which may

contribute to the evolution of DCM-degrading bacteria even in pristine environments.

The natural production of trace amounts of DCM from geothermal activity (Gribble, 2010) may also provide clues to the long-term evolution of anaerobic DCM-metabolising bacteria. The Wood-Ljungdahl pathway is generally agreed to be one of the oldest, if not the first, metabolisms on Earth (reviewed in Fuchs, 2011; Martin, 2012), and the geothermally active early Earth could plausibly have been rich in organic molecules such as DCM, fuelling the evolution of metabolism. Furthermore, whilst reductive dehalogenases in organohalide respiring bacteria show signs of rapid evolution in response to organochlorine pollution (Maillard *et al.*, 2005; West *et al.*, 2008; McMurdie *et al.*, 2009, 2011; Tang *et al.*, 2016), the three anaerobic DCM-degrading bacteria lack similar evidence of genomic plasticity (Figure 2.7). It is therefore suggested that anaerobic organisms capable of DCM degradation may in fact represent an ancient metabolism that has persisted from some of the first life forms on the planet to the present day, rather than a more recent response to anthropogenic pollution. As such, detailed study of their metabolism may provide valuable insight into life on early Earth.

2.5 Conclusions

Strain DCMF is a novel organism that harbours a large, relatively unique genome. Both long and short read genome sequencing technology were used to complement each other and assemble a single, circular chromosome for the organism, despite the low-level presence of other bacteria in the enrichment culture. Strain DCMF likely represents a novel genus within the family *Peptococcaceae* (NCBI taxonomy) or *Dehalobacteriaceae* (GTDB taxonomy), although this will be further discussed in the following chapter.

The strain DCMF genome suggests that it transforms DCM via the Wood-Ljungdahl pathway, as the other two anaerobic DCM-metabolising bacteria (*D. formicoaceticum* strain DMC and '*Ca.* Dichloromethanomonas elyunquensis' RM) are thought to do. Genomic analysis and comparison to these two species revealed that strain DCMF and strain DMC may also encode the ability to metabolise additional substrates, such as methylamines, glycine betaine, and ethanolamine.

Analysis of genomic traits including the core/pan genome of the anaerobic DCM-degraders, genomic synteny, and regions of putative horizontal gene transfer supported the 16S rRNA and whole genome phylogenetic analysis finding that strain DCMF is more similar to strain DMC than to strain RM. Examined as a whole, the genomes of these organisms raise interesting questions about their evolution and ecological niches, including the possibility that anaerobic DCM utilisation is one of the oldest metabolisms on the planet.

2.6 Acknowledgements

We are grateful to Dr Bat-Erdene Jugder at the University of New South Wales for his assistance with the DNA extraction for PacBio sequencing. The LABGeM (CEA/Genoscope & CNRS UMR8030), the France Génomique and French Bioinformatics Institute national infrastructures (funded as part of Investissement d'Avenir program managed by Agence Nationale pour la Recherche, contracts ANR-10-INBS-09 and ANR-11-INBS-0013) are acknowledged for support within the MicroScope annotation platform.

3 Strain DCMF is a one carbon specialist that is able to grow on a variety of methylated compounds

3.1 Introduction

To date, strain DCMF, *Dehalobacterium formicoaceticum*, and '*Candidatus* Dichloromethanomonas elyunquensis' have all been described as obligate DCM-degrading bacteria that utilise the Wood-Ljungdahl pathway for metabolism of DCM (Mägli *et al.*, 1996; Kleindienst *et al.*, 2017; Holland *et al.*, 2019). Early work with ^{13}C-labelled DCM involving cell suspensions of *Dehalobacterium formicoaceticum* strain DMC showed that DCM was incorporated into tetrahydrofolate (THF) forming methylene-THF, which was further transformed to formate and acetate in a 2:1 molar ratio (Mägli *et al.*, 1998), while '*Candidatus* Dichloromethanomonas elyunquensis' strain RM completely mineralises DCM to H_2, CO_2 and Cl^- (Chen *et al.*, 2020).

As discussed in the previous Chapter, strain DCMF contains genomic hints that it may be able to utilise a wider range of substrates, including an abundance of methyltransferases, particularly from the MttB superfamily, and a number of glycine/betaine/sarcosine reductase complex gene clusters. No anaerobic bacteria able to metabolise both DCM and quaternary amines have previously been identified, although there is precedent amongst aerobic methylotrophs capable of utilising DCM, methylated amines, and methanol for one-carbon metabolism (Brunner *et al.*, 1980; Doronina *et al.*, 2000). The metabolism of methylated amines and glycines is closely linked, particularly in coastal subsurface environments (King, 1984, 1988a). The dominant route of choline and glycine betaine metabolism produces trimethylamine, utilised almost exclusively by methanogens and thus linked to emission of the greenhouse gas methane into the atmosphere (Oremland *et al.*, 1982; King *et al.*, 1983; Gibb *et al.*, 1999). Methanol is also a known methanogenic substrate, and therefore understanding microbial cycling of choline, glycine betaine, and methanol is important for accurately quantifying methane flux from the subsurface.

The aim of this chapter was to investigate the growth of strain DCMF on DCM and other substrates. Work with ^{13}C-labelled compounds confirmed that the Wood-Ljungdahl pathway is central to DCM metabolism, and it was found that strain DCMF is in fact a one carbon (C1) specialist, i.e. an anaerobic methylotroph. The bacterium is able to utilise choline, glycine betaine (trimethylglycine), dimethylglycine, sarcosine (methylglycine) + H_2, and methanol for growth. A genome-based metabolic model for the transformation of each of these substrates is suggested and mass balances provide support for the putative metabolic pathways, showing that the Wood-Ljungdahl pathway is central to metabolism of all substrates by strain DCMF. The bacterium is proposed as a novel genus and species within the family *Peptococcaceae*, '*Candidatus* Formamonas warabiya'.

3.2 Materials and Methods

3.2.1 Culture medium

DFE cultures were grown in minimal mineral salt medium as described in Section 2.2.1. The dilution to extinction principle was utilised in both liquid medium and semi-solid agar shakes (0.6% low-melting agarose, w/v) in attempts to isolate strain DCMF. Attempts were also made to isolate strain DCMF by streaking the culture onto anaerobic agar plates (1% agarose, w/v) with 5 mM glycine betaine.

To investigate the requirement for exogenous bicarbonate during DCM degradation, cultures were buffered with 3-morpholinopropane-1-sulfonic acid (MOPS, 4.2 g l-1) in place of $NaHCO_3$, either with or without 4 mM $NaHCO_3$.

To analyse the fate of the carbon from DCM, the substrate was replaced with ^{13}C-labelled DCM ([^{13}C]DCM). To analyse the assimilation of inorganic carbon, the culture was transferred into MOPS-buffered medium and amended with 5 mM $NaH^{13}CO_3$.

To test alternative growth substrates, DCM was replaced with the following compounds (5 mM unless stated otherwise): carbon monoxide (2 mM), choline, *N,N,N*-trimethylglycine (commonly known as and referred to herein as glycine betaine; tested with and without 10 mM H_2), *N,N*-dimethylglycine, *N*-methylglycine (commonly known as and referred to herein as sarcosine; tested with and without 10 mM H_2), methanol, trimethylamine. Cultures amended with choline, glycine

betaine, and trimethylamine were also amended with the following compounds as electron acceptors (15 mM unless otherwise stated): fumarate (80 mM, tested with trimethylamine only), $NaNO_2$, $NaNO_3$, Na_2SO_3 and Na_2SO_4.

3.2.2 Fluorescence in situ hybridisation (FISH) microscopy

Fluorescence in situ hybridisation (FISH) was carried out with a strain DCMF-specific oligonucleotide probe (Dcm623, 5'-/Cy3/CTCAAGTGCCATCTCCGA-3') designed using ARB (Ludwig *et al.*, 2004). An established probe (Eub338i, 5'-/6-FAM/GCTGCCTCCCGTAGGAGT-3') targeting all bacteria was also used (Amann *et al.*, 1990). FISH was carried out as per an established protocol for fixation on a polycarbonate membrane, using minimal volumes of reagents (Ferrari *et al*, 2008). Cells were fixed both with the protocol for Gram negative cell walls (Amann *et al.*, 1990) and Gram positive cell walls (Roller *et al.*, 1994). Hybridisation was carried out with a formamide-free buffer, as the Dcm623 probe was shown to be unique to strain DCMF when the nucleotide sequence was searched against all 16S rRNA genes found in the PacBio sequencing data reported in Chapter 2. Cells were counterstained with VECTASHIELD® Antifade Mounting Medium containing 1.5 µg ml^{-1} 4,6-diamidino-2-phenylindole (DAPI; Vector Laboratories, Burlingame, CA, USA). Prepared membranes were placed on glass microscope slides for examination on an Olympus BX61 microscope equipped with an Olymus DP80 camera. Images were captured and overlaid using Olympus cellSens Dimension software v2.1. Strain DCMF cell length and width was determined from a sample of 20 cells using the linear measurement tool within the program.

3.2.3 Analytical techniques

3.2.3.1 Gas chromatography

DCM was quantified using a Shimadzu GC-2010 gas chromatograph with flame ionisation detector (GC-FID). Gaseous headspace samples (100 µl) were withdrawn directly from culture flasks with a lockable, gas-tight syringe and injected directly into the GC. The inlet temperature was 250˚C, split ratio 1:10, FID temperature 250˚C. The GC was equipped with a GS-Q column (30 m × 0.32 mm; Agilent Technologies) and the carrier gas was helium (3 ml min^{-1}). The oven temperature was initially 150˚C and then raised by 30˚C min^{-1} to 250˚C. A minimum three-point calibration curve was used. DCM concentrations are reported as the nominal

concentration in each serum bottle, calculated from the headspace concentration using the Henry's Law dimensionless solubility constant as per the OSWER (Office of Solid Waste and Emergency Response) method: H^{cc} = 0.107 for DCM (US EPA, 2001).

Acetate, formate and methanol were analysed using the same GC-FID, equipped with a DB-FFAP column (30 m × 0.32 mm × 0.25 µm film thickness; Agilent Technologies). Inlet and detector parameters remained the same, but the carrier gas flow (helium) was 2 ml min^{-1} and the oven was held at 40˚C for 6 min. For acetate and formate quantification, the compounds were first derivatised to their ethyl esters by adding 500 µl liquid culture samples to a 10 ml screw cap glass vial containing ethanol (200 µl) and 1 M sulphuric acid (200 µl). Samples for methanol analysis (200 µl liquid culture) were not derivatised but transferred directly to a 10 ml screw cap glass vial. Samples were agitated at 80˚C for 5 min before 250 µl headspace was injected from a PAL LHS2-xt-Shim headspace autosampler (Shimadzu). Quantification was via comparison to a five-point standard curve ranging from 0.5 – 15 mM (acetate and formate) or 1 – 15 mM (methanol).

Trimethylamine was quantified using the same GC-FID with a DB-5 column (30 m × 0.32 mm × 0.25 µm film thickness; Agilent Technologies) and the same inlet and detector parameters as above. Liquid culture samples (200 µl) were first alkalized by adding them to a 10 ml screw-cap glass vial with MilliQ water (600 µl) and 4M NaOH (200 µl). Vials were agitated at 80˚C for 2 min before 2 ml headspace was injected with a PAL LHS2-xt-Shim headspace autosampler (Shimadzu). The oven temperature was initially 60˚C then increased at 5˚C min^{-1} to 80˚C and the carrier gas flow was 2 ml min^{-1} helium. Samples were quantified by comparison to a five-point standard curve ranging from 0.1 – 5 mM trimethylamine.

Hydrogen was quantified on a Shimadzu 2010 GC with pulse discharge detector (GC-PDD) equipped with a HP-PLOT Molesieve column (30 m × 0.32 mm × 0.25 µm film thickness; Agilent Technologies). Headspace samples (20 µl) were withdrawn directly from culture flasks with a lockable, gas-tight syringe and injected into the GC. Inlet temperature 250˚C; split ratio 1:10; carrier gas helium (3 ml min^{-1}) oven held at 50˚C for 1.2 min; detector temperature 150˚C. Samples were quantified by comparison to a six-point standard curve (0 – 16.63 mM).

Bicarbonate (as gaseous CO_2) was quantified on the same GC-PDD with a HP-PLOT Q column (30 m × 0.32 mm; Agilent Technologies). Liquid samples (50 µl) were acidified in 1.5 ml screw-cap glass vials with 25% HCl (20 µl). Samples were left to equilibrate at room temperature for 2 h before 100 µl headspace was injected manually with a lockable, gas-tight syringe. GC-PDD parameters were as above except the oven, which was held at 50˚C for 1 min, then raised by 3.5˚C min^{-1} to 54.5˚C.

3.2.3.2 Liquid chromatography with tandem mass spectrometry

Choline and glycine betaine were quantified using liquid chromatography with tandem mass spectrometry (LC-MS/MS). The Agilent Technologies 1200 Series LC was fitted with a Luna C18(2) column (150 × 4.6 mm, 5 µm; Phenomenex). The mobile phases were 0.5 mM ammonium acetate in water (A) and 100% methanol (B). The column was equilibrated with a ratio of 95:5 (A:B) for 5 min, before samples (5 µl) were eluted with a linear gradient from 95:5 (A:B) to 0:100 (A:B) over 10 min, then held at 0:100 (A:B) for 1 min. The LC was coupled to QTRAP 4000 quadrupole mass spectrometer (Applied Biosystems SCIEX,) equipped with a TurboIonSpray source. Electrospray ionization was performed in the positive mode. The machine was operated in multiple reaction monitoring mode and the following precursor-product ion transitions were used for quantification: *m/z* 104.0 → 59.0 (choline) and *m/z* 118.0 → 57.7 (glycine betaine). Samples were quantified by comparison to a six-point standard curve ranging from 0.1 – 10 mM.

3.2.3.3 Gas chromatography with triple quadrupole mass spectrometry

Labelled and unlabelled acetate was quantified via gas chromatograph with triple quadrupole mass spectrometry (GC-TQMS) performed with a 7890A GC system (Agilent Technologies) containing a DB-FFAP column (30 m × 0.32 mm × 0.25 µm film thickness; Agilent Technologies). Culture samples (180 µl) were acidified with 10% formic acid (20 µl) in 2 ml screw cap glass vials. Liquid samples (2 µl) were injected manually. The oven was held at 60˚C for 1 min, then raised by 15˚C min^{-1} to 250˚C. Carrier gas was helium (2 ml min^{-1}). The TQMS was operated in MRM mode and the following precursor-product ion transitions were used for quantification: *m/z* 43 → 15.2 (unlabelled acetate), *m/z* 44 → 15.1 (1-^{13}C acetate), *m/z* 44 → 16 (2-^{13}C acetate), *m/z* 45 → 16.1 (1,2-^{13}C acetate.

Labelled and unlabelled HCO_3^- (as gaseous CO_2) was also quantified with GC-TQMS. Sample preparation and instrument parameters were the same as for unlabelled quantification with GC-PDD (Section 3.2.3.1). The TQMS was operated in MRM mode using the following transitions for quantification: *m/z* 45 → 29 ($^{13}CO_2$), *m/z* 44 → 28 ($^{12}CO_2$).

Culture samples being analysed for dimethylamine, methylamine, sarcosine, and glycine were derivatised based on the method by Villas-Bôas *et al* (2003). Briefly, liquid samples (50 µl) were combined with 1% sodium hydroxide solution (250 µl), 10 mM alanine (5 µl, as an internal standard), absolute ethanol (250 µl), and pyridine (50 µl) in a 1.5 ml screw-cap glass vial. Ethyl chloroformate (20 µl) was added to begin the reaction and the mixture was shaken for 20 s before adding a second aliquot of ethyl chloroformate and shaking again. DCM (1 ml) was added and the mixture was shaken for 10 s before the aqueous upper layer was discarded. Any remaining aqueous phase was removed by the addition of a small portion of anhydrous sodium sulphate. The organic solution was transferred to a fresh vial and capped for analysis.

Derivatised methylated amine and glycine samples were analysed on the same GC-TQMS system containing a DB-5 Column (30 m × 0.32 mm × 0.25 µm film thickness; Agilent Technologies). Liquid samples (1 µl) were injected with a 7693 autosampler (Agilent Technologies). The carrier gas was 1 ml min^{-1} helium. Inlet temperature was 200˚C; oven temperature was held at 50˚C for 2 min, raised by 10˚C min^{-1} to 180˚C. The TQMS was operated in multiple reaction monitoring (MRM) mode and the quantifying and qualifying precursor-product ion transitions are listed in Table 3.1. Samples were quantified by comparison to a minimum four-point standard curve, ranging from 1 – 15 mM.

Table 3.1. Compound-specific GC-TQMS method details for detection of methylated amines and glycines.

Compound	Quantifying transition (m/z)	Qualifying transition (m/z)	Elution time (min)	Collision energy (eV)
methylamine	103.2 → 74.9	103.2 → 74.0	8.00	5
dimethylamine	117.2 → 89.1	117.2 → 87.9	8.35	5
alanine	116.2 → 44.1	116.2 → 72.1	14.56	10
glycine	102 → 30.1	102 → 58.1	14.68	5
sarcosine	116.2 → 44.1	116.2 → 72.1	14.63	10

3.2.4 DNA extraction

Cells were harvested from 2 ml liquid culture by centrifugation at 10,000 rcf for 15 min at 4˚C. Supernatant (1,700 µl) was removed and the cell pellet was resuspended in the remaining 300 µl liquid. Samples were stored at -20˚C until required. Genomic DNA was extracted as described in Section 2.2.3.

3.2.5 Quantitative real-time PCR

Strain DCMF 16S rRNA genes were quantified via quantitative real-time PCR (qPCR) with primers Dcm775F (5'-AAGGCGACTTTCTGGACTGA-3') and Dcm930R (5'-GCGGGGTACTTATTGCGTTA-3') (Wong, 2015). Total bacterial 16S rRNA genes were quantified using the universal bacterial primers Eub1048F (5'-GTGSTGCAYGGYTGTCGTCA-3') and Eub1194R (5'-ACGTCRTCCMCACCTTCCTC-3') (Maeda *et al.*, 2003). qPCR reactions contained template DNA (2 µl), 2X SsoFast™ EvaGreen® Supermix (Bio-Rad, 5 µl), 100 nM each of forward and reverse primers (0.1 µl each of 10 mM stocks), 10 mg ml^{-1} bovine serum albumin (0.1 µl), and molecular grade water (2.7 µl), and were performed in triplicate for each sample. qPCR reactions were carried out in triplicate for each sample on a CFX96 thermal cycler (Bio-Rad) and the data was analysed with CFX Maestro v1.0 software (Bio-Rad). Technical repliates were only accepted if their standard deviation was <0.1, otherwise the qPCR process was repeated.

Standard curves were prepared by making serial 10-fold dilutions of plasmid DNA carrying cloned strain DCMF 16S rDNA or *Dehalococcoides* sp. 16S rDNA (for total bacterial quantification). Plasmids were constructed by cloning amplified genes into the pCR™2.1-TOPO® vector with TOPO TA Cloning Kit (Life Technologies) as per manufacturer's instructions. Vectors were inserted into One Shot® TOP10 *Escherichia coli* cells (Life Technologies). Plasmid DNA was extracted from overnight cutures of transformed cells using the PureYield™ Plasmid Miniprep System (Promega). The standard curve concentration ranged from 10^4 – 10^9 copies ml^{-1}. Strain DCMF 16S rRNA gene copy numbers were converted to cell numbers by dividing by four – the number of 16S rRNA genes in the genome.

3.3 Results

3.3.1 Morphological description and dominance of strain DCMF

FISH microscopy enabled selective visualisation of strain DCMF cells, which appeared to numerically dominate the DFE culture when amended with DCM (Figure 3.1C, F). Strain DCMF cells occurred singly or in end-to-end chains with a rod-shaped morphology (Figure 3.1A, D). On average, strain DCMF cells were 1.69 ± 0.27 μm long and 0.64 ± 0.12 μm wide. Counting the Cy3- and 6-FAM-labelled cells in three overlaid FISH images showed that strain DCMF represented 71 ± 3.8% of the total cells (Figure S2).

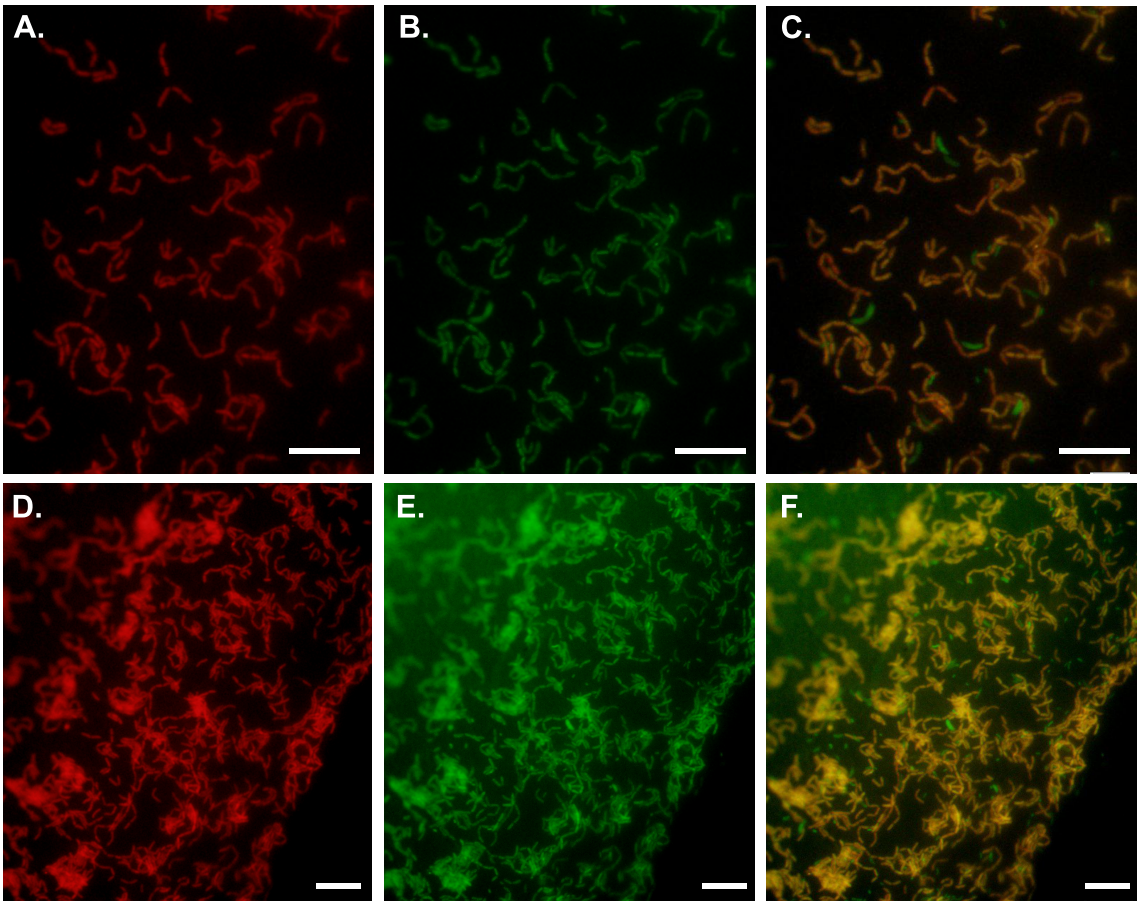


Figure 3.1 Strain DCMF is the dominant organism in DCM-amended cultures during exponential growth phase. Fluorescence in situ hybridisation (FISH) microscopy images with strain DCMF cells stained red with the Cy3-labelled Dcm623 probe (A and D), all bacterial cells stained green with the 6-FAM-labelled Eub338i probe (B and E), and the overlay of Cy3- and 6-FAM-labelling in these two pairs of images (C and E). The scale bars represent 10 μM.

3.3.2 Strain DCMF requires bicarbonate for growth with dichloromethane

Although the ability of strain DCMF to grow on DCM as a sole source of electrons has been reported in a previous thesis (Wong, 2015), this work was repeated in tandem with growth experiments on other substrates for the purpose of simultaneous comparison. In the present study, strain DCMF consumed 1.9 ± 0.0 mM DCM within 35 days, yielding 3.7 ± 2.2 × 10^8 cells ml^{-1} (Figure 3.2). The product of DCM fermentation was acetate (1.4 ± 0.1 mM), which was not observed in abiotic controls. Formate could be detected at low levels (37 ± 9.0 μM) at all stages of growth but did not accumulate with repeated DCM amendment (data not shown).

Growth on DCM yielded 2.0 ± 1.2 × 10^{14} cells per mole of substrate consumed (Table 3.2). This yield was converted into dry cell weight per mole substrate consumed by calculating the cell volume from the dimensions of strain DCMF cells reported in Section 3.3.1. It was assumed that the cells were cylindrical, that they had the same density as water, and that water constituted 80% of their mass. These calculations resulted in 24 ± 19 g dry strain DCMF cell material per mole DCM consumed (Table 3.2).

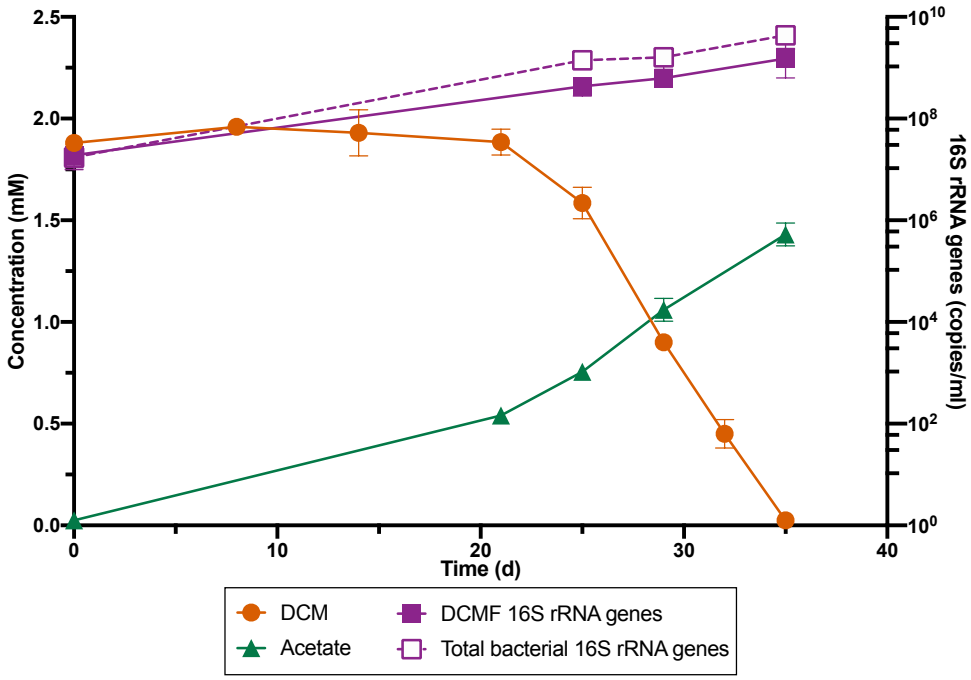


Figure 3.2 The consumption of DCM is concomitant with the production of acetate and an increase in strain DCMF 16S rRNA gene copies. Substrates and products are quantified on the left y-axis (linear scale), while strain DCMF and total bacterial 16S rRNA gene copy numbers are on the right y-axis ($\log_{10}$ scale). Error bars represent standard deviation (n = 2).

When transferred from bicarbonate-buffered to MOPS-buffered medium, strain DCMF required an exogenous source of bicarbonate in order to dechlorinate DCM (Figure 3.3). However, there was no significant decrease in bicarbonate concentration over the 65 days that DCM consumption was monitored (two-tailed unpaired t test between day 0 and day 65, $p = 0.11$), indicating that the culture likely produces bicarbonate in approximate stoichiometric equivalence to what it requires. No DCM consumption or substantial change in bicarbonate concentration was observed in bicarbonate-free abiotic controls (Figure 3.3).

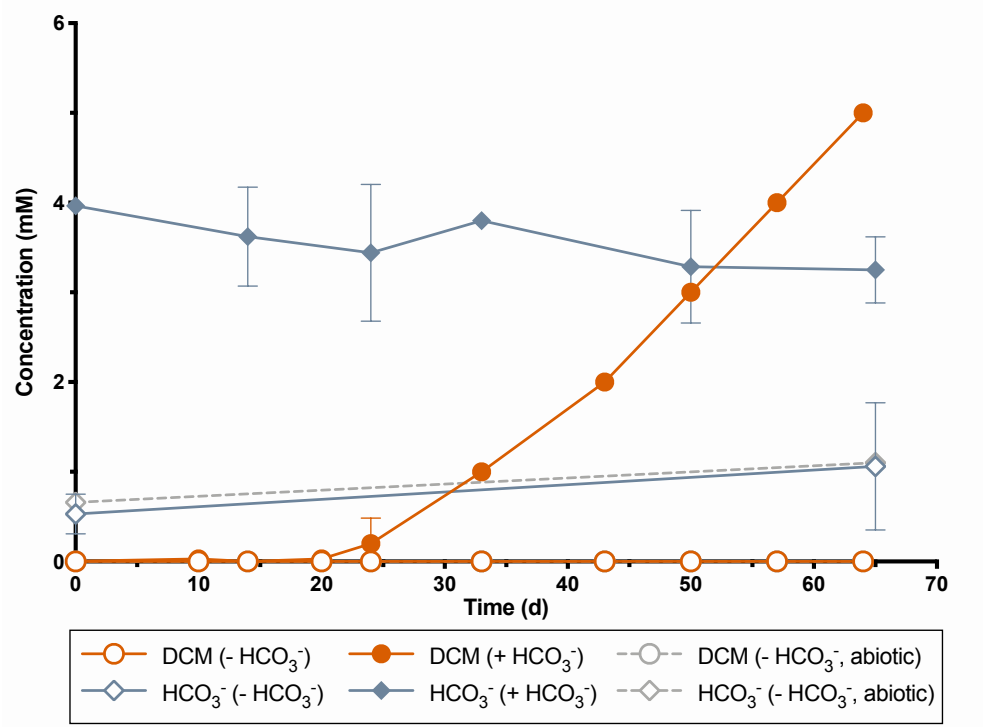


Figure 3.3 Strain DCMF requires an exogenous source of bicarbonate. The DFE culture was transferred into MOPS-buffered medium either with (filled shapes) or without (empty shapes) exogenous bicarbonate. DCM consumption (circles) was only observed in cultures amended with bicarbonate, although there was no significant change in bicarbonate concentration (diamonds) in these cultures over time. No DCM consumption or change in bicarbonate concentration was observed in abiotic, bicarbonate-free controls (dashed lines). Error bars represent standard deviation (n = 3 for biotic, n = 1 for abiotic cultures).

3.3.3 Carbon assimilation in strain DCMF

To ascertain the fate of DCM carbon, triplicate DFE cultures were amended with [^{13}C]DCM. After inoculation the initial concentration of acetate was 49 ± 11 μM, all of which was unlabelled. After 111 days, when 2,700 ± 328 μM DCM had been consumed, 666 ± 160 μM of acetate was produced (Figure 3.4A), of which 47.1 ± 5.5% was unlabelled, 30.4 ± 2.8% was labelled on the methyl group ([2-^{13}C]acetate), and 22.5 ± 4.3% was labelled on both the methyl and carboxyl groups ([1,2-^{13}C]acetate; Figure 3.4C).

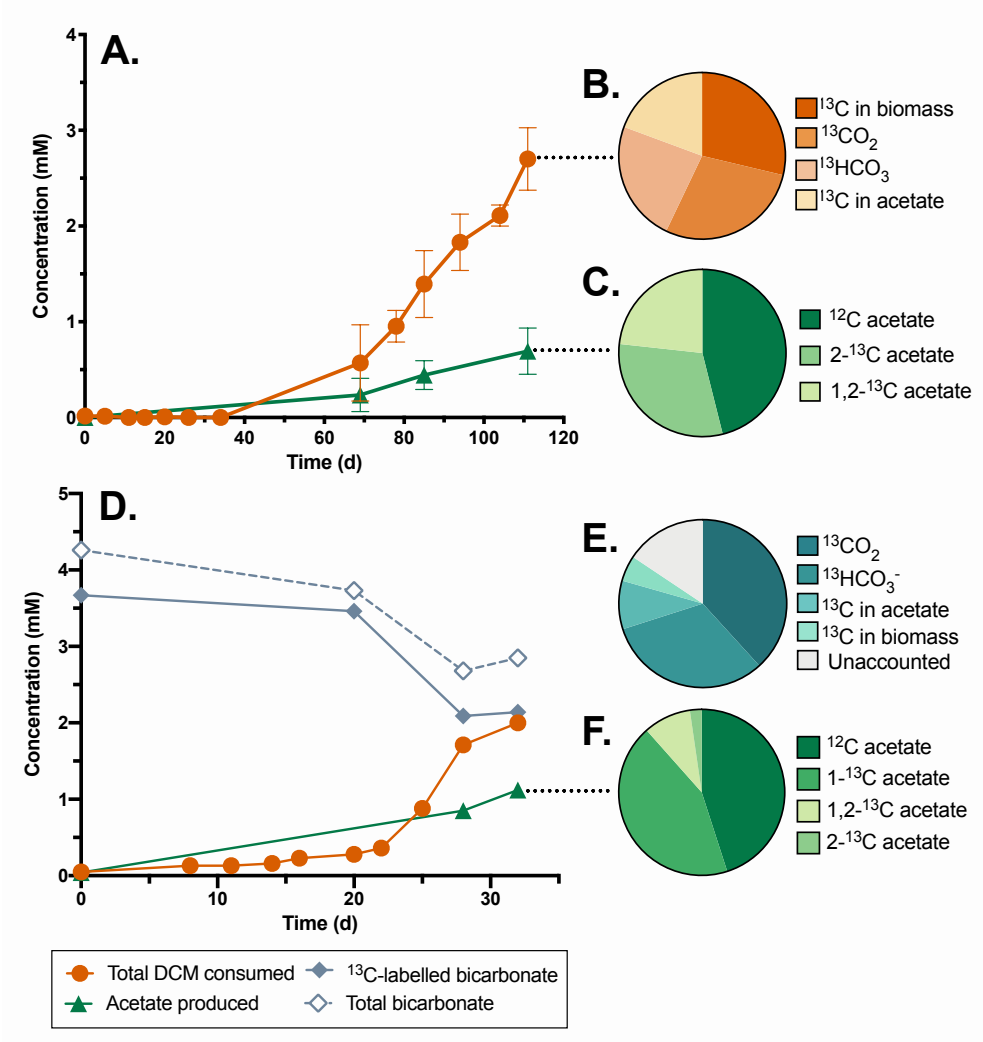


Figure 3.4 Strain DCMF assimilates carbon from DCM and bicarbonate to form acetate. **A.** Cumulative [^{13}C]DCM consumption with concomitant with acetate production. Error bars represent standard deviation, n = 3. **B.** The ^{13}C mass balance from [^{13}C]DCM showed 994 ± 121 µM ^{13}C in biomass, 982 ± 144 µM in $^{13}CO_2$, 815 ± 120 µM in $H^{13}CO_3^-$, and 670 ± 289 µM ^{13}C in acetate (128 ± 8.2% ^{13}C recovery). **C.** Of the total acetate produced from [^{13}C]DCM, 47.1 ± 5.5% was unlabelled, 30.4 ± 2.8% was [2-^{13}C]acetate, and 22.5 ± 4.3% was [1,2-^{13}C]acetate. **D.** Cumulative DCM consumption and acetate production in cultures amended with $H^{13}CO_3^-$; total (labelled and unlabelled) aqueous HCO_3^- is also shown (i.e. gaseous CO_2 is not accounted for here). Values are from a single representative culture. All triplicates had similar dechlorination rates and product concentrations but began dechlorinating at different times. **E.** ^{13}C mass balance from the $H^{13}CO_3^-$ amended cultures showed 2740 ± 204 µM $^{13}CO_2$, 2280 ± 170 µM $H^{13}CO_3^-$, 710 ± 9.74 µM ^{13}C in biomass, and 600 ± 84.9 µM ^{13}C in acetate, totalling 84.5 ± 7.0% ^{13}C recovery. **F.** Of the total acetate produced in DCM and $NaH^{13}CO_3$-amended cultures, 45.0 ± 2.3% was unlabelled, 43.5 ± 1.8% was [1-^{13}C]acetate, 9.3 ± 0.1% was [1,2-^{13}C]acetate, and 2.2 ± 1.3% was [2-^{13}C]acetate. All pie charts represent the average of triplicate cultures.

The concentration of ^{13}C-labelled HCO_3^- in the [^{13}C]DCM-amended cultures was also quantified. After 111 days, 820 ± 120 μM $H^{13}CO_3^-$ was detected in the cultures, although not in DCM-free or sterilised abiotic controls. Unlabelled HCO_3^- was in excess, as the culture was grown in 30 mM bicarbonate-buffered media. The concentration of $^{13}CO_2$ in the headspace of the cultures was calculated from this data using the Henry's Law dimensionless volatility constant (H_{cc} = 1.20 at 25°C, based on the mean H^{cp} reported in (Sander, 2015)). Additionally, the concentration of ^{13}C-labelled acetate equivalents in biomass were calculated based on data for strain DCMF in Table 3.2. A ^{13}C mass balance was achieved by summing the measured concentrations of ^{13}C-labelled carbon in acetate (670 ± 289 μM) and $H^{13}CO_3^-$ (815 ± 120 μM) with the calculated concentrations of $^{13}CO_2$ in the flask headspace (982 ± 144 μM and [^{13}C]acetate equivalents in biomass (994 ± 121 μM; Figure 3.4 B). This amounted to 128 ± 8.2% recovery of the labelled carbon amended via [^{13}C]DCM (2700 ± 328 μM). In summary, the ^{13}C label from DCM was found in [2-^{13}C]acetate, [1,2-^{13}C]acetate, and bicarbonate, indicating transformation of DCM via the Wood-Ljungdahl pathway. The near-complete recovery of the ^{13}C label also indicated no unknown fate of DCM in the DFE culture.

Analogous work was then carried out with unlabelled DCM in MOPS buffered medium amended with 7170 ± 441 μM ^{13}C-labelled bicarbonate (sum of aqueous bicarbonate shown in Figure 3.4 D and gaseous $^{13}CO_2$, calculated from the Henry's Law Constant as above). The culture consumed 2000 μM DCM and 2150 ± 492 μM ^{13}C from bicarbonate, producing 973 ± 140 μM acetate (Figure 3.4 D). Of the acetate produced, 45.0 ± 2.3% was unlabelled, 43.5 ± 1.8% was labelled on the carboxyl group ([1-^{13}C]acetate), 2.2 ± 1.3% was labelled on the methyl group, and 9.3 ± 0.1% was labelled on both carbons (Figure 3.4 F). A ^{13}C mass balance was achieved as before by summing the total labelled carbon in acetate (600 ± 84.9 μM) with the remaining $H^{13}CO_3^-$ (2280 ± 170 μM) and $^{13}CO_2$ (2740 ± 204 μM), and theoretical [^{13}C] in biomass (710 ± 9.74 μM). This amounted to 84.5 ± 7.0% recovery of the labelled carbon amended via $H^{13}CO_3^-$ (Figure 3.4 E). This definitively showed that strain DCMF incorporates carbon from CO_2 to form the carboxyl group of acetate and confirms operation of the Wood-Ljungdahl pathway in both oxidative and reductive directions, as both doubly labelled ([1,2-^{13}C]acetate) and unlabelled acetate were formed.

3.3.4 Strain DCMF can also grow on quaternary amines and methanol

Choline and glycine betaine were the first additional substrates found to support growth of strain DCMF. In the absence of any electron acceptor, the enrichment culture consumed choline (4.8 ± 0.2 mM) within 25 days, and 15 ± 0.6 mM acetate plus 6.0 ± 1.1 mM methylamine were produced (Figure 3.5 A). Glycine betaine (4.7 ± 0.3 mM) was consumed within 21 days with production of 11 ± 0.4 mM of acetate and 4.5 ± 0.6 mM methylamine (Figure 3.5 B). Trimethylamine, dimethylamine, sarcosine, and glycine were not detectable at any stage of growth. Neither acetate nor methylamine were detected in abiotic controls, and the latter was also absent from cultures grown with DCM. Minor amounts of acetate (0.61 ± 0.06 mM) and formate (0.50 ± 0.04 mM) were formed in duplicate no electron donor control cultures (data not shown).

Strain DCMF cell proliferation aligned with the consumption of these two substrates, yielding an increase of $1.4 \pm 0.4 \times 10^9$ and $5.3 \pm 0.4 \times 10^8$ cells per ml in choline- and glycine betaine-amended cultures, respectively, as determined by qPCR (Figure 3.5; strain DCMF 16S rRNA gene copies were divided by four – the number of 16S rRNA genes identified in the genome). These cell yields correspond to $3.0 \pm 0.9 \times 10^{14}$ cells per mole of choline utilised, and $1.1 \pm 0.1 \times 10^{14}$ cells per mole of glycine betaine utilised (Table 3.2). The high proportion of strain DCMF 16S rRNA gene copies to total bacterial 16S rRNA gene copies was consistent with strain DCMF as the dominant organism in the cultures at all stages of substrate consumption (Figure 3.5).

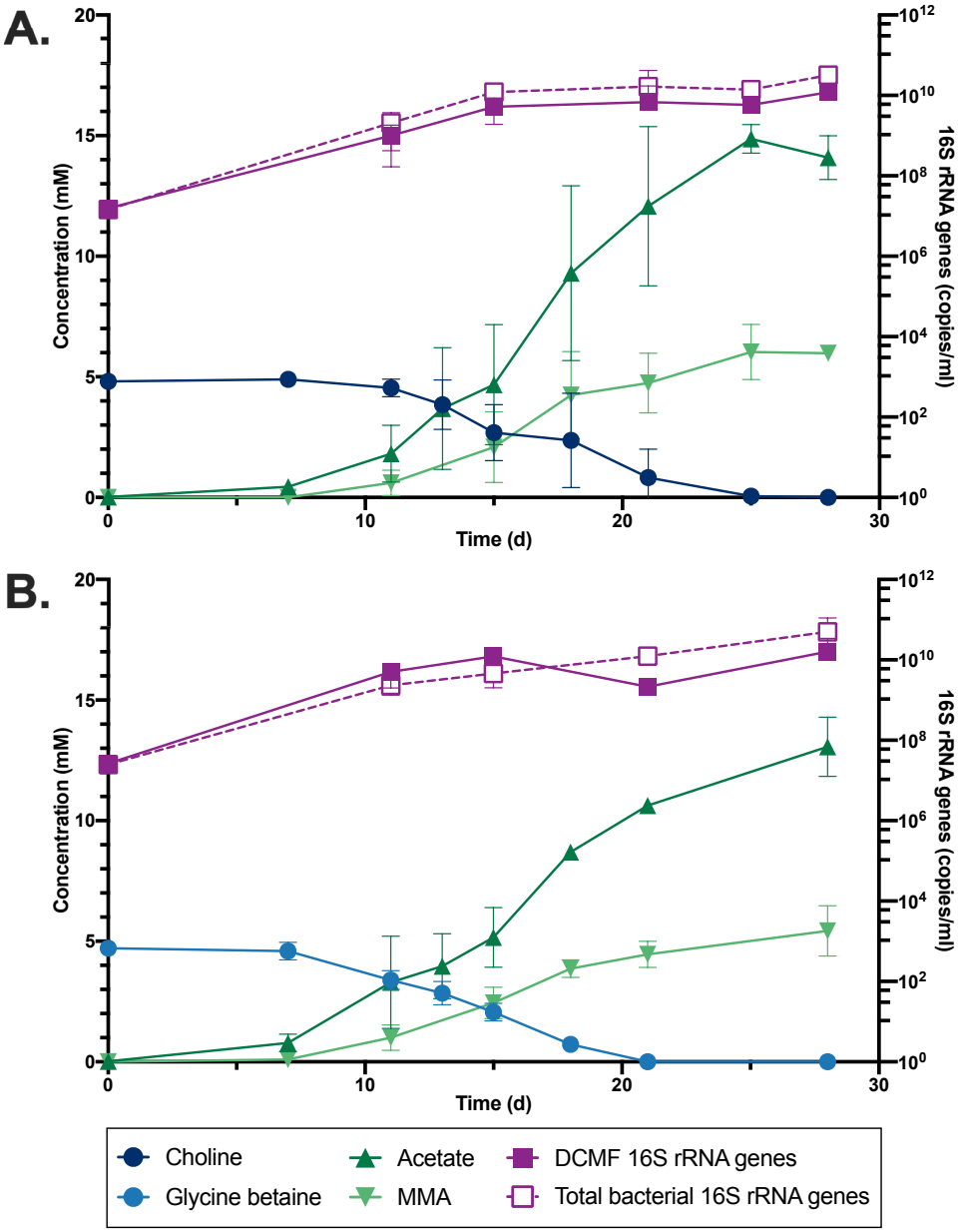


Figure 3.5 Strain DCMF can metabolise the quaternary amines choline and glycine betaine. In cultures amended with 5 mM **A.** choline or **B.** glycine betaine, substrate depletion was concomitant with an increase in acetate and methylamine (left y-axis), as well as strain DCMF and total bacterial 16S rRNA gene copies (right y-axis, $\log_{10}$ scale). Error bars represent standard deviation (n = 3).

DFE cultures amended with the putative quaternary amine metabolic pathway intermediates dimethylglycine and sarcosine (+ H_2) also demonstrated production of acetate and methylamine, which once again aligned with strain DCMF cell proliferation (Figure S3). Following the observation of strain DCMF growth and methylamine production in cultures amended with sarcosine + H_2, DFE cultures

were also set up with glycine betaine + H_2 to determine whether glycine betaine could be reductively cleaved to trimethylamine. These cultures consumed all glycine betaine (4.4 ± 0.4 mM) and hydrogen (7.9 ± 0.9 mM) within 28 days, producing 14.9 ± 0.6 mM acetate and 4.0 ± 0.4 mM methylamine, but no trimethylamine (Figure 3.6). This was concomitant with a strain DCMF cell yield of 4.0 ± 2.8 × 10^8 cells per ml, similar in magnitude to the yield on glycine betaine as the sole energy source.

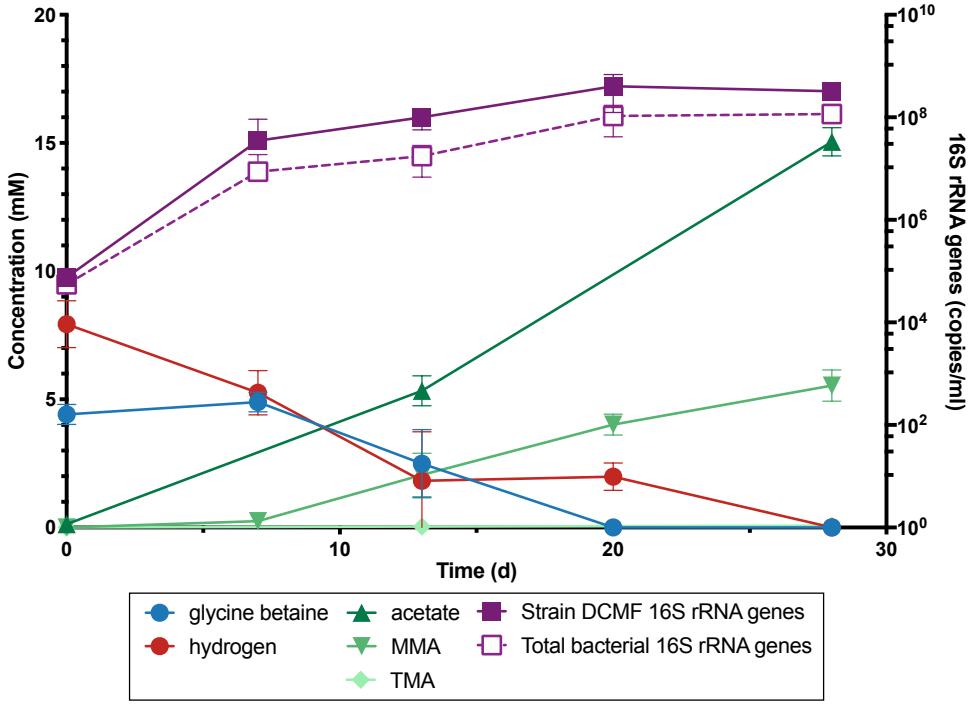


Figure 3.6 Strain DCMF does not produce trimethylamine from glycine betaine and H_2. Cultured amended with glycine betaine and hydrogen produced acetate and methylamine (left y-axis), concomitant with an increase in strain DCMF and total bacterial 16S rRNA gene copies (right y-axis, $\log_{10}$ scale). Error bars represent standard deviation (n = 3).

In addition to the quaternary amines, strain DCMF was found to grow on methanol as a sole source of electrons. The culture consumed 4.3 ± 0.2 mM methanol over 30 days, yielding 3.1 ± 0.1 mM acetate and 2.4 ± 0.6 × 10^9 strain DCMF cells per ml (5.7 ± 1.4 × 10^{14} cells per mole substrate utilised, corresponding to 66.7 ± 38.4 g dry cell weight) (Figure 3.7, Table 3.2). No methanol depletion was observed in the abiotic (cell-free) control, nor cell increase in the methanol-free control.

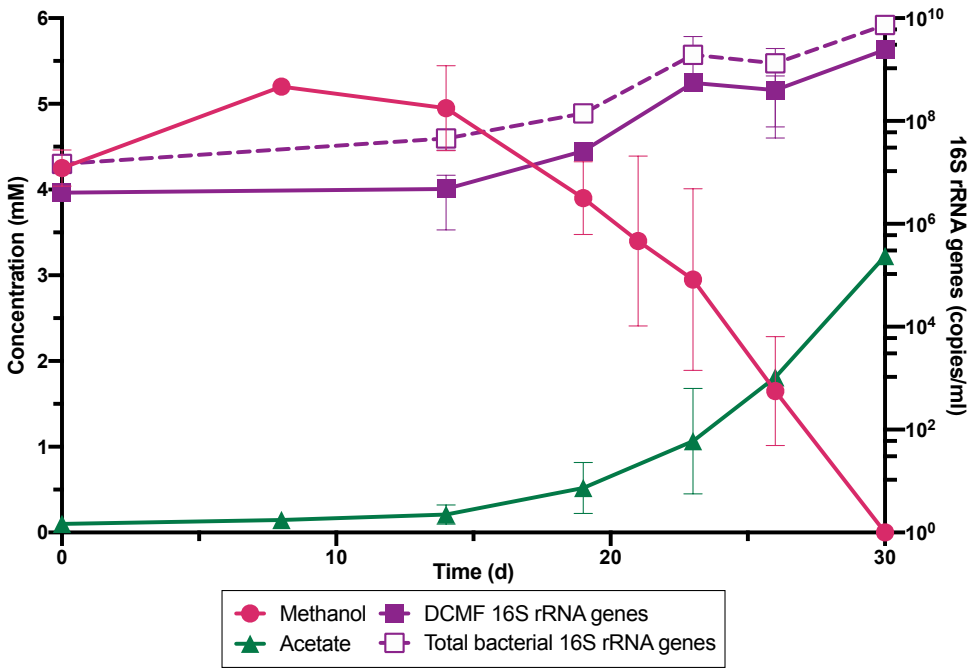


Figure 3.7 Acetate was the sole product of methanol consumption. The culture consumed methanol to produce acetate (left y-axis), concomitant with an increase in strain DCMF and total bacterial 16S rRNA gene copies (right y-axis, $\log_{10}$ scale). Error bars represent standard deviation (n = 2).

Strain DCMF was unable to utilise CO, ethanol, sarcosine or trimethylamine as sole energy sources, and unable to use any of the tested pairs of electron donors (choline, glycine betaine, lactate, trimethylamine) with electron acceptors (fumarate, Na_2SO_4, Na_2SO_3, $NaNO_2$, and $NaNO_3$).

3.4 Discussion

3.4.1 Strain DCMF growth on dichloromethane

Amongst the two other anaerobic DCM-dechlorinating bacteria (*Dehalobacterium formicoaceticum* and '*Candidatus* Dichloromethanomonas elyunquensis'), strain DCMF is unique in producing solely acetate as a fermentation product (Figure 3.2). *D. formicoaceticum* produced both formate and acetate in a 2:1 molar ratio (Mägli *et al.*, 1996), whilst '*Ca.* Dichloromethanomonas elyunquensis' mineralised DCM completely to H_2, CO_2 and Cl^- (Chen *et al.*, 2020). When supplied with DCM, strain DCMF yielded 2.0 ± 1.2× 10^{14} cells per mole substrate consumed (Table 3.2). This is an order of magnitude higher than the cell yields reported for '*Ca.* Dichloromethanomonas elyunquensis' (5.25 ± 1.00 × 10^{13} cells ml^{-1}; Kleindienst et

al., 2017) and *D. formicoaceticum* (3.73 ± 0.277 × 10^{13} cells ml^{-1}, Table 3.2) (Chen *et al.*, 2020). It is not yet clear why the cell yield for strain DCMF is higher than for the other two anaerobic DCM-dechlorinating bacteria.

Attempts to generate an axenic culture of strain DCMF were unsuccessful. The organism was unable to form colonies in soft agar shakes amended with DCM or on anaerobic agar plates amended with glycine betaine, leaving serial dilution to extinction in liquid medium as option of last resort. This was ultimately unsuccessful, implying that strain DCMF may require some as-yet unidentified cofactors from one or more of the culture cohabitants. How the cohabiting organisms in the DFE culture gain their energy is not clear, however the fermented yeast extract that is supplied as an undefined nutrient solution can be excluded, as it has been depleted of energy prior to its applications and serves as a source of co-factors only. The cohabiting organism are suspected to use cellular detritus resulting from expired strain DCMF cells as an energy source, as has recently been described in environmental *Spirochaetes* (Dong *et al.*, 2018). The growth and nature of the DFE culture cohabitants is explored in greater detail in Chapter 5.

3.4.2 ^{13}C-labelled carbon experiments support the use of the Wood-Ljungdahl pathway for dichloromethane transformation

Removal of bicarbonate from the culture medium precluded DCM dechlorination (Figure 3.3), a phenomenon that has also been observed in culture RM and axenic cultures of *D. formicoaceticum* (Kleindienst *et al.*, 2017; Chen *et al.*, 2020). The ensuing work with ^{13}C-labelled DCM and bicarbonate confirmed that strain DCMF is a mixotroph, i.e. it assimilates carbon from both DCM and CO_2 (Figure 3.4). Mixotrophy has also been demonstrated in *D. formicoaceticum* (Chen *et al.*, 2020) and is common among C1-utilising homoacetogens and methanogens (Schuchmann and Müller, 2016; Jones *et al.*, 2016; Yin *et al.*, 2019). This differs from the bicarbonate requirement of culture RM, in which CO_2 is required by the acetogenic and methanogenic organisms that consume H_2 produced by '*Ca.* Dichloromethanomonas elyunquensis', ensuring that DCM mineralisation remains thermodynamically favourable (Chen *et al.*, 2020).

Growth experiments using [^{13}C]DCM provide compelling evidence that strain DCMF employs the Wood-Ljungdahl pathway via incorporation of dehalogenated carbon into methylene-tetrahydrofolate (CH_2=THF) (Eq. 7). The ^{13}C label was found in HCO_3^-, [1-^{13}C]acetate, and [1,2-^{13}C]acetate (Figure 3.4C). The production of labelled HCO_3^- suggests that CH_2=THF is disproportionated into the Wood-Ljungdahl pathway where it is oxidised to HCO_3^- (Eq. 8, Figure 3.8). The electrons released then reduce the remaining CH_2=THF into the methyl group of acetate (Eq. 9). However, the production of unlabelled acetate (47%) indicates that the excess unlabelled HCO_3^- (30 mM) in the medium is an alternative electron acceptor to CH_2=THF for acetogenesis (Eq. 10). The reduction of HCO_3^- to acetate requires twice as many electrons for acetate synthesis than CH_2=THF (i.e. eight vs. four). Taking this ratio into account, along with ~1:1 ratio of unlabelled to labelled acetate suggests that approximately 67% of electrons derived from DCM oxidation were directed toward HCO_3^- reduction and 33% to CH_2=THF.

$$4\ CH_2Cl_2 + 4\ FH_4 \rightarrow 4\ CH_2{=}FH_2 + 8\ H^+ + 8\ Cl^- \qquad \text{Eq. 7}^2$$

$$3\ CH_2{=}FH_2 + 9\ H_2O \rightarrow 3\ HCO_3^- + 12\ e^- + 3\ FH_4 + 15\ H^+ \qquad \text{Eq. 8}$$

$$CH_2{=}FH_2 + 4\ e^- + HCO_3^- + 4\ H^+ \rightarrow CH_3COO^- + H_2O + FH_4 \qquad \text{Eq. 9}$$

$$2\ HCO_3^- + 9\ H^+ + 8\ e^- \rightarrow CH_3COO^- + 4\ H_2O \qquad \text{Eq. 10}$$

The production of [1,2-^{13}C]acetate is consistent with the reduction of HCO_3^- outlined above, as the DCM-fermenter could use $H^{13}CO_3^-$ produced from [^{13}C]DCM, via [^{13}C]formate (Figure 3.8). However, the proportion (22.5%) was surprisingly high, given the relatively small contribution that labelled $H^{13}CO_3^-$ from 2.7 mM [^{13}C]DCM would make to the 30 mM unlabelled HCO_3^- present in the culture medium. It is possible that co-localisation of Wood-Ljungdahl pathway proteins in the cytoplasm may cause the reduction of $H^{13}CO_3^-$ at a higher ratio than expected (i.e. 9%).

[^{13}C]DCM experiments carried out with cell extracts and axenic cultures of *D. formicoaceticum* showed the ^{13}C label was detected in formate and the methyl group of acetate ([2-^{13}C]acetate), as well as in methanol and glycine (Mägli *et al.*, 1998; Chen *et al.*, 2020). The lack of [1,2-^{13}C]acetate in *D. formicoaceticum* cultures is

[2] N.B. F = folate (not fluorine).

congruent with the observation that it produces formate as an end product of DCM degradation, and cannot further transform it into CO_2 (Mägli *et al.*, 1996), and hence also the lack of $^{13}CO_2$ observed in the more recent study carried out by Chen *et al* (Chen *et al.*, 2020). Another study found approximately 2.5-fold more [2-^{13}C]acetate than [1,2-^{13}C]acetate produced from [^{13}C]DCM, in a DCM-degrading mixed culture containing *Dehalobacterium* (Trueba-Santiso *et al.*, 2020).

DFE cultures amended with unlabelled DCM and ^{13}C-labelled HCO_3^- in MOPS buffered medium showed an analogous proportion of acetate labelled on the carboxyl group. A similar proportion of acetate (45.0%) observed in the [^{13}C]DCM work was unlabelled, in this case evidently formed using unlabelled HCO_3^- produced from DCM, while 43.5% of the acetate was labelled on the carboxyl group ([1-^{13}C]acetate; Figure 3.4F). The production of a small proportion (2.2%) of [2-^{13}C]acetate suggests that the Wood-Ljungdahl pathway operates in the reductive direction, to a small degree, as $H^{13}CO_3^-$ could be reduced via formate through the carbonyl branch to form $^{13}CH_2$=THF, which could then be reduced with unlabelled HCO_3^- produced from DCM (Figure 3.8).

The ^{13}C-labelling experiment supports the hypothesis that DCM metabolism involves the Wood-Ljungdahl pathway. This is congruent with the oxidation of formate to HCO_3^-, likely catalysed by a cytoplasmic formate dehydrogenase (112876, 112877, 112878s80), which was identified in the proteome. The production of HCO_3^- from formate balances with its uptake during acetogenesis, congruent with a net flux of approximately zero. This is despite the observation from that the organism requires HCO_3^- for growth (Figure 3.3). In light of these results, DCM is proposed to be transformed as per Equation 11.

$$2\ CH_2Cl_2 + 2H_2O \rightarrow CH_3COO^- + 5\ H^+ + 4\ Cl^- \qquad \text{Eq. 11}$$

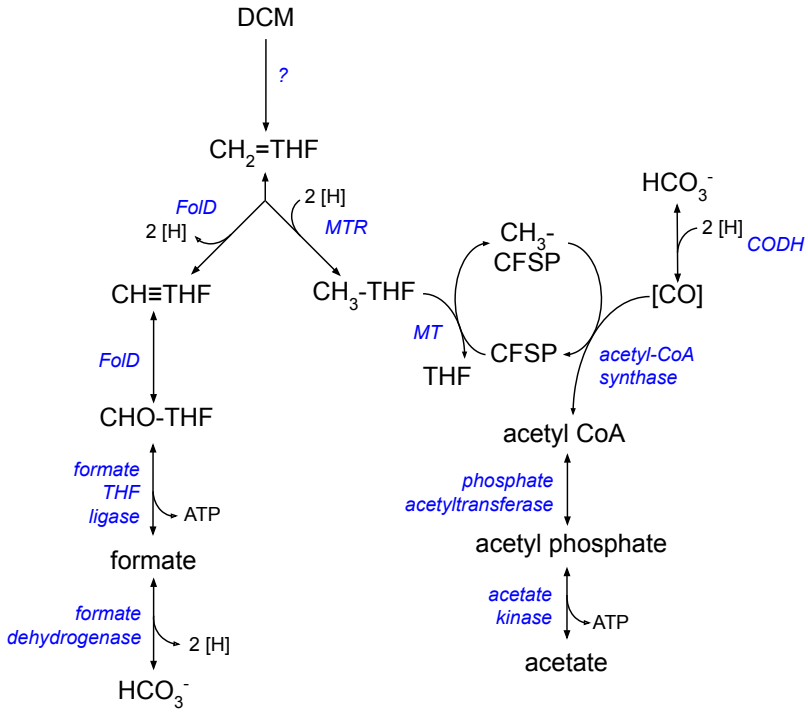


Figure 3.8 Putative DCM transformation pathway in strain DCMF. Enzymes are written in blue italics and have all been identified in the genome (Dataset S1), except for the unknown catalytic step between DCM and CH_2=THF. Abbreviations: CH_2=THF, 5,10-methylenetetrahydrofolate; CH_3-THF, methyltetrahydrofolate; CFSP, corrinoid iron-sulphur protein; CODH, carbon monoxide dehydrogenase; DCM, dichloromethane; FolD, methenyltetrahydrofolate cyclohydrolase/5,10-methylenetetrahydrofolate dehydrogenase; MT, methyltetrahydrofolate-corrinoid iron-sulphur protein Co-methyltransferase; MTR, 5,10-methylenetetrahydrafolate reductase; THF, tetrahydrofolate.

Formate oxidation distinguishes strain DCMF from *D. formicoaceticum*, in which DCM oxidation was shown to stop at formate (Eq. 12) (Mägli *et al.*, 1996, 1998). Indeed, it is unique amongst the three DCM-fermenting bacteria in producing only acetate. In '*Ca.* Dichloromethanomonas elyunquensis', DCM has recently been shown to be completely mineralised to H_2 and CO_2 (Eq. 13), which are only then utilised by methanogens and homoacetogens in the mixed culture to produce acetate and methane (Chen *et al.*, 2020). Accordingly, no labelled acetate was produced in ^{13}C[DCM] experiments with the latter culture. The ^{13}C labelling experiments described here confirm that, although strain DCMF is also present in a mixed culture, it produces acetate directly in DFE cultures.

$$3\ CH_2Cl_2 + 4\ H_2O + CO_2 \rightarrow CH_3COO^- + 2\ HCOO^- + 9\ H^+ + 6\ Cl^- \qquad \text{Eq. 12}$$

$$CH_2Cl_2 + 2\ H_2O \rightarrow CO_2 + 2\ H_2 + 2\ Cl^- + 2\ H^+ \qquad \text{Eq. 13}$$

3.4.3 A genome-based model for quaternary amine transformation in strain DCMF

The ability of strain DCMF to grow on choline, glycine betaine and methanol sets it apart from *D. formicoaceticum* and '*Ca.* Dichloromethanomonas elyunquensis', which have thus far been described as obligate DCM metabolising bacteria (Mägli *et al.*, 1996; Kleindienst *et al.*, 2017). The abundance of MttB superfamily genes (annotated as trimethylamine methyltransferases) in the strain DCMF genome initially led to trimethylamine being tested as substrate, both as sole source of electrons and as an electron donor paired with various electron acceptors. When no growth was observed with trimethylamine, glycine betaine was tested, as Ticak *et al* (2014) demonstrated that one non-pyrrolysine-containing "trimethylamine" methyltransferase in *Desulfitobacterium hafniense* Y51 was in fact a glycine betaine methyltransferase (*mtgB*).

3.4.3.1 Choline catabolism

Due to the importance of glycine betaine for protection against osmotic stress, the enzymes required to transform the more widely available compound choline into glycine betaine via betaine aldehyde are near ubiquitous in both terrestrial and aquatic bacteria (Wargo, 2013). Accordingly, strain DCMF could also utilise choline for growth and encodes both the choline dehydrogenase (Ga0180325_11215) and betaine aldehyde dehydrogenase (Ga0180325_114191) required for its transformation to glycine betaine (Figure 3.9).

Two other previously reported pathways for anaerobic choline metabolism are unlikely to be utilised by strain DCMF. It is possible for choline to be cleaved into trimethylamine and acetate via a choline-trimethylaminelyase (CutC) (Craciun and Balskus, 2012). However, while the strain DCMF genome does encode a putative glycyl radical enzyme similar to CutC (Ga0180325_112585), it contains only three of the six conserved residues predicted to be necessary for catalytic activity in other bacteria (Craciun and Balskus, 2012). Additionally, the lack of trimethylamine observed at any stage of growth makes catabolism of choline via the *cut* cluster of genes highly unlikely. Secondly, direct demethylation of choline to dimethylethanolamine has thus far only been reported in methanogenic Archaea from the genus *Methanococcoides* (Watkins *et al.*, 2012). The enzyme catalysing this

reaction has not yet been reported but could reasonably be assumed to be a methyltransferase within the MttB superfamily, which strain DMCF encodes in abundance. Further stepwise demethylation of dimethylethanolamine would yield ethanolamine, which can be transformed to acetaldehyde and ammonium within a bacterial microcompartment also encoded in the strain DCMF genome (Dataset S1). However, as all nitrogen in the provided choline was recovered within MMA (127% ± 19% N recovery), this pathway seems less likely than transformation of choline to glycine betaine in strain DCMF.

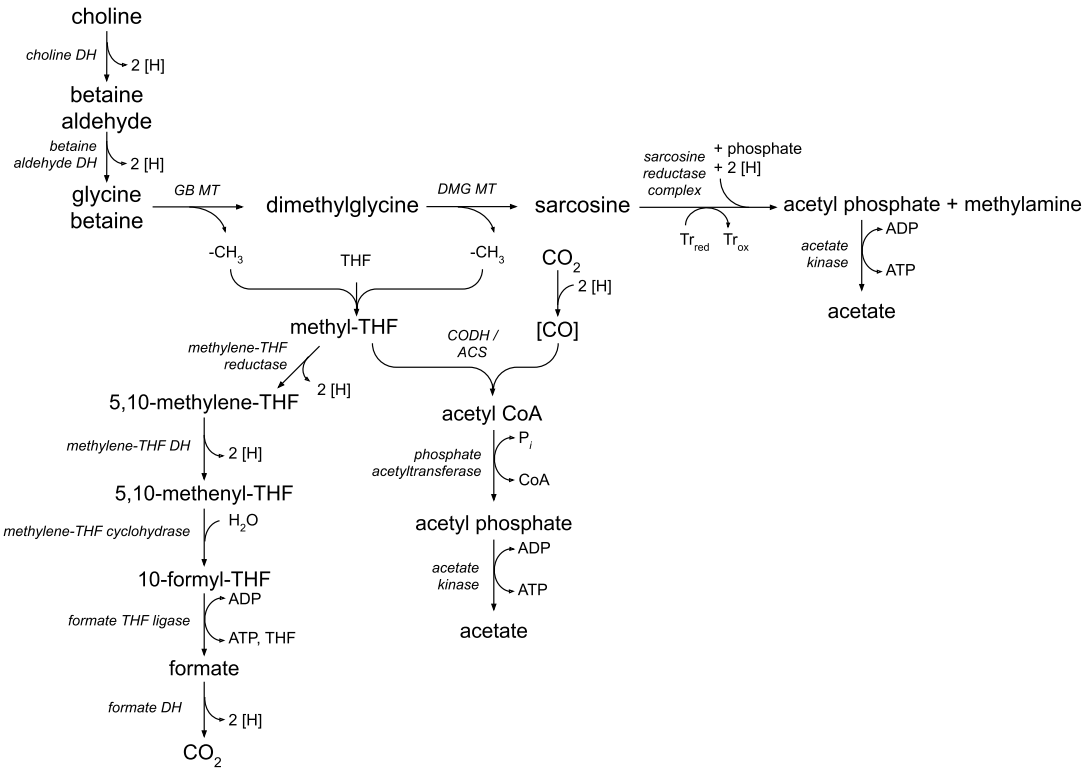


Figure 3.9 Metabolic model for strain DCMF growth on the quaternary amine compounds choline and glycine betaine. Enzymes are written in *italics* and have all been identified in the strain DCMF genome (Table S3). Redox cofactors are represented as electron equivalents ([H]) entering or leaving reactions. Acetate and methylamine are the primary products of both choline and glycine betaine degradation. Abbreviations: CH_3-THF, methyl-THF; CH≡THF, methenyl-THF; CODH/ACS, carbon monoxide dehydrogenase/acetyl coenzyme A synthase; DH, dehydrogenase; DMG MT, dimethylglycine methyltransferase; GB MT, glycine betaine methyltransferase; THF, tetrahydrofolate; Tr, thioredoxin.

3.4.3.2 Demethylation of glycine betaine

Glycine betaine (whether derived from choline or amended as substrate) can then be either demethylated to dimethylglycine or reductively cleaved to trimethylamine

and acetyl phosphate. The absence of trimethylamine in the cultures and accumulation of methylamine and acetate suggests that strain DCMF carries out a combination of demethylation and reductive cleavage, a variation thus far only observed in *Sporomusa* species (Möller *et al.*, 1984; Visser *et al.*, 2016). *Sporomusa ovata* demethylated a small fraction of glycine betaine to dimethylglycine and then sarcosine (methylglycine). Oxidation of the removed methyl groups provided reducing power for reductive cleavage of the majority of glycine betaine to trimethylamine (Möller *et al.*, 1984). The accumulation of methylamine in strain DCMF cultures suggests that it stepwise demethylates glycine betaine to dimethylglycine and then sarcosine, liberating electrons that can be used to reductively cleave sarcosine into methylamine and acetyl phosphate (Figure 3.9). This represents a novel metabolic pathway within the family *Peptococcaceae*.

Glycine betaine methyltransferases were first discovered in *D. hafniense* (Ticak *et al.*, 2014), but have since also been identified in *Sporomusa ovata* (Visser *et al.*, 2016) and *A. woodii* (Lechtenfeld *et al.*, 2018). The process involves a glycine betaine methyltransferase (MT_1, MtgB) that transfers a methyl group to a cognate corrinoid protein, and a methyl-tetrahydrofolate (THF) methyltransferase (MtgA, MT_2) that transfers the methyl group from the corrinoid protein to an accepting compound, THF (Ticak *et al.*, 2014). In *S. ovata* strain An4, the same two *mtgB* genes were suggested to carry out demethylation of both glycine betaine and dimethylglycine, forming sarcosine (Visser *et al.*, 2016), whilst in *A. woodii*, the protein appears to be specific to glycine betaine only, as there was no subsequent demethylation from dimethylglycine to sarcosine (Lechtenfeld *et al.*, 2018).

The MtgB homologs in strain DCMF with the highest identity to the proteins in *S. ovata*, *A. woodii*, and *D. hafniense* were all annotated as trimethylamine methyltransferases (consistent with their grouping in the MttB superfamily), and clustered together in a distinct clade in the MttB phylogenetic tree (Figure 2.3). Ga0180325_115483 had the highest percentage amino acid identity (54-55%) to the MT_1 proteins from these three species. However, it sits isolated on the opposite strand of DNA to all surrounding genes, thus making it an unlikely candidate. The second highest identity MT_1 homolog to all bar one species was Ga0180325_114740 (54-55% identity), which sits in a neighbourhood with three methyl-

tetrahydrofolate methyltransferases (shown to act as MT_2 proteins in *S. ovata* (Visser *et al.*, 2016)), two further MttB superfamily MT_1s, a cognate corrinoid protein, and two betaine/choline/carnitine transporter genes. This genetic neighbourhood contains all the requisite components for the demethylation system and is thus a good candidate for glycine betaine metabolism in strain DCMF.

The *Sporomusa* and *Acetobacterium* spp. glycine betaine MT_2 proteins had a lower percentage identity to the nearest homolog in strain DCMF (44-45% to Ga018325_111809), and even lower again to the second-best hit (32-36% to Ga018325_111232). This makes it difficult to determine any definitive candidates for a glycine betaine or dimethylglycine MT_2 gene in strain DCMF. Nonetheless, relevant MT_1s and methyltransferase cognate corrinoid proteins once again surround these lower identity homologs in the strain DCMF genome, making them further possible candidates for glycine betaine and/or dimethylglycine demethylation in strain DCMF.

3.4.3.3 Reductive cleavage

In anoxic environments, reductive cleavage of glycine betaine or sarcosine to trimethylamine or methylamine, respectively, is an alternative to demethylation. The strain DCMF genome encoded glycine/betaine/sarcosine reductases that were predicted to be specific to both glycine betaine (Ga0180325_115251, 115252s54) and sarcosine (Ga0810325_114802, 114803s), based on the alignment of the B component amino acid sequences to those of known substrate specificity from other organisms (Figure 2.8). However, production of trimethylamine was not observed when H_2 was provided as an electron donor for reductive cleavage of glycine betaine (Figure 3.6), suggesting that the putative glycine betaine reductase in strain DCMF may not be functional. Alternatively, it could be because transformation of glycine betaine to methylamine and acetate (Eq. 14, ΔG_f^0 = -403 kJ mol^{-1}) is energetically favourable while transformation into trimethylamine and acetate is not (Eq. 15; ΔG_f^0 = 12 kJ mol^{-1}).

$$(CH_3)_3N^+CH_2COO^- + H_2O + 0.5\ HCO_3^- \rightarrow 2.25\ CH_3COO^- + (CH_3)NH_3^+ + 0.75\ H^+ \quad \text{Eq. 14}$$

$$(CH_3)_3N^+CH_2COO^- + 2\ H^+ + 2\ e^- \rightarrow (CH_3)_3NH^+ + CH_3COO^- \quad \text{Eq. 15}$$

DFE cultures amended with sarcosine and H_2 were set up to help verify sarcosine as a pathway intermediate of choline and glycine betaine catabolism, as it could not be

observed at any stage of growth. The production of methylamine, acetate, and strain DCMF cells supported the organism's ability to reductively cleave sarcosine (Figure S3 B). While H_2 was required for the metabolism of sarcosine, the apparent ability of strain DCMF to utilise H_2 as an electron donor for reductive cleavage was seemingly at odds with its inability to grow with the classic acetogenic substrates H_2 + CO_2 (Wong, 2015). The genome does contain a putative membrane-bound NiFe hydrogen uptake hydrogenase (HyaABCD, Ga0180325_111497-9, 111503) which may be utilised to provide reducing equivalents for the sarcosine reductase.

In cultures amended with glycine betaine + H_2, the hydrogen was still consumed despite the lack of trimethylamine produced (Figure 3.6). It is not yet clear which organisms in the culture were utilising the hydrogen, as previous DFE cultures amended with H_2 + CO_2 demonstrated no growth or acetogenesis (Wong, 2015). Observations of the hydrogen only (i.e. glycine betaine and sarcosine-free) control culture in this experiment were in agreement with this: there was no significant decrease in hydrogen concentration or increase in acetate concentration (linear regression slope was not significantly non-zero, p-value > 0.05 for both compounds; data not shown). Concurrently, there was slightly higher acetate production observed in the glycine betaine + H_2 cultures (15 ± 0.6 mM), compared to the hydrogen-free glycine betaine cultures (11 ± 0.4 mM; Figure 3.5 B). It may be possible that CO_2 reduction by strain DCMF is enabled in the presence of glycine betaine and/or sarcosine, once other metabolic components (i.e. the Wood-Ljungdahl pathway) are in use. A proteomic analysis of hydrogen-amended cultures may shed further light on the ability of strain DCMF to utilise hydrogen as an electron donor.

Table 3.2 Cell yield and total acetate calculations for strain DCMF and other bacteria. nd = not described.

Organism + substrate	Cell length (µm)	Cell width (µm)	Cell volume (cm^3)[a]	Per mole substrate consumed: Cell increase (cells ml^{-1})	Cell wet weight (g)[b]	Cell dry weight (g)[c]	Acetate equivalents in biomass (mol)[d]	Acetate produced (mol)	Total acetate (mol)	Reference
Strain DCMF + DCM	1.7 ± 0.3	0.6 ± 0.1	$5.9 \pm 3.0 \times 10^{-13}$	$2.0 \pm 1.2 \times 10^{14}$	119 ± 93	24 ± 19	0.5 ± 0.4	0.8 ± 0.0	1.2 ± 0.4	This study
Strain DCMF + choline	1.7 ± 0.3	0.6 ± 0.1	$5.9 \pm 3.0 \times 10^{-13}$	$3.0 \pm 0.9 \times 10^{14}$	178 ± 105	36 ± 21	0.7 ± 0.4	3.1 ± 0.1	3.9 ± 0.5	This study
Strain DCMF + glycine betaine	1.7 ± 0.3	0.6 ± 0.1	$5.9 \pm 3.0 \times 10^{-13}$	$1.1 \pm 0.1 \times 10^{14}$	66 ± 35	13 ± 7.0	0.3 ± 0.1	2.3 ± 0.1	2.5 ± 0.2	This study
Strain DCMF + methanol	1.7 ± 0.3	0.6 ± 0.1	$5.9 \pm 3.0 \times 10^{-13}$	$5.7 \pm 1.4 \times 10^{14}$	334 ± 190	67 ± 38	1.4 ± 0.8	0.7 ± 0.0	2.1 ± 0.8	This study
'*Ca.* Dichloromethanomonas elyunquensis' + DCM	4.0 ± 0.8	0.4 ± 0.1	$4.6 \pm 1.2 \times 10^{-13}$	$5.3 \pm 1.0 \times 10^{13}$	23.9 ± 1.2	4.8 ± 0.2	0.1 ± 0.0	0[e]	nd	(Kleindienst *et al.*, 2017; Chen *et al.*, 2020)
Dehalobacterium formicoaceticum + DCM	1.8	1.1	1.7×10^{-12}	$3.7 \pm 0.3 \times 10^{13}$	62.6	12.5	0.26	0.16	0.42	(Mägli *et al.*, 1996; Chen *et al.*, 2020)
Eubacterium limosum 11A + glycine betaine	3.3	0.8	1.4×10^{-12}	nd	45.0	9.0	0.19	0.20	0.39	(Müller *et al.*, 1981)
Sporomusa sphaeroides E + glycine betaine	3.0	0.7	1.0×10^{-12}	nd	nd	nd	nd	0.86	nd	(Möller *et al.*, 1984)
Acetobacterium woodii NZva16 + methanol	1.9	0.9	1.2×10^{-12}	nd	17.5	3.5	0.07	0.68	0.75	(Tschech and Pfennig, 1984; Bache and Pfennig, 1981)

[a] Cells are assumed to be cylinders.
[b] Cells are assumed to have the same density as water.
[c] Cells are assumed to be 80% water.
[d] 1 mg dry cell weight assumed to be equal to 20.6 µmol acetate (Schink and Pfennig, 1982)
[e] No acetate is produced directly by '*Ca.* Dichloromethanomonas elyunquensis'. Rather, it is produced by acetogenic bacteria utilising the mineralisation products H_2 and CO_2.

3.4.3.4 A summary of quaternary amine metabolism in strain DCMF

Product formation and cell yields from the growth experiments with choline and glycine betaine were drawn together with the genomic information presented in Chapter 2 to generate a theoretical energy balance for consumption of these substrates. To this end, biomass produced was converted into acetate equivalents, assuming 1 mg dry cell weight equates to 20.6 μmol acetate (Schink and Pfennig, 1982), and summed with the quantified acetate in solution in order to calculate the "total acetate" produced by DFE cultures (Table 3.2).

The oxidation of two methyl groups from glycine betaine would yield 12 electrons (Eq. 16), of which two can be directed to reductive cleavage of sarcosine to yield one acetate and methylamine (Eq. 17). Given that sarcosine was not observed at any stage of growth, it is presumably rapidly cleaved. As eight electrons are required for acetate synthesis from bicarbonate (Eq. 18), the remaining 10 electrons equate to 1.25 acetate equivalents via bicarbonate reduction (Eq. 19), totalling 2.25 mol acetate equivalents and 1 mol methylamine per mole glycine betaine (Eq. 20; Figure 3.9). This approximately accords with the observed total acetate equivalent (2.3 ± 0.1 mM) and methylamine (0.9 ± 0.1 mM) concentrations in glycine betaine-amended cultures.

The methylamine yield in choline-amended cultures (1.3 ± 0.2 mM per mole choline utilised), was also close to the theoretical yield based on the above equations. The metabolism of choline into glycine betaine liberates four electrons (Eq. 21), which equate to an additional 0.625 mol acetate for each mol of choline metabolised to glycine betaine (Eq. 22). Combining the choline to glycine betaine, and glycine betaine to acetate and methylamine equations results in a theoretical yield of 2.75 mol acetate equivalents per mole choline (Eq. 23), which is within one standard deviation of the observed 3.1 ± 0.4 mM acetate equivalents.

$$(CH_3)_3N^+CH_2COO^- + 6\ H_2O \rightarrow (CH_3)NH_2^+CH_2COO^- + 2\ HCO_3^- + 14\ H^+ + 12\ e^- \quad \text{Eq. 16}$$

$$(CH_3)NH_2^+CH_2COO^- + 2\ H^+ + 2\ e^- \rightarrow (CH_3)NH_3^+ + CH_3COO^- \quad \text{Eq. 17}$$

$$2\ HCO_3^- + 9\ H^+ + 8\ e^- \rightarrow CH_3COO^- + 4\ H_2O \quad \text{Eq. 18}$$

$$2.5\ HCO_3^- + 11.25\ H^+ + 10\ e^- \rightarrow 1.25\ CH_3COO^- + 5\ H_2O \quad \text{Eq. 19}$$

$$(CH_3)_3N^+CH_2COO^- + H_2O + 0.5\ HCO_3^- \rightarrow 2.25\ CH_3COO^- + (CH_3)NH_3^+ + 0.75\ H^+ \quad \text{Eq. 20}$$

$$(CH_3)_3N^+CH_2CH_2OH + H_2O \rightarrow (CH_3)_3N^+CH_2COO^- + 5\ H^+ + 4\ e^- \quad \text{Eq. 21}$$

$$HCO_3^- + 4.5\ H^+ + 4\ e^- \rightarrow 0.5\ CH_3COO^- + 2\ H_2O \quad \text{Eq. 22}$$

$$(CH_3)_3N^+CH_2CH_2OH + 1.5\ HCO_3^- \rightarrow CH_3NH_3^+ + 2.75\ CH_3COO^- + 1.25\ H^+ \quad \text{Eq. 23}$$

The conversion of choline to acetate and methylamine (Eq. 23) is thermodynamically favourable, with a ΔG_f of -149 kJ mol^{-1}, as is the conversion of glycine betaine to acetate and methylamine (ΔG_f of -403 kJ mol^{-1}; Eq. 20). The cell dry weight yield of strain DCMF with glycine betaine (13 ± 7.0 g per mol substrate consumed) is higher than the previously reported value for *Eubacterium limosum* strain 11A (9.0 g per mol substrate consumed; Table 3.2). However, this organism produced dimethylglycine and acetate as end products of glycine betaine fermentation, indicating only a single demethylation took place (Müller *et al.*, 1981). This would yield fewer reducing equivalents for additional acetate and biomass production and is also in accordance with the less thermodynamically favourable ΔG_f of -183 kJ mol^{-1}. In fact, given that demethylation from glycine betaine to dimethylglycine yields six electrons, and complete conversion of glycine betaine to methylamine and acetate yields 18 electrons, *E. limosum* had a higher growth yield per mole electrons released than strain DCMF (1.5 g cell dry weight compared to 0.72 g, respectively).

3.4.4 A genome-based model for methanol metabolism in strain DCMF

Methanol catabolism in strain DCMF is proposed to be carried out through a similar three-component methyltransferase system to that proposed above for glycine betaine demethylation. While such methanol methyltransferase systems are relatively well-described in methanogenic archaea (van der Meijden et al., 1983a,b 1984; Burke & Krzycki, 1995, 1997; Sauer & Thauer, 1997; Sauer, Harms & Thauer, 1997; Hagemeier et al., 2006), there are only a few reports from acetogenic bacteria, namely in *Moorella thermoacetica* (Das *et al.*, 2007), *S. ovata* (Visser *et al.*, 2016) and *A. woodii* (Kremp *et al.*, 2018). The strain DCMF genome encodes a number of methanol-specific methyltransferases and associated corrinoid proteins (Table S4). One of these, a methanol:corrinoid methyltransferase (Ga0180325_112644), is also the closest homolog when the MT_1 (MtaB) proteins from both *S. ovata* and *A. woodii* were searched against the strain DCMF genome. It resides in a cluster containing a

methanol-specific MT_2 homolog (Ga0180325_112641) and a *mtbC* CoP homolog (Ga0180325_112642 and 112645). As in both *S. ovata* and *A. woodii*, the putative MT_2 gene is a methyl-THF methyltransferase, rather than the MtaA methanol methyltransferase found in methanogens (Visser *et al.*, 2016; Kremp *et al.*, 2018).

In *A. woodii*, the methanol-specific methyltransferase complex transfers the methyl group onto THF, forming methyl-THF, which can then be transformed via the Wood-Ljungdahl pathway (Kremp *et al.*, 2018). Strain DCMF is expected to follow a similar metabolic route, leading to the formation of acetate as the sole product. The removed methyl group can enter the Wood-Ljungdahl pathway as methyl-THF, from which metabolism can proceed in a similar manner to that outlined in Figure 3.9 for methyl groups removed from glycine betaine and dimethylglycine. In *A. woodii*, ~25% of the methyl-THF to CO_2 in the methyl branch of the Wood-Ljungdahl pathway produces the electron equivalents that are used to reduce CO_2 to CO in the carbonyl branch, through which the remaining ~75% of the methyl-THF is converted into acetate (Kremp *et al.*, 2018). In order to maintain redox balance of this pathway, *A. woodii* utilises a soluble electron-bifurcating hydrogenase to convert H_2 into reduced ferredoxin and NADH, and hydrolyses ATP to pump sodium ions into the periplasmic space so that they can re-enter the cytoplasm via the Rnf complex, which generates additional reduced ferredoxin at the expense of NADH (Kremp *et al.*, 2018). Strain DCMF harbours all genes for an F_1F_0-type ATPase (Ga0180325_113152-60) and the Rnf complex (Ga0180325_113065-70) in its genome. Furthermore, it contains genes homologous to those encoding the electron-bifurcating hydrogenase in *A. woodii* (Ga0180325_111565 – 111569; 39 – 69% amino acid identity). Thus, it is possible that methanol catabolism in strain DCMF uses a similar process for redox balance as that previously described in *A. woodii*.

Strain DCMF produced 4.9×10^{14} cells per mole methanol utilised, the largest growth yield out of the four substrates (DCM, choline, glycine betaine; Table 3.2). When this growth was converted into biomass equivalents as demonstrated in Section 3.4.3.4, it corresponded to 1.4 ± 0.8 mol acetate equivalents per mol methane utilised (Table 3.2). Given that only 0.7 ± 0.0 mol acetate was measured per mole methanol utilised, this corresponds to 65% of all potential acetate being transformed into biomass – a relatively large amount compared to strain DCMF

growth on DCM, choline, or glycine betaine (Table 3.2). In cultures of *A. woodii*, the cell yield on methanol was more than double that from H_2 + CO_2 (Bache and Pfennig, 1981; Tschech and Pfennig, 1984). This has since been attributed to the yield of 0.83 mol ATP per mol acetate formed from the catabolic pathway, the highest reported so far for an acetogenic C1 substrate (Kremp *et al.*, 2018). Indeed, even compared to multi-carbon substrates such as lactate (Weghoff *et al.*, 2015), ethanol (Bertsch *et al.*, 2016), butanediol (Hess *et al.*, 2015), and ethylene glycol (Trifunović *et al.*, 2016), the ATP yield per mol acetate formed from methanol was second only to that of fructose (Schuchmann and Müller, 2016). This may explain the higher strain DCMF cell yield on methanol than the more complex substrates choline and glycine betaine.

3.4.5 Quaternary amine and methanol metabolism have important implications for carbon and nitrogen cycling in the environment

The ability of strain DCMF to utilise choline, glycine betaine and methanol suggests that its environmental relevance extends beyond DCM contaminated sites. Coastal salt marshes and intertidal mudflats represent significant sources of methane from the demethylation of trimethylamine, which is in turn, derived from quaternary amines (Jameson *et al.*, 2019; King, 1984, 1988a) (Figure 3.10). Both trimethylamine and methanol are non-competitive methane precursors (i.e. not typically used by non-methanogenic microbes in anoxic environments; Figure 3.10), which may allow large methanogen populations to develop in environments where sulphate reduction would typically dominate (Oremland *et al.*, 1982; Oremland and Polcin, 1982). Indeed, trimethylamine is responsible for 60 - 90% of methane production in coastal salt marshes and intertidal sediments (Oremland *et al.*, 1982; King, 1984). Concurrently, areas of higher salt concentration (where glycine betaine is highly prevalent as an osmotic regulator) are typically more inhibitory to methanotrophs than methanogens (Iversen, 1996; Cohen *et al.*, 1994; Denier-VanderGon and Neue, 1995), which can further fuel an increase in atmospheric methane flux from these environs. The transformation of quaternary amines to methylamine by strain DCMF provides a pathway of lower methanogenic potential that could operate in coastal environments (Figure 3.10). Strain DMCF does produce acetate as a major end

product, which can be utilised by acetoclastic methanogens. However, unlike methylated amines, methanogens have to compete with more thermodynamically favourable processes such as sulphate reduction for this substrate. Overall, the abundance of strain DCMF, its role in quaternary amine and methanol turnover, and effect on methanogenesis in coastal environments has yet to be investigated.

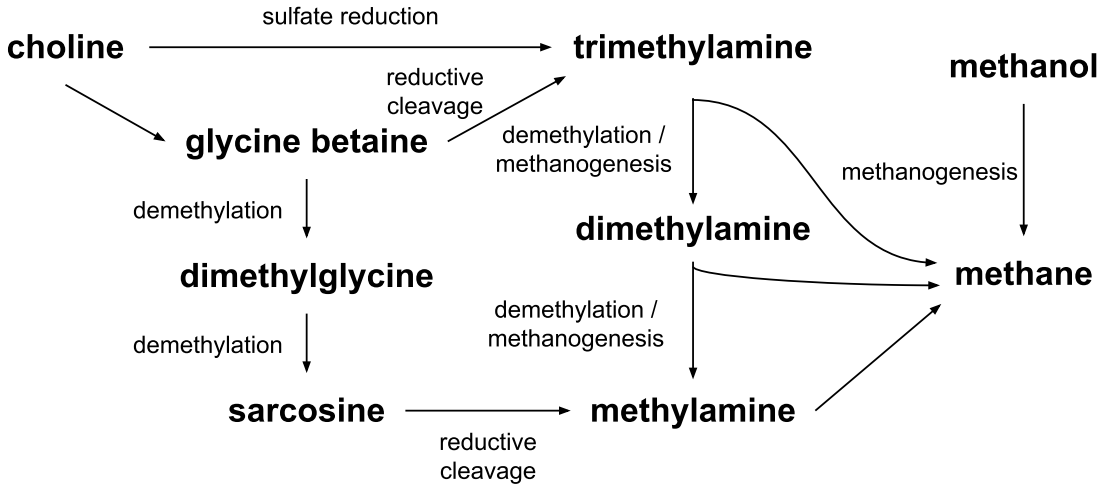


Figure 3.10 Overview of the metabolic processes involving quaternary amines and methanol in anaerobic, coastal environments.

3.4.6 Classification of strain DCMF as '*Candidatus* Formamonas warabiya' gen. nov. sp. nov.

Based on the genomic information in Chapter 2, it seemed likely that strain DCMF represented a novel genus within the family *Peptococcaceae*. By measures of 16S rRNA gene phylogeny, whole genome analysis of universally conserved marker genes, and amino acid identity, its closest relative was shown to be *Dehalobacterium formicoaceticum* strain DMC. However, the physiological information presented in this Chapter distinguishes strain DCMF from the sole representative of the genus *Dehalobacterium*, which has thus far only proved capable of growth on DCM (Mägli *et al.*, 1996). The metabolic repertoire of strain DCMF is a unique range of one-carbon compounds (DCM, methanol) or substrates from which it can utilise methyl groups (choline, glycine betaine, dimethylglycine, sarcosine + H_2). Strain DCMF also harbours a significantly larger genome than *D. formicoaceticum* (6.44 Mb for the former, 3.77 Mb for the latter) (Chen *et al.*, 2017b).

Thus, multiple lines of evidence support the placement of strain DCMF within a novel genus within the family *Peptococcaceae*. As strain DCMF is not yet represented in pure culture despite intensive efforts to isolate the organism, we propose it be classified in the *Candidatus* category. This is suggested for putative taxa for which there is enough evidence to justify classification, without being able to meet all requirements of the International Code of Nomenclature of Bacteria (Murray and Stackebrandt, 1995).

3.4.6.1 Description of '*Candidatus* Formamonas gen. nov.

'*Candidatus* Formamonas [Form.a.mon'as. L. adj. *formicum* relating to formic acid or, more generally, one-carbon compounds; Gr. n. *monas* unit; ML. n. *Formamonas* the one-carbon utilising unit.

'*Candidatus* Formamonas' is strictly anaerobic and metabolises one-carbon and methylated compounds including DCM, methanol and quaternary amines glycine. Methylene/methyl groups are metabolised via incorporation into the Wood-Ljungdahl pathway. The type species is '*Candidatus* Formamonas warabiya'.

3.4.6.2 Description of '*Candidatus* Formamonas warabiya sp. nov.

'*Candidatus* Formamonas warabiya [war.a.bi'ya N.L. n. *warabiya* the Dharawal name for the area between Botany Bay and Bunnerong, honouring the Traditional Custodians of the land where this bacterium was sampled from]. Permission was granted from the Dharawal Language Program for use of this placename as the bacterial species name (Appendix A).

Utilises DCM, methanol, choline, glycine betaine, dimethylglycine as sole sources of electrons under anoxic conditions. Can also utilise the electron donor and acceptor pair H_2 and sarcosine. The aforementioned substrates plus CO_2 are carbon sources. The primary fermentation product is acetate. Methylamine is also produced from choline, glycine betaine, dimethylglycine, and sarcosine + H_2. The Wood-Ljungdahl pathway is likely used for carbon fixation and metabolism of the methyl groups removed from substrates. Cells are rod shaped (1.69 × 0.27 μm).

Type strain DCMF^T is not available in pure culture. The source of inoculum was contaminated sediment from the Botany Sands aquifer, adjacent to Botany Bay, Sydney, Australia. The type material is the finished genome of '*Candidatus*

Formamonas warabiya' strain DCMF, which is 6.44 Mb and has a G+C content of 46.4% (GenBank accession number CP017634.1; IMG genome ID 2718217647).

3.5 Conclusions

Strain DCMF is a novel member of the family *Peptococcaceae* that is capable of fermenting the toxic groundwater pollutant DCM into the innocuous end product acetate. During growth on DCM, strain DCMF was the most abundant organism in the DFE culture, as visualised by FISH and qPCR data comparing strain DCMF to total bacterial 16S rRNA genes. Stable isotope (^{13}C) and bicarbonate-free culture experiments demonstrated the role of the Wood-Ljungdahl pathway in DCM transformation and confirmed that strain DCMF is a mixotroph, assimilating carbon from DCM and CO_2 for acetogenesis.

In contrast to *D. formicoaceticum* and '*Ca.* Dichloromethanomonas elyunquensis', which have thus far been described as obligate DCM-metabolising organisms, strain DCMF is also capable of utilising quaternary amines (choline and glycine betaine) and methanol for growth. Physiological observations of quaternary amine metabolism were supported with genomic data, which show the presence of glycine betaine methyltransferases and glycine/betaine/sarcosine reductases in the strain DCMF genome. As such, it is proposed that strain DCMF converts choline into glycine betaine. This is then demethylated in a stepwise manner to dimethylglycine and then sarcosine (methylglycine), which can then be cleaved by a sarcosine reductase to form methylamine and acetyl phosphate. The genome also encodes methanol methyltransferases likely utilised to convert methanol into methyl-THF, which can then be metabolised to acetate via the Wood-Ljungdahl pathway. Based on the novelty of the strain DCMF genome (including 16S rRNA gene sequence) and the organism's physical traits, we propose that it represents a new *Candidatus* genus, and propose the name '*Candidatus* Formamonas warabiya' for this unique organism.

3.6 Acknowledgements

I gratefully acknowledge Kate Montgomery in the School of Biotechnology and Biomolecular Sciences (UNSW) for her assistance with FISH microscopy, and Dr James McDonald for his assistance with LC-MS/MS.

4 A variety of methyltransferases are expressed by '*Ca.* Formamonas warabiya' during anaerobic dichloromethane and quaternary amine metabolism

4.1 Introduction

Under anaerobic conditions, DCM is metabolised in the Wood-Ljungdahl pathway by DCM-dechlorinating bacteria (Mägli *et al.*, 1998; Kleindienst *et al.*, 2019; Chen *et al.*, 2020). However, the enzyme(s) that catalyse the presumed dechlorination of DCM prior to its entry into the Wood-Ljungdahl pathway have still not been definitively identified. A proteogenomic study of '*Ca.* Dichloromethanomonas elyunquensis' found that four methyltransferases were amongst the most abundant proteins expressed during growth on DCM (Kleindienst *et al.*, 2019). Two reductive dehalogenases were also present in the '*Ca.* Dichloromethanomonas elyunquensis' proteome during growth on DCM, yet '*Candidatus* Formamonas warabiya' strain DCMF and *D. formicoaceticum* do not encode these enzymes at all.

In the current study, the abundance of methyltransferases encoded in the strain DCMF genome supports their hypothesised involvement in DCM dechlorination. The work in the previous chapter, demonstrating that strain DCMF can utilise substrates other than DCM, allowed a comparative proteomics experiment to be conducted, comparing strain DCMF cells grown on DCM, choline, glycine betaine, and methanol. The primary aim of this work was to identify proteins involved in DCM dechlorination. A cluster of methyltransferases were highly abundant in DCM-amended cells, homologous to gene clusters identified in *D. formicoaceticum* and '*Ca.* Dichloromethanomonas elyunquensis'. Proteomics also provided validation of the putative metabolic pathways for choline, glycine betaine, and methanol transformation that were hypothesised in Chapter 3 – confirming the involvement of a glycine betaine methyltransferase, a sarcosine reductase complex, and a methanol methyltransferase. This work illustrates how strain DCMF utilises a range of methyltransferases for growth on a variety of C1 and methylated substrates.

4.2 Materials and Methods

4.2.1 Analytical techniques

DCM and methanol were quantified with GC-FID as described in Section 3.2.3.1. Choline and glycine betaine were quantified with LC-MS/MS as described in Section 3.2.3.2.

4.2.2 Cultures for proteomics

Cultures were grown in minimal mineral salt medium as described in Section 3.2.1. One of: DCM (2 mM), choline chloride (5 mM), glycine betaine (5 mM) or methanol (10 mM) was added as the electron donor. DFE cultures amended with DCM, choline or glycine betaine were grown in 200 ml aliquots within 250 ml capacity glass serum bottles. DFE cultures amended with methanol were grown in 140 ml aliquots within 160 ml capacity glass serum bottles.

Cells were harvested in triplicate when approximately 80% of the initial substrate pulse was depleted, i.e. when the cultures were in late exponential growth phase. Cells were collected from either 400 ml (DCM, choline, glycine betaine) or 140 ml (methanol) liquid culture via centrifugation at 8,000 rcf for 30 min at 4˚C. Supernatant was removed and the remaining cell pellets were resuspended in ~1 ml of their respective supernatant before being transferred to 2 ml cryotubes and stored at -80˚C until the crude protein extraction.

4.2.3 Crude protein extraction and quantification

The concentrated cell extracts from Section 4.2.2 were thawed on ice before being centrifuged at 10,000 rcf for 10 min at 4˚C. The supernatant was removed and each cell pellet was resuspended in 120 µl protein extraction buffer, which consisted of 50 mM MOPS (pH 7), 4% sodium dodecylsulfate (SDS), 50 mM NaCl, 100 µM ethylenediamine tetraacetic acid (EDTA), 100 µM $MgCl_2$. The mixtures were transferred to 2 ml screw-cap tubes containing 0.06 g glass beads (150-212 µm, Sigma) and a ¼" ceramic sphere (MP Bio) before the cells were lysed via bead-beating at speed setting 1800 for 5 min on a PowerLyzer 24 Homogenizer (Qiagen). The tubes were centrifuged at 16,000 rcf for 10 min at 15˚C and the supernatants were transferred to fresh, 1.5 ml microfuge tubes. The centrifugation step was

repeated and the supernatant (the crude protein extract) was once again transferred to a fresh 1.5 ml microfuge tube.

Protein yield in the crude extract was quantified using the Micro BCA Protein Assay Kit (Thermo Fisher Scientific) as per manufacturer's instructions for the microplate assay. Crude protein extracts were diluted 1:250 in MilliQ water prior to quantification. Bovine serum albumin was used to create a seven-point standard curve ranging from 0.5 – 40 µg ml^{-1}.

Protein samples were diluted to a concentration of 1 µg $µl^{-1}$ in 50 mM NH_4HCO_3 and these diluted samples were again quantified via Micro BCA assay to ensure accurate dilution.

4.2.4 Filter-aided sample preparation (FASP)

FASP was used to prepare the crude protein extract for proteomic analysis, following the method outlined by Wiśniewski et al (2009; 2017). A total of 15.8 µg protein from each sample (22 µl total volume, using 50 mM NH_4HCO_3 to dilute where necessary) was transferred to a fresh 1.5 ml microfuge tube. Dithiothreitol (5 mM) was added to each tube prior to incubation at 37˚C for 30 min, to reduce the cysteine disulphide bonds in the peptides.

Samples were then loaded onto Amicon Ultra-0.5 30 kDa centrifugal filter units (Merck) with 200 µl UA solution (8 M urea in 100 mM Tris-HCl, pH 8.5). Filters were centrifuged at 14,000 rcf for 15 min before a further 200 µl UA was added to each filter and the centrifugation step was repeated. Peptides were alkylated by addition of 100 µl iodoacetamide solution (50 mM iodoacetamide in UA) and mixing at 600 rpm for 1 min prior to incubating statically in the dark for 20 min. Filters were centrifuged at 14,000 rcf for 10 min. UA (100 µl) was added to each filter and it was centrifuged at 14,000 rcf for 15 min; this was repeated once more. Then, 50 mM NH_4HCO_3 (100 µl) was added to each filter before centrifuging at 14,000 rcf for 10 min; this was repeated twice more. Protein digestion was performed by addition of trypsin (1:100 enzyme:protein ratio, i.e. 0.8 µl of a 200 ng $µl^{-1}$ trypsin solution) in a further 40 µl of 50 mM NH_4HCO_3 and mixing at 600 rpm for 1 min. Filters were incubated in a wet chamber at 37˚C overnight and then transferred to fresh collection tubes and centrifuged at 14,000 rcf for 10 min. A final 40 µl of 50 mM

NH_4HCO_3 as added to each filter and they were centrifuged at 14,000 rcf for 10 min. The eluent was stored at -80˚C prior to further use.

4.2.5 Proteomic analysis via LC-MS/MS

Sample analysis by liquid chromatography with tandem mass spectrometry (LC-MS/MS) was carried out at the Bioanalytical Mass Spectrometry Facility (BMSF) at the University of New South Wales. Peptides were separated by nanoLC on an UltiMate™ 3000 RSLCnano ultra performance liquid chromatograph (UPLC) and autosampler system (Dionex). Samples (2.5 µl) were concentrated and desalted onto a micro C18 precolumn (300 µm x 5 mm, Dionex) with H_2O:CH_3CN (98:2, 0.2 % TFA) at 15 µl min^{-1}. After a 4 min wash the pre-column was switched (Valco 10 port UPLC valve, Valco, Houston, TX) into line with a fritless nano column (75µ x ~15cm) containing C18AQ media (1.9µ, 120 Å Dr Maisch). Peptides were eluted using a linear gradient of H_2O:CH_3CN (98:2, 0.1% formic acid) to H_2O:CH_3CN (64:36, 0.1% formic acid) at 200 nl min^{-1} over 30 min. High voltage 2000 V was applied to low volume Titanium union (Valco) and the tip positioned ~0.5 cm from the heated capillary (T=275°C) of an Orbitrap Fusion Lumos (Thermo Electron) mass spectrometer. Positive ions were generated by electrospray and the Fusion Lumos operated in data dependent acquisition mode.

A survey scan *m/z* 350 – 1,750 was acquired in the orbitrap (resolution = 120,000 at *m/z* 200, with an accumulation target value of 400,000 ions) and lockmass enabled (*m/z* 445.12003). Data-dependent tandem MS analysis was performed using a top-speed approach (cycle time of 2 s). MS2 spectra were fragmented by HCD (NCE=30) activation mode and the ion-trap was selected as the mass analyser. The intensity threshold for fragmentation was set to 25,000. A dynamic exclusion of 20 s was applied with a mass tolerance of 10 ppm.

4.2.6 Proteomic data analysis

MS/MS spectra .raw files were searched against a custom database of all predicted proteins in the manually curated IMG-annotated strain DCMF genome using MaxQuant version 1.6.7.0 (Cox *et al.*, 2014). Enzyme specificity was trypsin/P with a maximum of two missed cleavages. Fixed (carbamidomethylation of cysteine) and variable (oxidation of methionine and N terminal acetylation) modifications were selected. Minimum peptide length was set as seven amino acids and maximum

peptide mass as 4,600 Da. 'LFQ' and 'Match between runs' were selected. All other settings were left as default.

Statistical analysis of the MaxQuant output was performed in Perseus v1.6.7.0 (Tyanova *et al.*, 2016). Proteins identified by site, reverse sequences and potential contaminants were removed. Proteins were filtered to retain only those identified by two or more unique peptides, and present in all three replicates of at least one substrate condition. Label free quantitative (LFQ) intensities were $\log_2$ transformed and missing values were imputed from a Gaussian distribution (down shift 1.8 and width 0.3, relative to the standard deviation of each column). Triplicate-averaged values were Z-score transformed within each substrate column to determine protein abundance relative to overall expression with each substrate.

The DCM- and glycine betaine-amended samples were directly compared via multiple t-tests to create a volcano plot. T-test parameters were: S_0 = 0.1, 250 randomizations, substrate grouping not preserved in randomizations. Proteins were considered differentially abundant if they had a False Discovery Rate (FDR) < 0.01.

4.3 Results

4.3.1 Protein expression patterns

Label-free quantitative (LFQ) proteomics was carried out with DFE cultures amended with DCM, glycine betaine, choline, and methanol in order to investigate the protein expression patterns of strain DCMF on each substrate. In total, 1,480 unique proteins were identified across the four substrate conditions (Dataset S2). Following imputation of missing values, principle components analysis (PCA) showed that triplicate samples from each substrate clustered tightly together. Within the substrate groups, samples from the DCM-amended cultures showed the most difference to all others (Figure 4.1). From the total number of identified proteins, 408 were significantly differentially abundant between DCM- and glycine betaine-grown cells (FDR < 0.01; Table S5).

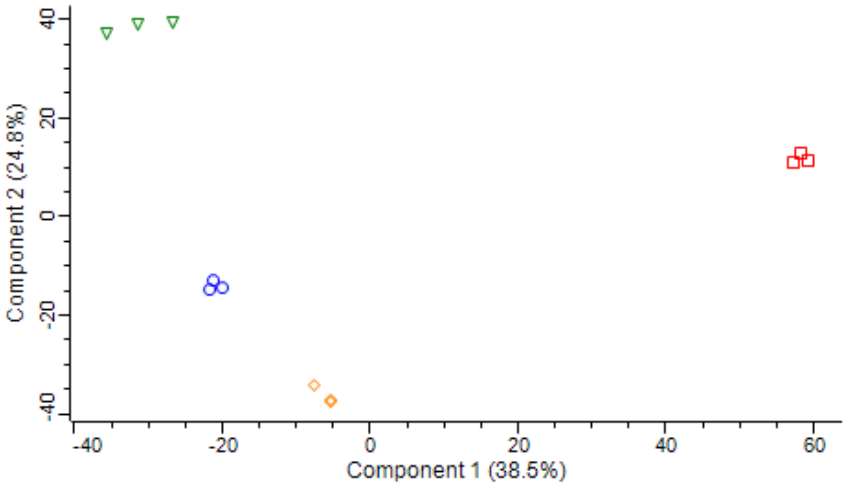


Figure 4.1 Principle components analysis plot shows triplicate proteomics samples cluster together by substrate. Samples amended with DCM (red squares) showed the most difference to samples amended with glycine betaine (blue circles), choline (green triangles), or methanol (orange diamonds).

4.3.1.1 Wood-Ljungdahl pathway proteins

Proteins from the Wood-Ljungdahl pathway were amongst the most abundant under all four substrate conditions (Figure 4.2). All proteins from the pathway were present, including formate dehydrogenase (112876 [3], 112877, 112878s80), formate-tetrahydrofolate (THF) ligase (114722), 5,10-methylene-THF dehydrogenase/ methenyl-THF cyclohydrase (11378), 5,10-methylene-THF reductase (111198), 5-methyl-THF Co Fe-S protein methyltransferase (111199), CO dehydrogenase (111205), and acetyl-CoA synthase (111204). Proteins for the conversion of acetyl-CoA to acetate were also present (phosphate propanoyltransferase, 112835, and acetate kinase, 111561; Figure 4.2).

4.3.1.2 Energy generation proteins

Proteins for an F_0F_1-type ATP synthase were also amongst the most abundant under all conditions, including all F_1 subcomplex units and one of the three F_0 subcomplex units (113152 – 113157; Figure 4.2). Additionally, the proteome contained three proteins from the Rnf complex (RnfBCG, 113065, 113070, 113068) and almost all proteins for a NADH:ubiquinone oxidoreductase (complex I). NuoGEFBCDHIJLM

[3] All strain DCMF gene numbers in this chapter refer to the IMG Gene Locus and are prefaced by "Ga0180325_".

(11330, 11791-92, 115678-83, 115685-86) were present, whilst NuoAKN (115677, 115684, 115687) were not identified, but are encoded in the genome. Two proteins for a putative K^+ or Na^+-stimulated pyrophosphate-energised sodium pump (113285, 113311) were also highly abundant with all substrates. The presence of all these proteins suggests that strain DCMF may generate energy through a chemiosmotic mechanism, as well as substrate level phosphorylation in the Wood-Ljungdahl pathway.

4.3.1.3 Methyltransferases

In total, the strain DCMF proteome contained 79 components identified either as a methyltransferase 1 (MT_1, responsible for transferring a methyl group from a substrate onto the corrinoid protein), methyltransferase 2 (MT_2, responsible for transferring the methyl group from the corrinoid protein to a receiving compound) or cognate corrinoid protein (CoP) (Figure 4.3; Dataset S2). This included 29 of the 82 MttB superfamily methyltransferases (Dataset S2). Due to the large quantity of methyltransferases expressed in the proteome, particular attention was given to those with high abundance (Z-score > 1), especially in one substrate more than others.

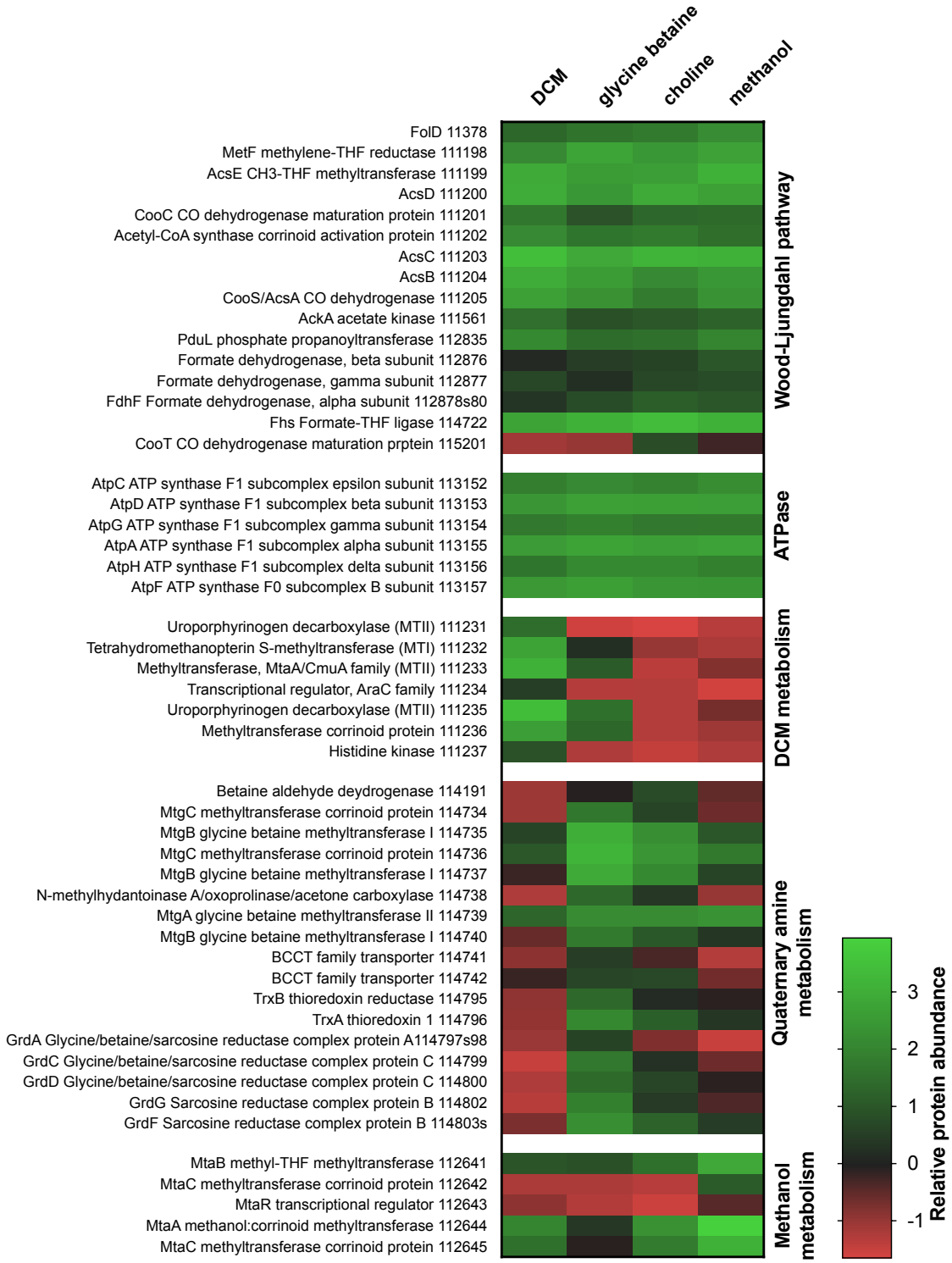


Figure 4.2 Heatmap depicting the abundance of proteins of interest in DCM-, glycine betaine-, choline- and methanol-amended culture. Proteins that were highly abundant across all four conditions (Wood-Ljungdahl pathway and ATPase) or with specific substrates are shown. Colours represent Z-scores of $\log_2$-transformed LFQ values, i.e. green/higher values represent higher abundance of that protein relative to overall protein expression with that substrate.

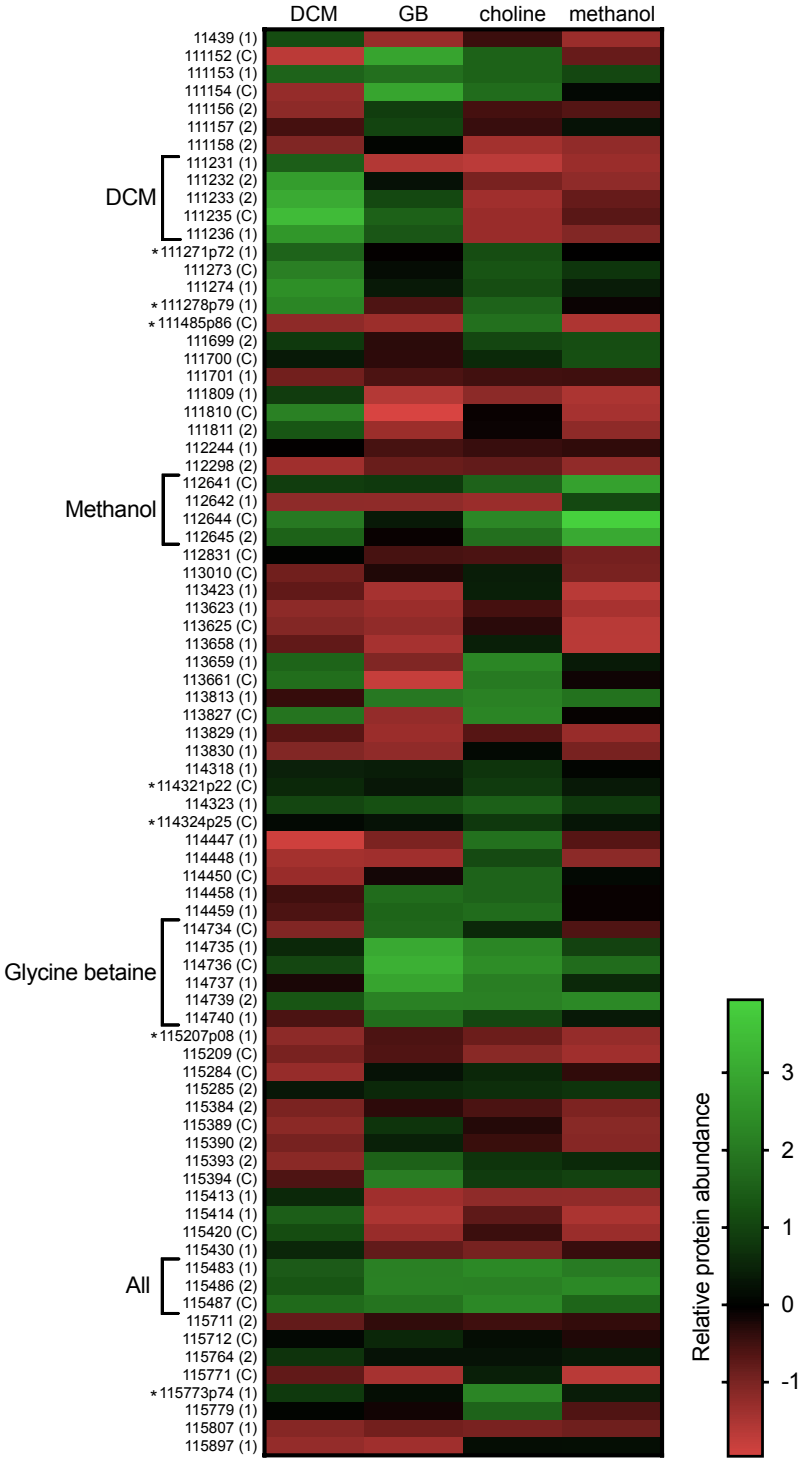


Figure 4.3 Heatmap of all corrinoid-dependent methyltransferase system proteins identified in the '*Ca.* Formamonas warabiya' proteome. Methyltransferase systems that were highly abundant with DCM, methanol, glycine betaine, and all substrates are indicated. Locus tags are followed by the likely role of each protein in the methyltransferase system: the methyltransferase 1 (1), methyltransferase 2 (2), or corrinoid protein (C). Methyltransferases that include a pyrrolysine residue are marked with an asterisk (*). Colours represent Z-scores of $\log_2$-transformed LFQ values, i.e. green/higher values represent higher abundance of that protein relative to overall protein expression with that substrate.

A cluster of methyltransferases spanning loci 111231 – 111236 were highly abundant in cells amended with DCM only, while methyltransferases spanning loci 111152 – 111158 and 114734 – 114740 were highly abundant in cell grown with glycine betaine or choline relative to cells grown with the other substrates (Figure 4.3). In methanol-amended samples, it was methyltransferases spanning 112641 – 112645 that stood out for their uniquely high abundance (Figure 4.3). One cluster of methyltransferases (115483, 115486, 115476) was highly abundant across all four growth conditions (Figure 4.3).

Seven of the expressed methyltransferases are pyrrolysine-containing methylated amine methyltransferases (111271p72, 111278p79, 111485-86, 114321p22, 114324p25, 115207p08, 115773p74; Figure 4.3). Accordingly, proteins for the biosynthesis and incorporation of the non-canonical amino acid pyrrolysine were also identified in the proteome (PylBCDSc; 112912, 116769, 114723, 115767; Dataset S2).

4.3.1.4 S-layer, motility, and chemotaxis proteins

With all four substrates, an S-layer homology domain-containing protein was the most abundant protein (113134 with DCM, glycine betaine and choline) or the second-most abundant protein (113124 with methanol). In total, 14 of the 43 S-layer homology domain-containing proteins encoded in the strain DCMF genome were identified in the proteome, with varying levels of abundance (Dataset S2).

Proteins for a flagellar and chemotaxis were also identified, indicating that strain DCMF is motile and can respond to environmental cues. Of the 39 flagellar-related genes encoded in the organism's genome, 21 were found in the proteome (Dataset S2). Flagellin (112818), in particular, was highly abundant across all four substrates. The proteome included 24 enzymes associated with chemotaxis, including the two-component regulatory system sensor histidine kinase CheA (113319) and the response regulator CheY (112765), as well as numerous methyl-accepting chemotaxis proteins (Dataset S2). None of the gas vesicle genes identified in the genome were expressed in the proteome.

4.3.2 Proteomics identified a methyltransferase gene cluster linked to dichloromethane metabolism

Investigation of the genetic neighbourhood surrounding the methyltransferase genes that were highly abundant in DCM-amended cells revealed a cluster of eight genes on the negative DNA strand spanning loci 111230 – 111237 (Figure 4.4). The genes are annotated by IMG as a two-component regulator system with sensor histidine kinase and AraC family response receiver, a tetrahydromethanopterin S-methyltransferase subunit H, two uroporphyrinogen decarboxylases, an MtaA/CmuA family methyltransferase, a methyl-THF methyltransferase, and a Kef-type cation exchange protein (Table 4.1). All eight proteins were comparatively more highly abundant in cells grown on DCM than any other substrate (Figure 4.2). The DCM- and glycine betaine-amended cultures were directly compared to create a "volcano plot" showing the $\log_2$-fold change (LFC) and associated statistical significance for each protein (Figure 4.5; Table S5). All eight proteins were significantly more abundant (FDR 0.01) in DCM-amended cells, with an average LFC of +5.29 compared to glycine betaine-grown cells. The largest LFC of +8.40 was for a uroporphyrinogen decarboxylase (111231) (Figure 4.5).

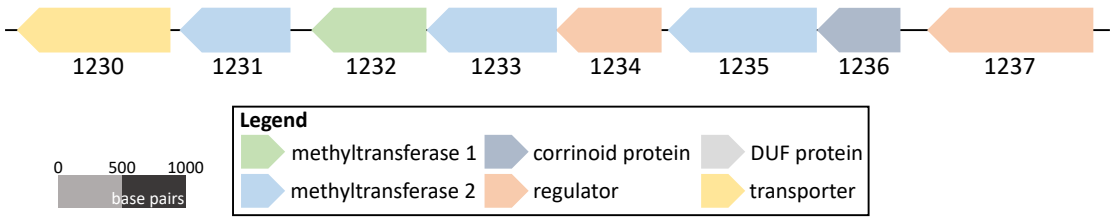


Figure 4.4 The DCM-associated methyltransferase gene cluster in '*Ca*. Formamonas warabiya'. IMG gene loci (each prefaced by Ga0180325_11) are written below each gene. Genes and intergenic regions are drawn to the scale shown.

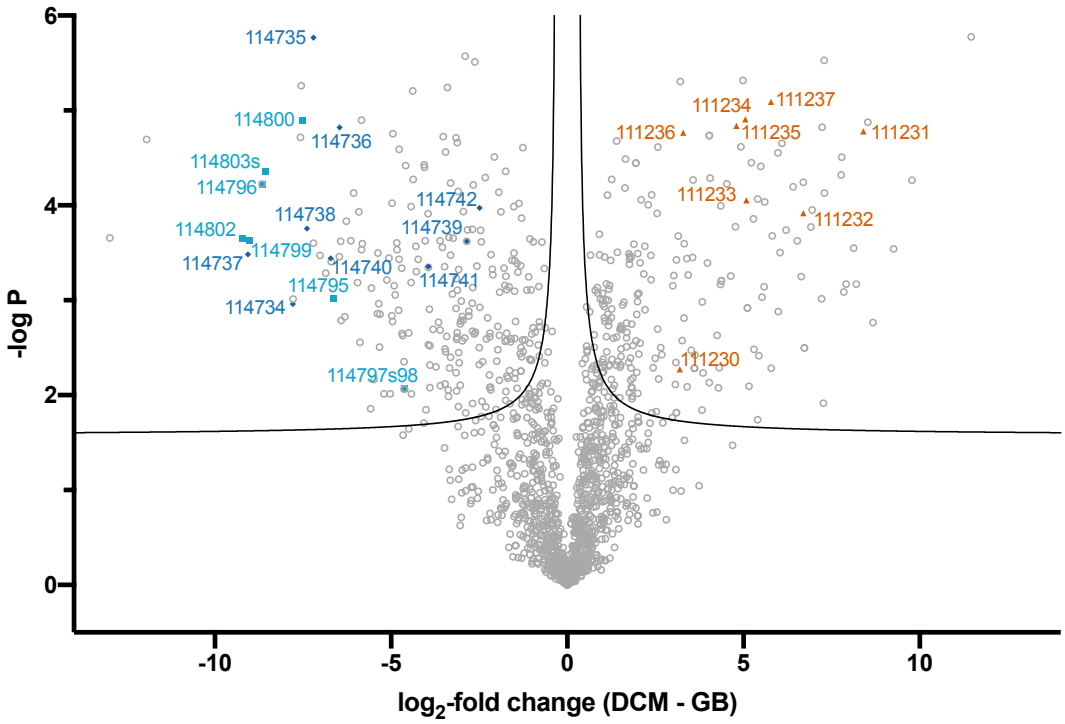


Figure 4.5 Volcano plot of '*Ca.* Formamonas warabiya' protein expression with DCM and glycine betaine. Log_2-fold change (LFC) in abundance is the difference between DCM and glycine betaine, i.e. proteins further to the right were more abundant with DCM, while those to the left were more abundant with glycine betaine. A false discovery rate (FDR) of 0.01 was the significance boundary, indicated by two black lines on the graph (i.e. the upper left and upper right portions contain proteins with significantly differential abundance). Proteins in the DCM-associated methyltransferase cluster are labelled as orange triangles, those from the putative glycine betaine methyltransferase cluster by dark blue diamonds, and those from the putative sarcosine reductase cluster by light blue squares. IMG gene loci (prefaced by 'Ga018035_') are indicated for these proteins of interest. A full list of all significantly differentially abundant proteins is included in Table S5.

Table 4.1 Functional annotation of the DCM-associated methyltransferase gene cluster identified in strain DCMF. nd, not described.

IMG Locus	Putative function	IMG Annotation	Length (AA)	LFC[a]	KEGG Orthology Term	EggNOG Classification	pfam Domains	Subcellular localisation[b]
111237	Two-component transcriptional regulator, histidine kinase	Histidine kinase	430	5.78	nd	08HRJ histidine kinase	PF10114 PocR domain; PF06580 Signal transduction histidine kinase, internal region	Cytoplasmic membrane
111236	Methyltransferase corrinoid protein	5-methyl-THF S-homocysteine methyltransferase	201	3.29	K00548 metH, MTR 5-methyl-THF-homocysteine methyltransferase [EC:2.1.1.13]	08IRG B12-binding_2	PF02310 B12-binding; PF02607 B12-binding_2 (cap domain)	Cytoplasm
111235	MT_2	Uroporphyrinogen decarboxylase	343	4.80	K01599 hemE, UROD uroporphyrinogen decarboxylase [EC:4.1.1.37]	07US9 uroporphyrinogen decarboxylase	PF01208 uroporphyrinogen decarboxylase	Cytoplasm
111234	Two-component transcriptional regulator, receiver	2 component transcriptional regulator, AraC family	272	5.05	nd	07VQW regulator	PF12833 DNA binding HTH domain, AraC type; PF00072 signal transduction response regulator, receiver domain	Cytoplasm

IMG Locus	Putative function	IMG Annotation	Length (AA)	LFC[a]	KEGG Orthology Term	EggNOG Classification	pfam Domains	Subcellular localisation[b]
111233	MT_2	Methyltransferase, MtaA/CmuA family	337	5.08	nd	07SA1 uroporphyrinogen-III decarboxylase	PF01208 uroporphyrinogen decarboxylase	Cytoplasm
111232	MT_1	tetrahydromethanopterin S-methyltransferase, subunit H (mtrH)	299	6.70	K00584 mtrH tetrahydromethano-pterin S-methyltransferase subunit H [EC:2.1.1.86]	08QZ0 tetrahydromethanopterin S-methyltransferase	PF02007 tetrahydromethanopterin S-methyltransferase subunit H / methyltransferase Mtx subunit H	Cytoplasm
111231	MT_2	uroporphyrinogen decarboxylase	288	8.40	nd	08RI4 uroporphyrinogen decarboxylase	PF01208 uroporphyrinogen decarboxylase	Cytoplasm
111230	Cation transporter	Kef-type K+ transport system	398	3.19	nd	04RUZ Na H antiporter	PF00999 Cation/H+ exchanger	Cytoplasmic membrane

[a] Log_2 fold change in protein abundance in DCM-amended cells compared to glycine betaine-amended cells.
[b] As determined by PsortB.

4.3.3 Proteomic identification of a putative glycine betaine methyltransferase

In cells grown with glycine betaine, a cluster of methyltransferases and genes spanning 114734 – 114742 were among the most abundant proteins (Figure 4.2, Figure 4.3). This nine-gene cluster is located on the positive strand and, according to IMG annotation, included three methyl-THF methyltransferases, three MttB superfamily methyltransferases (trimethylamine-corrinoid protein Co-methyltransferase), an N-methylhydantoinase A/oxoprolinase/acetone carboxylase, and two glycine betaine transporters (Table 4.2). In comparison to cells from DCM-amended cultures, these nine proteins were all significantly more abundant when grown with glycine betaine (FDR 0.01). The methyltransferase system components had LFCs of +6.46 to +9.06 in glycine betaine compared to DCM-grown cells (Figure 4.5, Table 4.2).

Functional annotation revealed B_{12}-binding domains in two of the methyl-THF methyltransferases (114734, 114736), indicating they may act as CoP (MtgC) in a methyltransferase system (Table 4.2). The three genes annotated as trimethylamine-corrinoid Co-methyltransferases sit within the MttB superfamily and are likely the MT_1 component of a glycine betaine methyltransferase system (MtgB), while the remaining methyl-THF methyltransferase (114739) lacks any B_{12}-binding domains and is a putative MT_2 component of the same system (MtgA; Figure 4.6A; Table 4.2). Amongst the methyl-THF methyltransferases, EggNOG functional annotation classed 114739 separately to 114734 and 114736, supporting their differentiation into putative MT_2 and CoP roles, respectively (Table 4.2). Both transporters fall within the betaine/choline/carnitine (BCCT) superfamily, with KEGG classification as the glycine betaine transporter OpuD (Table 4.2), further supporting the assignation of this gene cluster as a glycine betaine methyltransferase system (Figure 4.6A).

Another cluster of methyltransferase system genes (111152 – 111158) were also highly abundant in glycine betaine-amended cells. This cluster of genes, however, contained only putative MT_1, MT_2 and CoPs, with no other genes related to glycine betaine metabolism (e.g. transcriptional regulators or transporters) in the genetic vicinity.

Table 4.2 Functional annotation of genes in the putative glycine betaine methyltransferase gene cluster. nd, not described.

IMG Locus	Gene / putative function	IMG Annotation	Length (AA)	LFC[a]	KEGG Orthology Term	EggNOG Classification	pfam Domains	Subcellular localisation[b]
114734	MtgC / CoP	5-methyl-THF-homocysteine methyltransferase	210	7.79	K00548 metH, MTR 5-methyl-THF-homocysteine methyltransferase EC:2.1.1.13	07XHN domain protein (COG5012 predicted cobalamin binding protein)	PF02310 B12-binding PF02607 B12-binding_2	Cytoplasm
114735	MtgB / MT_1	trimethylamine--corrinoid protein Co-methyltransferase	485	7.20	K14083 mttB trimethylamine-corrinoid protein Co-methyltransferase EC:2.1.1.250	06EUB trimethylamine methyltransferase	PF06253 MTTB	Cytoplasm
114736	MtgC / CoP	5-methyl-THF-homocysteine methyltransferase	210	6.46	K00548 metH, MTR 5-methyl-THF-homocysteine methyltransferase EC:2.1.1.13	07XHN domain protein (COG5012 predicted cobalamin binding protein)	PF02310 B12-binding PF02607 B12-binding_2	Cytoplasm

IMG Locus	Gene / putative function	IMG Annotation	Length (AA)	LFC[a]	KEGG Orthology Term	EggNOG Classification	pfam Domains	Subcellular localisation[b]
114737	MtgB / MT_1	trimethylamine--corrinoid protein Co-methyltransferase	488	9.06	K14083 mttB trimethylamine-corrinoid protein Co-methyltransferase EC:2.1.1.250	06EUB trimethylamine methyltransferase	PF06253 MTTB	Cytoplasm
114738	Unknown	N-methylhydantoinase A/oxoprolinase/acetone carboxylase, beta subunit	668	7.39	nd	07RYM hydantoinase oxoprolinase	PF01968 hydantoinase_A PF05378 Hydant_A_N	Cytoplasm
114739	MtgA / MT_2	5-methyl-THF-homocysteine methyltransferase	265	2.85	K00548 metH, MTR 5-methyl-THF-homocysteine methyltransferase EC:2.1.1.13	07V3B methyl-THF corrinoid iron-sulphur protein methyltransferase	PF00809 pterin_bind	Cytoplasm
114740	MtgB / MT_1	trimethylamine--corrinoid protein Co-methyltransferase	471	6.71	K14083 mttB trimethylamine-corrinoid protein Co-methyltransferase EC:2.1.1.250	05F5V trimethylamine methyltransferase	PF06253 MTTB	Cytoplasm

IMG Locus	Gene / putative function	IMG Annotation	Length (AA)	LFC[a]	KEGG Orthology Term	EggNOG Classification	pfam Domains	Subcellular localisation[b]
114741	OpuD / Transporter	glycine betaine transporter	540	3.94	K05020 opuD, betL glycine betaine transporter	05C94 transporter	PF02028 BCCT, betaine/carnitine/choline family transporter	Cytoplasmic membrane
114742	OpuD / Transporter	glycine betaine transporter	538	2.49	K05020 opuD, betL glycine betaine transporter	05C94 transporter	PF02028 BCCT, betaine/carnitine/choline family transporter	Cytoplasmic membrane

[a] Log_2 fold change in protein abundance in glycine betaine-amended cells compared to DCM-amended cells.

[b] As determined by PsortB.

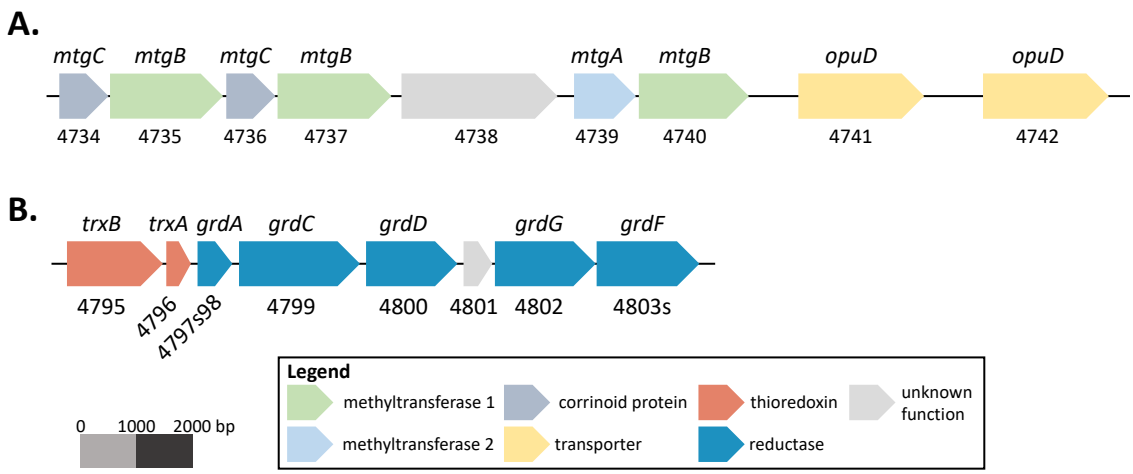


Figure 4.6 The putative glycine betaine methyltransferase gene cluster (A) and sarcosine reductase complex gene cluster (B) in '*Ca*. Formamonas warabiya'. Gene symbols are written above and IMG gene loci (each prefaced by Ga0180325_11) are written below each gene. Genes and intergenic regions are drawn to the scale shown.

Of the five glycine/betaine/sarcosine reductase gene clusters identified in the strain DCMF genome, only one was observed in the proteome. Genes 114795 – 114803s encode a thioredoxin reductase (*trxB*), thioredoxin I (*trxA*), glycine/betaine/sarcosine reductase complex selenoprotein A (*grdA*), protein C (*grdCD)*, and a presumed sarcosine-specific protein B (*grdGF*) (Figure 4.6B). A hypothetical protein also sits in the middle of the cluster (114801) but was not identified in the proteomic data. Excluding the hypothetical protein, all proteins in the cluster were significantly more abundant (FDR 0.01) in glycine betaine-amended cells, with LFCs ranging from +4.62 for GrdA (11497s98) to +9.21 for the putative GrdG (114802, compared to DCM-amended cells (Figure 4.5).

4.3.4 Protein expression with choline

The putative glycine betaine methyltransferase and sarcosine reductase gene clusters outlined in Section 4.3.3 were also abundant (albeit not to the same extent) in cells from choline-amended cultures (Figure 4.2). This supports the hypothesis from Chapter 3 that choline metabolism proceeds via glycine betaine (Figure 3.9). The choline dehydrogenase (11215) that was predicted to catalyse the first reaction of choline into betaine aldehyde was not identified in the proteome. However, the betaine aldehyde dehydrogenase (114191) that could transform betaine aldehyde

into glycine betaine was detected, and at a higher abundance with choline than any other substrate (Figure 4.2).

Interestingly, the majority of proteins for a putative ethanolamine bacterial microcompartment (BMC) were expressed in the proteome, with higher abundance in choline-amended cultures than any other. These included the shell proteins BMC-H (112352, 112353, 112372), BMC-T (112355, 112361) and BMC-P (112357), plus core enzymes EutE (112351) and PduL (112350).

4.3.5 A putative methanol methyltransferase system

The most abundant protein in methanol-amended cultures was a methanol:corrinoid methyltransferase (112644). Other methyltransferases nearby were also highly expressed (Figure 4.3), forming a likely methanol methyltransferase gene cluster spanning loci 112641 – 112645 on the negative strand (Figure 4.2, Figure 4.7). As well as the methanol:corrinoid methyltransferase (*mtaB*; 112644), the cluster includes two methyltransferase corrinoid proteins (*mtaC*; 112645, 112642), a AraC family transcriptional regulator (*mtaR;* 112643), and a methyl-THF methyltransferase (acting as *mtaA*; 112641; Figure 4.7, Table 4.3). The regulator contains a PocR sensory domain and a DNA-binding helix-turn-helix domain (Table 4.3).

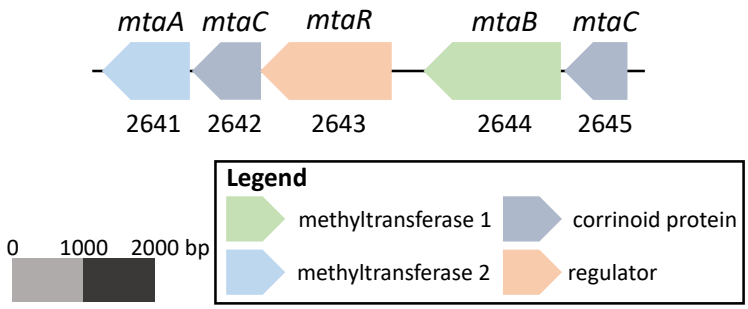


Figure 4.7 The putative methanol methyltransferase gene cluster in '*Ca.* Formamonas warabiya'. Putative gene symbols or functions are written above, and IMG gene loci (prefaced by Ga0180325_11) are written below each gene. Genes and intergenic regions are to the scale shown.

In general, overall methyltransferase abundance was far lower in methanol-amended cells than with other substrates (Figure 4.3). The average methyltransferase system component Z-score was -0.11 with methanol, compared to 0.25, 0.15 or 0.55 with DCM, glycine betaine or choline, respectively.

Table 4.3Functional annotation of the putative methanol methyltransferase gene cluster.

IMG Locus	Gene / putative function	IMG Annotation	Length (AA)	KEGG Orthology Term	EggNOG Classification	pfam Domains	Subcellular localisation[a]
112645	MtaC / CoP	5-methyl-THF-homocysteine methyltransferase	210	K00548 metH, MTR 5-methyl-THF-homocysteine methyltransferase EC:2.1.1.13	07XHN domain protein	PF02310 B12-binding PF02607 B12-binding_2	Unknown
112644	MtaB / MT_1	Methanol:corrinoid methyltransferase	458	K04480 mtaB methanol-5-hydroxybenzimidazolylcobamide Co-methyltransferase EC:2.1.1.90	07WH9 methyltransferase	PF12176 MtaB	Cytoplasm
112643	MtaR / Regulator	Transcriptional regulator, AraC family	438	K07720 yesN two-component system, response regulator YesN	08T7U transcriptional regulator AraC family	PF10114 PocR PF12833 Helix-turn-helix domain, AraC type	Cytoplasm
112642	MtaC / CoP	5-methyl-THF-homocysteine methyltransferase	230	K00548 metH, MTR 5-methyl-THF-homocysteine methyltransferase EC:2.1.1.13	07Y7V cobalamin B12-binding domain-containing protein	PF02310 B12-binding PF02607 B12-binding_2	Cytoplasm
112641	MtaA / MT_2	5-methyl-THF-homocysteine methyltransferase	292	K00548 metH, MTR 5-methyl-THF-homocysteine methyltransferase EC:2.1.1.13	07V3B methyl-THF corrinoid Fe-S protein methyltransferase	PF00809 pterin binding	Cytoplasm

[a] As determined by PsortB.

4.3.6 Strain DCMF thrives with an exogenous cobalamin source

As the methyltransferases implicated in DCM, glycine betaine, and methanol metabolism all require a corrinoid cofactor to function, the exogenous corrinoid requirement of strain DCMF was investigated. Triplicate, DCM-amended cultures were set up with or without cyanocobalamin (50 μg L^{-1}). In the first generation of cultures, those with exogenous cyanocobalamin dechlorinated 5.52 ± 0.71 mM DCM, while those without dechlorinated only 2.72 ± 1.00 mM DCM (Figure 4.8A). The deleterious effect of the absence of cobalamin was more pronounced in the second generation of these cultures, in which the triplicates with and without cyanocobalamin dechlorinated 6.42 ± 1.86 mM and 1.50 ± 1.00 mM DCM, respectively (Figure 4.8B).

Strain DCMF encodes a complete cobalamin biosynthesis pathway in its genome (Dataset S1), although it lacks the *bza* genes for synthesising the lower corrinoid ligand, 5,6-dimethylbenzimidazole. In total, 14 of the 25 proteins involved in cobalamin biosynthesis were detected in the proteome (despite the presence of 50 μg l^{-1} cyanocobalamin in the culture medium). These were primarily the enzymes involved in corrin ring formation, corrin ring amination, and the initiation and completion of the lower nucleotide loop (Dataset S2). The enzymes involved in cobalt active transport and insertion were not detected.

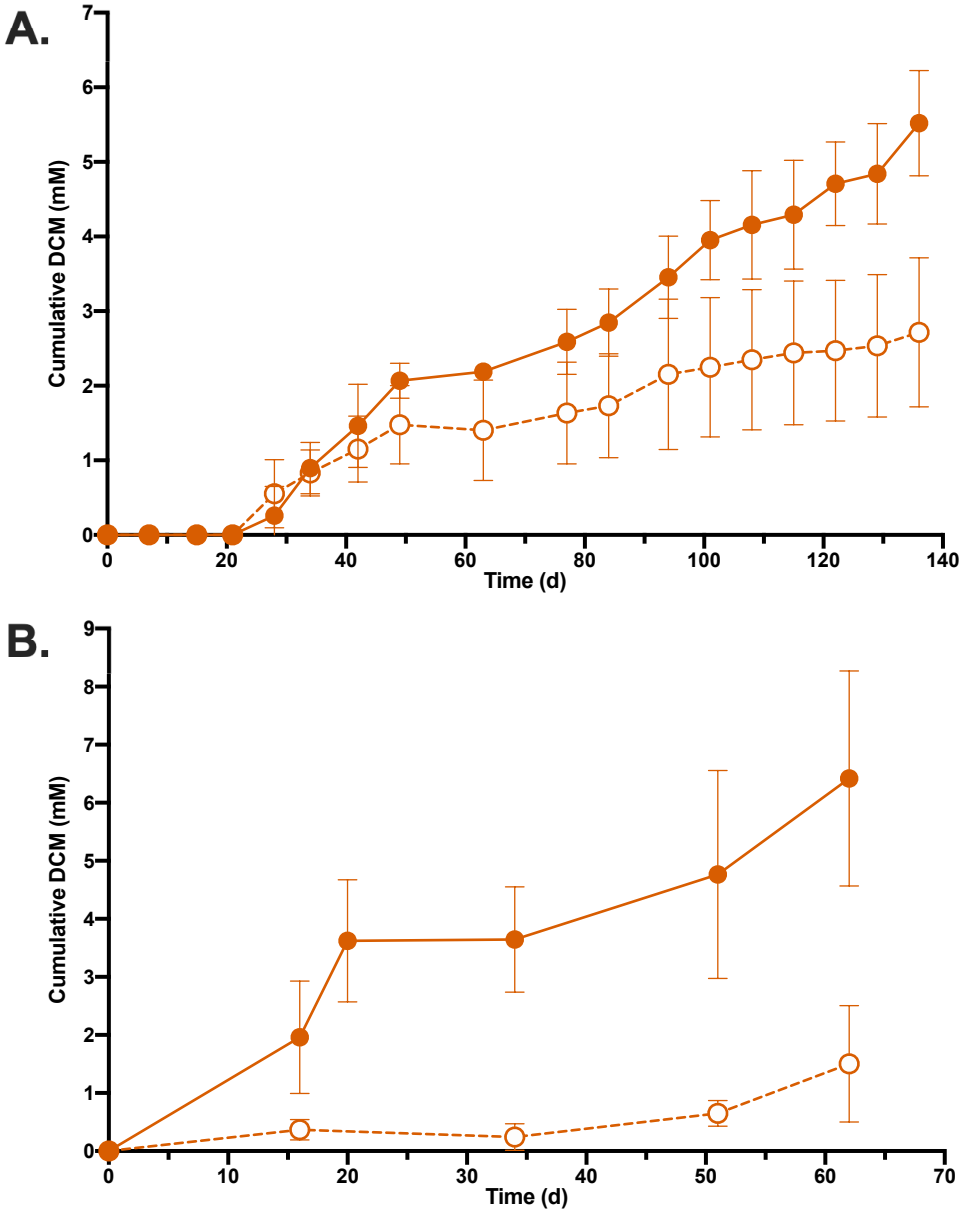


Figure 4.8 Strain DCMF likely requires an exogenous cobalamin source. Cumulative DCM consumption in the first (**A**) and second (**B**) generation of DFE cultures with (filled orange circles) and without (empty orange circles) exogenous cyanocobalamin (VB_{12}) added to the medium. Cumulative DCM consumption was higher in cultures amended with cyanocobalamin, and the second generation of cyanocobalamin-free cultures exhibited nearly no DCM dechlorination. Error bars represent standard deviation, n = 3.

4.4 Discussion

Despite renewed interest in anaerobic microbial DCM transformation in recent years (Kleindienst *et al.*, 2016, 2017; Chen *et al.*, 2017a, 2017b; Kleindienst *et al.*, 2019; Chen *et al.*, 2020), the enzyme(s) responsible for the dechlorination of DCM remain unknown. Thus far, *D. formicoaceticum* and '*Ca.* Dichloromethanomonas elyunquensis' have both been described as obligate anaerobic DCM-degrading bacteria, which has limited the sole proteomic study (Kleindienst *et al.*, 2019) to reporting the presence or absence of proteins in cells grown with DCM. The ability of strain DCMF to grow on additional substrates allowed a comparative proteomic experiment to be carried out, directly comparing protein abundance in strain DCMF cells grown on DCM, glycine betaine, choline, and methanol. The primary aim of this was to identify proteins that were significantly more abundant in DCM-grown cells, and thus implicated in DCM dechlorination. It also enabled verification of the genome-based metabolic models for quaternary amine and methanol metabolism proposed in Chapter 3.

4.4.1 '*Ca.* Formamonas warabiya' is a one-carbon specialist with metabolism is underpinned by methyltransferases

Proteins for the complete Wood-Ljungdahl pathway were highly abundant under all four conditions (Dataset S2), supporting the genomic and physiological metabolic models as well as work with ^{13}C-labelled compounds. However, the presence of numerous methyltransferases amongst the most abundant proteins observed with each substrate complicated interpretation of the data (Figure 4.3). One of the most highly expressed methyltransferase gene clusters under all four conditions (115483, 115486-7; Figure 4.3) in fact contained the closest homolog (by percentage amino acid identity) to a glycine betaine methyltransferase, MtgB, from *Acetobacterium woodii*, *Sporomusa ovata*, and *Desulfitobacterium hafniense*. The abundance of these putative glycine betaine methyltransferase proteins in strain DCMF across all four substrate conditions implies that they are constitutively expressed.

Interestingly, seven of the MttB superfamily methyltransferases detected in the proteome contain the unusual pyrrolysine residue (Figure 4.3). Although non-pyrrolysine members of the MttB superfamily are widespread in Bacteria and

Archaea, pyrrolysine-MttB proteins have thus far only been associated with methanogenesis from methylated amines in Archaea. A limited number of other bacterial genera also encode MttB superfamily genes with the pyrrolysine residue, but proteomic expression has not previously been observed. *Desulfitobacterium hafniense* Y51 also encodes a pyrrolysine-containing MttB protein, although its transcription was not monitored during growth on glycine betaine (Ticak *et al.*, 2014). Further experimental work to elucidate the function of these proteins in strain DCMF could provide insight into a potentially novel role for Pyl-MttB proteins, given that the organism did not appear to consume the only currently known substrate, trimethylamine.

4.4.2 Identification of a novel methyltransferase system linked to dichloromethane metabolism

A cluster of genes including methyltransferase and regulatory components were amongst the most abundant proteins in cells grown with DCM (111230 – 111238; Figure 4.2, Figure 4.4). These proteins were all significantly more abundant in cells grown with DCM than glycine betaine (FDR 0.01; Figure 4.5) and represent a putative novel methyltransferase system linked to DCM metabolism. Protein 111232 is likely the MT_1 in this system, which could putatively catalyse methyl (or methylene) transfer from DCM (or a derivative) to a cognate corrinoid protein. Both EggNOG and pfam tools for functional annotation classed protein 111232 in the tetrahydromethanopterin S-methyltransferase MtrH/MtxH family (Table 4.1), which also includes the glycine betaine MT_1 from *Desulfitobacterium hafniense* (Ticak *et al.*, 2014), supporting a MT_1 role for this protein in strain DCMF.

Another of the methyltransferases (111233) was annotated as a member of the MtaA/CmuA family, which includes methanol (MtaA) and chloromethane (CmuA) MT_2 proteins (Vannelli *et al.*, 1999; van der Meijden *et al.*, 1984b). While the enzyme involved in anaerobic chloromethane dechlorination has not yet been identified, protein 111233 shared 21% amino acid identity with CmuA from the aerobic chloromethane dechlorinator, *Methylorubrum extorquens* CM4. It is therefore feasible that it could act upon a similar chlorinated C1 compound such as DCM. Protein 111233 was one of three in the DCM-associated cluster that was classified

as a uroporphyrinogen decarboxylase by both eggNOG and pfam, suggesting that all three may be MT_2 proteins for a putative DCM methyltransferase system.

A B_{12}-binding domain present in 111236 indicated that it acts as the corrinoid protein and, therefore, that the methyltransfer reaction is corrinoid dependent. Light reversible inhibition of DCM dechlorination with propyl iodide and a requirement for substoichiometric amounts of ATP in cell extract reactions mixtures suggested that a corrinoid-dependent protein was also involved in DCM dechlorination by *D. formicoaceticum* (Mägli *et al.*, 1996, 1998). The four abundant methyltransferases in the proteome of DCM-grown '*Ca.* Dichloromethanomonas elyunquensis' cells were also all corrinoid dependent (Kleindienst *et al.*, 2019). Consistent with these observations, strain DCMF thrived in medium supplemented with exogenous cyanocobalamin (Figure 4.8). However, strain DCMF also encoded a complete pathway for *de novo* corrinoid biosynthesis in its genome, and a number of these genes were expressed in the proteome of DCM-grown cells (Dataset S2).

It is unclear why DCM dechlorination appeared to stall in the second generation of cultures without exogenous cobalamin amendment. Although there have been previous examples of cohabiting bacteria in mixed or co-culture producing cobalamins which meet the needs of auxotrophic community members (e.g. Yan *et al.*, 2013), this appeared not to be the case in the DFE culture. For strain DCMF, the energetic benefits to importing exogenous cobalamin rather than synthesising it *de novo* may have caused the discrepancy between the two sets of cultures. Alternatively, strain DCMF may require the addition of lower ligand precursors (e.g. 5,6-dimethylbenzimidazole) in the medium in order to synthesise its own cobalamin cofactors, as the genes for corrinoid lower ligand synthesis were not identified in the strain DMCF genome. Many dehalogenating bacteria require corrinoids with specific lower ligands as cofactors for growth (Yan *et al.*, 2012, 2013, 2016) and the ability of strain DCMF to dechlorinate DCM when supplied with exogenous cyanocobalamin (which contains 5,6-dimethylbenzimidazole as a lower ligand) supports its use of this specific lower ligand moiety.

The putative DCM methyltransferase cluster contains two proteins that may act as a two-component transcriptional regulator directly influencing gene transcription in response to DCM – a histidine kinase sensor (111237) and a YesN/AraC family

response regulator (111234). The histidine kinase contains a PocR domain (Table 4.1), which has been suggested to bind small hydrocarbons such as 1,3-propanediol. Its architecture is similar to that of the MEDS domain, which was typified in the aerobic DCM dehalogenase regulatory protein DcmR (Anantharaman and Aravind, 2005). The authors suggest that both the MEDS and the PocR domains could feasibly be utilised to bind DCM or other small hydrocarbons (Anantharaman and Aravind, 2005), indicating that it may affect transcription of the cluster directly in response to DCM.

Finally, the gene annotated as a Kef-type K+ transporter (111230) in the DCM-linked cluster is likely a cation/proton antiporter in the CPA-2 family. Pfam analysis revealed a cation/H^+ exchanger domain, while EggNOG classed it as an Na^+/H^+ antiporter (Table 4.1), leaving the cation specificity unclear. It is less clear whether this protein has a direct role in the gene cluster, although it had an LFC of +3.19 in DCM-grown cells compared to those with glycine betaine (Figure 4.5).

It is not yet clear how a putative DCM methyltransferase system would function, or whether additional reactions are involved in dechlorination, given that DCM is most likely transformed into methylene-THF prior to entry into the Wood-Ljungdahl pathway. In the glycine betaine and methanol methyltransferase systems, the methyl group is transferred from the corrinoid protein onto THF, resulting in methyl-THF (Stupperich and Konle, 1993; Ticak *et al.*, 2014). Studies with *D. formicoaceticum* clearly showed the evolution of methylene-THF from DCM though, not methyl-THF (Mägli *et al.*, 1998). Further biochemical characterisation of the DCM-linked methyltransferase proteins may provide insight into their specific catabolic function within the cell.

A number of proteins involved in electron transport and energy metabolism were widely expressed across all four substrate conditions, including subunits for an F_0F_1-type ATPase, a Rnf complex, and NADH:ubiquinone oxidoreductase. Each of these protein complexes may be involved in balancing the pool of reduced cofactors required for the function of the Wood-Ljungdahl pathway. Studies on glycine betaine and methanol metabolism in *A. woodii* have previously shown how the ATPase and Rnf complex can act in the "reverse" direction to normal (i.e., the ATPase converting ATP to ADP and pumping ions out of the cytoplasm and the Rnf complex

pumping the ions back into the cytoplasm to generate reduced ferredoxin from NADH) in order to provide enough reduced ferredoxin for CO_2 reduction (Lechtenfeld *et al.*, 2018; Kremp *et al.*, 2018). The presence of these protein subunits in the strain DCMF proteome could allow for a similar system to operate in this organism.

4.4.3 The novel methyltransferase gene cluster is conserved amongst anaerobic dichloromethane-metabolising bacteria

When the proteins in the DCM-linked methyltransferase cluster were searched against the NCBI non-redundant database with blastp, their closest homologs came from only three species: the anaerobic DCM-degrading organisms *Dehalobacterium formicoaceticum* strain DMC and '*Candidatus* Dichloromethanomonas elyunquensis' strain RM, and the trichloromethane-respiring bacterium *Dehalobacter* sp. UNSWDHB. Further analysis revealed that each of these organisms contains a highly similar cluster of genes in very close synteny (although interrupted by contig boundaries in '*Ca.* Dichloromethanomonas elyunquensis'; Figure 4.9). They span loci CEQ75_RS03275–30 in *D. formicoaceticum*, AWM53_02086–85 and AWM53_01378–83 in '*Ca.* Dichloromethanomonas elyunquensis', and UNSWDHB_2315–2270 in strain UNSWDHB. The proteins share high amino acid sequence percentage identities (75 – 94%) to those in strain DCMF, indicating a high degree of sequence conservation (Figure 4.9). No other species contained proteins with similarly high amino acid sequence percentage identities to the strain DCMF proteins in this cluster.

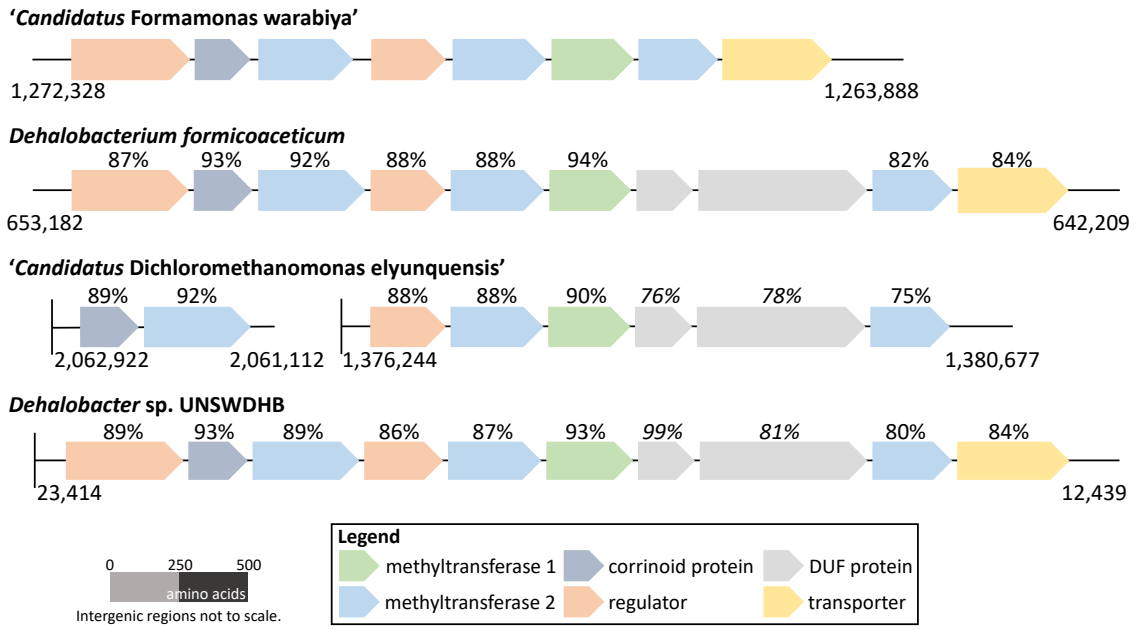


Figure 4.9 The DCM-associated methyltransferase gene cluster in '*Ca.* Formamonas warabiya' is conserved amongst DCM-fermenting bacteria *D. formicoaceticum* and '*Ca.* Dichloromethanomonas elyunquensis', and also present in *Dehalobacter* sp. UNSWDHB. Genes are colour-coded with their putative function as per the legend. Loci shown are (from left to right) '*Ca.* Formamonas warabiya' strain DCMF: Ga0180325_111238 – 111230 (reverse strand); *D. formicoaceticum* strain DMC: CEQ75_RS03275 – 30 (reverse strand); '*Ca.* Dichloromethanomonas elyunquensis' strain RM AWM53_02086 – 85 (reverse strand) and AWM53_01378 – 83 (positive strand); Dehalobacter sp. strain UNSWDHB: UNSWDHB_2315 – 2270 (reverse strand). Percentage amino acid sequence identity to strain DCMF is shown above each locus; percentage identity above the DUF proteins (*italicised*) is to strain DMC as these genes are absent from the cluster in strain DCMF. Genes are drawn to scale; intergenic regions are not.

Strain DCMF, *D. formicoaceticum* and '*Ca.* Dichloromethanomonas elyunquensis' are the only three anaerobic DCM-metabolising bacteria that have been characterised and genome sequenced thus far (Mägli *et al.*, 1996; Kleindienst *et al.*, 2016; Chen *et al.*, 2017b; Kleindienst *et al.*, 2017; Holland *et al.*, 2019). Proteomic studies with '*Ca.* Dichloromethanomonas elyunquensis' showed that all eight genes in the cluster were present in the cells grown with DCM, but none of the other tested substrates. Furthermore, the methyltransferase components, including the MT_1 (AWM53_01380) and MT_2 (AWM53_01379) on the larger contig and the CoP (AWM53_02086) and MT_2 (AWM53_02085) on the smaller contig were among the most abundant proteins (Kleindienst *et al.*, 2019). The high abundance of these

proteins during DCM metabolism in strain DCMF and '*Ca.* Dichloromethanomonas elyunquensis', and presence of this highly conserved cluster amongst all three DCM-degrading bacteria lends strong support to its involvement in the early steps of DCM metabolism. Proteomic studies with *D. formicoaceticum* have not been carried out but could provide further support for the involvement of this highly conserved gene cluster in DCM dechlorination.

In contrast, *Dehalobacter* sp. strain UNSWDHB is an organohalide-respiring bacterium capable of growth with trichloromethane, 1,1,1- and 1,1,2-trichloroethane (Wong *et al.*, 2016). While there are two other reports of uncharacterised *Dehalobacter* species likely fermenting DCM (Justicia-Leon *et al.*, 2012; Lee *et al.*, 2012), strain UNSWDHB cannot dechlorinate DCM (Wong *et al.*, 2016) and the genes in the DCM-associated cluster were not expressed in response to trichloromethane, except for the methylcobalamin methyltransferase (MT_2, UNSWDHB_RS02305)(Jugder *et al.*, 2016b). *Dehalobacter* sp. UNSWDHB and strain DCMF were enriched from the same mixed culture, CFEVO (Lee *et al.*, 2012), providing the possibility that the gene cluster may have spread between them via a horizontal gene transfer event. Indeed, following the final transporter gene in the methyltransferase gene cluster in strain UNSWDHB are two partial transposases (UNSWDHB_RS02265 and UNSWDHB_RS15345), whilst further down the contig are three genes encoding recombinase proteins (UNSWDHB_RS02225, UNSWDHB_RS02220, UNSWDHB_RS02210). The methyltransferase gene cluster in *D. formicoaceticum* is also flanked by identical IS66 family transposases (CEQ75_RS03190 – 200 and CEQ75_RS03285 – 95), raising questions of its genetic mobility, although it has thus far been described as an obligate DCM-fermenter and has been maintained as an axenic culture since its discovery (Mägli *et al.*, 1996). The function of these genes in *Dehalobacter* sp. UNSWDHB (if any) remains to be determined.

In *D. formicoaceticum*, '*Ca.* Dichloromethanomonas elyunquensis' and *Dehalobacter* sp. UNSWDHB, the gene cluster encodes two additional proteins not present in strain DCMF – a DUF1638- and a DUF4445-domain containing protein (Figure 4.9). The predicted amino acid sequences of these proteins remain highly conserved amongst the three organisms (76 – 99% to the *D. formicoaceticum* proteins; Figure

4.9). The DUF4445-domain containing protein may function as a reductive activase for the corrinoid protein in the methyltransferase system. It contains all four characteristic domains of reductive activators of corrinoid-dependent enzyme (RACE) proteins – a 2Fe-2S cluster and RACo linker, middle and C-terminal domains. Although the gene cluster in strain DCMF does not encode this protein, there are two homologs elsewhere in the genome that also include all four RACE protein domains (112542 and 114747, 36.6% and 43.95% amino acid identity with the *D. formicoaceticum* protein, respectively) and may fulfil this function, the latter of which was expressed in the proteome.

The role of the DUF1638-domain containing protein is less clear, as there is little information available about its function. The strain DCMF genome encodes eight proteins (112241, 112895, 112973, 113617, 114749, 114932, 115323, and 115628) with DUF1638 domains that could fulfil whatever role it may have in the conserved methyltransferase cluster. Only 115323 and 115628 were found in the proteome – the former was more abundant in DCM and glycine betaine-grown cells, while the latter was below average abundance in all four conditions. However, these 13 proteins have a far lower amino acid sequence identity (~20%) to the *D. formicoaceticum* DUF1638 domain-containing protein, which is uncharacteristic given the high homology of all other proteins in this cluster between the four species. Further study may reveal the role of this poorly characterised protein in the methyltransferase gene cluster.

4.4.4 Proteomic data supported the suggested model for quaternary amine and methanol metabolism in strain DCMF

Proteomic data supported the glycine betaine and methanol methyltransferase gene clusters suggested in Chapters 2 and 3 based on sequence homology to methyltransferases of known substrate specificity. Functional annotation of proteins in these gene clusters revealed putative MT_1, MT_2 and CoP components in both of them (Figure 4.6, Figure 4.7). The MtgB protein (MT_2, 114740) in the putative glycine betaine methyltransferase system gene cluster in strain DCMF was the second-closest homolog (54 – 55% amino acid identity) to MtgB proteins found in *Acetobacterium woodii*, *Sporomusa ovata*, and *Desulfitobacterium hafniense* (Lechtenfeld *et al.*, 2018; Visser *et al.*, 2016; Ticak *et al.*, 2014). The putative glycine

betaine methyltransferase gene cluster of strain DCMF, also included a 'N-methylhydantoinase A/oxoprolinase/acetone carboxylase, beta subunit', with unclear function. A similar hydantoinase is also encoded in the glycine betaine methyltransferase cluster from *Desulfitobacterium hafniense* Y51 (Ticak *et al.*, 2014), implying it may have some yet-to-be determined function for glycine betaine demethylation.

The proteomic data also supported the identification of a putative sarcosine reductase gene cluster from Chapter 2. Only one of the five clusters of glycine/betaine/sarcosine reductase genes encoded in the genome were identified in the proteome. This cluster included the reduction complex B proteins that clustered mostly closely with the known sarcosine reductase GrdGH proteins from *Sporomusa* sp. An4 (86.5% and 83.9% amino acid identity, respectively; Figure 2.8). The presence of these putative sarcosine-specific reductase proteins in turn supports the quaternary amine metabolic pathway suggested in Figure 3.9. As discussed in Chapter 3, we hypothesise that choline is transformed into glycine betaine and this is demethylated to sarcosine, which is then reductively cleaved to form monomethylamine and acetate (via acetyl phosphate). The electron equivalents released from the stepwise demethylation of glycine betaine into dimethylglycine and then sarcosine can be used for HCO_3^- reduction to acetate and, to a lesser extent, in the sarcosine reductase reaction. With the exception of choline dehydrogenase, hypothesised to transform choline into betaine aldehyde, all proteins for this putative pathway were identified in the proteome and particularly abundant in cultures amended with glycine betaine and choline (Figure 4.2).

The presence of proteins for an ethanolamine BMC, particularly in choline-amended cells, was an unexpected observation, given the presumption that choline is transformed into glycine betaine in strain DCMF. Ethanolamine may be produced from the stepwise demethylation of choline (also known as *N,N,N*-trimethylethanolamine). This metabolic pathway has previously only been demonstrated in methanogenic Archaea (Watkins *et al.*, 2012), but is congruent with the abundance of MttB superfamily methyltransferases expressed by strain DCMF during growth with choline. Within BMCs, ethanolamine is transformed into acetaldehyde and ammonia by ethanolamine ammonia lyase (EutBC; 1112363-64,

1112367-68). Acetaldehyde can then be further converted to either ethanol via an alcohol dehydrogenase (EutG; absent from the strain DCMF genome) (Stojiljkovic *et al.*, 1995) or into acetyl-CoA by aldehyde dehydrogenase (EutE; 112351), which can then be further transformed into acetate (Roof and Roth, 1988, 1989). However, the approximately stoichiometric 1:1 production of MMA from choline reported in Chapter 3 make the metabolism of choline via ethanolamine unlikely, as the nitrogen in ethanolamine is instead released as ammonium.

Metabolism of choline into trimethylamine via the glycyl radical enzyme CutC is also known to occur in BMCs, but shown not to be the case in strain DCMF under physiological conditions (i.e. trimethylamine was not a product of choline or glycine betaine metabolism). The CutC homolog encoded in the strain DCMF genome (112585) contains only three of the six residues postulated to be critical for catalysis, is ~50 codons shorter than characterised CutC proteins, and is not located within a BMC gene cluster. Accordingly, it was not identified in the proteome. It may be possible that, as a structurally analogous compound, the presence of choline caused an upregulation of ethanolamine BMC proteins. It is also possible that a very minor proportion of choline was directly demethylated to di-, mono-, or unmethylated ethanolamine, triggering the translation of the ethanolamine BMC proteins. Ultimately, we have yet to elucidate whether a functional BMC is indeed present in choline-amended strain DCMF cells, and what role it may have.

Finally, proteomic data was congruent with the putative methanol metabolic pathway suggested in Section 3.4.4. The putative methanol methyltransferase gene cluster (Figure 4.7) is smaller than those implicated in DCM and glycine betaine demethylation, but still contains a likely MT_1 (*mtaB*, 112644), MT_2 (*mtaA*, 112641), CoP (*mtaC*, 112645, 112642) and transcriptional regulator (*mtaR*, 112643) gene. The putative methanol methyltransferase cluster in strain DCMF was compared to those in *Acetobacterium dehalogenans* DSM 11527, *Acetobacterium woodii* DSM 1030, *Eubacterium limosum* KIST612, *Moorella thermoacetica* ATCC 39075, and *Sporomusa ovata* H1/DSM 2662. Strain DCMF is closest to the *M. thermoacetica* methanol methyltransferase gene cluster, sharing the highest percentage amino acid identity with the MtaA (62%), MtaB (54%), and one of the MtaC (61%) protein sequences from this organism. Additionally, both strain DCMF and *M. thermoacetica*

lack the *mtaWXY* genes (thought to be involved in the cobalamin cofactor synthesis) and the more conserved genomic jksynteny seen between the other four bacteria (Kremp *et al.*, 2018). Nonetheless, there was still a relatively high level of amino acid identity between strain DCMF and the homologous proteins from *A. dehalogenans*, *A. woodii*, *E. limosum* and *S. ovata* (51-62% for MtaA, 46-53% for MtaB, 34-57% for MtaC1, 41-49% for MtaC2, and 31-44% for MtaR).

4.5 Conclusions

The proteomic experiment supported genomic predictions and physiological observations of DCM, choline, glycine betaine, and methanol metabolism by strain DCMF. A cluster of methyltransferase genes that was highly abundant in DCM-grown cells is highly conserved amongst the two other anaerobic DCM-degrading bacteria, strongly suggesting its involvement in the funnelling of dechlorinated DCM into the Wood-Ljungdahl pathway via methylene-THF. Electron equivalents released during this transformation are channelled into the reduction of both methylene-THF and CO_2 to acetate. Putative glycine betaine and methanol methyltransferase gene clusters were also identified in strain DCMF. Overall, the data depicts on organism that is a C1 specialist, with a metabolism underpinned by a wide variety of methyltransferases. Coupled with proteins for energy conservation via chemiosmotic mechanism (F_1F_0-type ATP synthase, Rnf complex, NADH:ubiquinone oxidoreductase), these allow it to generate energy from a range of methylated and C1 substrates, including the toxic environmental contaminant DCM, the quaternary amine choline and glycine betaine, and methanol.

4.6 Acknowledgements

Mass spectrometric analysis for this work was carried out at the Bioanalytical Mass Spectrometry Facility, UNSW and was supported in part by infrastructure funding from the New South Wales Government as part of its co-investment in the National Collaborative Research Infrastructure Strategy. Many thanks to Dr Ling Zhong for her guidance and assistance with the proteomic mass spectrometry and data analysis.

5 Dichloromethane-fed community diversity supported by necromass recycling

5.1 Introduction

Despite ongoing efforts to generate an axenic culture of strain DCMF, it has proven recalcitrant to isolation and exists within the DCM-Fermenting Enrichment (DFE) culture. Growth experiments carried out in Chapter 3 of this thesis showed that the abundance of strain DCMF increased concomitant with degradation of DCM, choline, glycine betaine, and methanol, and comparison of strain DCMF to total bacterial 16S rRNA gene abundance showed that it is the most abundant community member during substrate consumption. Generally, the DFE culture has been maintained on 1 mM DCM for many tens of transfers over the past six years. In light of this, the question remains as to how the other members of the DFE community persist, given that they do not appear to utilize the sole source of energy provided to the culture (DCM).

A similar phenomenon was observed by Chen *et al* (2020) in the DCM-fermenting enrichment culture RM, which contains '*Candidatus* Dichloromethanomonas elyunquensis'. Culture RM has been maintained with DCM as the sole external electron donor for seven years and over 80 transfers, yet the diverse community remains (Chen *et al.*, 2020). Investigating how these communities are sustained with only a simple, chlorinated one-carbon (C1) compound offers insight into microbial community dynamics and may shed light on how dechlorinating populations act in the environment. A better understanding of these microbial communities and interactions, particularly the role of the non-dechlorinating community, may facilitate the curation of more robust commercial cultures for bioremediation applications. It could also inform the manipulation of conditions to enhance *in situ* bioremediation, particularly with respect to the production of hydrogen by fermentative partners, which can be used as an electron donor by the majority of organohalide-respiring bacteria (Smidt and de Vos, 2004) but inhibits fermentation and mineralisation of DCM at high partial pressures (Lee *et al.*, 2012; Holland *et al.*, 2019; Chen *et al.*, 2020).

To this end, the scope of this chapter was to investigate the cohabiting organisms of the DFE community. Illumina 16S rRNA gene amplicon sequencing demonstrated temporal shifts in the community profile, in line with substrate utilization by strain DCMF, while the generation of "strain DCMF-free" enrichment cultures definitively excluded a number of the DFE culture cohabitants from metabolism of DCM, choline, and glycine betaine. Finally, previously reported genomic (Chapter 2) and proteomic (Chapter 4) data was re-analysed with a focus on the wider community to generate a metaproteogenomic model for the persistence of non-dechlorinating organisms. We propose that the cohabiting bacteria in the DFE culture are predominantly supported by necromass recycling, utilizing proteins and carbohydrates released during biomass turnover from strain DCMF.

5.2 Materials and Methods

5.2.1 Culture medium

Cultures were grown in minimal mineral salt medium as described in Section 2.2.1. Cultures for Illumina 16S rRNA gene amplicon sequencing are the same as those described in the growth experiments in Chapter 3, amended with: 2 mM DCM, 5 mM choline chloride, 5 mM glycine betaine, or 5 mM methanol as energy source.

5.2.2 Analytical techniques

DCM and methanol were quantified with GC-FID as per Section 3.2.3.1. Choline and glycine betaine were quantified with LC-MS as per Section 3.2.3.2.

5.2.3 16S rRNA gene identification and phylogeny

In order to identify any non-strain DCMF 16S rRNA genes from the PacBio sequencing data reported in Chapter 2 (NCBI Sequence Reads Archive identifier SRR5179548), all contigs from all of the attempted assemblies were combined into a single file. Strain DCMF contigs were identified and removed and the remaining contigs reduced to a set that were non-redundant at the level of 99% global sequence identity for the shorter contig, using GABLAM (Davey *et al.*, 2006). Where redundancy was identified, the longer contig was retained. In total, 20,201 contigs were reduced to 1,538 non-redundant non-strain DCMF contigs, hereon referred to as "NR Cohabitants". Cohabitant bacteria 16S rRNA gene sequences were identified

from NR Cohabitants using barrnap v0.9 (implementing HMMer v3.2.1 and bedtools v2.27.1). Sequences were mapped to taxa using the SILVA Alignment, Classification and Tree Service (Pruesse *et al.*, 2012) with default values.

5.2.4 Enrichment of DFE cohabitant bacteria

To eliminate strain DCMF and enrich the remaining DFE community members, dilution to extinction cultures (20 ml) were set up in 30 ml glass serum bottles (Figure S4). These were prepared with the standard minimal mineral salt medium amended with one of: casamino acids (5 g l^{-1}), ethanol (10 mM), glucose (10 mM), peptone (5 g l^{-1}), 1-propanol (10 mM), yeast extract (5 g l^{-1}). After two rounds of dilution to extinction, DNA was extracted from the lowest active dilution culture as per Section 2.2.3 and used for qPCR targeting the strain DCMF 16S rRNA gene (as per Section 3.2.5). In all cases, this gene was below the qPCR limit of detection and so the dilution culture was analysed with Illumina 16S rRNA gene amplicon sequencing and used to inoculate triplicate flasks of minimal mineral salt medium amended with one of: 1 mM DCM, 5 mM choline chloride, or 5 mM glycine betaine (Figure S4). An abiotic (uninoculated) control was also included for each of the DCM-, glycine betaine- and choline-amended cultures. These cultures were monitored fortnightly for any change in substrate concentration over eight weeks.

5.2.5 Community analysis via Illumina 16S rRNA gene amplicon sequencing

DNA was extracted from cultures during exponential growth phase, as per Section 2.2.3. The 16S rRNA gene was amplified with primers 515F (5′-TCGTCGGCAGCGTCAGATGTGTATAAGAGACAG-[GTGYCAGCMGCCGCGGTAA]-3′) and 806R (5’-GTCTCGTGGGCTCGGAGATGTGTATAAGAGACAG-[GGACTACNVGGGTWTCTAAT -3′) (Caporaso *et al.*, 2011). Reaction mixtures (total volume 40 µl) contained: 2 µl template DNA (diluted to between 1 – 5 ng ul^{-1}), 20 µl EconoTaq PLUS Green 2X master mix (Lucigen), 0.4 µl each forward and reverse primers (10 µM stock concentration; IDT), 17.2 µl molecular biology grade water (Sigma). Reactions were performed on an MJ Mini Bio-Rad thermocycler. The PCR products were verified on a 1% agarose gel before processing and sequencing with Illumina MiSeq technology at the Next Generation Sequencing Facility (University of Western Sydney).

16S rRNA gene amplicon forward and reverse reads were manually inspected with FastQC to determine where they should be truncated (i.e. where the Phred scores fell below 20). Reads were then processed and OTUs were assigned in QIIME2 (Bolyen *et al.*, 2019). Using the dada2 pipeline (Callahan *et al.*, 2016), forward and reverse reads were trimmed and joined, chimeras were removed, and samples were rarefied to the lowest sequencing depth. Taxonomy was assigned to genus level using a Naïve Bayes classifier trained on a full-length 16S rRNA gene SILVA database (release 133) and low abundance organisms were filtered out by removing the lowest 1% abundant reads. Alpha diversity was assessed with Shannon's diversity index and compared pairwise between samples and substrate consumption time points with a Kruskal-Wallis test. Two-dimensional PCA plots were created from the weighted Unifrac distance matrix both with and without the methanol-amended culture samples. As each substrate was consumed at different rates, samples were not compared by timepoints. Rather, each sample was classed into one of six groups describing the state of substrate consumption, 'start (the inoculum and all day 0 samples), 'pre', 'early', 'mid', 'late', or 'post' (Table S5).

5.2.6 Metagenome assembly and annotation

The NR Cohabitants contigs were assembled using Unicycler v0.4.7 (Wick *et al.*, 2017) with default parameters. The assembled contigs were then frameshift-corrected via DIAMOND v0.9.24 (Buchfink *et al.*, 2014) and MEGAN Community Edition v6.16.4 (Huson *et al.*, 2016), as previously described (Arumugam *et al.*, 2019). Although taxon bins were generated by MEGAN, the final set of taxonomic bins were generated separately by Maxbin v2.2.4 (Wu *et al.*, 2016). CheckM v1.1.2 (Parks *et al.*, 2015) was used to remove reads that were outliers by all three measures of G+C content, coding density, and tetranucleotide frequency from certain bins, and then to assess the completeness and contamination of the outlier-refined bins, henceforth referred to as MAGs (metagenome-assembled genomes). Taxonomy was assigned to the MAGs with GTDB-tk v1.0.2 (Chaumeil *et al.*, 2019) and gene calling was performed with Prokka v1.13.3 (Seemann, 2014). Nucleotide files for the five MAGs are available online (Dataset S3). The full list of Prokka-annotated proteins are available online in Dataset S4.

Prokka-annotated proteins were further annotated with KEGG orthology groups and taxonomically classified with GhostKOALA (Kanehisa *et al.*, 2016). Metabolic pathways and processes were predicted using the inbuilt 'Reconstruct pathway' tool based on this KEGG annotation. PSORTb (Yu *et al.*, 2010) was used to predict the sub-cellular localization of proteins in each metagenome bin. dbCAN2 (Zhang *et al.*, 2018) was used to predict Carbohydrate Active enZymes (CAZymes) via searches of the protein annotations against the dbCAN CAZyme domain HMM database (Lombard *et al.*, 2014), DIAMOND (Buchfink *et al.*, 2014) search against the CAZy database (Lombard *et al.*, 2014), and Hotpep search against the peptide pattern recognition library (Busk *et al.*, 2017). Hydrogenase catalytic subunits were classified using HydDB (Søndergaard *et al.*, 2016).

5.2.7 Metaproteome analysis

The proteomics data from Chapter 4 was re-analysed to investigate the metaproteome of the DFE community. The MS/MS spectra .raw files generated from the analysis of DFE culture samples amended with DCM, choline, or glycine betaine in Section 4.2.5 were this time searched against two custom databases in MaxQuant v1.6.10.43 (Cox *et al.*, 2014): the Prokka-annotated MAGs described above and the strain DCMF genome as before. Search parameters were as described in Section 4.2.6.

Statistical analysis of the MaxQuant output was performed in Perseus v1.6.7.0 (Tyanova *et al.*, 2016). Proteins identified by site, reverse sequences and potential contaminants were removed. Proteins were filtered to retain only those identified by two or more unique peptides, and present in at least two replicates of one substrate condition (a more permissive filter than for the strain DCMF proteins in Chapter 4, due to the expected lower abundance of cohabiting bacterial proteins in the samples). Label free quantitative (LFQ) intensities were $\log_2$ transformed and triplicate values were averaged. For analysis purposes, ambiguously identified proteins (i.e. in cases where peptides matched multiple proteins in one or more taxonomic bins) were counted as present in all possible taxa. Proteins identified from strain DCMF (either solely or ambiguously in strain DCMF and ≥1 cohabiting organism) were removed from the dataset. Protein abundance data is included in the full list of Prokka-annotated proteins (Dataset S4).

5.3 Results

5.3.1 16S rRNA gene phylogeny identified five phylotypes in the DFE community

To investigate the members of the DFE community other than strain DCMF, the PacBio sequencing data generated in Chapter 2 was re-analysed with a focus on the non-strain DCMF reads. The non-redundant PacBio sequencing reads that did not align to the final strain DCMF genome (referred to herein as "NR Cohabitants") were searched for 16S rRNA gene sequences. In total, 17 16S rRNA gene sequences were identified. Based on SILVA classification, these sequences were clustered into five distinct phylotypes (identified here by their lowest classified taxonomic rank): *Synergistaceae*, *Spirochaetaceae*, *Desulfovibrio*, *Ignavibacteria*, and *Lentimicrobiaceae*. These were supported by clear clades within the unrooted SILVA tree (Figure 5.1).

The 16S rRNA genes were searched against the NCBI prokaryotic 16S rRNA database with blastn. The closest relatives to the five *Desulfovibrio* sequences were *Desulfovibrio oryzae* (94.74 – 99.54% identity), *Desulfovibrio oxamicus* strain DSM 1925 (94.26 - 98.81%) and *Desulfovibrio longreachensis* strain 16910a (94.01 - 98.88%). The single *Ignavibacteria* sequence appears to be distantly related to currently known organisms, sharing only 85.63% identity with its closest relative, *Ignavibacterium album* strain JCM 16511, followed by 83.65% with *Melioribacter roseus* strain P3M-2 and 82.44% with *Chlorobaculum limnaeum* strain DSM 1677. The longer of the two *Lentimicrobiaceae* sequences was also relatively far removed from its closest relatives: *Lentimicrobium saccharophilum* strain TBC1 (91.86%), *Alistipes finegoldii* strain DSM 17242 (86.06%), and *Owenweeksia hongkongensis* strain DSM 17367 (86.00%). The closest relatives to the four *Spirochaetaceae* sequences were: *Rectinema cohabitans* strain HM (93.22 – 95.05%), *Treponema caldarium* strain DSM 7334 (88.24 – 90.14%), and *Treponema stenostreptum* strain DSM 2028 (87.99 – 89.45%); while those for the four *Synergistaceae* sequences were: *Thermovirga lienii* strain DSM 17291 (89.22 - 89.28%), *Thermanaerovibrio velox* strain Z-9701 (88.54 - 88.60%), and *Thermanaerovibrio acidaminovorans* DSM 6589 (87.70 - 87.76%). The NR Cohabitants 16S rRNA genes were placed in a phylogenetic tree with these most closely related sequences (Figure 5.2).

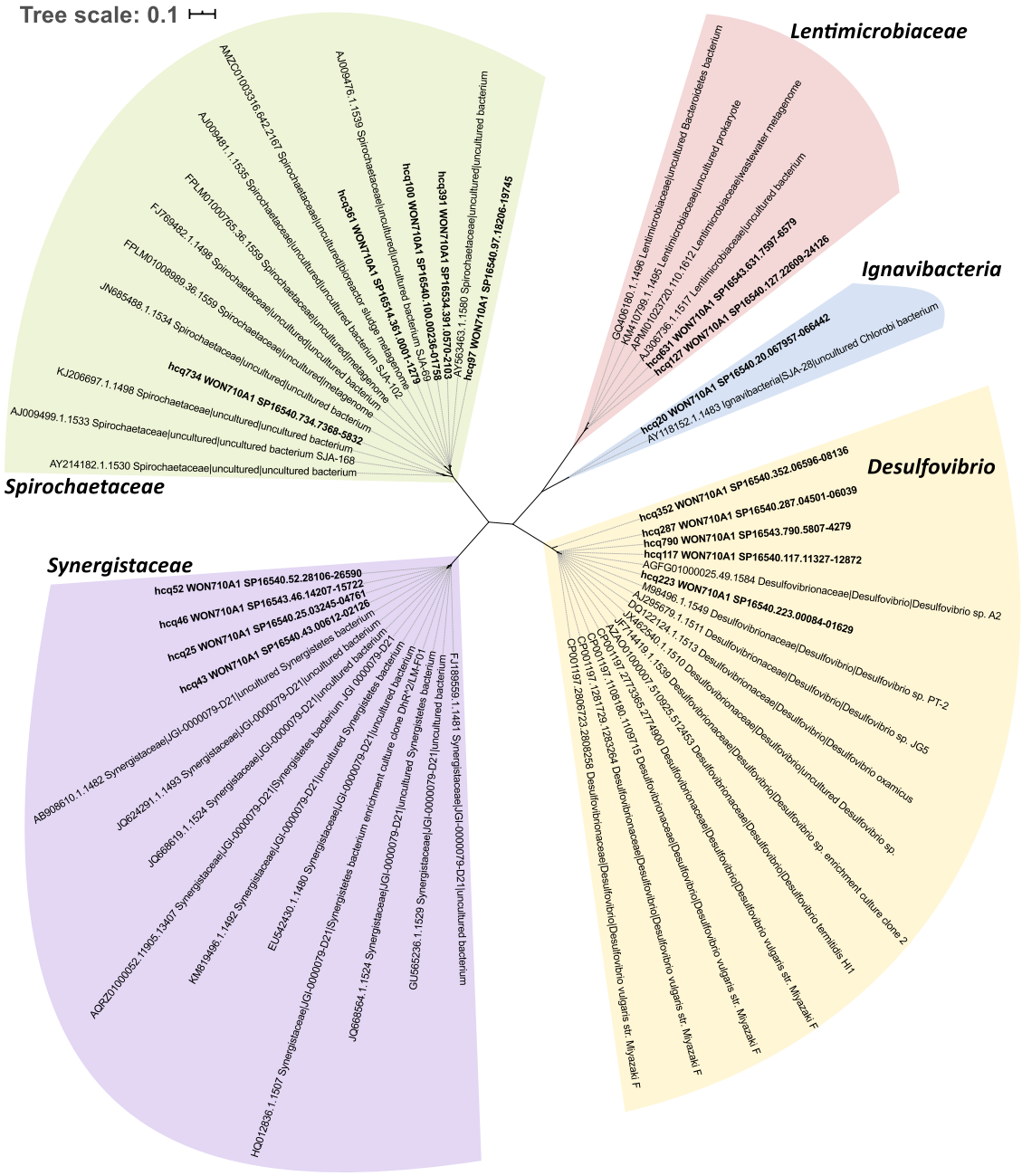


Figure 5.1 The 16S rRNA genes identified in the DFE metagenome cluster into five distinct clades. Phylogenetic tree of all non-strain DCMF 16S rRNA genes identified in the DFE culture via PacBio sequencing, plus their nearest relatives. The DFE culture cohabitants cluster into five distinct taxonomic groups: *Spirochaetaceae* (green), *Lentimicrobiaceae* (red), *Ignavibacteria* (blue), *Desulfovibrio* (yellow), and *Synergistaceae* (purple). Sequences were mapped to taxa using the SILVA Alignment, Classification and Tree service with default values (i.e., the 10 closest relatives with >95% identity to each sequence are included in this tree). Sequences from the DFE culture are written in **bold**. All other sequences are identified from family level down as per SILVA taxonomy.

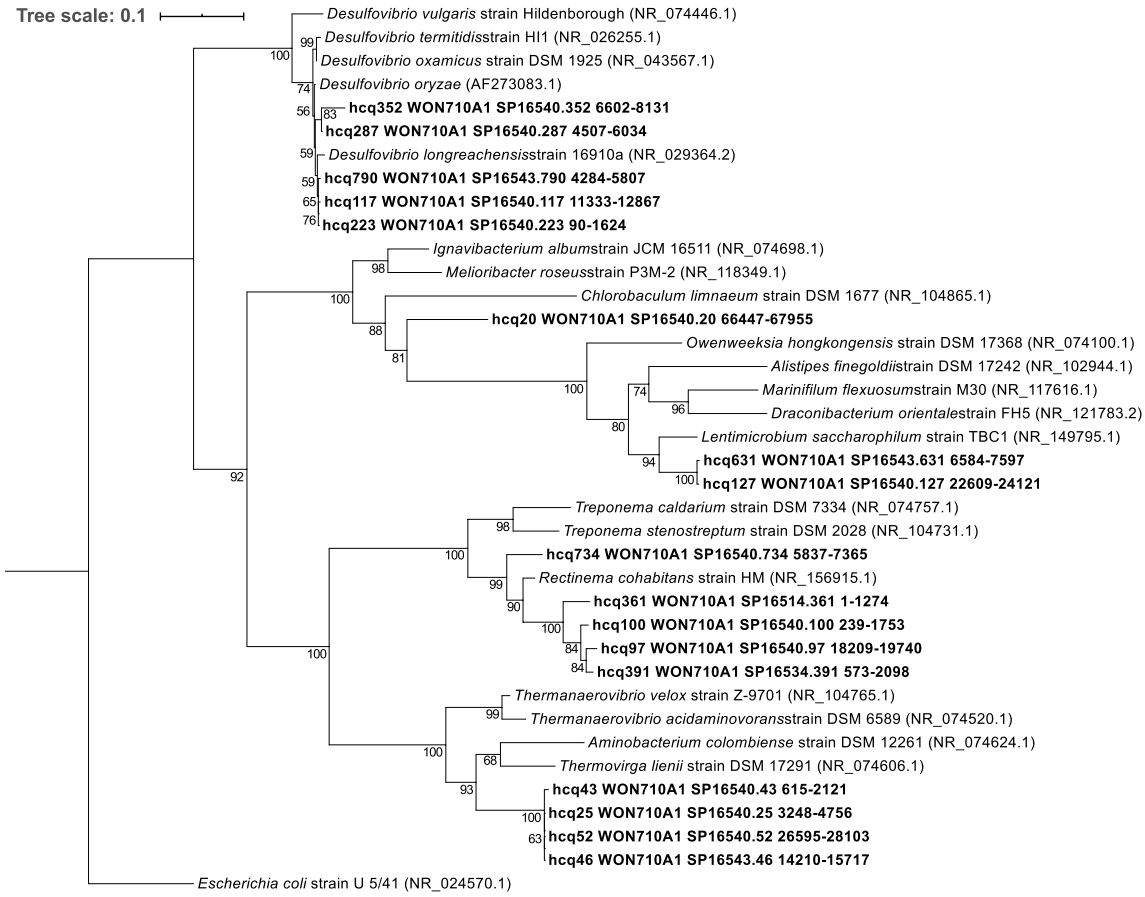
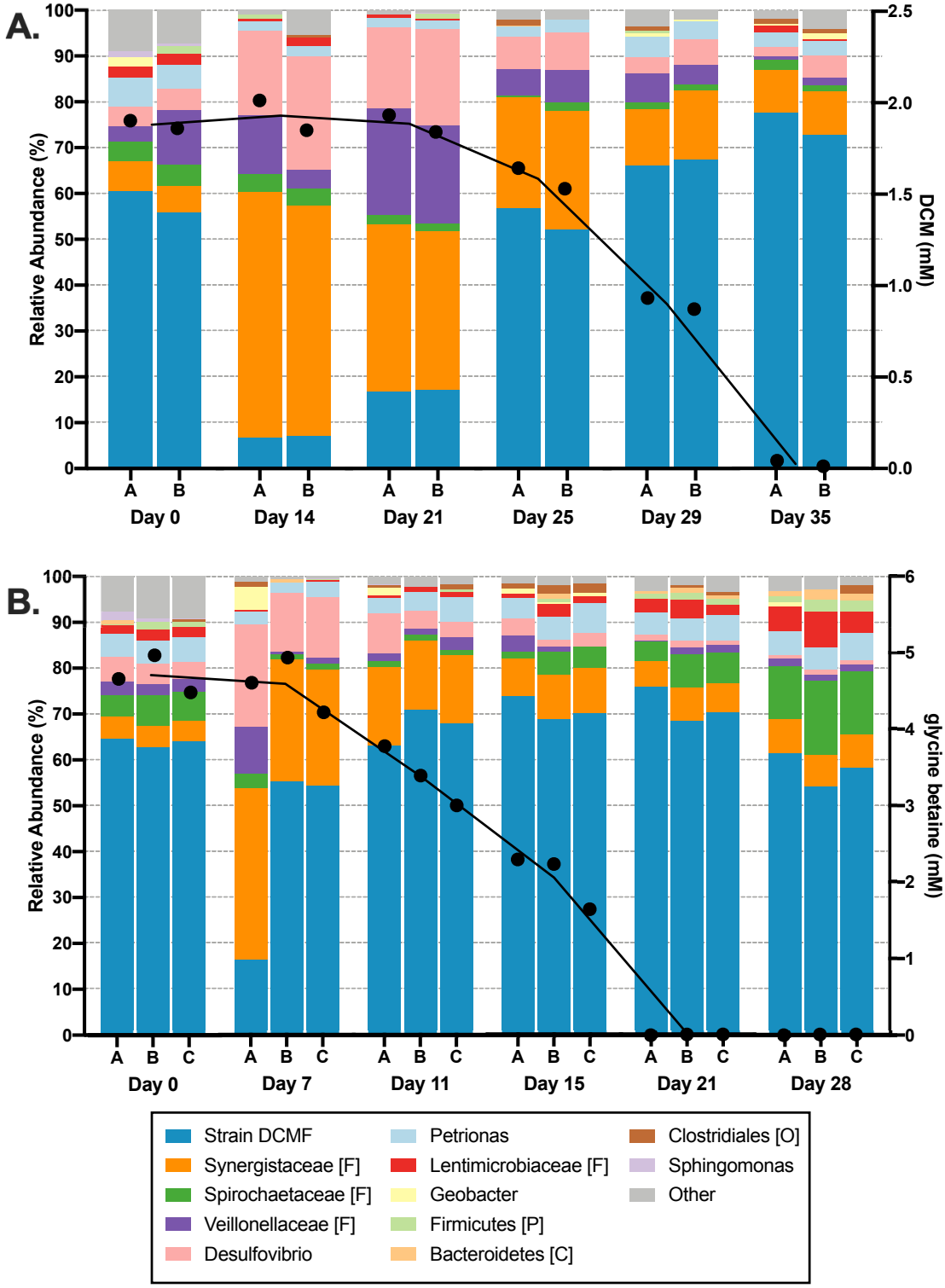


Figure 5.2 Phylogenetic tree of the 16S rRNA sequences found in the DFE metagenome and their closest relatives. 16S rRNA sequences from the DFE metagenome are in **bold**. Closest relatives were identified by blastp search against the nr and 16S ribosomal RNA databases. The top three results for formally named taxa only (i.e., "Uncultured bacterium" sequences were discarded) were included. Numbers indicate percentage of branch support from 1000 bootstraps. *Escherichia coli* strain U 5/41 was used to root the tree.

5.3.2 Shifts in DFE community composition in response to substrate consumption

Community profiling with Illumina 16S rRNA gene amplicon sequencing showed that the DFE community did not change substantially in cultures with a common, DCM-fed inoculum amended with either DCM, glycine betaine, or choline. Methanol-amended cultures however, which had been maintained on this substrate for the previous two generations, had a distinct community profile (Figure 5.3). All methanol-amended cultures lacked the taxa *Synergistaceae* and *Bacteroidetes*, and the majority of these samples also lacked *Clostridiales*, *Firmicutes*, *Geobacter*, *Lentimicrobiaceae* and *Sphingomonas* (Figure 5.3).


A.
Relative Abundance (%)
DCM (mM)
Day 0
Day 14
Day 21
Day 25
Day 29
Day 35
B.
Relative Abundance (%)
glycine betaine (mM)
Day 0
Day 7
Day 11
Day 15
Day 21
Day 28
Strain DCMF
Synergistaceae [F]
Spirochaetaceae [F]
Veillonellaceae [F]
Desulfovibrio
Petrionas
Lentimicrobiaceae [F]
Geobacter
Firmicutes [P]
Bacteroidetes [C]
Clostridiales [O]
Sphingomonas
Other

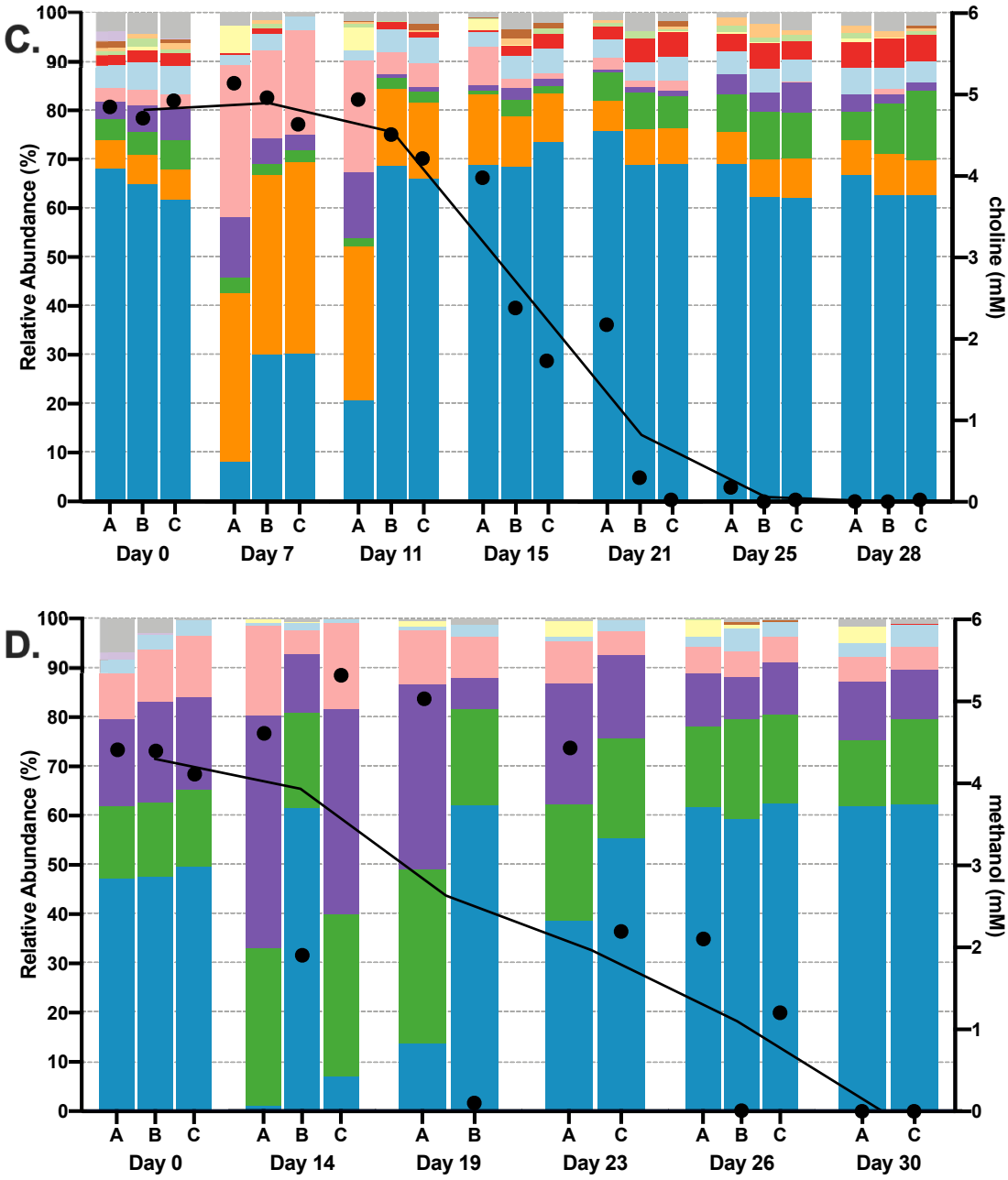


Figure 5.3 The DFE community is subject to temporal shifts in composition. Illumina 16S rRNA amplicon sequencing was used to determine DFE community composition (left y axis) in the growth experiment cultures (reported in Chapter 3) amended with DCM (A), glycine betaine (B), choline (C) or methanol (D). Taxa are reported down to genus level where possible, otherwise taxonomic level is indicated in the legend ([F] = family, [P] = phylum, [C] = class, [O] = order). Reads with <1% abundance were filtered out in QIIME2. Unassigned reads and taxa with <2% relative abundance in every sample were classed together as 'Other'. Substrate concentration (black circles, right y axis) and a line connecting the mean substrate concentration at each time point is overlaid on the community composition graphs. These are aligned with the time points written on the x axis, not drawn to scale.

The difference in the methanol-amended cultures was supported by a PCA plot of the weighted Unifrac distance matrix (a measure of diversity that takes both the phylogenetic distance between taxa and their relative abundance into account) in which they appear separated from the other samples (Figure 5.4A). Methanol-amended cultures had a significantly higher degree of evenness (adjusted p-value <0.01 in pairwise Kruskal-Wallis analysis of methanol compared to all other substrates) compared to cultures on the other three substrates, reflective of the lower relative abundance of strain DCMF in these cultures compared to DCM-, glycine betaine-, or choline-amended cultures (Figure S5).

Nonetheless, at the time of inoculation and during substrate consumption, strain DCMF dominated the DFE community with each of the four substrates. During the lag phase, there was a large increase in the relative abundance of *Synergistaceae* (except in methanol-amended cultures, where this taxon was absent), and a moderate increase in *Desulfovibrio* and *Veillonellaceae*. Indeed, the relative abundance of strain DCMF dropped to just 0.96% in one methanol-amended replicate at day 14 (Figure 5.3 D). Other taxa such as *Spirochaetaceae* and, to a lesser extent, *Lentimicrobiaceae*, tended to increase in relative abundance towards the end of substrate consumption and once all substrate had been depleted (Figure 5.3). Given that both strain DCMF and total bacterial cell numbers continued to increase over the course of substrate degradation in all cultures (Figure 3.2, Figure 3.5, Figure 3.7), this increase in relative abundance could reasonably be interpreted as an increase in actual abundance of the cohabiting bacteria as a whole (as opposed to a decrease in strain DCMF causing an artificial increase in relative abundance of the cohabitants).

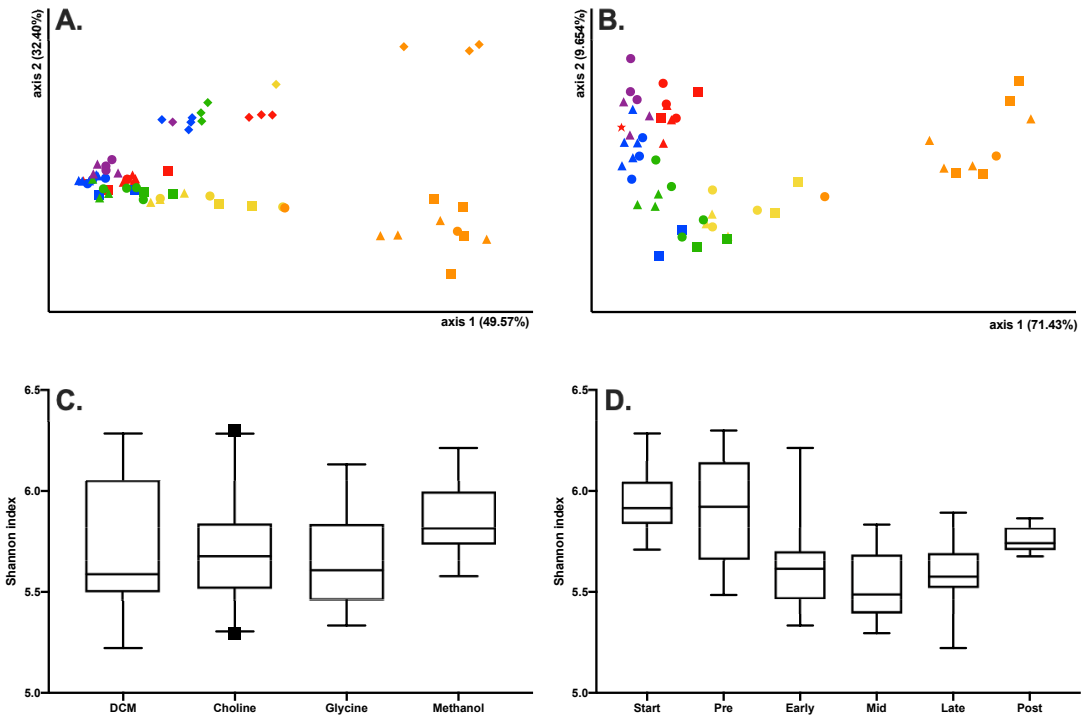


Figure 5.4 The DFE culture community is most different during the culture lag phase. PCA plots of the weighted Unifrac distance matrices of samples including (A) or excluding (B) those from methanol-amended cultures show that the Illumina 16S rRNA amplicon sequencing samples tend to group together based on substrate consumption (colours) more than substrate (shapes). The exception to this was the methanol-amended community (diamonds), which had a different inoculum and was clearly separate from the DCM- (square), choline- (triangle), or glycine betaine-amended (circle) communities (A). The inoculum for the latter three sets of cultures is represented by a star. The time series samples were grouped according to substrate consumption: inoculum/day 0 samples (red), pre (orange), early (yellow), mid (green), late (blue), or post (purple) (Table S5). There was no significant difference in the Shannon diversity index between samples grouped by substrate (C), but significant differences between samples when grouped by substrate consumption (D).

When samples were classed in relation to substrate consumption (start, pre, early, mid, late, or post substrate consumption; Table S5), clear groups emerged in the weighted Unifrac PCA plots (Figure 5.4 A and B). Clear differences were observed in the DFE community before, during, and after substrate consumption, particularly noticeable when the less-related methanol-amended samples were omitted from the PCA analysis Figure 5.4 B). The 'pre' substrate consumption samples showed the highest degree of difference to all others. While there was no significant difference in the Shannon diversity index between the samples when grouped by substrate

(Kruskal-Wallis p-value 0.0976; Figure 5.4 C), there was a highly significant difference between all groups when clustered by substrate consumption (Kruskal-Wallis p-value <0.00001 Figure 5.4 D).

5.3.3 Exclusion of some DFE cohabitants from dichloromethane, choline, or glycine betaine fermentation

Strain DCMF did not form colonies on solid or semi-solid medium, which limited isolation attempts to successive rounds of dilution to extinction cultures. This proved ineffective as the cohabiting organisms persisted in the DFE community. In lieu of generating an axenic culture of strain DCMF to demonstrate that it is the singular organism able to directly metabolise DCM, glycine betaine, choline, and methanol, the cohabiting bacteria in the DFE community were enriched to the exclusion of strain DCMF and their inability to transform these substrates in its absence was demonstrated. Dilution to extinction cultures amended with casamino acids, glucose, peptone, and yeast extract all demonstrated robust growth. No growth was observed in cultures amended with ethanol or 1-propanol over a period of eight weeks (data not shown).

Strain DCMF 16S rRNA genes were below the limit of detection in the lowest active enrichment culture after two rounds of dilution to extinction. These cultures were then subject to Illumina 16S rRNA gene amplicon sequencing to determine the bacterial community. Cultures amended with casamino acids resulted in a community dominated by *Veillonellaceae* (relative abundance 79.4%), as well as *Desulfovibrio* (8.3%), *Geobacter* (7.5%), and *Bacillus* (4.9%) (Figure 5.5 E). Both the glucose and yeast extract-amended cultures almost exclusively enriched *Petrimonas*, with a low proportion of *Desulfovibrio* remaining in the cultures (1.1% and 3.1% relative abundance, respectively; Figure 5.5 F and H). The peptone-amended cultures enriched *Desulfovibrio* to apparent purity (100% relative abundance; Figure 5.5 G). However, some phylotypes with a significant relative abundance in the original DFE community (Figure 5.3) were not present in the strain DCMF-free enrichments (*Spirochaetaceae*, *Synergistaceae*).

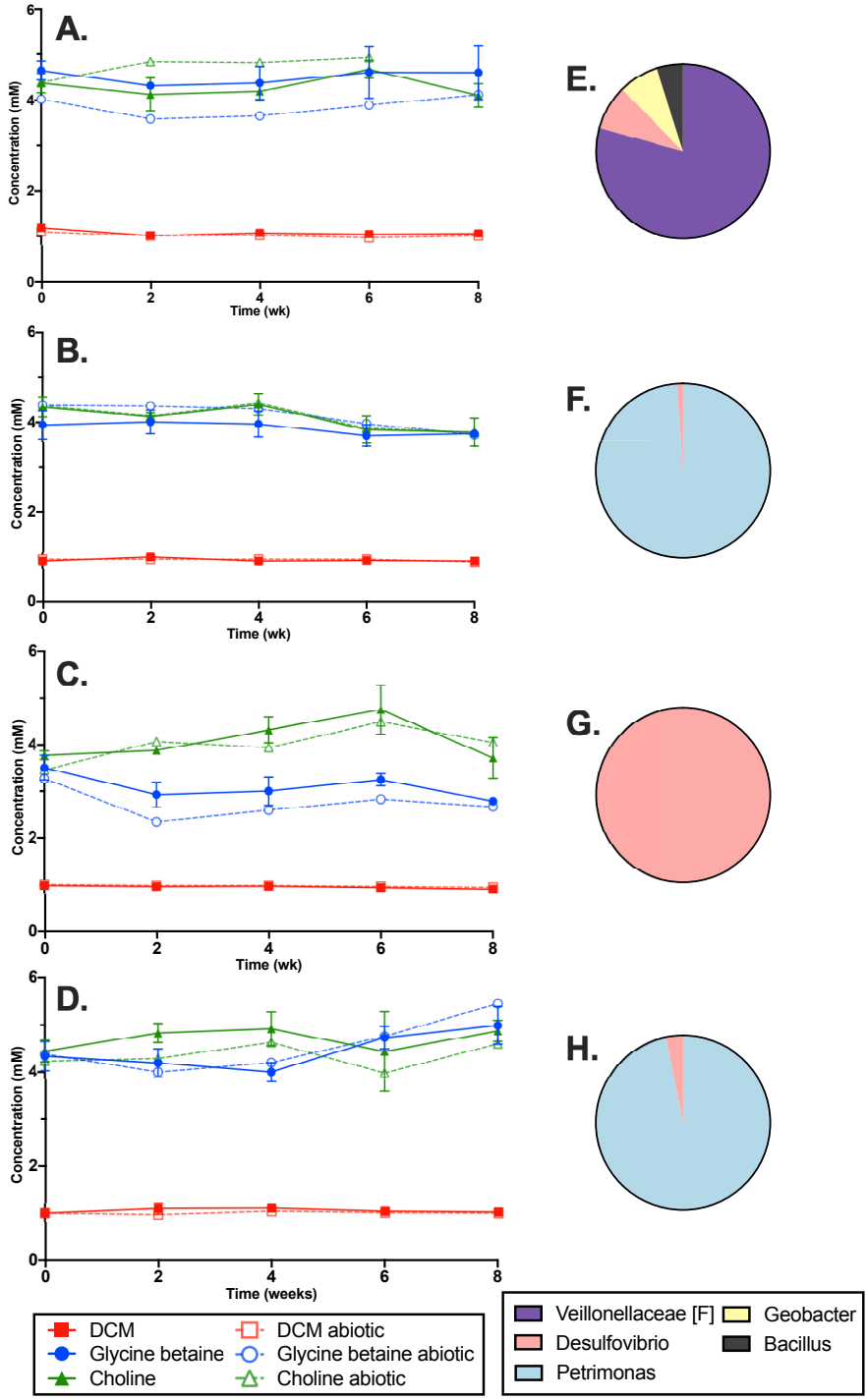


Figure 5.5 DFE community enrichments excluding strain DCMF cannot consume DCM, glycine betaine, or choline. DFE cohabitants were enriched on casamino acids (A), glucose (B), peptone (C), yeast extract (D) to the exclusion of strain DCMF, then transferred back into medium containing the typical strain DCMF substrates (DCM, glycine betaine, choline). Active cultures n = 3; abiotic n = 1; error bars represent standard deviation. Community profiling was carried out via Illumina 16S rRNA gene amplicon sequencing with the lowest active dilution culture in the second round of dilution-to-extinction on casamino acids (E), glucose (F), peptone (G), yeast extract (H); i.e., the cultures used as inoculum for the cultures depicted in (A) to (D), respectively.

These strain DCMF-free dilution cultures were then used to inoculate triplicate cultures amended with DCM (1 mM), glycine betaine (5 mM) or choline (5 mM; see Figure S4 for cultivation strategy). No significant substrate depletion was observed in any of the cultures (linear regression showed no significantly non-linear slope; Figure 5.5), with two exceptions: choline depletion in the glucose-enriched community cultures (slope = -0.07 [95% CI -0.12 to -0.016], p-value = 0.0143; Figure 5.5 B) and DCM depletion in the peptone-enriched community cultures (slope = -0.01 [95% CI -0.012 to -0.006], p-value < 0.0001; Figure 5.5 C). However, the choline measurements in the glucose-enriched community closely mimicked the abiotic control and the slope in the DCM-amended peptone-enriched community was so slight as to have negligible effect after an eight-week incubation, while DFE cultures containing strain DCMF typically consume 1 mM DCM within the first three weeks (including the lag phase). There is therefore no evidence of DCM, choline, or glycine betaine degradation by the *Bacillus*, *Desulfovibrio*, *Geobacter*, *Petrimonas*, or *Veillonellaceae* phylotypes in the DFE culture.

5.3.4 Genome-centric metagenomics of the DFE community

In order to further investigate the role of the cohabiting bacteria in the DFE culture and provide a taxa-specific platform for metaproteomic analysis, metagenomic assembly of the NR Cohabitants reads was carried out. A total of 195,364 long reads (732,500,489 bp) assembled into 398 contigs (total size 13,752,672 bp). This included 395 linear contigs and three putative circularized contigs. Of the NR Cohabitants reads, 33.6% mapped back to the assembled metagenome, which had an average of 18.6x coverage depth (Figure S6).

The reads were sorted into five distinct bins, with all bar two contigs binned (Table 5.1). The N50 for each bin varied from 17,998 – 292,358 bp (Table 5.1), and full Nx plots were also produced (Figure S7). Bins one and two had a relatively high level of completion (87.38% and 95.30%, respectively), with very little contamination (Table 5.1). Bin three was the only one to contain any measure of strain heterogeneity (16.67%). Based on their completeness, contamination, and presence of rRNA genes, bins one to four represent medium quality draft metagenome-assembled genomes (MAGs), while bin five is a low-quality MAG (Bowers *et al.*, 2017). Nucleotide sequences for each MAG are available in Datasets S3.

Table 5.1 Summary of the bins identified from the DFE metageonme assembly.

Bin	Total contigs (putative circular contigs)	Size (bp)	N50 (bp)	Coverage depth	Completeness (%)[a]	Contamination (%)[a]	Strain Heterogeneity (%)[a]
1	22	3,197,875	292,359	26.6	87.38	1.75	0
2	62	2,496,147	54,174	20.2	95.30	0.00	0
3	131 (2)	3,734,740	35,707	15.0	70.78	1.92	16.67
4	99 (1)	2,312,344	28,719	13.4	64.87	8.05	0
5	80	1,394,855	17,998	10.8	25.78	0.00	0

[a] Determined by CheckM.

The classification of each MAG roughly matched the taxa identified via the 16S rRNA gene search. Based on GTDB-tk classification (identified here down to their lowest formally-named taxonomic rank), the five bins were classified as belonging to the class *Ignavibacteria*, the order *Synergistales*, the family *Lentimicrobiaceae*, the order *Treponematales*, and the genus *Desulfovibrio* (Table 5.2). Only Bin 5 could be classified down to genus level, with the others appearing to represent novel organisms at higher taxonomic levels. It appears that the *Ignavibacteria* phylotype was lost from the DFE culture between the time of PacBio sequencing (2015) and Illumina 16S rRNA amplicon analysis (2019). A set of qPCR primers specific to the 16S rRNA gene classed as *Ignavibacteriaceae* also failed to produce a product in DNA samples from the DFE culture in 2017 (unpublished data). Therefore, no further metagenomic or metaproteomic analysis of *Ignavibacteria* was carried out. The four remaining phylotypes are henceforth referred to as SYN (*Synergistales*), LEN (*Lentimicrobiaceae*), SPI (*Treponematales*), and DSV (*Desulfovibrio*; Table 5.2).

Table 5.2 Summary of the MAGs retrieved from the DFE community metagenome.

Bin	Code	GTDB-tk Taxonomic Classification	Size (Mb)[a]	GC Content (%)	CDS	rRNA	tRNA
1	IGN	d__Bacteria;p__Bacteroidota; c__Ignavibacteria;o__SJA-28; f__B-1AR;g__;s__	3.20	43.0	2,901	2	42
2	SYN	d__Bacteria;p__Synergistota; c__Synergistia;o__Synergistales; f__79-D21;g__79-D21;s__	2.50	60.0	2,368	3	48
3	LEN	d__Bacteria;p__Bacteroidota; c__Bacteroidia;o__Bacteroidales; f__Lentimicrobiaceae;g__UBA4417;s__	3.73	44.0	3,346	0	41
4	SPI	d__Bacteria;p__Spirochaetota; c__Spirochaetia;o__Treponematales; f__UBA8932;g__UBA8932;s__	2.30	56.3	2,187	0	40
5	DSV	d__Bacteria;p__Desulfobacterota_A; c__Desulfovibrionia; o__Desulfovibrionales; f__Desulfovibrionaceae; g__Desulfovibrio_A;s__Desulfovibrio_A sp000226255	1.39	67.3	1,200	1	16

[a] Total size of bin including circular contigs

Annotation with Prokka revealed 1,200 coding sequences (CDS) in the smallest and least complete MAG (DSV) and up to 3,346 CDS in the largest MAG (LEN). Of the 17 16S rRNA gene sequences identified in Section 5.3.1, only two were present in the assembled metagenome bins (DFE_IGN_01820 and DFE_SYN_01601), indicating that the majority of reads containing these sequences were not used in the metagenome assembly. Protein localization, as predicted by PSORTb, showed an above-average proportion of extracellular proteins in LEN (2.60%; Figure S8). These proteins were possible candidates for extracellular degradation of larger

molecules that might be required for bacterial necromass fermentation. Putative carbohydrate active enzymes (CAZymes) were then predicted, with the greatest number by far also found in LEN (116 proteins). Of the predicted CAZymes, two were also predicted to be extracellular – DFE_LEN_00583, a putative glycoside hydrolase, and DFE_LEN_00300, a trehalose synthase/amylase. SPI and SYN both encoded 23 putative CAZymes, while DSV did not contain any.

5.3.5 Phylotype-resolved metaproteomics

Proteomic data generated in Chapter 3 was this time searched against both the predicted strain DCMF proteome and the predicted DFE culture metaproteome databases (see Section 5.2.7 for details), resulting in identification of 234 unique proteins from the DFE cohabitants under the three investigated substrate conditions (DCM, glycine betaine, and choline; Dataset S4). This included 15 proteins that were ambiguously identified and assigned to >1 cohabiting bacteria. Ambiguously assigned proteins were kept in the dataset, bringing the working total to 249 proteins (Dataset S4), but treated with caution during assignment of function to any one specific taxon. The methanol-amended culture samples were not used for metaproteomics due to the absence of two of the four taxa (*Treponematales/Spirochaetes* and *Synergistales*; Figure 5.3 D) in the DFE metagenome search database from these samples.

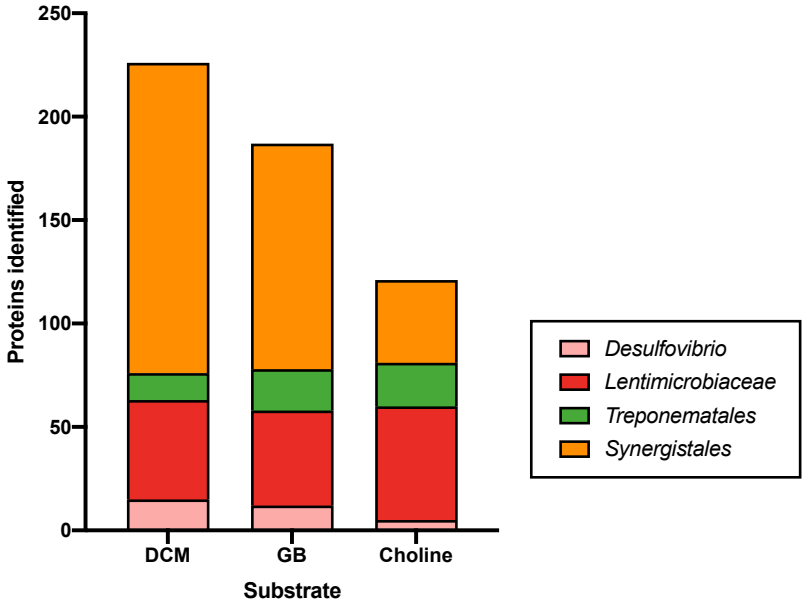


Figure 5.6 The number of proteins assigned to each taxa in cultures amended with DCM, glycine betaine [GB], or choline. Proteins that were ambiguously identified (i.e. the peptides could belong to proteins in multiple taxa) are represented under each of those taxa.

The vast majority (151) of proteins were expressed by SYN, followed by 56 from LEN, 22 from SPI, and 16 from DSV (Figure 5.6). Four proteins were assigned to the *Ignavibacteria* phylotype, but these were all ambiguously assigned to other taxa as well, further confirming the disappearance of *Ignavibacteria* from the DFE culture. These four proteins were not counted towards the listed totals. The DCM-amended cultures yielded the highest number of proteins in the metaproteome (226), followed by the glycine betaine- (187) and then choline-amended cultures (121; Figure 5.6).

Many of the most highly abundant proteins had roles in crucial cell functions such as DNA replication and repair, e.g. DNA chaperonin (DFE_SYN_00508), DNA-binding protein HU (DFE_SYN_00992 and DFE_LEN_01906), and elongation factor Tu (DFE_LEN_01289; Table 5.3). However, other highly abundant proteins provided some insight into what substrates the cohabiting organisms may metabolise.

Table 5.3 The top 20 most abundant proteins identified in the DFE metaproteome. The top blastp result is included in brackets for hypothetical proteins or those where the results significantly differ from the Prokka annotation. nd = not detected.

Locus tag	Prokka annotation	Role	log_2 LFQ value		
			DCM	GB	CHO
DFE_LEN_01134	NAD-specific glutamate dehydrogenase	Nitrogen metabolism	29.0	28.4	30.5
DFE_SYN_00508	60 kDa chaperonin	Protein folding	31.4	29.3	26.3
DFE_SYN_00625	hypothetical protein (ABC transporter substrate-binding protein)	Transport of branched chain amino acids into the cell	30.1	27.7	nd
DFE_SYN_01900	hypothetical protein (ABC transporter substrate-binding protein)	Transport of branched chain amino acids into the cell	32.0	28.9	25.4
DFE_SYN_00992	DNA-binding protein HU	DNA replication, recombination and repair	28.7	30.0	27.5
DFE_LEN_01906	DNA-binding protein HU-beta	DNA replication, recombination and repair	27.7	28.9	29.3
DFE_SYN_01677	hypothetical protein (C4-dicarboxylate ABC transporter)	Transport of dicarboxylates into the cell	32.5	30.4	22.9
DFE_SYN_02216	Glutamate dehydrogenase	Nitrogen metabolism	31.6	29.3	23.7
DFE_SYN_00924	Selenocysteine-containing peroxiredoxin PrxU	Antioxidant	29.5	26.2	nd
DFE_SYN_02067	putative L-lysine-epsilon aminotransferase	Amino acid metabolism	28.3	27.2	nd
DFE_LEN_01289	Elongation factor Tu	Translation	28.1	27.5	27.7
DFE_SYN_00844	Ectoine-binding periplasmic protein TeaA (C4-dicarboxylate ABC transporter)	Transport of dicarboxylates into the cell	30.9	28.1	24.0
DFE_SYN_01114; DFE_SYN_01119	hypothetical protein (bacteriocin)	Exported antibacterial compound	28.7	27.8	26.3

Locus tag	Prokka annotation	Role	log_2 LFQ value		
			DCM	GB	CHO
DFE_SYN_01280	hypothetical protein (SYNERG-CTERM sorting domain-containing protein)	Unknown	28.9	26.3	nd
DFE_LEN_02184	Vitamin B12 transporter BtuB (SusC/RagA family TonB-linked outer membrane protein)	Transport of large substrates into the cell	27.8	27.0	28.0
DFE_SYN_00756	2,3-dimethylmalate dehydratase small subunit	Amino acid metabolism	30.4	27.0	25.1
DFE_SYN_01695	Glycine reductase complex component B subunits alpha and beta	Metabolism of glycine, glycine betaine, or sarcosine	27.5	nd	nd
DFE_DSV_00655	Sulphate adenylyltransferase	Sulphate metabolism	29.1	27.4	25.9
DFE_SYN_01548	hypothetical protein (DUF883 family protein)	Stress response?	28.5	26.5	nd
DFE_LEN_02395	hypothetical protein (DUF4440-domain containing protein; nuclear transport factor 2 family protein)	Protein transport into the nucleus?	28.0	26.5	27.9

5.3.6 Metaproteogenomic insight into the DFE community

The presence of numerous electron transport chain genes in the LEN MAG, including those for complex I (DFE_LEN_1878, DFE_LEN_02378-82, 2904 and 3085), cbb3-type cytochrome c oxidase (DFE_LEN_02413-14), cytochrome bd complex (DFE_LEN_00410 and 01765-66), and an F-type ATP synthase (DFE_LEN_00088-90, 00092-93 and 00139-40), suggest that LEN is capable of anerobic respiration. There was no obvious terminal reductase encoded in the MAG, however. LEN also encodes the *nrfAH* genes for dissimilatory nitrite reduction to ammonia (DFE_LEN_02008-9), as well as one of the genes for denitrification (*norB*, DFE_LEN_00689). While no electron transport chain genes were found in the DSV MAG (although it was only 25.78% complete), all genes for dissimilatory sulphate reduction were expressed in the metaproteome (Sat, AprAB, and DsrAB). DSV also encoded the *nif* genes for nitrogen fixation (DFE_DSV_00940-41 and 00944).

LEN also encoded the largest number of predicted CAZymes, although none of these 116 genes were identified in the metaproteome. Nonetheless, LEN and SPI contain a higher proportion of genes involved in starch and sucrose metabolism than SYN, and these genes were almost completely absent from DSV.

Lacking genes for an electron transport chain and no obvious terminal reductase, SPI and SYN appear to have a fermentative metabolism. SPI encoded the greatest number and variety of ABC transporters, including the complete set of genes for import of spermidine/putrescine, raffinose/stachyose/melibiose. D-methionine, multiple sugars, nucleosides, ribose/D-xylose, rhamnose, *myo*-inositol, branched chain amino acids, oligopeptides, and lipoproteins. Two LivK proteins, components of the branched chain amino acid ABC transporter, were found in the metaproteome (DFE_SPI_00328 and DFE_SPI_00738). Both SPI and SYN also encode genes for a V-type ATP synthase (DFE_SPI_01064, 01066-67, 01444-49 and DFE_SYN_00166-68, 00329-30, 00332, 00335-37, 02095, 02098-100, 02169-70) indicating that they may also generate energy via a chemiosmotic mechanism alongside substrate level phosphorylation. No phylotype encoded a complete or near complete set of Wood-Ljungdahl pathway genes, concordant with the previous observation that the DFE culture is unable to generate biomass on H_2/CO_2 alone (Wong, 2015).

There is a possibility that SYN may participate in the metabolism of quaternary amines and their derivatives. The SYN MAG contained the *cutCD* genes for a choline trimethylamine lyase (DFE_SYN_01649) and its activating enzyme (DFE_SYN_00095), although neither were found in the metaproteome. It also encoded seven (incomplete) glycine/betaine/sarcosine reductase gene clusters. Of these, just one set of component B alpha/beta and gamma subunits were expressed in the proteome (DFE_SYN_01698 and 1695, respectively). Sarcosine oxidase genes *soxAB* (DFE_SYN_00816, 1401, 2113, 2117) were also found in the SYN MAG and a serine hydroxymethyltransferase *glyA* (DFE_SYN_00806) was expressed in the metaproteome.

5.4 Discussion

5.4.1 Persistence of non-dechlorinating bacteria on a chlorinated substrate

Despite repeated efforts, we have been unable to generate a pure culture of the novel organism strain DCMF. The DFE culture appears to comprise a long-term stable-state community, making the extrication of the DCM fermenter highly challenging. After many tens of transfers, rather than devoting finite resources to isolation, attention was directed towards understanding how the cohabiting organisms persist despite being highly unlikely to utilize the primary substrate. Chen *et al* (2020) raised a similar question about the DCM-dechlorinating enrichment culture RM, in which cohabiting bacteria, the growth of which is not linked to DCM transformation, remain present despite over 80 transfers in seven years. How do these bacteria persist when the sole energy source is a simple, chlorinated C1 compound?

While much of microbiology has been focused on the activities of isolates in pure cultures, this is at best an approximation of how these organisms survive and thrive in the environment. Axenic cultures provide many benefits to the researcher attempting to provide direct evidence of a function within a specific organism, and yet can be prone to artefacts and detrimental to the organism itself. Cultures of the phototroph *Synechococcus* had longer and more robust lifespans when so-called "contaminant" bacteria were present. In the absence of the necromass-scavenging

Roseobacter cohabitant, a build-up of nitrogen-rich organic matter from *Synechococcus* was directly linked to its death in axenic culture (Christie-Oleza *et al.*, 2017). In an example from a methanogenic community, auxotrophs relied on others for generation of essential amino acids (Embree *et al.*, 2015). The more members in a community, the more complex these webs of interaction and interdependence can become. Reliance on the cohabitant members for nutrient recycling or production of essential amino acids or other cofactors may prevent the isolation of strain DCMF in the absence of a more nuanced culture medium.

The DFE community appears to be relatively stable. Four of the five phylotypes identified from the metagenomic data, generated from DNA extracted in 2015, were still present when the Illumina community analysis was performed in 2019. Even earlier than this, the methanogenic DCMP and DCMD cultures from which DFE arose also contained *Desulfovibrio*, *Geobacter*, *Treponema* (a Spirochaete), and Synergistaceae spp. (Lee *et al.*, 2012). It was suspected that the cohabiting organisms in the DFE culture were persisting via consumption of components of lysed strain DCMF cells, i.e. necromass recycling. As such, previously generated genomic and proteomic data was investigated from a "meta -omics" standpoint to see if any clues could be obtained into the metabolic mechanisms at play.

5.4.2 A preliminary model for microbial interactions in the DFE community

The growth curves presented in Chapter 3 showed only a slight increase in strain DCMF and total bacterial cells between the time of inoculation and initial substrate degradation (Figure 3.2, Figure 3.5, Figure 3.7), but community profiling revealed dramatic shifts in relative abundance of other taxa in the DFE culture during this lag phase. Relative abundance of strain DCMF dropped to near zero in the most extreme cases (Figure 5.3 D), while taxa such as *Synergistaceae*, *Veillonellaceae*, and *Desulfovibrio* increased (Figure 5.3). These phylotypes perhaps represent those organisms better able to capitalize on the initial necromass transferred with the inoculum, and/or the nutrients available with the transfer into fresh medium (i.e. R strategists). Some increase in relative abundance here may also be due to death of strain DMCF cells prior to the onset of substrate consumption. Phylotypes whose relative abundance increased more in the "late" and "post" substrate consumption

stages (*Spirochaetaceae* and *Lentimicrobiaceae*) may be slower growing (K strategists) or rely on the production of fermentation products or nutrients from other organisms before cell replication becomes logarithmic. The concomitant increase in strain DCMF and total bacterial 16S rRNA gene copies in these phases of the cultures suggest this is more likely due to actual increase in the abundance of these organisms, rather than an increase in relative abundance at the expense of other taxa.

5.4.2.1 Synergistaceae

Synergistaceae are widely distributed in nature, being found in diverse environments from soil to anaerobic digesters, to a number of polluted environments (Hugenholtz *et al.*, 1998; Vartoukian *et al.*, 2007). Indeed, members of the phylum *Synergistetes* were found in 90% of anaerobic environments surveyed by Godon *et al* (2005), although they are typically only a minor constituent of the community (Vartoukian *et al.*, 2007). Contrary to this, they comprise the second-most relatively abundant taxa in the DFE community at most stages of growth on DCM, glycine betaine, and choline (Figure 5.3 A, B, C).

Synergistaceae are well-known amino acid fermenters (Jumas-Bilak *et al.*, 2007; Pitluck *et al.*, 2010) and some species are also saccharolytic (Vartoukian *et al.*, 2007), meaning they likely capitalize on the amino acids and dicarboxylates present in the DFE culture (Figure 5.7). It was expected that they would be enriched on casamino acids, but they appear to have been outcompeted by *Veillonellaceae* and other taxa (Figure 5.5 D). This supports the suggestion that these taxa only increase in relative abundance towards the end of substrate degradation as they are slower-growing K-strategists. It is also possible they are reliant on synergistic interactions with other taxa which were removed from the casamino acid dilution to extinction cultures.

Nonetheless, ABC transporters for branched chain amino acids and C4-dicarboxylate were among the most highly abundant proteins expressed by SYN (Table 5.3), supporting their role as amino acid fermenters in the DFE culture. The phylotype also encoded many genes for the tricarboxylic acid cycle, and a glutamate dehydrogenase – important for linking amino acids with the tricarboxylic acid cycle – was highly abundant in the metaproteome. A number of other genes for metabolism of amino acids were expressed in the metaproteome, including

tryptophanase (DFE_SYN_01892), methylmalonyl-CoA mutase (DFE_SYN_00423), and L-lysine 6-transaminase (DFE_SYN_02067). Additionally, SYN encodes subunits for both a [FeFe] group A and a [NiFe] Group 3d hydrogenase; subunits from both were found in the metaproteome. The former was unable to be more specifically classified, and therefore could be responsible for fermentative hydrogen evolution or electron bifurcation, while the latter is a bidirectional NAD-coupled enzyme (Søndergaard *et al.*, 2016), likely acting to catalyse fermentative NADH-dependent hydrogen production in this context (Figure 5.7).

Alpha/beta (DFE_SYN_01698) and gamma (DFE_SYN_001695) subunits of the B proteins of a glycine/betaine/sarcosine reductase complex, were also identified in the metaproteome. While in strain DCMF these reductases putatively act upon sarcosine (see Section 3.4.3.3), it seems most likely that they are glycine-specific in SYN. There are no published reports of any *Synergistetes* or genera in the family *Synergistaceae* utilizing glycine betaine or sarcosine. Therefore, attempts to determine the protein's substrate specificity via an unrooted phylogenetic tree (as per Section 2.2.8 for strain DCMF) were unsuccessful. The GrdB and GrdE proteins from SYN all clustered more closely with each other than any of the reductases of varied substrate specificity from the *Firmicutes* (data not shown). Thus, it seems likely that these proteins act upon the amino acid glycine in SYN, which would generate acetyl-phosphate and NH_3.

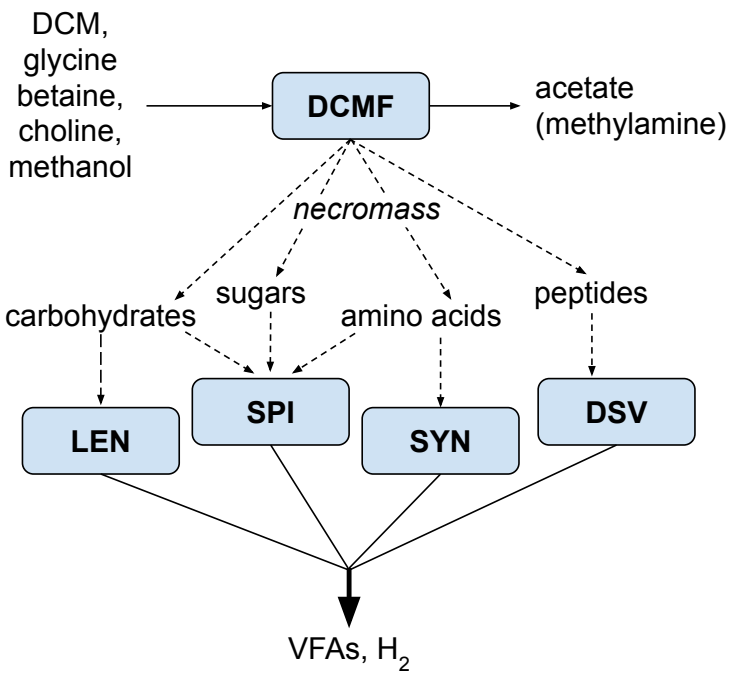


Figure 5.7 Conceptual model of potential interactions between strain DCMF and the cohabitant bacteria in the DFE community. Necromass, primarily composed of expired strain DCMF cells, is a source of carbohydrates, sugars and amino acids fermented by the cohabiting organisms, likely producing volatile fatty acids (VFAs) and/or H_2. Abbreviations: DCM; dichloromethane; DCMF, strain DCMF; LEN, *Lentimicrobiaceae*; SPI, *Spirochaetaceae/Treponematales*; SYN, *Synergistales*; DSV, *Desulfovibrio*.

5.4.2.2 *Spirochaetaceae / Treponematales*

Spirochaetes are commonly detected in anoxic environments contaminated with hydrocarbons and organohalides (Einsiedl *et al.*, 2015; Tan *et al.*, 2015) and a recent report demonstrated their role as necromass recyclers in such ecosystems (Dong *et al.*, 2018). They were anticipated to proliferate through fermentative metabolism within the DFE culture, fuelled by sugars and carbohydrates (Figure 5.7). As such, it was expected that they would be enriched with glucose, however they appear to have been outcompeted by the fast-growing *Petrimonas* (Figure 5.5 F). *Petrimonas* thrived in the high nutrient conditions (5 g l^{-1} glucose or yeast extract), whereas the *Spirochaetaceae* appear to be K strategists that prefer more oligotrophic conditions, which is congruent with their putative role as a dedicated necromass recycler.

Within the DFE culture, SPI appeared to be a versatile scavenger, encoding transport proteins for a wide range of substrates. A number of these ABC transporters for branched chain amino acids were expressed in the proteome. The MAG also encodes

subunits of an NADP-dependent [FeFe] Group A hydrogenase (DFE_SPI_00311, 1420 [found in the metaproteome], 1422, 1423), suggesting that it produces hydrogen from fermentation of amino acids and other necromass constituents (Dong *et al.*, 2018). Fermentation and hydrolysis of these compounds may also produce alcohols and short-chain fatty acids (Dong *et al.*, 2018), which can in turn be utilized by other cohabitant bacteria (Figure 5.7).

Spirochaetaceae and *Synergistaceae* could not be definitively excluded from direct metabolism of DCM, glycine betaine, or choline due to their absence in the strain DCMF-free enrichment cultures (Figure 5.5E, F, G, H). However, the metaproteogenomic data did not reveal any obvious mechanisms by which these organisms may be directly involved in the metabolism of these substrates. Rather, they appear to exist on fermentation of necromass.

5.4.2.3 Lentimicrobiaceae

Members of the phylum *Bacteroidetes* are widely distributed in the environment and particularly prevalent in anoxic ecosystems rich in organic material. *Lentimicrobiaceae* is a relatively new family within this phylum and its sole cultured representative, *Lentimicrobium saccharophilum*, was isolated from starch-based organic wastewater and shown to ferment a range of carbohydrates (Sun *et al.*, 2016). Despite the presence of most components of an electron transport chain encoded in the MAG, there was no obvious terminal reductase and there are no previous reports of anaerobic respiration by *Lentimicrobiaceae* in the literature.

The LEN phylotype in the DFE culture is likely to utilize carbohydrates, distinguishing it from the amino acid-fermenting SPI and SYN. The MAG is rich in CAZymes and encodes the highest proportion of proteins predicted to be extracellular (Figure S8), which is important for degrading large biopolymers inherent in microbial biomass (Christie-Oleza *et al.*, 2015). This may assist other organisms in the DFE culture by making larger molecules more metabolically available. The expression of multiple TonB-family proteins that are homologous to RagB/SusC nutrient uptake outer membrane proteins from *Pedobacter* and *Bacteroidetes* species indicates that LEN is also likely able to import larger molecules resulting from the extracellular degradation of proteins (Hall *et al.*, 2005) and carbohydrates (Cho and Salyers, 2001) (Figure 5.7). The presence of an NADP-

reducing FeFe hydrogenase in the LEN MAG (DFE_LEN_02656-57, 2663-65), one subunit of which was found in the metaproteome, indicates that it may produce hydrogen during fermentation, like SPI (Figure 5.7).

5.4.2.4 Desulfovibrio

The sulphate-respiring genus *Desulfovibrio* has a broad substrate range and has been associated with contaminated environments both as a primary degrader (e.g. Löffler, Sanford & Ritalahti, 2005; Kleinsteuber et al., 2008) and a synergistic cohabitant consuming fermentation products of other organisms (e.g. Müller et al., 2009; Taubert et al., 2012; Tan et al., 2015). Although DSV represented the smallest and least complete MAG (25.78% complete), a number of its proteins were still identified in the metaproteome. Most salient were the full set of proteins for dissimilatory sulphate reduction (sat, DFE_DSV_00655; AprA, DFE_DSV_00306; AprB, DFE_DSV_00305; DsvA, DFE_DSV_00210; DsvB, DFE_DSV_00209) and a [NiFe] hydrogenase (DFE_DSV_00560, 561, 564). The latter was classified as a Group 1b unidirectional hydrogen-uptake hydrogenase by HydDB (Søndergaard *et al.*, 2016). Also expressed was a menaquinone reductase (DFE_DSV_01182) and two formate dehydrogenase subunits (DFE_DSV_01146-7). While this suggests that DSV could be utilising hydrogen produced by other fermenters in the culture as an electron donor for sulphate reduction, there is no obvious source of sulphate in the culture. It is possible that the proteins for sulphate reduction are constitutively expressed in DSV, even in the absence of sulphate, as has been shown for *Desulfovibrio desulfuricans* strain 27774 (Marietou *et al.*, 2009).

Given that DSV was enriched to apparent purity on peptone (Figure 5.5 C), it is clearly also capable of a proteolytic fermentative metabolism. A dipeptide transport protein (DppA, DFE_DSV_00351) was also identified in the proteome in DCM-amended cultures, indicating potential import of peptides for growth. DSV was the only phylotype that was present in all four strain DCMF-free enrichments – casamino acids, glucose, peptone, and yeast extract (Figure 5.5). It appears to be a robust member of the DFE community that is able to survive and grow under a range of different metabolic conditions. This is in agreement with the broad substrate range of the genus.

The combined community profiling, strain DCMF-free enrichment culturing, and metaproteogenomic data support strain DCMF as the sole substrate user in the DFE cultures. None of the cohabiting taxa increase in relative abundance over the course of substrate utilization. Rather, their growth patterns and phylogenetic contexts suggest that their growth is driven by necromass oxidation. Strain DCMF appears to be a foundation species in the DFE culture, responsible for degrading the amended substrate (DCM, choline, glycine betaine, methanol) and thus providing the majority of energy to the overall community, while the cohabiting organisms are dependent on its exponential growth and biomass turnover to produce the necromass components for fermentation. Necromass may be completely oxidised to hydrogen or, more likely, partially oxidised to form various volatile fatty acids.

5.4.3 Necromass recycling has important implications for contaminated site remediation

The role of microbial necromass recycling is increasingly being considered as an important contributor to nutrient cycling within subsurface communities, particularly in oligotrophic ecosystems that are otherwise limited in organic carbon and nitrogen (Simpson *et al.*, 2007; Liang *et al.*, 2011). In contaminated sites in particular, necromass recycling may contribute to enhancing bioremediation at the contaminant plume fringe, as the production of hydrogen, acetate, and ethanol can both stimulate microbial blooms or serve as secondary substrates for co-metabolism (Horvath, 1972; Wrighton *et al.*, 2014). Consumption of dead biomass can also liberate important nutrients including nitrogen, phosphorous, and trace elements (Head *et al.*, 2006; Christie-Oleza *et al.*, 2017).

Phylotypes identified in the DFE culture – *Desulfovibrio*, *Bacteroidetes*, *Spirochaetes/Treponematales*, *Synergistetes* – have long been associated with hydrocarbon and organohalide-degrading cultures (e.g. Duhamel & Edwards, 2006; Kleinsteuber et al., 2008; Strąpoć et al., 2011; Taubert et al., 2012; Dong et al., 2018), although their abundance is not linked to degradation of the primary substrate (with the exception of some *Desulfovibrio* species). While previous reports have suggested that these organisms persist on necromass recycling (Kleinsteuber *et al.*, 2012; Lee *et al.*, 2012; Taubert *et al.*, 2012; Dong *et al.*, 2018), little research has been done into the specific mechanisms utilised and what effect this has on the pollutant-

degrading organism(s). Understanding microbial interactions in the DFE culture, in which a chlorinated one-carbon compound has sustained a stable, diverse community for several years, could reveal new insights into syntrophic community dynamics. This in turn may lead to the development of more robust mixed cultures for *in situ* bioremediation applications, as well as a broader understanding of carbon and nitrogen flux in contaminated subsurface environments.

5.5 Conclusions

The DFE community is a long-term stable-state consortium that has been sustained on DCM for many years. The dominant organism in the culture, the novel DCM-fermenter strain DCMF, appears to be the sole taxon capable of metabolizing the amended substrate, while biomass turnover from expired cells fuels the metabolism of the cohabiting organisms. Cohabitant organisms increase in relative abundance both before and after consumption of the primary substrate by strain DCMF, and strain DCMF-free enrichments directly excluded a number of these phylotypes from involvement in DCM, glycine betaine, or choline transformation. Metaproteogenomic analysis provided information about four key cohabitant phylotypes in the DFE culture – *Desulfovibrio*, *Lentimicrobiaceae*, *Synergistaceae*, and *Treponematales* – the latter three of which appear to be novel lineages at the genus, or even family level, based on 16S rRNA gene identity. Necromass recycling in the DFE culture requires an intricate web of interdependencies and interactions. Further investigation into these organisms and their roles may shed light on the important role of necromass recycling in anoxic, contaminated environments.

5.6 Acknowledgements

Thank you to Dr Richard Edwards for generating the “NR Cohabitants” set of PacBio reads and advice on metagenomic assembly.

6 General Discussion and Conclusions

6.1 Summary of findings

Despite the importance of anaerobic DCM degradation for the remediation of contaminated sites worldwide, information on microbial metabolism of DCM in anoxic environments remains scarce. In particular, the enzyme(s) responsible for carbon-chlorine bond cleavage remain unknown, hindering monitoring of DCM bioremediation *in situ*. This thesis reports the genomic, proteomic, and physiological characterisation of a novel, anaerobic DCM-fermenting bacterium, '*Candidatus* Formamonas warabiya' strain DCMF, present in the enrichment culture DFE.

Whole genome sequencing of strain DCMF produced a single, circularised chromosome that was 6.44 Mb in length and contained 5,772 CDS. Extensive manual curation showed that it encodes a full Wood-Ljungdahl pathway and an unusually high number of MttB superfamily methyltransferases (82 CDS). Genomic comparison to *Dehalobacterium formicoaceticum* DMC and '*Candidatus* Dichloromethanomonas elyunquensis' RM revealed a relatively small core genome amongst the three anaerobic, DCM-degrading bacteria, reflective of the phylogenetic differences between the three genera. Contrary to reports of high genomic plasticity in organohalide-respiring bacteria, particularly in areas encoding reductive dehalogenases (Liang *et al.*, 2012), there was no evidence of significant horizontal gene transfer between the anaerobic DCM-degrading bacteria. Strain DCMF and *D. formicoaceticum* had more in common with one another than with '*Ca.* Dichloromethanomonas elyunquensis', including the apparent potential to metabolise methylated amines, glycine betaine, and ethanolamine.

Strain DCMF is the first anaerobic DCM-degrading bacterium reported to metabolise substrates other than DCM. Choline, glycine betaine, and methanol degradation was observed concomitant with increases in strain DCMF cell numbers and the accumulation of acetate and (in quaternary amine-amended cultures) methylamine. In contrast to the fermentative metabolism reported thus far, strain DCMF could also utilise sarcosine + H_2. Genome-based metabolic models for growth on quaternary amines suggested that choline was first metabolised to glycine betaine, then demethylated to sarcosine, which was subsequently reductively cleaved to

methylamine and acetate (via acetyl phosphate). Cultures amended with ^{13}C-labelled DCM or bicarbonate also demonstrated the action of the Wood-Ljungdahl pathway in DCM fermentation, showing that it operates in both oxidative and reductive directions for acetogenesis. Due to its genetic and physiological novelty, strain DCMF was proposed to belong to a novel genus in an underrepresented family (*Peptococcaceae*) and assigned the name '*Candidatus* Formamonas warabiya'.

Proteomic analysis of strain DCMF revealed a putative DCM methyltransferase gene cluster, enzymes of which were significantly more abundant in cells grown with DCM than those with glycine betaine and highly conserved in the two other anaerobic DCM-degrading bacteria. The DCM-associated methyltransferase gene cluster is likely responsible for DCM entry into the Wood-Ljungdahl pathway. Proteomics also supported the genome-based metabolic models for quaternary amine and methanol metabolism, confirming the involvement of a glycine betaine methyltransferase and sarcosine reductase, and a methanol methyltransferase, respectively. This expands the relatively small pool of bacterial taxa able to metabolise the environmentally important compounds in anoxic environments.

Finally, the role of the wider DFE community was assessed. Cohabiting organisms in the DFE culture were identified via 16S rRNA genes in the metagenomic data and Illumina 16S rRNA gene amplicon sequencing. Their relative abundance was shown to change dramatically in response to substrate consumption by strain DCMF, and a number of cohabiting taxa were definitively ruled out from contributing to the metabolism of DCM, choline, and glycine betaine via their enrichment to the exclusion of strain DCMF. Metagenomic assembly of the PacBio sequencing data reported in Chapter 2 produced five MAGs, belonging to the phylotypes *Desulfovibrio*, *Lentimicrobiaceae*, *Spirochaetaceae* (*Treponematales*), *Synergistaceae*, and *Ignavibacteria* (no longer present in the culture). A metaproteogenomic approach revealed the presence of genes involved in the metabolism of carbohydrates, sugars, and amino acids in these phylotypes, supporting the hypothesis that the cohabiting bacteria persist in the DFE culture by consuming expired strain DCMF cells, i.e. they are necromass recyclers.

6.2 The evolution and ecological niches of anaerobic DCM degrading bacteria

The Wood-Ljungdahl pathway is central to the metabolism of DCM in all three anaerobic DCM-degrading bacteria (Mägli *et al.*, 1998; Kleindienst *et al.*, 2019; Chen *et al.*, 2020, Chapters 3 and 4 of this thesis). While '*Ca.* Dichloromethanomonas elyunquensis' appears to also utilise reductive dehalogenases for DCM dechlorination (Kleindienst *et al.*, 2019), neither strain DCMF nor *D. formicoaceticum* encode any such enzymes. In many organohalide-respiring bacteria, the reductive dehalogenases that catalyse dehalogenation have been subject to intense evolutionary pressure and are often encoded within regions of high genomic plasticity (Maillard *et al.*, 2005; West *et al.*, 2008; McMurdie *et al.*, 2009, 2011; Tang *et al.*, 2016). Conversely, there was little evidence of horizontal gene transfer between the three anaerobic DCM-degrading bacteria (Figure 2.7).

Natural sources of DCM, particularly in the marine environment (Gribble, 2010), may have contributed to the evolution of '*Ca.* Dichloromethanomonas elyunquensis', which was isolated from pristine river sediments in Puerto Rico (Kleindienst *et al.*, 2017). Indeed, although volcanic sources represent only a small fraction of naturally produced DCM today (Gribble, 2010), the early earth was highly geothermally active and DCM may therefore have been one of the first organic molecules present in abundance. Coupled with this, is the antiquity of the Wood-Ljungdahl pathway, which may the oldest metabolism on earth (Fuchs, 1989; Wood, 1991; Russell and Martin, 2004; Berg *et al.*, 2010). It is therefore feasible that DCM fermentation is one of the oldest metabolisms on the planet, with its study illuminating microbial metabolism on early earth. Furthermore, strain DCMF, *D. formicoaceticum* and '*Ca.* Dichloromethanomonas elyunquensis' each utilise the Wood-Ljungdahl pathway slightly differently – producing acetate, formate and acetate, or H_2 and CO_2 as end products, respectively (Mägli *et al.*, 1996; Chen *et al.*, 2020). Evolution of multiple variations of this pathway for DCM metabolism suggest that these microbes adapted to DCM long before the large influx of anthropogenic organochlorines into the environment (Chen *et al.*, 2020).

Further to this, is the ability of strain DCMF to utilise quaternary amines as well as DCM, and presence of genes for methylated amine and glycine utilisation in the *D.*

formicoaceticum genome (Chapter 2). The osmoprotectant glycine betaine and its precursor, choline, are widespread in marine and coastal environments. In combination with substantial oceanic sources of DCM (Kolusu *et al.*, 2017, 2018), it is worth considering whether strain DCMF and *D. formicoaceticum* are in fact marine organisms, as raised in the Discussion of Chapter 2.

An alternative explanation may be found in the gut. Whilst the marine environment and the human gut may seem like disparate ecological niches at first glance, they do share some common attributes, such as osmotic stress, low oxygen, and high productivity (Jameson *et al.*, 2016a). As well as its osmoprotectant role in the marine context, glycine betaine is present in the gut via ingestion in various foods or via the oxidation of choline. Jameson *et al* (2016) reported on the presence of various organisms and pathways responsible for the production of trimethylamine in both the gut and the marine environment, including the reductive cleavage of glycine betaine. Whilst they found that different microorganisms tended to be prevalent in the two environments, there was an overlap of the genus *Clostridium* in the reductive cleavage pathway between the two (Jameson *et al.*, 2016a). In recent years, members of the *Dehalobacteriaceae* have been reported in vertebrate gut microbiomes (e.g. Goodrich *et al.*, 2014; Reed *et al.*, 2017; Arrazuria *et al.*, 2018; Nemoto *et al.*, 2019; Zhan *et al.*, 2019).

Furthermore, strain DCMF and *D. formicoaceticum* also encode putative bacterial microcompartment (BMCs) for ethanolamine (both) and propanediol (strain DCMF only) utilisation (Figure 2.6). Whilst these two compounds are used extensively in industry (and therefore could be reasonably expected as "co-pollutants" at DCM-contaminated sites), they are also naturally occurring. Ethanolamine is readily available from the breakdown of phosphatidylethanolamine (Larson *et al.*, 1983; Proulx and Fung, 1969), an abundant phospholipid in eukaryotic and prokaryotic cell membranes (Randle *et al.*, 1969; White, 1973), while propanediol is a metabolic intermediate of the naturally occurring deoxyhexose sugars fucose and rhamnose (Badía *et al.*, 1985; Petit *et al.*, 2013). In a human context, there is a growing association between BMCs and pathogenicity, as compounds metabolised within them (ethanolamine, 1,2-propanediol, choline) are all prevalent in the gut. Degradation of 1,2-propanediol has been linked to enteric pathogenesis by

numerous studies (Conner *et al.*, 1998; Buchrieser *et al.*, 2003; Joseph *et al.*, 2006; Thiennimitr *et al.*, 2011), as has ethanolamine utilisation (reviewed in Garsin, 2010). Indeed, the ethanolamine-utilising BMC gene cluster is present in three of the most dangerous pathogens responsible for food poisoning: *Salmonella enterica*, *Listeria monocytogenes* and *Clostridium perfringens* (Korbel *et al.*, 2005). In light of the genomically-predicted substrate range of strain DCMF and *D. formicoaceticum*, it is worth questioning whether these two organisms in fact have origins as bacterial scavengers, or even opportunistic pathogens, rather than merely viewing them as adapted to life in polluted environments.

6.3 The importance of community function over individual ability

Within the DFE community, strain DCMF appears to be the sole direct consumer of the amended substrate, DCM, while all other organisms persist via necromass recycling. As such, strain DCMF represents the foundation species, as its presence and consumption of the sole external source of electrons has the largest influence on overall community structure (Figure 5.4). Although the relative abundance of taxa in the DFE community fluctuates dependent on the growth of strain DCMF, it maintains an overall steady-state system with little gain or loss of taxa over time. Similar community structures in which a foundational, organohalide degrading bacterium co-exists with less abundant, non-dehalogenating taxa have been reported in culture RM, dominated by the DCM-mineralising bacterium '*Ca.* Dichloromethanomonas elyunquensis' (Kleindienst *et al.*, 2017; Chen *et al.*, 2020), and KB-1, dominated by the chlorinated ethene-respiring *Dehalococcoides* (Duhamel and Edwards, 2006). However, the role of the non-dechlorinating community has not been well explored.

Within the DFE culture, the cohabiting bacteria appear to be entirely dependent on strain DCMF for growth, as it is highly unlikely that they can metabolise DCM and instead rely on necromass components from dead cells. The relationship between strain DCMF and these cohabitants could be classed as metabiosis – an indirect dependency in which one organism (strain DCMF) creates a suitable environment for a second. Community relationships within the DFE culture appear to be more

complex than merely these two trophic levels, however, as some cohabiting bacteria may produce, for example, VFAs from fermentation of necromass components, that can in turn be used by other cohabitants for growth. These processes fall under the concept of "intermediary ecosystem metabolism" put forward by Drake *et al* (2009), encompassing all the intermediate reactions in an ecosystem that link the input to the overall output, ultimately driving methanogenesis (as the terminal process in anoxic ecosystems). Although methanogens were previously removed from the DFE community (Holland *et al.*, 2019), detailed understanding of the processes involved in intermediary ecosystem metabolism within a simplified community such as DFE can help shape our understanding of processes that lead to methane flux from anoxic subsurface environments.

Strain DCMF may also rely on the cohabiting bacteria for some as-yet undetermined reason, meaning the relationship between strain DCMF and some (or even all) DFE cohabitants is closer to mutualism than commensal metabiosis. The recalcitrance of strain DCMF to isolation despite repeated series of dilution to extinction cultures suggests that its growth is dependent on the cohabiting organisms. Based on relative abundance, strain DCMF dominated the DFE community during the late exponential growth phase (Figure 5.3) and this is typically when dilution cultures were transferred. Yet DCM metabolism was only maintained in higher dilution (i.e. 10^{-1} to 10^{-5}, data not shown) cultures that apparently preserved cohabiting organisms.

Benefits of cohabiting bacteria to keystone or foundational species in microbial consortia can be varied – production of amino acids for which the other is auxotrophic (e.g. Embree *et al.*, 2015), or essential cofactors such as cobalamins (e.g. Yan *et al.*, 2012), or the removal of nitrogen-rich wastes via necromass consumption (e.g. Christie-Oleza *et al.*, 2017) have all been reported. Communities of dehalogenating bacteria and non-dehalogenating but beneficial cohabitants may form what Pascual-García *et al* (2019) term Metabolically Cohesive Consortium (MeCoCo). MeCoCos are subgroups within natural environments where the consortium as a whole stabilises ecosystem dynamics, buffering the community against external resource fluctuations by mutualistic interactions. This forms a positive feedback loop and can enable the whole MeCoCo to be more competitive than any of the component species could be individually. The authors hypothesise

that MeCoCos are common in natural communities and application of this paradigm to dechlorinating consortia may help explain why there is a commonality to certain non-dechlorinating taxa that are frequently observed in contaminated sites or pollutant-enriched microcosms (e.g. *Ignavibacteria/Chlorobi*, *Spirochaetes*, and deltaproteobacteria such as *Desulfovibrio*). In the context of anaerobic dechlorinating communities, artificially constructed MeCoCos of hydrogenotrophic organohalide-respiring bacteria with hydrogenogenic necromass fermenters could produce robust cultures for *in situ* applications. Constant cell turnover could ensure a steady supply of fermentation substrate to the hydrogenotrophic organisms, and thus a stable supply of hydrogen for the pollutant degrading bacteria.

Much of microbiological research to date has focussed on pure cultures, both inside and outside the world of microbial dehalogenation. While axenic cultures provide clear advantages for elucidating specific functions, determining mass balances, and unquestionably identifying molecular mechanisms, they are not representative of the natural state of microorganisms in the environment. In an age of "meta-omics", perhaps it is time to widen our scope and think more about dechlorinating communities, rather than dechlorinating species. Whether seeking to understand mechanisms of natural bioattenuation occurring at a polluted site, stimulating indigenous populations of dehalogenating organisms, or augmenting lab-grown cultures, understanding the microbial ecology of contaminated aquifers as a whole offers the opportunity to improve remediation outcomes.

Assessing microbial community diversity and function in subsurface environments can be challenging but is a valuable endeavour for deeper understanding of the environmental processes at work. Molecular analyses can be hindered by low biomass or attachment of microbial communities to sediment particles or rock surfaces (Griebler and Lueders, 2009). Therefore, simply sampling groundwater can provide biased results, but taking cores is invasive and requires elaborate equipment. DNA or RNA stable isotope probing can provide greater insight into the taxa responsible for substrate consumption and flow of carbon through a system, but further consumption of degradation products by other community members (i.e. intermediary ecosystem metabolism) can confound results. Recent advances in "meta-omics" technologies have provided a growing ability to determine the profiles

and functions of microbial communities *in situ* and metaproteogenomics approaches are now rightly considered the gold standard.

Cycling of carbon and nitrogen underpins the formation of microbial communities in the environment. Soil organic matter is the largest organic carbon and nitrogen pool in the terrestrial biosphere (Schmidt *et al.*, 2011) and a significant portion (>50%) of it is presumed to be microbial necromass (Simpson *et al.*, 2007; Liang *et al.*, 2011; Miltner *et al.*, 2012). This necromass is an important source of organic material, particularly in otherwise oligotrophic subsurface environments, e.g. Arctic soils (Bradley *et al.*, 2016). Given that soil stores at least three-fold more carbon than either the atmosphere or terrestrial plants (Fischlin *et al.*, 2007), understanding how much of this is sequestered, recycled into other organic molecules, or mineralised has implications for CO_2 emissions, with flow-on effects for global climate.

6.4 Future perspectives

This thesis represents the first report of a putative DCM methyltransferase gene cluster found in strain DCMF and conserved among anaerobic DCM degrading bacteria. Future work should focus on verifying the involvement of these genes in DCM dechlorination and biochemical characterisation. It is possible that native polyacrylamide gel electrophoresis could be used to definitively link DCM dechlorination to the putative methyltransferase, as previously demonstrated in the identification of reductive dehalogenases (Adrian *et al.*, 2007; Wong *et al.*, 2016). The involvement of multiple proteins in the dechlorination reaction could hinder this approach, if they are not all contained within the same gel band. Should the putative methyltransferase proteins co-elute as a single complex however, activity would likely be retained. Quantification of metabolic intermediates of DCM dechlorination could also shed light on the initial metabolic fate of DCM in strain DCMF, particularly identification of methylene-THF. However, this can also be difficult due to the highly labile nature of THF intermediates and their presumed very low concentration.

Kleindienst *et al* (2019) suggested two different hypotheses for how '*Ca.* Dichloromethanomonas elyunquensis' might dechlorinate DCM, based on the

abundance of both reductive dehalogenases and methyltransferases in the proteome of DCM-grown cells. Given the absence of reductive dehalogenases in strain DCMF and *D. formicoaceticum*, different chemical mechanisms are likely operating in these two species. Compound specific isotope analysis of C—Cl bond breakage in '*Ca.* Dichloromethanomonas elyunquensis' and *D. formicoaceticum* was congruent with different dechlorination mechanisms in these two organisms (Chen *et al.*, 2017a) and similar work with strain DCMF would provide further insight into the putative C—Cl bond breakage mechanism it employs.

In order to determine whether the DCM-associated methyltransferase cluster is a complete operon, reverse transcription PCR could be employed. Primers designed to hybridise with the intergenic regions of mRNA transcripts would show whether the eight genes in the cluster are transcribed together as a single operon. Ultimately, cloning of the putative DCM methyltransferase operon into a non-dechlorinating bacterium would directly link the genes to the DCM dechlorinating phenotype. The model acetogen *Acetobacterium woodii* could be a good candidate for gathering this direct evidence, as it already encodes the full Wood-Ljungdahl pathway and utilises similar methyltransferase systems for metabolism of methanol (Kremp *et al.*, 2018) and glycine betaine (Lechtenfeld *et al.*, 2018). Successful cloning of the putative DCM methyltransferase cluster could also lead to purification of the methyltransferase enzymes for direct biochemical characterisation and energetics experiments.

Identification of the DCM dechlorinating enzyme(s) has important implications for assessing bioremediation of DCM-contaminated sites. While molecular analyses of taxa (e.g. via qPCR of their 16S rRNA gene) can be used to assess the potential for natural attenuation at a site, quantification of the functional genes involved in dechlorination provide a more accurate assessment. Quantitative PCR of a DCM dechlorinating enzyme could also be used for monitoring bioaugmentation with DCM degrading cultures. Targeted qPCR in pristine and contaminated environments could also help assess the existing abundance of the putative DCM methyltransferase. This could be combined with database searches of existing metagenomic data to investigate the innate environmental potential for anaerobic DCM metabolism in a range of ecosystems. As a recently recognised greenhouse gas

(Hossaini *et al.*, 2017), a more detailed picture of DCM sinks in anoxic environments would improve the accuracy of models exploring the cycling of this compound.

6.5 Concluding remarks

Strain DCMF is an unusual bacterium capable of metabolising a novel range of substrates including DCM, methanol, the quaternary amines choline and glycine betaine, and their metabolic intermediates. The discovery of an anaerobic DCM-degrading bacterium that is able to metabolise additional substrates allowed for the first comparative, label-free quantitative proteomics experiment to be carried out, which identified a putative novel DCM methyltransferase gene cluster that was significantly more abundant in cells grown with DCM than glycine betaine. This work paves the way for identification of the enzyme(s) catalysing the C—Cl bond cleavage in DCM within anoxic environments, is of interest to the remediation industry for bioremediation and monitoring purposes. The overall substrate range of strain DCMF (DCM, quaternary amines, methylated glycines, and methanol) has important implications for the cycling of climate-active trace gasses and highlights a number of novel or uncommon metabolic pathways present in the anoxic subsurface.

References

Abby SS, Néron B, Ménager H, Touchon M, Rocha EPC. (2014). MacSyFinder: A Program to Mine Genomes for Molecular Systems with an Application to CRISPR-Cas Systems. *PLoS One* **9**: e110726.

Adrian L, Manz W, Szewzyk U, Görisch H. (1998). Physiological characterization of a bacterial consortium reductively dechlorinating 1,2,3- and 1,2,4-trichlorobenzene. *Appl Environ Microbiol* **64**: 496–503.

Adrian L, Rahnenführer J, Gobom J, Hölscher T. (2007). Identification of a chlorobenzene reductive dehalogenase in *Dehalococcoides* sp. strain CBDB1. *Appl Environ Microbiol* **73**: 7717–7724.

Agency for Toxic Substances and Disease Registry. (2000). Toxicological profile for methylene chloride. Atlanta, Georgia.

Almeida J, Schobesberger S, Kürten A, Ortega IK, Kupiainen-Määttä O, Praplan AP, *et al.* (2013). Molecular understanding of sulphuric acid-amine particle nucleation in the atmosphere. *Nature* **502**: 359–363.

Amann RI, Binder BJ, Olson RJ, Chisholm SW, Devereux R, Stahl DA. (1990). Combination of 16S rRNA-targeted oligonucleotide probes with flow cytometry for analyzing mixed microbial populations. *Appl Environ Microbiol* **56**: 1919–25.

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**: 2805–2811.

Andreesen JR. (2004). Glycine reductase mechanism. *Curr Opin Chem Biol* **8**: 454–461.

Arndt D, Grant JR, Marcu A, Sajed T, Pon A, Liang Y, *et al.* (2016). PHASTER: a better, faster version of the PHAST phage search tool. *Nucleic Acids Res* **44**: W16–W21.

Arrazuria R, Pérez V, Molina E, Juste RA, Khafipour E, Elguezabal N. (2018). Diet induced changes in the microbiota and cell composition of rabbit gut associated lymphoid tissue (GALT). *Sci Rep* **8**: 1–11.

Arumugam K, Bağcl C, Bessarab I, Beier S, Buchfink B, Górska A, *et al.* (2019). Annotated bacterial chromosomes from frame-shift-corrected long-read metagenomic data. *Microbiome* **7**: 1–13.

Atashgahi S, Häggblom MM, Smidt H. (2017). Organohalide respiration in pristine environments: implications for the natural halogen cycle. *Environ Microbiol* **20**: 934–948.

Australian Government - Department of the Environment and Energy. (2019). 2017/2018 data within Australia - Dichloromethane from All Sources. *Natl Pollut Invent*. http://www.npi.gov.au/.

Axen SD, Erbilgin O, Kerfeld CA. (2014). A taxonomy of bacterial microcompartment loci constructed by a novel scoring method. *PLoS Comput Biol* **10**: e1003898.

Bache R, Pfennig N. (1981). Selective isolation of *Acetobacterium woodii* on methoxylated aromatic acids and determination of growth yields. *Arch Microbiol* **130**: 255–261.

Bader R, Leisinger T. (1994). Isolation and characterization of the *Methylophilus* sp. strain DM11 gene encoding dichloromethane dehalogenase/glutathione S-transferase. *J Bacteriol* **176**: 3466–3473.

Badía J, Ros J, Aguilar J. (1985). Fermentation mechanism of fucose and rhamnose in *Salmonella typhimurium* and *Klebsiella pneumoniae*. *J Bacteriol* **161**: 435–437.

Baeseman JL, Novak PJ. (2001). Effects of various environmental conditions on the transformation of chlorinated solvents by *Methanosarcina thermophila* cell exudates. *Biotechnol Bioeng* **75**: 634–641.

Bagley DM, Gossett JM. (1995). Chloroform degradation in methanogenic methanol enrichment cultures and by *Methanosarcina barkeri* 227. *Appl Environ Microbiol* **61**: 3195–3201.

Baker JM, Sturges WT, Sugier J, Sunnenberg G, Lovett AA, Reeves CE, *et al.* (2001). Emissions of CH_3Br, organochlorines, and organoiodines from temperate macroalgae. *Chemosph - Glob Chang Sci* **3**: 93–106.

Balk M, Weijma J, Goorissen HP, Ronteltap M, Hansen TA, Stams AJM. (2007). Methanol utilizing *Desulfotomaculum* species utilizes hydrogen in a methanol-fed

sulfate-reducing bioreactor. *Appl Microbiol Biotechnol* **73**: 1203–1211.

Bauchop T. (1967). Inhibition of rumen methanogenesis by methane analogues. *J Bacteriol* **94**: 171–175.

Beale R, Dixon JL, Arnold SR, Liss PS, Nightingale PD. (2013). Methanol, acetaldehyde, and acetone in the surface waters of the Atlantic Ocean. *J Geophys Res Ocean* **118**: 5412–5425.

Beale R, Liss PS, Dixon JL, Nightingale PD. (2011). Quantification of oxygenated volatile organic compounds in seawater by membrane inlet-proton transfer reaction/mass spectrometry. *Anal Chim Acta* **706**: 128–134.

Beers JR. (1967). The species distribution of some naturally occurring quaternary ammonium compounds. *Comp Biochem Physiol* **21**: 11–21.

Berg IA, Kockelkorn D, Ramos-Vera WH, Say RF, Zarzycki J, Hügler M, *et al.* (2010). Autotrophic carbon fixation in archaea. *Nat Rev Microbiol* **8**: 447–460.

Bertelli C, Laird MR, Williams KP, Lau BY, Hoad G, Winsor GL, *et al.* (2017). IslandViewer 4: Expanded prediction of genomic islands for larger-scale datasets. *Nucleic Acids Res* **45**: W30–W35.

Bertsch J, Siemund AL, Kremp F, Müller V. (2016). A novel route for ethanol oxidation in the acetogenic bacterium *Acetobacterium woodii*. *Environ Microbiol* **18**: 2913–2922.

Biegel E, Muller V. (2010). Bacterial Na^+-translocating ferredoxin:NAD^+ oxidoreductase. *Proc Natl Acad Sci* **107**: 18138–18142.

Biegel E, Schmidt S, Müller V. (2009). Genetic, immunological and biochemical evidence for a Rnf complex in the acetogen *Acetobacterium woodii*. *Environ Microbiol* **11**: 1438–1443.

Blunden G, Gordon SM, McLean WFH, Guiry MD. (1982). The distribution and possible taxonomic significance of quaternary ammoniun and other Dragendorff-positive compounds in some genera of marine algae. *Bot Mar* **XXV**: 536–567.

Bobik TA. (2007). Bacterial microcompartments. *Microbe* **2**: 25–31.

Boch J, Kempf B, Bremer E. (1994). Osmoregulation in *Bacillus subtilis*: Synthesis of the osmoprotectant glycine betaine from exogenously provided choline. *J Bacteriol*

176: 5364–5371.

Bolyen E, Rideout JR, Dillon MR, Bokulich NA, Abnet CC, Al-Ghalith GA, *et al.* (2019). Reproducible, interactive, scalable and extensible microbiome data science using QIIME 2. *Nat Biotechnol* **37**: 852–857.

Bonventre J, Brennan O, Jason D, Henderson A, Bastos ML. (1977). Two deaths following accidental inhalation of dichloromethane and 1,1,1-trichloroethane. *J Anal Toxicol* **1**: 158–160.

Bowers RM, Kyrpides NC, Stepanauskas R, Harmon-Smith M, Doud D, Reddy TBK, *et al.* (2017). Minimum information about a single amplified genome (MISAG) and a metagenome-assembled genome (MIMAG) of bacteria and archaea. *Nat Biotechnol* **35**: 725–731.

Bradbeer C. (1965). The Clostridial fermentations of choline and ethanolamine. I. Preparation and properties of cell-free extracts. *J Biol Chem* **240**: 4669–4674.

Bradley JA, Arndt S, Šabacká M, Benning LG, Barker GL, Blacker JJ, *et al.* (2016). Microbial dynamics in a High Arctic glacier forefield: A combined field, laboratory, and modelling approach. *Biogeosciences* **13**: 5677–5696.

Brooke D, How P. (1994). Environmental hazard assessment: dichloromethane. Watford.

Brunner W, Staub D, Leisinger T. (1980). Bacterial degradation of dichloromethane. *Appl Environ Microbiol* **40**: 950–958.

Buchfink B, Xie C, Huson DH. (2014). Fast and sensitive protein alignment using DIAMOND. *Nat Methods* **12**: 59–60.

Buchrieser C, Rusniok C, Kunst F, Cossart P, Glaser P. (2003). Comparison of the genome sequences of *Listeria monocytogenes* and *Listeria innocua*: clues for evolution and pathogenicity. *FEMS Immunol Med Microbiol* **35**: 207–213.

Burke SA, Krzycki JA. (1995). Involvement of the 'A' isozyme of methyltransferase II and the 29-kilodalton corrinoid protein in methanogenesis from monomethylamine. *J Bacteriol* **177**: 4410–4416.

Burke SA, Krzycki JA. (1997). Reconstitution of monomethylamine:coenzyme M methyl transfer with a corrinoid protein and two methyltransferases purified from

Methanosarcina barkeri. *J Biol Chem* **275**: 29053–60.

Busk PK, Pilgaard B, Lezyk MJ, Meyer AS, Lange L. (2017). Homology to peptide pattern for annotation of carbohydrate-active enzymes and prediction of function. *BMC Bioinformatics* **18**: 1–9.

Byers HK, Sly LI. (1993). Toxic effects of dichloromethane on the growth of methanotrophic bacteria. *FEMS Microbiol Ecol* **12**: 35–38.

Callahan BJ, McMurdie PJ, Rosen MJ, Han AW, Johnson AJA, Holmes SP. (2016). DADA2: High-resolution sample inference from Illumina amplicon data. *Nat Methods* **13**: 581–583.

Camacho C, Coulouris G, Avagyan V, Ma N, Papadopoulos J, Bealer K, *et al.* (2009). BLAST+: architecture and applications. *BMC Bioinformatics* **10**: 421.

Caporaso JG, Lauber CL, Walters WA, Berg-Lyons D, Lozupone CA, Turnbaugh PJ, *et al.* (2011). Global patterns of 16S rRNA diversity at a depth of millions of sequences per sample. *PNAS* **108**: 4516–4522.

Carpenter LJ, Reimann S. (2015). Scientific assessment of ozone depletion: 2014. In: Engel A, Montzka SA (eds). *Global Ozone Research and Monitoring Project - Report No. 55*. World Meteotological Organization.

Chaisson MJ, Tesler G. (2012). Mapping single molecule sequencing reads using basic local alignment with successive refinement (BLASR): application and theory. *BMC Bioinformatics* **13**: 238.

Chaumeil P-A, Mussig AJ, Hugenholtz P, Parks DH. (2019). GTDB-Tk: a toolkit to classify genomes with the Genome Taxonomy Database. *Bioinformatics* btz848.

Chen G, Fisch AR, Gibson CM, Erin Mack E, Seger ES, Campagna SR, *et al.* (2020). Mineralization versus fermentation: evidence for two distinct anaerobic bacterial degradation pathways for dichloromethane. *ISME J* **14**: 959-970.

Chen G, Kleindienst S, Griffiths DR, Mack EE, Deger ES, Löffler FE. (2017a). Mutualistic interaction between dichloromethane- and chloromethane-degrading bacteria in an anaerobic mixed culture. *Environ Microbiol* **19**: 4784–4796.

Chen G, Murdoch RW, Mack EE, Seger ES, Löffler FE. (2017b). Complete genome sequence of *Dehalobacterium formicoaceticum* strain DMC, a strictly anaerobic

dichloromethane-degrading bacterium. *Genome Announc* **5**: 18–19.

Chen G, Shouakar-Stash O, Phillips E, Justicia-Leon SD, Gilevska T, Sherwood Lollar B, *et al.* (2018). Dual carbon-chlorine isotope analysis indicates distinct anaerobic dichloromethane degradation pathways in two members of *Peptococcaceae*. *Environ Sci Technol* **52**: 8607–8616.

Chen IMA, Chu K, Palaniappan K, Pillay M, Ratner A, Huang J, *et al.* (2019). IMG/M v.5.0: An integrated data management and comparative analysis system for microbial genomes and microbiomes. *Nucleic Acids Res* **47**: D666–D677.

Chin C-S, Alexander DH, Marks P, Klammer AA, Drake J, Heiner C, *et al.* (2013). Nonhybrid, finished microbial genome assemblies from long-read SMRT sequencing data. *Nat Methods* **10**: 563–569.

Cho KH, Salyers AA. (2001). Biochemical analysis of interactions between outer membrane proteins that contribute to starch utilization by *Bacteroides thetaiotaomicron*. *J Bacteriol* **183**: 7224–7230.

Choquet CG, Ahonkhai I, Klein M, Kushner DJ. (1991). Formation and role of glycine betaine in the moderate halophile *Vibrio costicola*. *Arch Microbiol* **155**: 153–258.

Christie-Oleza JA, Scanlan DJ, Armengaud J. (2015). 'You produce while I clean up', a strategy revealed by exoproteomics during *Synechococcus-Roseobacter* interactions. *Proteomics* **15**: 3454–3462.

Christie-Oleza JA, Sousoni D, Lloyd M, Armengaud J, Scanlan DJ. (2017). Nutrient recycling facilitates long-term stability of marine microbial phototroph–heterotroph interactions. *Nat Microbiol* **2**: 17100.

Cohen Y, Helman Y, Sigalevich P. (1994). Light-driven sulfate reduction and methane emission in hypersaline cyanobacterial mats. In: Stahl LJ, Caumette P (eds). *Microbial Mats. Structure, Development and Environmental Significance*. Springer-Verlag: Berlin, pp 421–427.

Conner CP, Heithoff DM, Julio SM, Sinsheimer RL, Mahan MJ. (1998). Differential patterns of acquired virulence genes distinguish *Salmonella* strains. *Proc Natl Acad Sci U S A* **95**: 4641–4645.

Couvin D, Bernheim A, Toffano-Nioche C, Touchon M, Michalik J, Néron B, *et al.*

(2018). CRISPRCasFinder, an update of CRISRFinder, includes a portable version, enhanced performance and integrates search for Cas proteins. *Nucleic Acids Res* **46**: W246–W251.

Cox J, Hein MY, Luber CA, Paron I, Nagaraj N, Mann M. (2014). Accurate proteome-wide label-free quantification by delayed normalization and maximal peptide ratio extraction, termed MaxLFQ. *Mol Cell Proteomics* **13**: 2513–2526.

Cox MP, Peterson DA, Biggs PJ. (2010). SolexaQA: At-a-glance quality assessment of Illumina second-generation sequencing data. *BMC Bioinformatics* **11**: 485.

Craciun S, Balskus EP. (2012). Microbial conversion of choline to trimethylamine requires a glycyl radical enzyme. *Proc Natl Acad Sci* **109**: 21307–21312.

Craig SAS. (2004). Betaine in human nutrition. *Am J Clin Nutr* **80**: 539–549.

Csonka LN. (1989). Physiological and genetic responses of bacteria to osmotic stress. *Microbiol Rev* **53**: 121–147.

Daniel SL, Hsu T, Dean SI, Drake HL. (1990). Characterization of the H_2- and CO-dependent chemolithotrophic potentials of the acetogens *Clostridium thermoaceticum* and *Acetogenium kivui*. *J Bacteriol* **172**: 4464–4471.

Das A, Fu Z-Q, Tempel W, Liu Z-J, Chang J, Chen L, *et al.* (2007). Characterization of a corrinoid protein involved in the C1 metabolism of strict anaerobic bacterium *Moorella thermoacetica*. *Proteins Struct Funct Bioinforma* **67**: 167–176.

Davey NE, Shields DC, Edwards RJ. (2006). SLiMDisc: Short, linear motif discovery, correcting for common evolutionary descent. *Nucleic Acids Res* **34**: 3546–3554.

Deeg R, Kriemler H-P, Bergmann K-H, Müller G. (1977). On cobyrinic acid biosynthesis. Novel methylated hydroporphyrins and their role in cobyrinic acid formation (author's transl). *Hoppe Seylers Z Physiol Chem* **358**: 339–352.

Denier-VanderGon HAC, Neue HU. (1995). Methane emission from a wetland rice field as affected by salinity. *Plant Soil* **170**: 307–313.

Ding C, Zhao S, He J. (2014). A *Desulfitobacterium sp.* strain PR reductively dechlorinates both 1,1,1-trichloroethane and chloroform. *Environ Microbiol* **16**: 3387–3397.

Dong X, Greening C, Brüls T, Conrad R, Guo K, Blaskowski S, *et al.* (2018).

Fermentative Spirochaetes mediate necromass recycling in anoxic hydrocarbon-contaminated habitats. *ISME J* **12**: 2039–2050.

Doronina N V., Trotsenko YA, Tourova TP, Kuznetsov BB, Leisinger T. (2000). *Methylopila helvetica* sp. nov. and *Methylobacterium dichloromethanicum* sp. nov. — Novel aerobic facultatively methylotrophic bacteria utilizing dichloromethane. *Syst Appl Microbiol* **23**: 210–218.

Doronina N V, Trotsenko YA, Krausova VI, Suzina NE. (1998). *Paracoccus methylutens* sp. nov. - a new aerobic facultatively methylotrophic bacterium utilizing dichloromethane. *Syst Appl Microbiol* **21**: 230–236.

Doronina N V, Trotsenko YA, Tourova TP, Kuznetsov BB, Leisinger T. (2001). *Albibacter methylovorans* gen. nov., sp. nov., a novel, aerobic, facultatively autotrophic and methylotrophic bacterium that utilizes dichloromethne. *Int J Syst Evol Microbiol* **51**: 1051–1058.

Drake HL, Horn MA, Wüst PK. (2009). Intermediary ecosystem metabolism as a main driver of methanogenesis in acidic wetland soil. *Environ Microbiol Rep* **1**: 307–318.

Duhamel M, Edwards EA. (2006). Microbial composition of chlorinated ethene-degrading cultures dominated by *Dehalococcoides*. *FEMS Microbiol Ecol* **58**: 538–549.

Duncan BN, Logan JA, Bey I, Megretskaia IA, Yantosca RM, Novelli PC, *et al.* (2007). Global budget of CO, 1988 - 1997: Source estimates and validation with a global model. *J Geophys Res Atmos* **112**: 1988–1997.

Egli C, Tschan T, Scholtz R, Cook AM, Leisinger T. (1988). Transformation of tetrachloromethane to dichloromethane and carbon dioxide by *Acetobacterium woodii*. *Appl Environ Microbiol* **54**: 2819–2824.

Eichler B, Schink B. (1984). Oxidation of primary aliphatic alcohols by *Acetobacterium carbinolicum* sp. nov., a homoacetogenic anaerobe. *Arch Microbiol* **140**: 147–152.

Einsiedl F, Pilloni G, Ruth-Anneser B, Lueders T, Griebler C. (2015). Spatial distributions of sulphur species and sulphate-reducing bacteria provide insights into sulphur redox cycling and biodegradation hot-spots in a hydrocarbon-

contaminated aquifer. *Geochim Cosmochim Acta* **156**: 207–221.

Embree M, Liu JK, Al-Bassam MM, Zengler K. (2015). Networks of energetic and metabolic interactions define dynamics in microbial communities. *Proc Natl Acad Sci U S A* **112**: 15450–15455.

Evans M V, Caldwell JC. (2010). Evaluation of two different metabolic hypotheses for dichloromethane toxicity using physiologically based pharmacokinetic modeling for in vivo inhalation gas uptake data exposure in female B6C3F1 mice. *Toxicol Appl Pharmacol* **244**: 280–290.

Fagin J, Bradley J, Williams D. (1980). Carbon monoxide poisoning secondry to inhaling methylene chloride. *Br Med J* **281**: 1980.

Fardeau M-L, Ollivier B, Patel BKC, Dwivedi P, Ragot M, Garcia J-L. (1995). Isolation and characterization of a thermophilic, sulfate-reducing bacterium, *Desulfotomaculum thermosapovorans* sp. nov. *Int J Syst Biol* **45**: 218–221.

Fenchel T, Blackburn TH. (1979). Bacteria and mineral cycling. Academic Press, Inc.: London, UK.

Ferguson DJ, Krzycki JA. (1997). Reconstitution of trimethylamine-dependent coenzyme M methylation with the trimethylamine corrinoid protein and the isozymes of methyltransferase II from *Methanosarcina barkeri*. *J Bacteriol* **179**: 846–852.

Ferguson DJJ, Gorlatova N, Grahame DA, Krzycki JA. (2000). Reconstitution of dimethylamine:coenzyme M methyl transfer with a discrete corrinoid protein and two methyltransferases purified from *Methanosarcina barkeri*. *J Biol Chem* **275**: 29053–29060.

Ferrari BC, Winsley T, Gillings M, Binnerup S. (2008). Cultivating previously uncultured soil bacteria using a soil substrate membrane system. *Nat Protoc* **3**: 1261–1269.

Fiebig K, Gottschalk G. (1983). Methanogenesis from choline by a coculture of *Desulfovibrio* sp. and *Methanosarcina barkeri*. *Appl Environ Microbiol* **45**: 161–168.

Finlayson-Pitts BJ, Pitts NJJ. (2000). Chemistry of the Upper and Lower Atmosphere. Academic Press: San Diego, CA.

Firsova J, Doronina N, Lang E, Spröer C, Vuilleumier S, Trotsenko Y. (2009). *Ancylobacter dichloromethanicus* sp. nov. – a new aerobic facultatively methylotrophic bacterium utilizing dichloromethane. *Syst Appl Microbiol* **32**: 227–232.

Firsova J, Doronina N V, Trotsenko YA. (2010). Analysis of the key functional genes in new aerobic degraders of dichloromethane. *Microbiology* **79**: 72–78.

Fischlin A, Midgley GF, Price JT, Leemans R, Gopal B, Turley CM, *et al.* (2007). Ecosystems their properties goods and services. Climate Change 2007: Impacts Adaptation and Vulnerability. In: Parry ML, Canziani OF, Palutikof JP, Van der Linden PJ, Hanson CE (eds). *Assessment Report of the Intergovernmental Panel on Climate Change, 4*. Cambridge University Press: Cambridge, UK, pp 211–272.

Freedman DL, Gossett JM. (1991). Biodegradation of dichloromethane and its utilization as a growth substrate under methanogenic conditions. *Appl Environ Microbiol* **57**: 113-2847–2857.

Freedman DL, Smith CR, Noguera DR. (1997). Dichloromethane biodegradation under nitrate-reducing conditions. *Water Environ Res* **69**: 115–122.

Fuchs G. (1989). Alternative pathways of autotrophic carbon dioxide fixation in autoptrophic bacteria. In: Schlegel HG (ed). *Biology of Autotrophic Bacteria*. Science Tech: Madison, WI, pp 365–382.

Fuchs G. (2011). Alternative pathways of carbon dioxide fixation: Insights into the early evolution of life? *Annu Rev Microbiol* **65**: 631–658.

Galbally IE, Kirstine W. (2002). The production of methanol by flowering plants and the global cycle of methanol. *J Atmos Chem* **43**: 195–229.

Galinski EA, Trüper HG. (1982). Betaine, a compatible solute in the extremely halophilic phototrophic bacterium *Ectothiorhodospira halochloris*. *FEMS Microbiol Lett* **13**: 357–360.

Gälli R, McCarty PL. (1989). Biotransformation of 1,1,1-trichloroethane, trichloromethane, and tetrachloromethane by a *Clostridium* sp. *Appl Environ Microbiol* **55**: 837–844.

Galperin MY, Brover V, Tolstoy I, Yutin N. (2016). Phylogenomic analysis of the

family *Peptostreptococcaceae* (*Clostridium* cluster xi) and proposal for reclassification of *Clostridium litorale* (Fendrich et al. 1991) and *Eubacterium acidaminophilum* (Zindel et al. 1989) as <i>Pept. *Int J Syst Evol Microbiol* **66**: 5506–5513.

Garsin DA. (2010). Ethanolamine utilization in bacterial pathogens: Roles and regulation. *Nat Rev Microbiol* **8**: 290–295.

Gaston MA, Jiang R, Krzycki JA. (2011). Functional context, biosynthesis, and genetic encoding of pyrrolysine. *Curr Opin Microbiol* **14**: 342–349.

Gibb SW, Mantoura RFC, Liss PS. (1999). Ocean-atmosphere exchange and atmospheric speciation of ammonia and methylamines in the region of the NW Arabian Sea. *Global Biogeochem Cycles* **13**: 161–178.

Godde JS, Bickerton A. (2006). The repetitive DNA elements called CRISPRs and their associated genes: Evidence of horizontal transfer among prokaryotes. *J Mol Evol* **62**: 718–729.

Godon JJ, Morinière J, Moletta M, Gaillac M, Bru V, Delgènes JP. (2005). Rarity associated with specific ecological niches in the bacterial world: The 'Synergistes' example. *Environ Microbiol* **7**: 213–224.

Goodrich JK, Waters JL, Poole AC, Sutter JL, Koren O, Blekhman R, *et al.* (2014). Human genetics shape the gut microbiome. *Cell* **159**: 789–799.

Goorissen HP, Boschker HTS, Stams AJM, Hansen TA. (2003). Isolation of thermophilic *Desulfotomaculum* strains with methanol and sulfite from solfataric mud pools, and characterization of *Desulfotomaculum solfataricum* sp. nov. *Int J Syst Evol Microbiol* **53**: 1223–1229.

Gribble GW. (2010). Naturally Occurring Organohalogen Compounds: A Comprehensive Update. In: Vol. 91. *Progress in the Chemistry of Organic Natural Products*. Springer Science & Business Media, pp 12–13.

Griebler C, Lueders T. (2009). Microbial biodiversity in groundwater ecosystems. *Freshw Biol* **54**: 649–677.

Grostern A, Duhamel M, Dworatzek S, Edwards EA. (2010). Chloroform respiration to dichloromethane by a *Dehalobacter* population. *Environ Microbiol* **12**: 1053–

1060.

Hagemeier CH, Krüer M, Thauer RK, Warkentin E, Ermler U. (2006). Insight into the mechanism of biological methanol activation based on the crystal structure of the methanol-cobalamin methyltransferase complex. *Proc Natl Acad Sci U S A* **103**: 18917–18922.

Hall AH, Rumack BH. (1990). Methylene chloride exposure in furniture-stripping shops: Ventilation and respirator use practices. *J Occup Med* **32**: 33–37.

Hall LMC, Fawell SC, Shi X, Faray-Kele MC, Aduse-Opoku J, Whiley RA, *et al.* (2005). Sequence diversity and antigenic variation at the rag locus of *Porphyromonas gingivalis*. *Infect Immun* **73**: 4253–4262.

Han M, Zmasek C. (2009). phyloXML: XML for evolutionary biology and comparative genomics. *BMC Bioinformatics* **10**: 356.

Hayward HR. (1960). Anaerobic degradation of choline: III. Acetaldehyde as an intermediate in the fermentation of choline by extracts of *Vibrio cholinicus*. *J Biol Chem* **235**: 3592–3596.

Hayward HR, Stadtman TC. (1959). Anaerobic degradation of choline: I. Fermentation of choline by an anaerobic, cytochrome-producing bacterium, *Vibrio cholinicus* N. Sp. *J Bacteriol* **78**: 557–561.

Head IM, Jones DM, Röling WFM. (2006). Marine microorganisms make a meal of oil. *Nat Rev Microbiol* **4**: 173–182.

Heidelberg JF, Nelson WC, Schoenfeld T, Bhaya D. (2009). Germ warfare in a microbial mat community: CRISPRs provide insights into the co-evolution of host and viral genomes. *PLoS One* **4**: e4169.

Heijthuijsen JHFG, Hansen TA. (1989). Betaine fermentation and oxidation by marine *Desulfuromonas* strains. *Appl Environ Microbiol* **55**: 965–969.

Heikes BG, Chang W, Pilson MEQ, Swift E, Singh HB, Guenther A, *et al.* (2002). Atmospheric methanol budget and ocean implication. *Global Biogeochem Cycles* **16**: 80-1-80–13.

Herring TI, Harris TN, Chowdhury C, Mohanty SK, Bobik TA. (2018). A bacterial microcompartment is used for choline fermentation by *Escherichia coli* 536. *J*

Bacteriol **200**: 1–13.

Hess V, Oyrik O, Trifunovic D, Müller V. (2015). 2,3-butanediol metabolism in the acetogen *Acetobacterium woodii*. *Appl Environ Microbiol* **81**: 4711–4719.

Hi Valley Chemical. (2020). Dichloromethane: All You Need To Know. https://hvchemical.com/dichloromethane-all-you-need-to-know/ (Accessed February 22, 2020).

Holland SI, Edwards RJ, Ertan H, Wong YK, Russell TL, Deshpande NP, *et al.* (2019). Whole genome sequencing of a novel, dichloromethane-fermenting *Peptococcaceae* from an enrichment culture. *PeerJ* **7**: e7775.

Horvath P, Coûté-Monvoisin AC, Romero DA, Boyaval P, Fremaux C, Barrangou R. (2009). Comparative analysis of CRISPR loci in lactic acid bacteria genomes. *Int J Food Microbiol* **131**: 62–70.

Horvath RS. (1972). Microbial co-metabolism and the degradation of organic compounds in nature. *Bacteriol Rev* **36**: 146–155.

Hossaini R, Chipperfield MP, Montzka SA, Leeson AA, Dhomse SS, Pyle JA. (2017). The increasing threat to stratospheric ozone from dichloromethane. *Nat Commun* **8**: 15962.

Hossaini R, Chipperfield MP, Montzka SA, Rap A, Dhomse S, Feng W. (2015a). Efficiency of short-lived halogens at influencing climate through depletion of stratospheric ozone. *Nat Geosci* **8**: 186–190.

Hossaini R, Chipperfield MP, Saiz-Lopez A, Harrison JJ, Von Glasow R, Sommariva R, *et al.* (2015b). Growth in stratospheric chlorine from short-lived chemicals not controlled by the Montreal Protocol. *Geophys Res Lett* **42**: 4573–4580.

Huang B, Lei C, Wei C, Zeng G. (2014). Chlorinated volatile organic compounds (Cl-VOCs) in environment - sources, potential human health impacts, and current remediation technologies. *Environ Int* **71**: 118–138.

Huerta-Cepas J, Forslund K, Coelho LP, Szklarczyk D, Jensen LJ, Von Mering C, *et al.* (2017). Fast genome-wide functional annotation through orthology assignment by eggNOG-mapper. *Mol Biol Evol* **34**: 2115–2122.

Huerta-Cepas J, Szklarczyk D, Forslund K, Cook H, Heller D, Walter MC, *et al.* (2016).

eggNOG 4.5: A hierarchical orthology framework with improved functional annotations for eukaryotic, prokaryotic and viral sequences. *Nucleic Acids Res* **44**: D286–D293.

Hugenholtz P, Goebel BM, Pace NR. (1998). Erratum: Impact of culture-independent studies on the emerging phylogenetic view of bacterial diversity (Journal of Bacteriology (1998) 180:18 (4765-4774)). *J Bacteriol* **180**: 6793.

Huson DH, Beier S, Flade I, Górska A, El-Hadidi M, Mitra S, *et al.* (2016). MEGAN Community Edition - Interactive exploration and analysis of large-scale microbiome sequencing data. *PLoS Comput Biol* **12**: 1–12.

Imhoff JF. (1986). Osmoregulation and compatible solutes in eubacteria. *FEMS Microbiol Rev* **39**: 57–66.

Intergovernmental Panel on Climate Change. (2013). Climate Change 2013: The Physical Science Basics.

International Agency for Research on Cancer. (1986). Dichloromethane. In: *IARC monograph on the evaluation of the carcinogenic risk of chemicals to humans*. Lyon, France, pp 43–85.

Iversen N. (1996). Methane oxidation in coastal marine environments. In: Murrell JC, Kelly DP (eds). *Microbiology of Atmospheric Trace Gases. Sources, Sinks and Global Change Processes*. Springer-Verlag: Berlin, pp 51–68.

Jameson E, Doxey AC, Airs R, Purdy KJ, Murrell JC, Chen Y. (2016a). Metagenomic data-mining reveals contrasting microbial populations responsible for trimethylamine formation in human gut and marine ecosystems. *Microb Genomics* **2**: e000080.

Jameson E, Fu T, Brown IR, Paszkiewicz K, Purdy KJ, Frank S, *et al.* (2016b). Anaerobic choline metabolism in microcompartments promotes growth and swarming of *Proteus mirabilis*. *Environ Microbiol* **18**: 2886–2898.

Jameson E, Stephenson J, Jones H, Millard A, Kaster A-K, Purdy KJ, *et al.* (2019). *Deltaproteobacteria* (*Pelobacter*) and *Methanococcoides* are responsible for choline-dependent methanogenesis in a coastal saltmarsh sediment. *ISME J* **13**: 277–289.

Johnson CL V, Pechonick E, Park SD, Havemann GD, Leal NA, Bobik TA. (2001).

Functional genomic, biochemical, and genetic characterization of the *Salmonella pduO* gene, an ATP:cob(I)alamin adenosyltransferase gene. *J Bacteriol* **183**: 1577–1584.

Jones HJ, Kröber E, Stephenson J, Mausz MA, Jameson E, Millard A, *et al.* (2019). A new family of uncultivated bacteria involved in methanogenesis from the ubiquitous osmolyte glycine betaine in coastal saltmarsh sediments. *Microbiome* **7**: 120.

Jones SW, Fast AG, Carlson ED, Wiedel CA, Au J, Antoniewicz MR, *et al.* (2016). CO_2 fixation by anaerobic non-photosynthetic mixotrophy for improved carbon conversion. *Nat Commun* **7**: 12800.

Joseph B, Przybilla K, Stühler C, Schauer K, Slaghuis J, Fuchs TM, *et al.* (2006). Identification of *Listeria monocytogenes* genes contributing to intracellular replication by expression profiling and mutant screening. *J Bacteriol* **188**: 556–568.

Jugder B-E, Ertan H, Bohl S, Lee M, Marquis CP, Manefield M. (2016a). Organohalide respiring bacteria and reductive dehalogenases: Key tools in organohalide bioremediation. *Front Microbiol* **7**: 1–12.

Jugder B-E, Ertan H, Wong YK, Braidy N, Manefield M, Marquis CP, *et al.* (2016b). Genomic, transcriptomic and proteomic analyses of *Dehalobacter* UNSWDHB in response to chloroform. *Environ Microbiol Rep* **8**: 814–824.

Jumas-Bilak E, Carlier JP, Jean-Pierre H, Citron D, Bernard K, Damay A, *et al.* (2007). *Jonquetella anthropi* gen. nov., sp. nov., the first member of the candidate phylum '*Synergistetes*' isolated from man. *Int J Syst Evol Microbiol* **57**: 2743–2748.

Justicia-Leon SD, Ritalahti KM, Mack EE, Löffler FE. (2012). Dichloromethane fermentation by a *Dehalobacter* sp. in an enrichment culture derived from pristine river sediment. *Appl Environ Microbiol* **78**: 1288–1291.

Kalyaanamoorthy S, Minh BQ, Wong TKF, von Haeseler A, Jermiin LS. (2017). ModelFinder: fast model selection for accurate phylogenetic estimates. *Nat Methods* **14**: 587–589.

Kameyama S, Tanimoto H, Inomata S, Tsunogai U, Ooki A, Takeda S, *et al.* (2010). High-resolution measurement of multiple volatile organic compounds dissolved in seawater using equilibrator inlet-proton transfer reaction-mass spectrometry (EI-

PTR-MS). *Mar Chem* **122**: 59–73.

Kanazawa S, Filip Z. (1986). Effects of trichloroethylene, tetrachloroethylene and dichloromethane on enzymatic activities in soil. *Appl Microbiol Biotechnol* **25**: 76–81.

Kanehisa M, Sato Y, Morishima K. (2016). BlastKOALA and GhostKOALA: KEGG tools for functional characterization of genome and metagenome sequences. *J Mol Biol* **428**: 726–731.

Katoh K, Misawa K, Kuma K, Miyata T. (2002). MAFFT: a novel method for rapid multiple sequence alignment based on fast Fourier transform. *Nucleic Acids Res* **30**: 3059–3066.

Kayser MF, Stumpp MT, Vuilleumier S. (2000). DNA Polymerase I is essential for growth of *Methylobacterium dichloromethanicum* DM4 with dichloromethane. *J Bacteriol* **182**: 5433–5439.

Kayser MF, Ucurum Z, Vuilleumier S. (2002). Dichloromethane metabolism and C1 utilization genes in *Methylobacterium* strains. *Microbiology* **148**: 1915–1922.

Kayser MF, Vuilleumier S. (2001). Dehalogenation of dichloromethane by dichloromethane dehalogenase/glutathione S-transferase leads to formation of DNA adducts. *J Bacteriol* **183**: 5209–5212.

Kelly WJ, Henderson G, Pacheco DM, Li D, Reilly K, Naylor GE, *et al.* (2016). The complete genome sequence of *Eubacterium limosum* SA11, a metabolically versatile rumen acetogen. *Stand Genomic Sci* **11**: 1–10.

Kerfeld CA, Aussignargues C, Zarzycki J, Cai F, Sutter M. (2018). Bacterial microcompartments. *Nat Rev Microbiol* **16**: 277–290.

Kiene RP. (1998). Uptake of choline and its conversion to glycine betaine by bacteria in estuarine waters. *Appl Environ Microbiol* **64**: 1045–1051.

King GM. (1988a). Distribution and metabolism of quaternary amines in marine sediments. In: Blackburn TH, Soorensen J (eds). *Nitrogen Cycling in Coastal Marine Environments*. John Wiley & Sons, pp 143–173.

King GM. (1984). Metabolism of trimethylamine, choline, and glycine betaine by sulfate-reducing and methanogenic bacteria in marine sediments. *Appl Environ*

Microbiol **48**: 719–725.

King GM. (1988b). Methanogenesis from methylated amines in a hypersaline algal mat. *Appl Environ Microbiol* **54**: 130–136.

King GM, Klug MJ, Lovley DR. (1983). Metabolism of acetate, methanol, and methylated amines in intertidal sediments of Lowes Cove, Maine. *Appl Environ Microbiol* **45**: 1848–1853.

Kleindienst S, Chourey K, Chen G, Murdoch RW, Higgins SA, Iyer R, *et al.* (2019). Proteogenomics reveals novel reductive dehalogenases and methyltrasnferases expressed during anaerobic dichloromethane metabolism. *Appl Environ Microbiol* **85**: 1–16.

Kleindienst S, Higgins SA, Tsementzi D, Chen G, Konstantinidis KT, Mack EE, *et al.* (2017). '*Candidatus* Dichloromethanomonas elyunquensis' gen. nov., sp. nov., a dichloromethane-degrading anaerobe of the *Peptococcaceae* family. *Syst Appl Microbiol* **40**: 150–159.

Kleindienst S, Higgins SA, Tsementzi D, Konstantinidis KT, Mack EE, Löffler FE. (2016). Draft genome sequence of a strictly anaerobic dichloromethane-degrading bacterium. *Genome Announc* **4**: e00037-16.

Kleinsteuber S, Schleinitz KM, Breitfeld J, Harms H, Richnow HH, Vogt C. (2008). Molecular characterization of bacterial communities mineralizing benzene under sulfate-reducing conditions. *FEMS Microbiol Ecol* **66**: 143–157.

Kleinsteuber S, Schleinitz KM, Vogt C. (2012). Key players and team play: Anaerobic microbial communities in hydrocarbon-contaminated aquifers. *Appl Microbiol Biotechnol* **94**: 851–873.

Klemps R, Cypionka H, Widdel F, Pfennig N. (1985). Growth with hydrogen, and further physiological characteristics of *Desulfotomaculum* species. *Arch Microbiol* **143**: 203–208.

Koenig J, Lee M, Manefield M. (2014). Aliphatic organochlorine degradation in subsurface environments. *Rev Environ Sci Biotechnol* **14**: 49–71.

Kohler-Staub D, Leisinger T. (1985). Dichloromethane dehalogenase of *Hyphomicrobium* sp. strain DM2. *J Bacteriol* **162**: 676–681.

Kolb S. (2009). Aerobic methanol-oxidizing Bacteria in soil. *FEMS Microbiol Lett* **300**: 1–10.

Kolusu SR, Schlünzen KH, Grawe D, Seifert R. (2017). Chloromethane and dichloromethane in the tropical Atlantic Ocean. *Atmos Environ* **150**: 417–424.

Kolusu SR, Schlünzen KH, Grawe D, Seifert R. (2018). Determination of chloromethane and dichloromethane in a tropical terrestrial mangrove forest in Brazil by measurements and modelling. *Atmos Environ* **173**: 185–197.

Konstantinidis KT, Tiedje JM. (2007). Prokaryotic taxonomy and phylogeny in the genomic era: advancements and challenges ahead. *Curr Opin Microbiol* **10**: 504–509.

Koonin E V, Wolf YI. (2008). Genomics of bacteria and archaea: The emerging dynamic view of the prokaryotic world. *Nucleic Acids Res* **36**: 6688–6719.

Korbel JO, Doerks T, Jensen LJ, Perez-Iratxeta C, Kaczanowski S, Hooper SD, *et al.* (2005). Systematic association of genes to phenotypes by genome and literature mining. *PLoS Biol* **3**: 0815–0825.

Kotsyurbenko OR, Simankova M V., Nozhevnikova AN, Zhilina TN, Bolotina NP, Lysenko AM, *et al.* (1995). New species of psychrophilic acetogens: *Acetobacterium bakii* sp. nov., *A. paludosum* sp. nov., *A. fimetarium* sp. nov. *Arch Microbiol* **163**: 29–34.

Kreimer S, Andreesen JR. (1995). Glycine reductase of *Clostridium litorale* - Cloning, sequencing, and molecular analysis of the grdAB operon that contains two in-frame TGA codons for selenium incorporation. *Eur J Biochem* **234**: 192–199.

Kremp F, Poehlein A, Daniel R, Müller V. (2018). Methanol metabolism in the acetogenic bacterium *Acetobacterium woodii*. *Environ Microbiol* **20**: 4369–4384.

Krzycki JA. (2004). Function of genetically encoded pyrrolysine in corrinoid-dependent methylamine methyltransferases. *Curr Opin Chem Biol* **8**: 484–491.

Kuehl J V, Price MN, Ray J, Wetmore KM, Esquivel Z, Kazakov AE, *et al.* (2014). Functional genomics with a comprehensive library of transposon mutants for the sulfate-reducing bacterium *Desulfovibrio alaskensis* G20. *MBio* **5**: e01041-14.

Kuraku S, Zmasek CM, Nishimura O, Katoh K. (2013). aLeaves facilitates on-demand exploration of metazoan gene family trees on MAFFT sequence alignment server

with enhanced interactivity. *Nucleic Acids Res* **41**: W22–W28.

Landfald B, Strom AR. (1986). Choline-glycine betaine pathway confers a high level of osmotic tolerance in *Escherichia coli*. *J Bacteriol* **165**: 849–855.

Larher F, Jolivet Y, Briens U, Goas U. (1982). Osmoregulation in higher plants halophytes: organic nitrogen accumulation in glycine, betaine, and proline during the growth of *Aster tripolinum* and *Sueda macrocarpa* under saline conditions. *Plant Sci Lett* **24**: 201–210.

Larson TJ, Ehrmann M, Boos W. (1983). Periplasmic glycerophosphodiester phosphodiesterase of *Escherichia coli*, a new enzyme of the glp regulon. *J Biol Chem* **258**: 5428–5432.

Laube JC, Engel A, Bönisch H, Möbius T, Worton DR, Sturges WT, *et al.* (2008). Contribution of very short-lived organic substances to stratospheric chlorine and bromine in the tropics - a case study. *Atmos Chem Phys* **8**: 7325–7334.

Lechtenfeld M, Heine J, Sameith J, Kremp F, Müller V. (2018). Glycine betaine metabolism in the acetogenic bacterium *Acetobacterium woodii*. *Environ Microbiol* **20**: 4512–4525.

Lee M, Low A, Zemb O, Koenig J, Michaelsen A, Manefield M. (2012). Complete chloroform dechlorination by organochlorine respiration and fermentation. *Environ Microbiol* **14**: 883–894.

Leedham Elvidge EC, Oram DE, Laube JC, Baker AK, Montzka SA, Humphrey S, *et al.* (2015). Increasing concentrations of dichloromethane, CH2Cl2, inferred from CARIBIC air samples collected 1998-2012. *Atmos Chem Phys* **15**: 1939–1958.

Leisinger T, Braus-Stromeyer SA. (1995). Bacterial growth with chlorinated methanes. *Environ Health Perspect* **103**: 33–36.

Letunic I, Bork P. (2019). Interactive Tree Of Life (iTOL) v4: recent updates and new developments. *Nucleic Acids Res* **47**: W256–W259.

Li H. (2013). Aligning sequence reads, clone sequences and assembly contigs with BWA-MEM. *BMC Bioinformatics* **11**: 485.

Liang B, Jiang J, Zhang J, Zhao Y, Li S. (2012). Horizontal transfer of dehalogenase genes involved in the catalysis of chlorinated compounds: evidence and ecological

role. *Crit Rev Microbiol* **38**: 95–110.

Liang C, Cheng G, Wixon DL, Balser TC. (2011). An Absorbing Markov Chain approach to understanding the microbial role in soil carbon stabilization. *Biogeochemistry* **106**: 303–309.

Liu Y, Karnauchow TM, Jarrell KF, Balkwill DL, Drake GR, Ringelberg D, *et al.* (1997). Description of two new thermophilic *Desulfotomaculum* spp., *Desulfotomaculum putei* sp. nov., from a deep terrestrial subsurface, and *Desulfotomaculum luciae* sp. nov., from a hot spring. *Int J Syst Bacteriol* **47**: 615–621.

Löffler FE, Sanford R a, Ritalahti KM. (2005). Enrichment, cultivation, and detection of reductively dechlorinating bacteria. *Methods Enzymol* **397**: 77–111.

Lombard V, Golaconda Ramulu H, Drula E, Coutinho PM, Henrissat B. (2014). The carbohydrate-active enzymes database (CAZy) in 2013. *Nucleic Acids Res* **42**: 490–495.

Lucchesi GI, Pallotti C, Lisa AT, Domenech CE. (1998). Constitutive choline transport in *Pseudomonas aeruginosa*. *FEMS Microbiol Lett* **162**: 123–126.

Ludwig W, Strunk O, Westram R, Richter L, Meier H, Yadhukumar A, *et al.* (2004). ARB: A software environment for sequence data. *Nucleic Acids Res* **32**: 1363–1371.

Lueders T. (2017). The ecology of anaerobic degraders of BTEX hydrocarbons in aquifers. *FEMS Microbiol Ecol* **93**: 1–13.

Lynd LH, Zeikus JG. (1983). Metabolism of H_2-CO_2, methanol, and glucose by *Butyribacterium methylotrophicum*. *J Bacteriol* **153**: 1415–1423.

Mack K. (1973). The problems of waste water purification in the chemical pharmaceutical industry. *Prog Water Technol GB* **3**: 329.

Maeda H, Fujimoto C, Haruki Y, Maeda T, Kokeguchi S, Petelin M, *et al.* (2003). Quantitative real-time PCR using TaqMan and SYBR Green for *Actinobacillus actinomycetemcomitans*, *Porphyromonas gingivalis*, *Prevotella intermedia*, *tetQ* gene and total bacteria. *FEMS Immunol Med Microbiol* **39**: 81–86.

Mägli A, Messmer M, Leisinger T. (1998). Metabolism of dichloromethane by the strict anaerobe *Dehalobacterium formicoaceticum*. *Appl Environ Microbiol* **64**: 646–650.

Mägli A, Rainey FA, Leisinger T. (1995). Acetogenesis from dichloromethane by a two-component mixed culture comprising a novel bacterium. *Appl Environ Microbiol* **61**: 2943–2949.

Mägli A, Wendt M, Leisinger T. (1996). Isolation and characterization of *Dehalobacterium formicoaceticum* gen. nov. sp. nov., a strictly anaerobic bacterium utilizing dichloromethane as source of carbon and energy. *Arch Microbiol* **166**: 101–108.

Maillard J, Regeard C, Holliger C. (2005). Isolation and characterization of Tn-Dha1, a transposon containing the tetrachloroethene reductive dehalogenase of *Desulfitobacterium hafniense* strain TCE1. *Environ Microbiol* **7**: 107–117.

Marietou A, Griffiths L, Cole J. (2009). Preferential reduction of the thermodynamically less favorable electron acceptor, sulfate, by a nitrate-reducing strain of the sulfate-reducing bacterium *Desulfovibrio desulfuricans* 27774. *J Bacteriol* **191**: 882–889.

Martin WF. (2012). Hydrogen, metals, bifurcating electrons, and proton gradients: The early evolution of biological energy conservation. *FEBS Lett* **586**: 485–493.

Martínez-del Campo A, Bodea S, Hamer HA, Marks JA, Haiser HJ, Turnbaugh PJ, *et al.* (2015). Characterization and detection of a widely distributed gene cluster that predicts anaerobic choline utilization by human gut bacteria. *MBio* **6**: e00042-15.

McMurdie PJ, Behrens SF, Müller JA, Göke J, Ritalahti KM, Wagner R, *et al.* (2009). Localized plasticity in the streamlined genomes of vinyl chloride respiring *Dehalococcoides*. *PLoS Genet* **5**: e1000714.

McMurdie PJ, Hug L a, Edwards E a, Holmes S, Spormann AM. (2011). Site-specific mobilization of vinyl chloride respiration islands by a mechanism common in *Dehalococcoides*. *BMC Genomics* **12**: 287.

Médigue C, Calteau A, Cruveiller S, Gachet M, Gautreau G, Josso A, *et al.* (2019). MicroScope—an integrated resource for community expertise of gene functions and comparative analysis of microbial genomic and metabolic data. *Brief Bioinform* **20**: 1071–1084.

van der Meijden P, Te Brömmelstroet BW, Poirot CM, van der Drift C, Vogels GD.

(1984a). Purification and properties of methanol:5-hydroxybenzimidazolylcobamide methyltransferase from *Methanosarcina barkeri*. *J Bacteriol* **160**: 629–635.

van der Meijden P, van der Drift C, Vogels GD. (1984b). Methanol conversion in *Eubacterium limosum*. *Arch Microbiol* **138**: 360–364.

van der Meijden P, Heythuysen HJ, Pouwels A, Houwen F, van der Drift C, Vogels GD. (1983a). Methyltransferases involved in methanol conversion by *Methanosarcina barkeri*. *Arch Microbiol* **134**: 238–242.

van der Meijden P, Jansen LPJM, Drift C, Vogels GD. (1983b). Involvement of corrinoids in the methylation of coenzyme M (2-mercaptoethanesulfonic acid) by methanol and enzymes from *Methanosarcina barkeri*. *FEMS Microbiol Lett* **19**: 247–251.

Melendez CR, Roman MD, Smith GB. (1993). Biodegradation of dichloromethane under denitrifying conditions by a waste water microbial community and by pure cultures of a *Hyphomicrobium* strain X. In: *Abstracts of the 93rd General Meeting, American Society for Microbiology*. p Session 126, 381.

Miele V, Penel S, Duret L. (2011). Ultra-fast sequence clustering from similarity networks with SiLiX. *BMC Bioinformatics* **12**: 116.

Mikesell MD, Boyd SA. (1990). Dechlorination of chloroform by *Methanosarcina* strains. *Appl Environ Microbiol* **56**: 1198–1201.

Millet DB, Jacob DJ, Turquety S, Hudman RC, Wu S, Fried A, *et al.* (2006). Formaldehyde distribution over North America: Implications for satellite retrievals of formaldehyde columns and isoprene emission. *J Geophys Res Atmos* **111**: 1–17.

Miltner A, Bombach P, Schmidt-Brücken B, Kästner M. (2012). SOM genesis: Microbial biomass as a significant source. *Biogeochemistry* **111**: 41–55.

Mitchell SC, Smith RL. (2001). Trimethylaminuria: the fish malodour syndrome. *Drug Metab Dispos* **29**: 517–521.

Möller B, Hippe H, Gottschalk G. (1986). Degradation of various amine compounds by mesophilic clostridia. *Arch Microbiol* **145**: 85–90.

Möller B, Oßmer R, Howard BH, Gottschalk G, Hippe H. (1984). *Sporomusa*, a new

genus of gram-negative anaerobic bacteria including *Sporomusa sphaeroides* spec. nov. and *Sporomusa ovata* spec. nov. *Arch Microbiol* **139**: 388–396.

Montzka SA, Reimann S, Engel A, Kruger K, O'Doherty S, Sturges WT, *et al.* (2011). Ozone-depleting substances (ODSs) and related chemicals. In: *Scientific Assessment of Ozone Depletion: 2010, Global Ozone Research and Monitoring Project - Report No. 52*. World Meteorological Organisation: Geneva, Switzerland, pp 1–112.

Moore RM. (2004). Dichloromethane in North Atlantic waters. *J Geophys Res C Ocean* **109**: 1–6.

Mori H, Evans-Yamamoto D, Ishiguro S, Tomita M, Yachie N. (2019). Fast and global detection of periodic sequence repeats in large genomic resources. *Nucleic Acids Res* **47**: e8.

Morris JT, Whiting GJ. (1986). Emission of gaseous carbon dioxide from salt-marsh sediments and its relation to other carbon losses. *Estuaries* **9**: 9–19.

Mouné S, Manac'h N, Hirschler A, Caumette P, Willison JC, Matheron R. (1999). *Haloanaerobacter salinarius* sp. nov., a novel halophilic fermentative bacterium that reduces glycine-betaine to trimethylamine with hydrogen or serine as electron donors; emendation of the genus *Haloanaerobacter*. *Int J Syst Biol* **49**: 103–112.

Müller E, Fahlbusch K, Walther R, Gottschalk G. (1981). Formation of N,N-dimethylglycine, acetic acid, and butyric acid from betaine by *Eubacterium limosum*. *Appl Environ Microbiol* **42**: 439–445.

Muller EEL, Bringel F, Vuilleumier S. (2011). Dichloromethane-degrading bacteria in the genomic age. *Res Microbiol* **162**: 870–876.

Müller S, Vogt C, Laube M, Harms H, Kleinsteuber S. (2009). Community dynamics within a bacterial consortium during growth on toluene under sulfate-reducing conditions. *FEMS Microbiol Ecol* **70**: 586–596.

Murray RGE, Stackebrandt E. (1995). Taxonomic Note: Implementation of the provisional status *Candidatus* for incompletely described procaryotes. *Int J Syst Bacteriol* **45**: 186–187.

Nanninga HJ, Gottschal JC. (1987). Properties of *Desulfovibrio carbinolicus* sp. nov. and other sulfate-reducing bacteria isolated from an anaerobic-purification plant.

Appl Environ Microbiol **53**: 802–809.

Naumann E, Hippe H, Gottschalk G. (1983). Betaine: New oxidant in the Stickland reaction and methanogenesis from betaine and L-alanine by a *Clostridium sporogenes-Methanosarcina barkeri* coculture. *Appl Environ Microbiol* **45**: 474–483.

Neill AR, Grime DW, Dawson RM. (1978). Conversion of choline methyl groups through trimethylamine into methane in the rumen. *Biochem J* **170**: 529–535.

Nemoto N, Takeda Y, Nara H, Araki A, Gazi MY, Takakubo Y, *et al.* (2020). Analysis of intestinal immunity and flora in a collagen-induced mouse arthritis model: differences during arthritis progression. *Int Immunol* **32**: 49–56.

Nguyen LT, Schmidt HA, Von Haeseler A, Minh BQ. (2015). IQ-TREE: A fast and effective stochastic algorithm for estimating maximum-likelihood phylogenies. *Mol Biol Evol* **32**: 268–274.

Nisman B. (1954). The Stickland Reaction. *Bacteriol Rev* **18**: 16–42.

Ohta-Fukuyama M, Miyake Y, Emi S, Yamano T. (1980). Identification and properties of the prosthetic group of choline oxidase from *Alcaligenes* sp. *J Biochem* **88**: 197–203.

Okonechnikov K, Golosova O, Fursov M, the UGENE team. (2012). Unipro UGENE: A unified bioinformatics toolkit. *Bioinformatics* **28**: 1166–1167.

Olvera-Bello AE, Estrada-Muñiz E, Elizondo G, Vega L. (2010). Susceptibility to the cytogenetic effects of dichloromethane is related to the glutathione S-transferase theta phenotype. *Toxicol Lett* **199**: 218–224.

Ooki A, Yokouchi Y. (2011). Dichloromethane in the Indian Ocean: Evidence for in-situ production in seawater. *Mar Chem* **124**: 119–124.

Oremland RS, Marsh LM, Polcin S. (1982). Methane production and simultaneous sulphate reduction in anoxic, salt marsh sediments. *Nature* **296**: 143–145.

Oremland RS, Polcin S. (1982). Methanogenesis and sulfate reduction: Competitive and noncompetitive substrates in estuarine sediments. *Appl Environ Microbiol* **44**: 1270–1276.

Parks DH, Chuvochina M, Waite DW, Rinke C, Skarshewski A, Chaumeil PA, *et al.* (2018). A standardized bacterial taxonomy based on genome phylogeny

substantially revises the tree of life. *Nat Biotechnol* **36**: 996.

Parks DH, Imelfort M, Skennerton CT, Hugenholtz P, Tyson GW. (2015). CheckM: assessing the quality of microbial genomes recovered from isolates, single cells, and metagenomes. *Genome Res* **25**: 1043–55.

Parther T. (2003). Die Peroxidase-Aktivität Selenocystein-haltiger Proteine des strikt anaeroben Bakteriums *Eubacterium acidaminophilum* [Translation: The peroxidase activity of selenocysteine-containing proteins of the strict anaerobic bacterium *Eubacterium acidaminophilum*]. Universität Halle.

Pascual-García A, Bonhoeffer S, Bell T. (2019). Microbial metabolically cohesive consortia and ecosystem functioning. *bioRxiv*. doi: 10.1101/859421.

Paul L, Ferguson DJ, Krzycki JA. (2000). The trimethylamine methyltransferase gene and multiple dimethylamine methyltransferase genes of Methanosarcina barkeri contain in-frame and read- through amber codons. *J Bacteriol* **182**: 2520–2529.

Petit E, LaTouf WG, Coppi M V., Warnick TA, Currie D, Romashko I, *et al.* (2013). Involvement of a bacterial microcompartment in the metabolism of fucose and rhamnose by *Clostridium phytofermentans*. *PLoS One* **8**: 1–12.

Picking JW, Behrman EJ, Zhang L, Krzycki JA. (2019). MtpB, a member of the MttB superfamily from the human intestinal acetogen *Eubacterium limosum*, catalyzes proline betaine demethylation. *J Biol Chem* **294**: 13697–13707.

Pitluck S, Yasawong M, Held B, Lapidus A, Nolan M, Copeland A, *et al.* (2010). Non-contiguous finished genome sequence of *Aminomonas paucivorans* type strain (GLU-3 T). *Stand Genomic Sci* **3**: 285–293.

Portillo MC, Gonzalez JM. (2009). CRISPR elements in the Thermococcales: evidence for associated horizontal gene transfer in *Pyrococcus furiosus*. *J Appl Genet* **50**: 421–430.

Proulx P, Fung CK. (1969). Metabolism of phosphoglycerides in *E. coli*. IV. The positional specificity and properties of phospholipase A. *Can J Biochem* **47**: 1125–1128.

Pruesse E, Peplies J, Glöckner FO. (2012). SINA: Accurate high-throughput multiple sequence alignment of ribosomal RNA genes. *Bioinformatics* **28**: 1823–1829.

Qatibi AI, Niviere V, Garcia JL. (1991). *Desulfovibrio alcoholovorans* sp. nov., a sulfate-reducing bactrium able to grow on glycerol, 1,2- and 1,3-propanediol. *Arch Microbiol* **155**: 143–148.

Randle CL, Albro PW, Dittmer JC. (1969). The phosphoglyceride composition of Gram-negative bacteria and the changes in composition during growth. *Biochim Biophys Acta* **187**: 214–220.

Reed S, Neuman H, Glahn RP, Koren O, Tako E. (2017). Characterizing the gut (*Gallus gallus*) microbiota following the consumption of an iron biofortified Rwandan cream seeded carioca (*Phaseolus Vulgaris* L.) bean-based diet. *PLoS One* **12**: 1–15.

La Roche SD, Leisinger T. (1991). Identification of *dcmR*, the regulatory gene governing expression of dichloromethane dehalogenase in *Methylobacterium* sp. strain DM4. *J Bacteriol* **173**: 6714–6721.

La Roche SD, Leisinger T. (1990). Sequence analysis and expression of the bacterial dichloromethane dehalogenase structural gene, a member of the glutathione-S-transferase supergene family. *J Bacteriol* **172**: 164–171.

Rodriguez-R LM, Konstantinidis KT. (2014). Bypassing cultivation to identify bacterial species. *Microbe* **9**: 111–117.

Rodriguez Martinez MF, Kelessidou N, Law Z, Gardiner J, Stephens G. (2008). Effect of solvents on obligately anaerobic bacteria. *Anaerobe* **14**: 55–60.

Roller C, Wagner M, Amann R, Ludwig W, Schleifer K-H. (1994). In situ probing of Gram-positive bacteria with high DNA G+C content using 23S rRNA-targeted oligonucleotides. *Microbiology* **140**: 2849–2858.

Roof DM, Roth JR. (1988). Ethanolamine utilization in *Salmonella typhimurium*. *J Bacteriol* **170**: 3855–3863.

Roof DM, Roth JR. (1989). Functions required for vitamin B_{12}-dependent ethanolamine utilization in *Salmonella typhimurium*. *J Bacteriol* **171**: 3316–3323.

Russell MJ, Martin W. (2004). The rocky roots of the acetyl-CoA pathway. *Trends Biochem Sci* **29**: 358–363.

Salvano MA, Lisa TA, Domenech CE. (1989). Choline transport in *Pseudomonas aeruginosa*. *Mol Cell Biochem* **85**: 81–89.

Sander R. (2015). Compilation of Henry's law constants (version 4.0) for water as solvent. *Atmos Chem Phys* **15**: 4399–4981.

Sanz JL, Rodriguez N, Amils R. (1997). Effect of chlorinated aliphatic hydrocarbons on the acetoclastic methanogenic activity of granular sludge. *Appl Microbiol Biotechnol* **47**: 324–328.

Sauer K, Harms U, Thauer RK. (1997). Methanol:coenzyme M methyltransferase from *Methanosarcina barkeri*: purification, properties and encoding genes of the corrinoid protein MT1. *Eur J Biochem* **243**: 670–677.

Sauer K, Thauer RK. (1997). Methanol:coenzyme M methyltransferase from *Methanosarcina barkeri*: Zinc dependence and thermodynamics of the methanol:cob(I)alamin methyltransferase reaction. *Eur J Biochem* **249**: 280–285.

Schink B. (1985). Fermentation of acetylene by an obligate anaerobe, *Pelobacter acetylenicus* sp. nov. *Arch Microbiol* **142**: 295–301.

Schink B, Pfennig N. (1982). Fermentation of trihydroxybenzenes by *Pelobacter acidigallici* gen. nov. sp. nov., a new strictly anaerobic, non-sporeforming bacterium. *Arch Microbiol* **133**: 195–201.

Schink B, Zeikus JG. (1980). Microbial methanol formation: A major end product of pectin metabolism. *Curr Microbiol* **4**: 387–389.

Schmidt MWI, Torn MS, Abiven S, Dittmar T, Guggenberger G, Janssens IA, *et al.* (2011). Persistence of soil organic matter as an ecosystem property. *Nature* **478**: 49–56.

Schnell S, Bak F, Pfennig N. (1989). Anaerobic degradation of aniline and dihydroxybenzenes by newly isolated sulfate-reducing bacteria and description of *Desulfobacterium anilini*. *Arch Microbiol* **152**: 556–563.

Scholtz R, Wackett LP, Egli C, Cook AM, Leisinger T. (1988). Dichloromethane dehalogenase with improved catalytic activity isolated from a fast-growing dichloromethane-utilizing bacterium. *J Bacteriol* **170**: 5698–5704.

Schuchmann K, Müller V. (2016). Energetics and application of heterotrophy in acetogenic bacteria. *Appl Environ Microbiol* **82**: 4056–4069.

Seemann T. (2014). Prokka: Rapid prokaryotic genome annotation. *Bioinformatics*

30: 2068–2069.

Shan H, Kurtz HD, Mykytczuk N, Trevors JT, Freedman DL. (2010). Anaerobic biotransformation of high concentrations of chloroform by an enrichment culture and two bacterial isolates. *Appl Environ Microbiol* **76**: 6463–6439.

Sheard NF, Zeisel SH. (1989). Choline: an essential dietary nutrient? *Nutrition* **5**: 1–5.

Sherry A, Grant RJ, Aitken CM, Jones DM, Head IM, Gray ND. (2014). Volatile hydrocarbons inhibit methanogenic crude oil degradation. *Front Microbiol* **5**: 1–9.

Shestakova M, Sillanpää M. (2013). Removal of dichloromethane from ground and wastewater: A review. *Chemosphere* **93**: 1258–1267.

Sikkema J, de Bont JAM, Poolman B. (1995). Mechanisms of membrane toxicity of hydrocarbons. *Microbiol Rev* **59**: 201–222.

Simpson AJ, Simpson MJ, Smith E, Kelleher BP. (2007). Microbially derived inputs to soil organic matter: Are current estimates too low? *Environ Sci Technol* **41**: 8070–8076.

Singh HB, Kanakidou M, Crutzen PJ, Jacob DJ. (1995). High concentrations and photochemical fate of oxygenated hydrocarbons in the global troposphere. *Nature* **378**: 50–54.

Singh HB, Tabazadeh A, Evans MJ, Field BD, Jacob DJ, Sachse G, *et al.* (2003). Oxygenated volatile organic chemicals in the oceans: Inferences and implications based on atmospheric observations and air-sea exchange models. *Geophys Res Lett* **30**: 1–5.

Smidt H, de Vos WM. (2004). Anaerobic microbial dehalogenation. *Annu Rev Microbiol* **58**: 43–73.

Snipes W, Keith A, Wanda P. (1974). Active transport of choline by a marine pseudomonad. *J Bacteriol* **120**: 197–202.

Søndergaard D, Pedersen CNS, Greening C. (2016). HydDB: A web tool for hydrogenase classification and analysis. *Sci Rep* **6**: 1–8.

Song W, Sun H-X, Zhang C, Cheng L, Peng Y, Deng Z, *et al.* (2019). Prophage Hunter: an integrative hunting tool for active prophages. *Nucleic Acids Res* **47**: W74–W80.

Sousa DZ, Visser M, van Gelder AH, Boeren S, Pieterse MM, Pinkse MWH, *et al.* (2018). The deep-subsurface sulfate reducer *Desulfotomaculum kuznetsovii* employs two methanol-degrading pathways. *Nat Commun* **9**: 239.

Srinivasan G. (2002). Pyrrolysine encoded by UAG in archaea: Charging of a UAG-decoding specialized tRNA. *Science (80-)* **296**: 1459–1462.

Starr TB, Matanoski G, Anders MW, Andersen ME. (2006). Workshop overview: Reassessment of the cancer risk of dichloromethane in humans. *Toxicol Sci* **91**: 20–28.

Stewart RD, Hake CL. (1976). Paint-Remover Hazard. *JAMA* **235**: 398–401.

Stojiljkovic I, Baumler AJ, Heffron F. (1995). Ethanolamine utilization in *Salmonella typhimurium*: Nucleotide sequence, protein expression, and mutational analysis of the *cchA cchB eutE eutJ eutG eutH* gene cluster. *J Bacteriol* **177**: 1357–1366.

Strąpoć D, Mastalerz M, Dawson K, Macalady J, Callaghan A V., Wawrik B, *et al.* (2011). Biogeochemistry of microbial coal-bed methane. *Annu Rev Earth Planet Sci* **39**: 617–656.

Stromeyer SA, Winkelbauer W, Kohler H, Cook AM, Leisinger T. (1991). Dichloromethane utilized by an anaerobic mixed culture: acetogenesis and methanogenesis. *Biodegradation* **2**: 129–137.

Stuckey DC, Owen WF, McCarty PL, Parkin GF. (1980). Anaerobic toxicity evaluation by batch and semi-continuous assays. *J (Water Pollut Control Fed* **52**: 720–729.

Stucki G, Gälli R, Ebersold H-R, Leisinger T. (1981). Dehalogenation of dichloromethane by cell extracts of *Hyphomicrobium* DM2. *Arch Microbiol* **130**: 366–371.

Studer A, Stupperich E, Vuilleumier S, Leisinger T. (2001). Chloromethane:tetrahydrofolate methyl transfer by two proteins from *Methylobacterium chloromethanicum* strain CM4. *Eur J Biochem* **268**: 2931–2938.

Stupperich E, Konle R. (1993). Corrinoid-dependent methyl transfer reactions are involved in methanol and 3,4-dimethoxybenzoate metabolism by *Sporomusa ovata*. *Appl Environ Microbiol* **59**: 3110–3116.

Sun L, Toyonaga M, Ohashi A, Tourlousse DM, Matsuura N, Meng XY, *et al.* (2016).

Lentimicrobium saccharophilum gen. nov., sp. nov., a strictly anaerobic bacterium representing a new family in the phylum *Bacteroidetes*, and proposal of *Lentimicrobiaceae* fam. nov. *Int J Syst Evol Microbiol* **66**: 2635–2642.

Szewzyk R, Pfennig N. (1987). Complete oxidation of catechol by the strictly anaerobic sulfate-reducing *Desulfobacterium catecholicum* sp. nov. *Arch Microbiol* **147**: 163–168.

Tan B, Jane Fowler S, Laban NA, Dong X, Sensen CW, Foght J, *et al.* (2015). Comparative analysis of metagenomes from three methanogenic hydrocarbon-degrading enrichment cultures with 41 environmental samples. *ISME J* **9**: 2028–2045.

Tanaka K, Pfennig N. (1988). Fermentation of 2-methoxyethanol by *Acetobacterium malicum* sp. nov. and *Pelobacter venetianus*. *Arch Microbiol* **149**: 181–187.

Tang S, Wang PH, Higgins SA, Löffler FE, Edwards EA. (2016). Sister *Dehalobacter* genomes reveal specialization in organohalide respiration and recent strain differentiation likely driven by chlorinated substrates. *Front Microbiol* **7**: 100.

Taubert M, Vogt C, Wubet T, Kleinsteuber S, Tarkka MT, Harms H, *et al.* (2012). Protein-SIP enables time-resolved analysis of the carbon flux in a sulfate-reducing, benzene-degrading microbial consortium. *ISME J* **6**: 2291–2301.

Tebo BM, Obraztsova AY. (1998). Sulfate-reducing bacterium grows with Cr(VI), U(VI), Mn(IV), and Fe(III) as electron acceptors. *FEMS Microbiol Lett* **162**: 193–198.

Thiel PG. (1969). The effect of methane analogues on methanogenesis in anaerobic digestion. *Water Res* **3**: 215–223.

Thiennimitr P, Winter SE, Winter MG, Xavier MN, Tolstikov V, Huseby DL, *et al.* (2011). Intestinal inflammation allows *Salmonella* to use ethanolamine to compete with the microbiota. *Proc Natl Acad Sci U S A* **108**: 17480–17485.

Ticak T, Kountz DJ, Girosky KE, Krzycki JA, Ferguson DJ. (2014). A nonpyrrolysine member of the widely distributed trimethylamine methyltransferase family is a glycine betaine methyltransferase. *Proc Natl Acad Sci U S A* **111**: E4668–E4676.

Trifunović D, Schuchmann K, Müller V. (2016). Ethylene glycol metabolism in the acetogen *Acetobacterium woodii*. *J Bacteriol* **198**: 1058–1065.

Trueba-Santiso A, Fernández-Verdejo D, Marco-Rius I, Soder-Walz JM, Casabella O, Vicent T, *et al.* (2020). Interspecies interaction and effect of co-contaminants in an anaerobic dichloromethane-degrading culture. *Chemosphere* **240**: 124877.

Trueba-Santiso A, Parladé E, Rosell M, Lliros M, Mortan SH, Martínez-Alonso M, *et al.* (2017). Molecular and carbon isotopic characterization of an anaerobic stable enrichment culture containing *Dehalobacterium* sp. during dichloromethane fermentation. *Sci Total Environ* **581**: 640–648.

Tschech A, Pfennig N. (1984). Growth yield increase linked to caffeate reduction in *Acetobacterium woodii*. *Arch Microbiol* **137**: 163–167.

Tyanova S, Temu T, Sinitcyn P, Carlson A, Hein MY, Geiger T, *et al.* (2016). The Perseus computational platform for comprehensive analysis of (prote)omics data. *Nat Methods* **13**: 731–740.

Tyson GW, Banfield JF. (2008). Rapidly evolving CRISPRs implicated in acquired resistance of microorganisms to viruses. *Environ Microbiol* **10**: 200–207.

United States Environmental Protection Agency. (2017). Fact Sheet: Methylene Chloride or Dichloromethane (DCM). *Assess Manag Chem under TSCA*. https://www.epa.gov/assessing-and-managing-chemicals-under-tsca/fact-sheet-methylene-chloride-or-dichloromethane-dcm-0#what.

Urakawa H, Martens-Habbena W, Stahl DA. (2010). High abundance of ammonia-oxidizing archaea in coastal waters, determined using a modified DNA extraction method. *Appl Environ Microbiol* **76**: 2129–2135.

US EPA. (2001). Fact Sheet: Correcting the Henry's Law Constant for Soil Temperature.

US National Library of Medicine. (2019). TOXMAP. https://toxmap.nlm.nih.gov/toxmap/app/ (Accessed November 30, 2019).

Vannelli T, Messmer M, Studer A, Vuilleumier S, Leisinger T. (1999). A corrinoid-dependent catabolic pathway for growth of a *Methylobacterium* strain with chloromethane. *Proc Natl Acad Sci U S A* **96**: 4615–4620.

Vargas C, Ahlert RC. (1987). Anaerobic degradation of chlorinated solvents. *Water Pollut Control Fed* **59**: 964–968.

Vartoukian SR, Palmer RM, Wade WG. (2007). The division '*Synergistes*'. *Anaerobe* **13**: 99–106.

Villas-Bôas SG, Delicado DG, Åkesson M, Nielsen J. (2003). Simultaneous analysis of amino and nonamino organic acids as methyl chloroformate derivatives using gas chromatography-mass spectrometry. *Anal Biochem* **322**: 134–138.

Visser M, Pieterse MM, Pinkse MWH, Nijsse B, Verhaert PDEM, de Vos WM, *et al.* (2016). Unravelling the one-carbon metabolism of the acetogen *Sporomusa* strain An4 by genome and proteome analysis. *Environ Microbiol* **18**: 2843–2855.

Vuilleumier S. (2002). Coping with a Halogenated One-Carbon Diet: Aerobic Dichloromethane-Mineralising Bacteria. In: Agathos S, Reineke W (eds). *Biotechnology for the Environment: Strategy and Fundamentals. Focus on Biotechnology vol 3A*. Springer: Dordrecht, pp 105–130.

Vuilleumier S, Chistoserdova L, Lee M-C, Bringel F, Lajus A, Zhou Y, *et al.* (2009). *Methylobacterium* genome sequences: a reference blueprint to investigate microbial metabolism of C1 compounds from natural and industrial sources. *PLoS One* **4**: e5584.

Wagner M, Sonntag D, Grimm R, Pich A, Eckerskorn C, Söhling B, *et al.* (1999). Substrate-specific selenoprotein B of glycine reductase from *Eubacterium acidaminophilum*. *Eur J Biochem* **260**: 38–49.

Walker BJ, Abeel T, Shea T, Priest M, Abouelliel A, Sakthikumar S, *et al.* (2014). Pilon: An integrated tool for comprehensive microbial variant detection and genome assembly improvement. *PLoS One* **9**: e112963.

Walsby AE. (1994). Gas vesicles. *Microbiol Rev* **58**: 94–144.

Wang Z, Klipfell E, Bennett BJ, Koeth RA, Levison BS, DuGar B, *et al.* (2011). Gut flora metabolism of phosphatidylcholine promotes cardiovascular disease. *Nature* **472**: 57–63.

Wargo MJ. (2013). Homeostasis and catabolism of choline and glycine betaine: Lessons from *Pseudomonas aeruginosa*. *Appl Environ Microbiol* **79**: 2112–2120.

Watkins AJ, Rousse EG, Parkes RJ, Sass H. (2014). Glycine betaine as a direct substrate for methanogens (*Methanococcoides* spp.). *Appl Environ Microbiol* **80**:

289–293.

Watkins AJ, Roussel EG, Webster G, Parkes RJ, Sass H. (2012). Choline and N,N-dimethylethanolamine as direct substrates for methanogens. *Appl Environ Microbiol* **78**: 8298–8303.

Weghoff MC, Bertsch J, Müller V. (2015). A novel mode of lactate metabolism in strictly anaerobic bacteria. *Environ Microbiol* **17**: 670–677.

Weiss MC, Sousa FL, Mrnjavac N, Neukirchen S, Roettger M, Nelson-Sathi S, *et al.* (2016). The physiology and habitat of the last universal common ancestor. *Nat Microbiol* **1**: 1–8.

West KA, Johnson DR, Hu P, DeSantis TZ, Brodie EL, Lee PKH, *et al.* (2008). Comparative genomics of '*Dehalococcoides ethenogenes*' 195 and an enrichment culture containing unsequenced '*Dehalococcoides*' strains. *Appl Environ Microbiol* **74**: 3533–3540.

White DA. (1973). Form and Function of Phospholipids. In: Ansell GB, Hawthorne JN, Dawson RMC (eds). Elsevier: New York, pp 441–482.

WHO Regional Office for Europe. (2000). Dichloromethane. In: *Air Quality Guidelines*. Copenhagen, Denmark.

Wick RR, Judd LM, Gorrie CL, Holt KE. (2017). Unicycler: Resolving bacterial genome assemblies from short and long sequencing reads. *PLoS Comput Biol* **13**: 1–22.

Williams J, Holzinger R, Gros V, Xu X, Atlas E, Wallace DWR. (2004). Measurements of organic species in air and seawater from the tropical Atlantic. *Geophys Res Lett* **31**: 1–5.

Wiśniewski JR. (2017). Filter-Aided Sample Preparation: The Versatile and Efficient Method for Proteomic Analysis. In: Shukla AK (ed) Vol. 585. *Methods in Enzymology. Proteomics in Biology, Part A.* Elsevier BV, pp 15–27.

Wiśniewski JR, Zougman A, Nagaraj N, Mann M. (2009). Universal sample preparation method for proteome analysis. *Nat Methods* **6**: 359–62.

Wolin EA, Wolin MJ, Wolfe RS. (1963). Formation of methane by bacterial extracts. *J Biol Chem* **238**: 2882–2886.

Wong YK. (2015). Isolation and characterization of chloroform, 1,1,2-

trichloroethane and dichloromethane degrading bacteria. University of New South Wales.

Wong YK, Holland SI, Ertan H, Manefield M, Lee M. (2016). Isolation and characterization of *Dehalobacter sp.* strain UNSWDHB capable of chloroform and chlorinated ethane respiration. *Environ Microbiol* **18**: 3092–3105.

Wood HG. (1991). Life with CO or CO_2 and H_2 as a source of carbon and energy. *FASEB J* **5**: 156–163.

Wrighton KC, Castelle CJ, Wilkins MJ, Hug LA, Sharon I, Thomas BC, *et al.* (2014). Metabolic interdependencies between phylogenetically novel fermenters and respiratory organisms in an unconfined aquifer. *ISME J* **8**: 1452–1463.

Wu SJ, Hu ZH, Zhang LL, Yu X, Chen JM. (2009). A novel dichloromethane-degrading *Lysinibacillus sphaericus* strain wh22 and its degradative plasmid. *Appl Microbiol Biotechnol* **82**: 731–740.

Wu SJ, Zhang LL, Wang JD, Chen JM. (2007). *Bacillus circulans* WZ-12 - A newly discovered aerobic dichloromethane-degrading methylotrophic bacterium. *Appl Microbiol Biotechnol* **76**: 1289–1296.

Wu Y-W, Simmons BA, Singer SW. (2016). MaxBin 2.0: an automated binning algorithm to recover genomes from multiple metagenomic datasets. *Bioinformatics* **32**: 605–607.

Yan J, Im J, Yang Y, Löffler FE. (2013). Guided cobalamin biosynthesis supports *Dehalococcoides mccartyi* reductive dechlorination activity. *Philos Trans R Soc Lond B Biol Sci* **368**: 20120320.

Yan J, Ritalahti KM, Wagner DD, Löffler FE. (2012). Unexpected specificity of interspecies cobamide transfer from *Geobacter* spp. to organohalide-respiring Dehalococcoides mccartyi strains. *Appl Environ Microbiol* **78**: 6630–6636.

Yan J, Simsir B, Farmer AT, Bi M, Yang Y, Campagna SR, *et al.* (2016). The corrinoid cofactor of reductive dehalogenases affects dechlorination rates and extents in organohalide-respiring *Dehalococcoides mccartyi*. *ISME J* **10**: 1092–1101.

Yancey PH, Clark ME, Hand SC, Bowlus RD, Somero GN. (1982). Living with water stress: Evolution of osmolyte systems. *Science (80-)* **217**: 1214–1222.

Yang M, Nightingale PD, Beale R, Liss PS, Blomquist B, Fairall C. (2013). Atmospheric deposition of methanol over the Atlantic Ocean. *Proc Natl Acad Sci U S A* **110**: 20034–20039.

Yao L, Garmash O, Bianchi F, Zheng J, Yan C, Kontkanen J, *et al.* (2018). Atmospheric new particle formation from sulfuric acid and amines in a Chinese megacity. *Science (80-)* **361**: 278–281.

Yarza P, Yilmaz P, Pruesse E, Glöckner FO, Ludwig W, Schleifer K-H, *et al.* (2014). Uniting the classification of cultured and uncultured bacteria and archaea using 16S rRNA gene sequences. *Nat Rev Microbiol* **12**: 635–645.

Yin X, Wu W, Maeke M, Richter-Heitmann T, Kulkarni AC, Oni OE, *et al.* (2019). CO2 conversion to methane and biomass in obligate methylotrophic methanogens in marine sediments. *ISME J* **13**: 2107–2119.

Yu NY, Wagner JR, Laird MR, Melli G, Rey S, Lo R, *et al.* (2010). PSORTb 3.0: Improved protein subcellular localization prediction with refined localization subcategories and predictive capabilities for all prokaryotes. *Bioinformatics* **26**: 1608–1615.

Yu Z, Smith GB. (2000). Inhibition of methanogenesis by C1- and C2-polychlorinated aliphatic hydrocarbons. *Environ Toxicol Chem* **19**: 2212–2217.

Zeisel SH. (2000). Choline: an essential nutrient for humans. *Nutrition* **16**: 669–671.

Zhan G, Hua D, Huang N, Wang Y, Li S, Zhou Z, *et al.* (2019). Anesthesia and surgery induce cognitive dysfunction in elderly male mice: The role of gut microbiota. *Aging (Albany NY)* **11**: 1778–1790.

Zhang H, Yohe T, Huang L, Entwistle S, Wu P, Yang Z, *et al.* (2018). DbCAN2: A meta server for automated carbohydrate-active enzyme annotation. *Nucleic Acids Res* **46**: W95–W101.

Zhang Y, Gladyshev VN. (2005). An algorithm for identification of bacterial selenocysteine insertion sequence elements and selenoprotein genes. *Bioinformatics* **21**: 2580–2589.

Zindel U, Freudenberg W, Rieth M, Andreesen JR, Schnell J, Widdel F. (1988). *Eubacterium acidaminophilum* sp. nov., a versatile amino acid-degrading anaerobe producing or utilizing H_2 or formate - Description and enzymatic studies. *Arch*

Microbiol **150**: 254–266.

Supplementary Information

List of Supplementary Data

Dataset S1 List of genes in strain DCMF, *Dehalobacterium formicoaceticum* strain DMC and '*Candidatus* Dichloromethanomonas elyunquensis' strain RM referred to in Chapter 2. Available at https://bit.ly/datasetS1.

Dataset S2 List of proteins expressed by strain DCMF during growth with DCM, choline, glycine betaine, and methanol, including putative functions, $\log_2$-normalised label free quantitative (LFQ) expression, and Z-score normalised $\log_2$ LFC values. Available at https://bit.ly/datasetS2.

Dataset S3 Nucleotide sequence fasta files for the five MAGs generated from the DFE metagenome. Available at https://bit.ly/datasetS3.

Dataset S4 List of genes, annotations, and protein expression data from the DSV, IGN, LEN, SPI, and SYN MAGs. Available at https://bit.ly/datasetS4.

List of Supplementary Figures

List of Supplementary Tables

Supplementary Figures

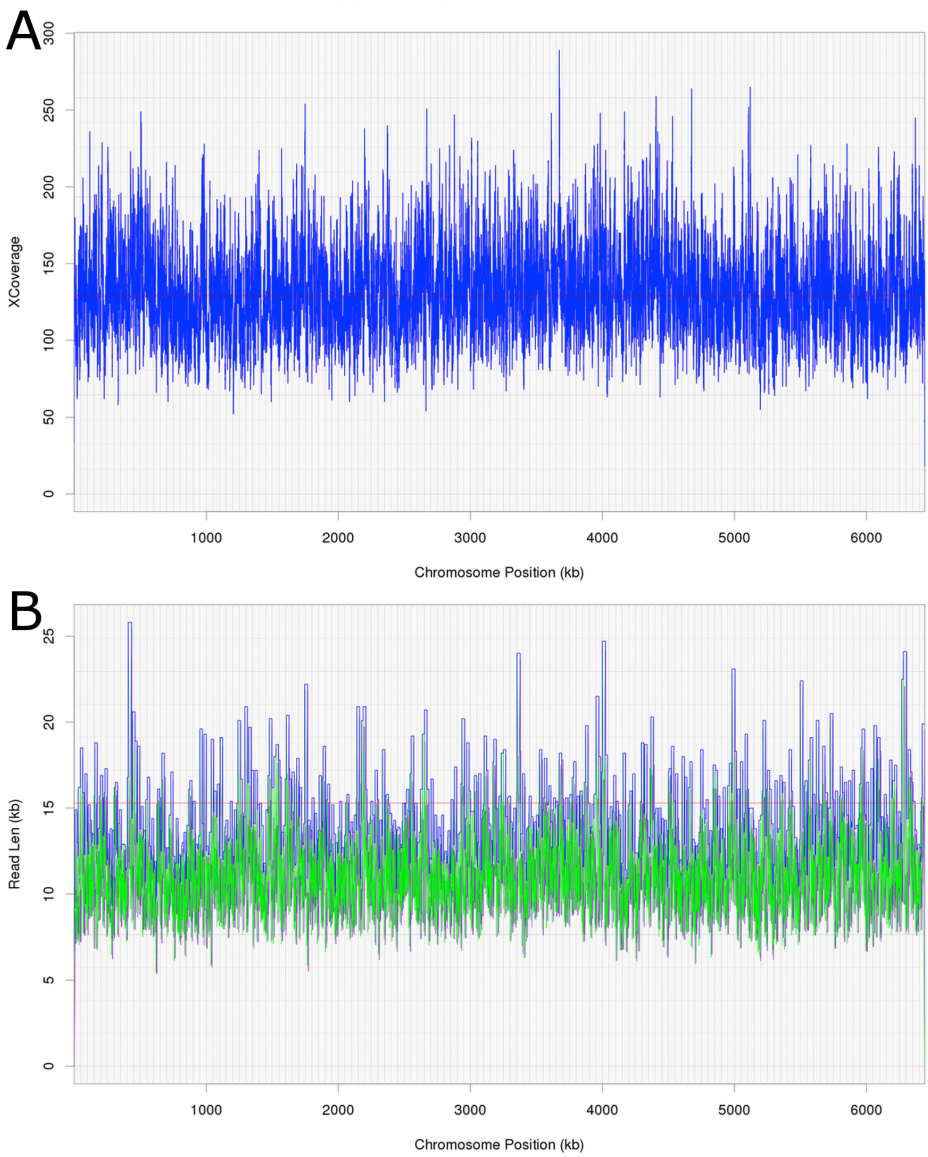


Figure S1 Average coverage depth and read length across the strain DCMF genome assembly. (A) PacBio read depth along the full strain DCMF chromosome. Horizontal lines mark median depth (132×), and gradations as 1/8 median depth. (B) The maximum length of a PacBio read (kb) spanning each base along the full strain DCMF chromosome. Horizontal lines mark median length (15.3 kb), and gradations as 1/8 median length. Colours indicate total read length (blue), longest 5' distance from base spanned by a single read (purple), and longest 3' distance from base spanned by a single read (green).

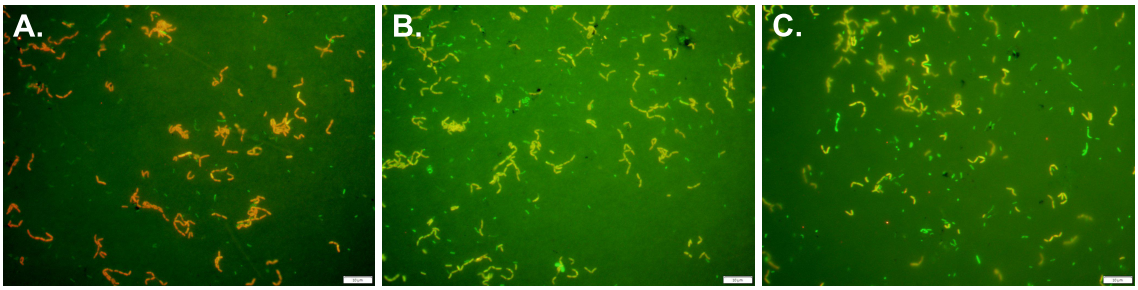


Figure S2 Overlaid Cy3- and 6-FAM-labelled fluorescence in situ hybridisation images used to count strain DCMF vs total bacterial cells in the DFE culture.

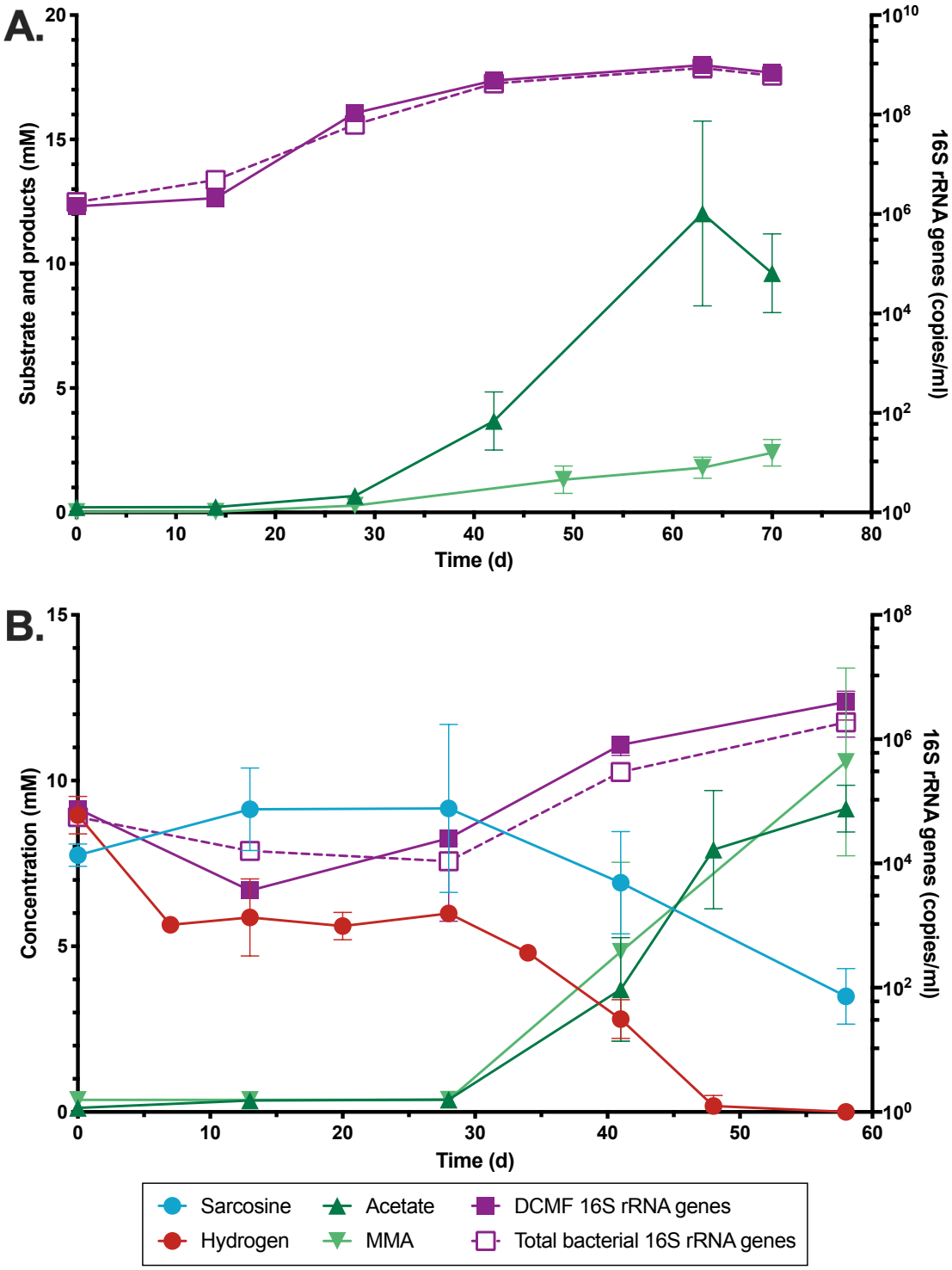


Figure S3 Strain DCMF is able to utilise dimethylglycine (A) and sarcosine (methylglycine) + H_2 (B) for growth.

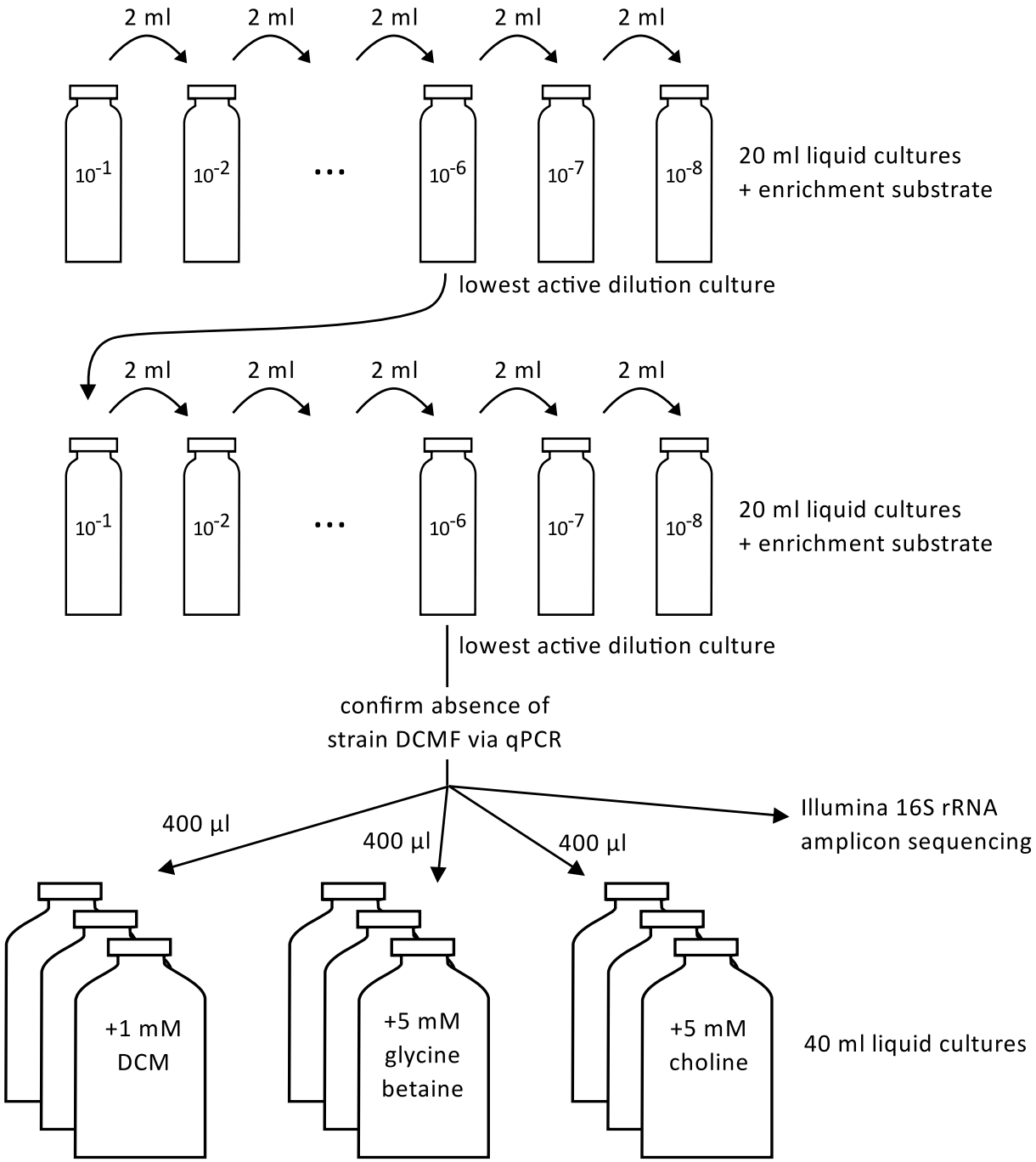


Figure S4 Overview of the method used to make the strain DCMF-free cohabitant enrichment cultures.

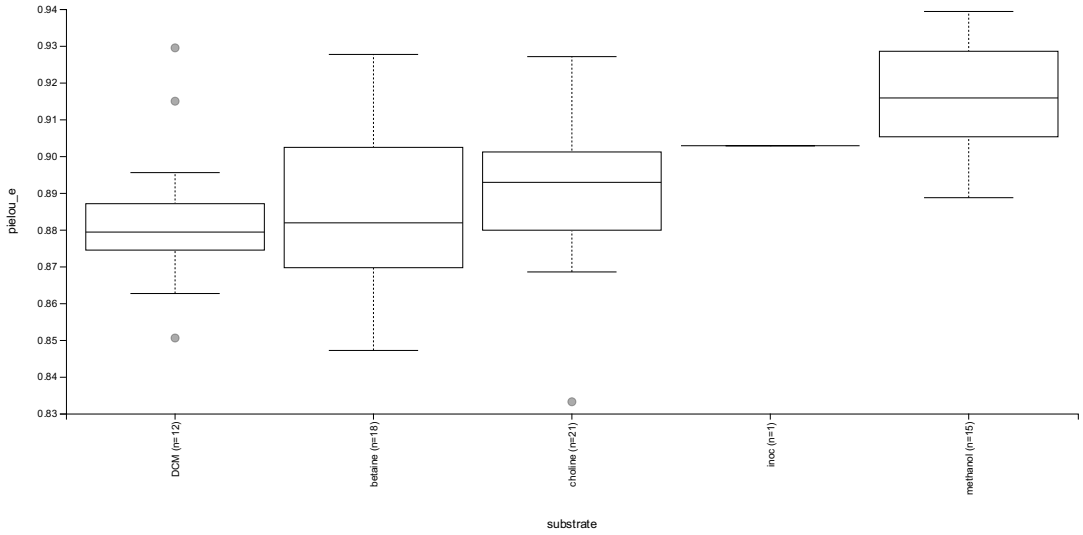


Figure S5 Box plot of Pielou's evenness index for the DFE community in cultures amended with DCM, glycine betaine, choline, or methanol.

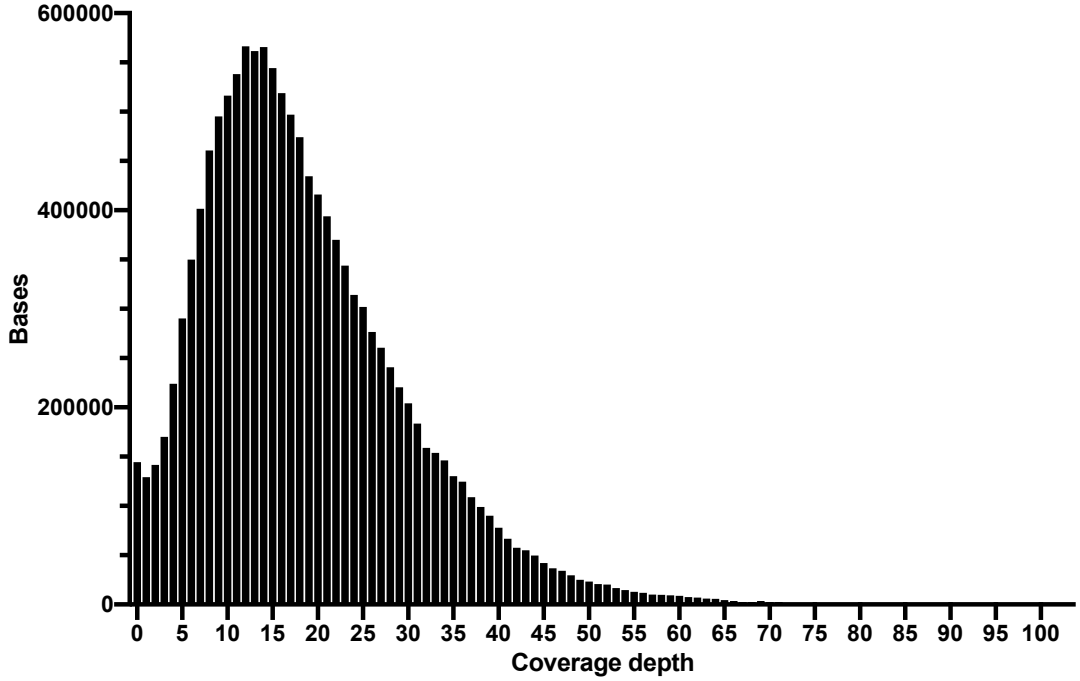


Figure S6 Coverage histogram for the DFE metagenome assembly.

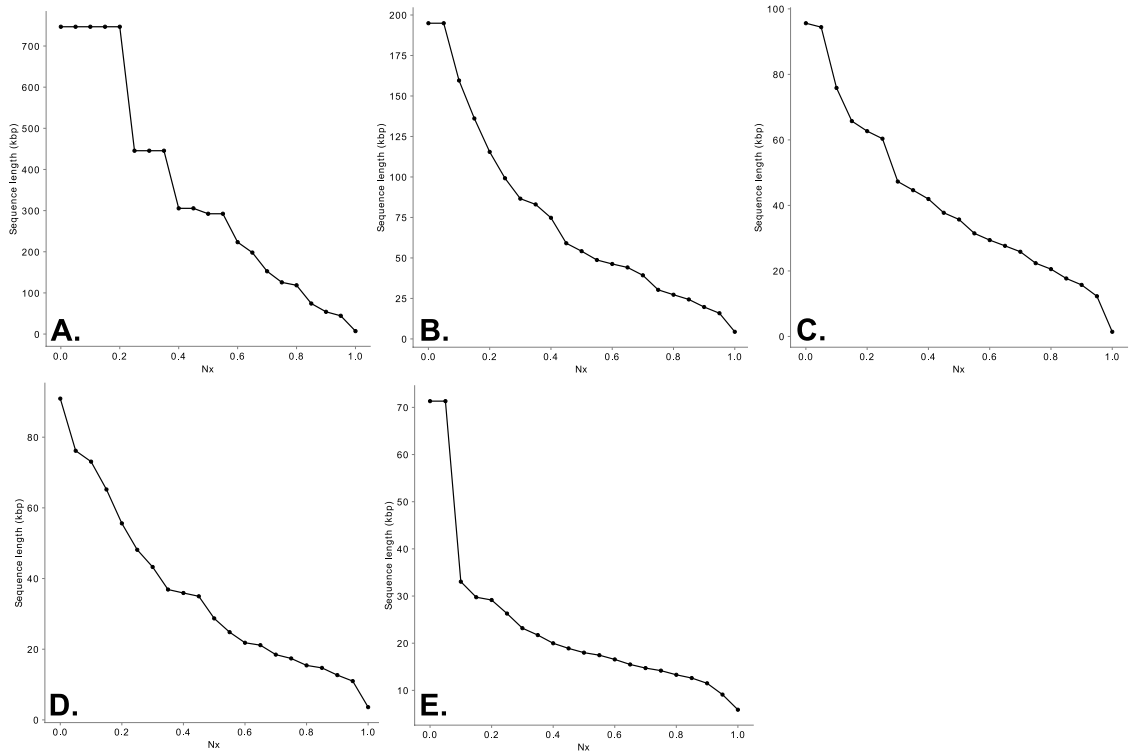


Figure S7 Nx plots for the five metagenome bins: bin 1/IGN (A), bin 2/SYN (B), bin 3/LEN (C), bin 4/SPI (D) and bin 5/DSV (E). Plots were generated by CheckM.

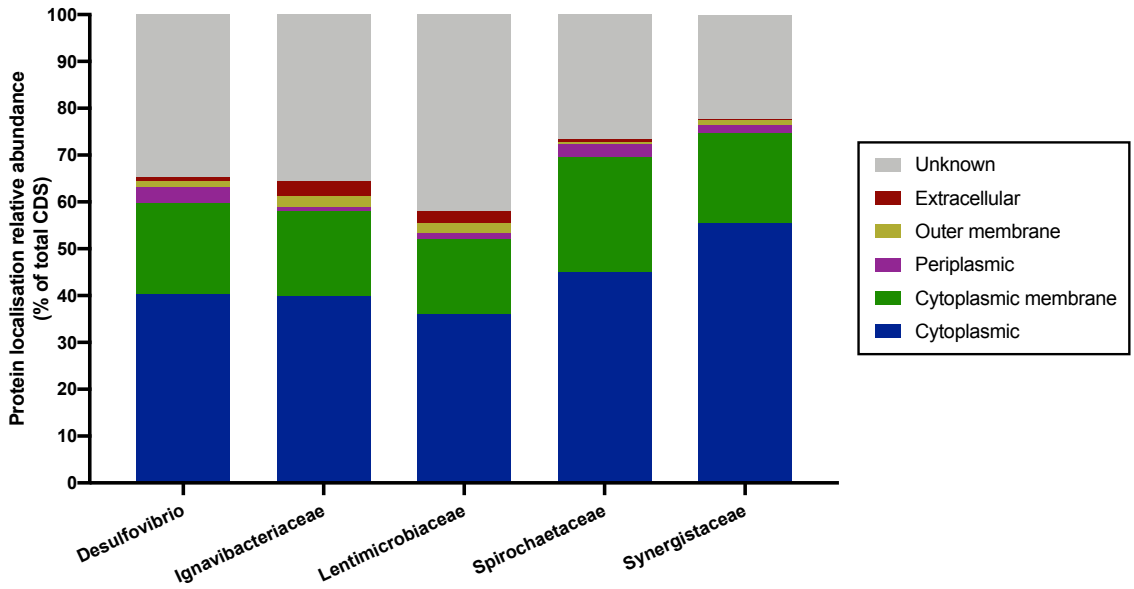


Figure S8 Predicted protein localization for the five MAGs assembled from the DFE metagenome. The unclassified *Ignavibacteriaceae* and unclassified *Lentimicrobiaceae* have a higher proportion of outer membrane and extracellular proteins than the other three taxa. Protein abundance in each location is normalized to the size of each MAG.

Supplementary Tables

Table S1 Assembly parameters tested on the strain DCMF genome, with relevant statistics.

Assembly ID	SMRT (N)	Assumed Genome Size (Mb)	Min read length (bp)	Seed Read cutoff	Min RQ	Min Correction Coverage Depth (×)	Contigs (N)	Total Assembly (bp)	Largest contig (bp)	Largest Contig coverage (×)	Circ Genome (kb)
WON710A1.SP16514.hcq	2	5	500	6,000	0.8	6	1,031	14,767,807	6,442,920	56.6	6,441
WON710A1.SP16515.hcq	2	5	4,287	8,300	0.86	6	971	12,147,540	3,133,739	35	3,134
WON710A1.SP16534.hcq	2	7	4,539	8,614	0.8	6	367	11,586,443	6,451,482	47.3	6,441
WON710A1.SP16535.hcq	2	7	2,368	8,065	0.84	6	219	10,403,067	6,451,597	48.6	6,441
WON710A1.SP16540.hcq	3	7	500	6,000	0.8	6	1,164	20,360,297	4,849,716	122.9	4,850
WON710A1.SP16543.hcq	3	7	4,010	8,003	0.86	6	1,307	16,089,549	6,450,325	49.2	6,446
WON710A1.SP16547.hcq	3	7	4,010	8,003	0.86	7	207	9,836,800	6,449,981	49.3	6,441
WON710A1.SP16546.hcq	3	7	4,010	8,003	0.86	8	213	9,307,776	6,449,974	49.3	6,441
WON710A1.SP16548.hcq	3	7	4,010	8,003	0.86	9	233	9,010,440	6,449,952	49.3	6,441
WON710A1.SP16550.hcq	3	7	4,010	8,003	0.86	10	276	8,801,861	6,449,940	49.3	6,441
WON710A1.SP16551.hcq	3	7	4,010	8,003	0.86	11	380	8,773,381	1,893,414	48.9	1,893

Table S2 Manually curated protein sequences. These genes were either fragmented or cut short in the original IMG annotation, due to the presence of a pyrrolysine or selenocysteine residue. All annotations have been publicly updated on IMG.

IMG Gene Locus	Product Name	DNA Coordinates	Replacing Genes	Amino Acid Sequence
Ga0180325_111271p72	dimethylamine: corrinoid methyltransferase	1309129 - 1310520	Ga0180325_111271 Ga0180325_111272	MGKFFSRMGDGAAVWLSEADIRADIEAGTLDAADRAKVPPLTGDEQKYLFELCTMPQQMVCV DRGKEIVTTGDSGATKLPYDGSIPMDRSTAVLVHERAFCVDTMEMGNIDYSYKQVKAILHEEKS ALETAQMHTVVPLLYGAMPNLGLYTKPDGPVDNWSELLPAGKIAEARAAQEEAVEHAVKDLV YCADGLYEAGADGINFDTVGASGDADVLATLNAVRILKEKHPDLPINLGMAGEFILGMHGQLYF DGVRLAGLYPHKQVKIAEQAGATIFGAVVNTNSSQSFPWNVARVCTLVKACTEAAHIPVHVNV GMGVGSVPLTLTPPVDAVSRADKAIIEIGKADGLOVGVGDPFGMEITHEMAAGMGGIRTAGDLV ARVQLSKAMKIDAAKQYVAEKLAISRAELADPIVMGELRADLDIGRVQPPDGAAIGIEAKFNIAR LLDIRINSVTKFMALARIK*
Ga0180325_111278p79	trimethylamine---corrinoid protein Co-methyltransferase	1315739 - 1317211	Ga0180325_111278 Ga0180325_111279	MRRNLLSGRNLWQGFGLQVFTDSELEMIHLATLEVLERTGIFVQSQEALSIFEKNGALVDRENR NVKIPAYLVDEAIRSAPQKIVLAGRDPQNDLVLESGRVNFCPFGEGIHVIDPYTGEHRKSTKKDIG DVALLVDALDQYDMLEYTVSPRDVHPKVAVYHTYEATVSNTTKHIPQSPEDKESAQVLIEMAA AVAGGKEKLRDRPIVSGAACPQSPLSLSEGCCEGIMEYARARLPMNVLSMAMAGGSSPVSLAGT LVTHNAEVLGGLVLAQLTAKGAPVIYGSSTTILDLRLASATVGCPELAMISAAVAKMAQFYQLPS YVAGGOGDSKVSDLQAGHEKTLTALLPALAGANMIYGLGMLEMGMTLSYGQIVADNEFAAMIR RVLEGIPVHDEDLAVDVIARVGARGNFLTEEHTLKHMRGYQSQPKLMDRNMREFWEAKGSTD MIARANEVARGILETYKPQPLADSVARTLRSLVEDAEKAHKIK*
Ga0180325_111485p86	trimethylamine---corrinoid protein Co-methyltransferase	1519657 - 1521168	Ga0180325_111485 Ga0180325_111486	LTVRGIHAGSFRKEGLGLNLFTEDDLYDIHLATLDVLWNVGVKVESEEAVEYFDGGGCQVDKKT HIVKIPAWLIEDALRSIPKTFRACGRNPENDWLNEGNRTGFVNFGEAVTMIDPYSYELRKPKKK DVDDVTRFCEAMDSVIVFERALGPSEVDPDVAQVHIAESFFNNCTKHAYIGMNRPENMRAVAK MAYLVCGGEDKFRERPIFSQSTDPVSPLVHSKSSTDTLIQAIRCGVPAKINPMGLAGGTTCVHLA GTLVTHNAEVLSMFVLAQLVKKGHPMVYGSSTAMMDLRTTLACVGSPELALFSAAVAKIAQFY LIPSWVAGGOTDSKVPDAQAAHEFTLTALLPALAGANMIYGLGMLEGGLTWDYAQLVMQNEM VKMILHCVKGIPVNDEKMAMEVIRSVGPGGEFISHEHTFRNFRLLSAPTLLDRHNRDGWKAAG SKDIVTKAYEKARDILENFQPTPLPDQMRDQIREIVKEAEAETAEIKAKEKEASRKGKL*

Ga0180325_111620p21	trimethylamine---corrinoid protein Co-methyltransferase	1664433 - 1665911	Ga0180325_111620 Ga0180325_111621	VKRNTNAGYASFSGASLQLFSEGDLHFIHLATLEVLERTGLNITDEEALEYFERGGAIVDHQKKV VKIPGYMVEEAIKSAPASVFLAGKDPKYDIILEEGRVYFCPFGLGIEIVDPYTGVLRETTKQDIADC ARIVDYLDEYDFCFDTMVARDVDPHVACIHGFEAPLTNTSKNVLASPENKQTAQILLEMGAAA AGGMDNLKERPIMMLGGCTISPLTIPESTVAATIEAAKARIPAMILSMAMSGGMGPVTLAGALV VMNAEILGALTLSQLVNRGTPFIYGSSTGTLDMRHNAAAMVGCPELGLISAGVAALCRMYNVPS LVAGGOTDSKVPDMQAGHEKTLTAILPILAGANMIYGPGMLDSGIIMSLGQVVADADFIRMFKT VLAGVPVNEETLAVDVIHSVGIKGQYLGEEHTFRLFKQLQSVPKIMDRNNREDWKSKGSKDMA ERAAEEAQRILEHHKPQVLPEEVITKIRQIVTNAEKVLATKK*
Ga0180325_111622p23	trimethylamine---corrinoid protein Co-methyltransferase	1665978 - 1667447	Ga0180325_111622 Ga0180325_111623	MKRNTAAGYNTWVGCGVQLFSDNDLEAIHSNTLEVMGKTGINVQSPKAMDIFEQGGAKVDRE KQNVKIPAYMVEEAVRTAPGKVLMAGRDPKNDVILEYGRVNFIPFGTGVMVVDENGEYRKSTK KDICQLATITDALDQMDFCFDTVIPRDVDQRTVCYHSFEGHINNTTKHVFTSPEDTHSAQVLIE MAGKVIGGKEKLKDRPIITGGGCPISPLSWSEGLCESMIEYAKAELPFLLVSMAMAGGTSPVTLA GTLVTQNAELLTGVVLSQLVQKGAPVIYGSSTTNLDLRKAMATVGSPELSLISSAVAKLAQFYNL PSFVAGGOADSKDSDLQAGHEKTLTYMLPMLAGANIIYGSGMLEMGMTFGYAQYVADNEMVR MLRRLLQGIPVNDESMAVDVIKQVGAGGHFLMEDHTMTHMKSAHVSPRLIDRTNRDAWIEQG KPNLIGKAKEEVLNILATQKPDPLPEKVASEIRSTIEQIEKELGIK*
Ga0180325_112547p48	monomethylamine: corrinoid methyltransferase	2702096 - 2703484	Ga0180325_112547 Ga0180325_112548	MALPKKVTVFDIYDRAKTGPKLEEKEWDTKVIPQTAAKLKQKYGIKMDKQVIVPTDQELIRHLF QAGLEMLVECGVYCMDTGRIIKYTEEEVLASLDAAPSKVMIGEGKDAVELACRSYHOGNETSLS AAGDPRPPIIQGGPTGAPCSEEHFLGIHQSYAQEPLVGTIVDGVLQTINGHDPVPGSPWEIAAVK SEAILVRAAQLRAGRSGMGLOGNETSLSAAGVIAADFPGGMRPSDSHEVSQLNELKIDVGALAF TAHYVLAGNIIMCEQMPIYGGYAGGLEETTIVDVATTINSFVMTQATWHLDGPVHVRWGITTA REALAVAGHCAMAIEANTHLMLGNQYYTMAGPCTVMCLLETAAQAITDTASGREIVSGVAAAK GVATNYTTGMEARMMAEAARAVAGMETEKVNEILNKLIALYEKEYKTAPKGKPYQECYDVVSL VPTQEYLDVYDEAVKILTGLGLTYWTK*

Ga0180325_112878s80	formate dehydrogenase	3045568 - 3048246	Ga0180325_112878 Ga0180325_112880	VDAVKLTIDGREVEVPAGTTILEAAEQIGIDIPRLCFDPELSLQGSCRLCVVEVEGAPLLSASCVTP VGRGMVVHTESPLVVETRRTILELLLANHPLDCLTCEKNGDCRLAEYCYRYGVKDSSFAGERHH YPIDDSNPFVLRDMNKCIQCGKCVRACAEIIGKDNIDYINRGFDRKVATYGDKPYVDSVCTFCGN CVAVCPTGALTEKPMQGKGRRWELDRVRTTCPFCGVGCNFDLCVKDGKLVGVLSNPDSPANG RALCVKGRFGWDFVNNEQRLTTPLIKREGKFEPASWEEAIDLISTRFTEIREKYGPKSFAALSSA RCTNEENFLMQKFTRAVMGTNNVDHCARTUHAPTVAGLATTFGSGAMTNSIAEISGAELLMLV GTNTTEAHPVIGYKMRQAKRRGAKLIVVDPRRIELAEEADYWLRLRSGTDIPLLNGLMHIIIKED LQDKNFIEERTENFAALKETVEKYTPEYVSRLTGIPVEDLYAVARLYAKTDKAMLFYTLGITEHV CGTSNVMSIANLAMLTGHLGRPHTGVNPIRGQNNVQGACDMGALPNVFSGYQRVVDPAARTK FETAWGVTLPAENGLMIPQMFEKANEGELKAMYILGENPVLTDPNTNEIRSGLEKLEFLVVQEL FLSETAQYADVVLPAASFAETDGTFTSTERRIQRVRKAINPLPGQANWQTIISISNAMGYPMNYT HPEEIWREMAALTPSMAGVSYPRLEEKGMQWPCPTADHPGTPYLHGKSFSRGLGLFQPSEHIP PVEIPDEQFPFLLSTGRILYHYNVTTPYSKGIQSMWPEEMAQVNPEDAARLGVGTGEKVKVISR RGEVTTRAQVTNKVPAGMIWMSFHYKESPTNVLTSHGLDPVTKTGEYKVCAVKIEKVG*
Ga0180325_113356p57	dimethylamine: corrinoid methyltransferase	3570932 - 3572323	Ga0180325_113356 Ga0180325_113357	MKKYFTRMGDGSAVWMSDEDIRWDLEEGMKDAADRGKIPELTDDEIEQLFEIITHPQKTVSCE RGNEAVVTFDAGTLKLPVRAGLPMDRTTTILTHERVFCSDTMELCATDYSYKALKNIVSEEAM AMERAQLNCIIPIFYGAMPNLGLYTKPDGPIDNWSELLPLAKIPEARAAQEEAVEHAVRDMVFI ASALYESGADGINFDTIGASGDGDYLAALKAAEILKDKYPGIPIEMGMAGEFVLGMHGQLKYDG VRLAGLYPHQQVKVCEKAGATIFGCVINTNSSMSFAWNLARTVTFAKACVEAADIPVHVNAGM GVGGIPLTNTSPTDATSRASKAIIEIAKADGLOIGHGDAYGMAVTHEVATGMGGIRTAGDLVAR MQMLKGMKLKDAKEYVAGKLGVSVFDLSDSNMMKEIREQLDIGHVLDRTKCANGMEAKFNIA RVLDIEINSVNRFKSKIHLV*
Ga0180325_113361p62	trimethylamine--- corrinoid protein Co-methyltransferase	3576747 - 3578225	Ga0180325_113361 Ga0180325_113362	VARNLHAGFNRMDGFALNTFSQDELYAIHCATLDVLHHVGVRVDSPEARDLFDGGGAFVDPKT NIVKIPPYLVEDAIRSTPGTLILAGRDPNKDYCMEANRVGFVNFGEGVNVIDPVTRKYRTTTKAD VANAARMSDYLSEMDISYRAVVAQDQPGHVQSLHNAEAIFPNTTKHFFIGADGVKNARKLIKM AQAVAGGKDQLRERPLITFNVCPTSLLKLIPECTDVVIEAARAGIAINIISMAMAGATSPVTLAGT LVTHNAEVLSTIVLNQLACKGAPCIYGSSTTIMDMRYTTAPVGAPELGMISASVAKMAQYYLLPS FVAGGOADSKIPDAQAAHEKTLTALLAAQGGANLIYGAGMLELGITFDFAQFVMDNEMYKMIR KAVGGISVTDANMAVDIIKEIGPGGEFISHAHTFENFKREQSQSKLIDRTMRETWLLNGGKDFT ERAYEEANHILSTHQVAPLAPGVEATIRSIVEEAEEEYGIKKK*

Ga0180325_114321p22	dimethylamine: corrinoid methyltransferase	4654504 - 4655895	Ga0180325_114321 Ga0180325_114322	LSKFFTRMGDGSGIYLSADDIRWDLEEGTKDAADRGKIPELTKEELDHLYDIITMPGIVVGVERG NEVVTTSDSGGCKITYHANIAIDRGTAVLIHEKVLGADSLDIGHIDYSYKAVKSVLHDEAAVMEL TQLNAVMPVLYGSMPNLGLYTKPDGPVDNWSELLPLGKIAEARAAQEEAVEHAVKDIVYVAGG MYESGADGMNIDTCGASGDADVLAALKAIELIKQKYPDLGIEMGMAGEFVIGMHGKLEYDGVR LAGLYPHKQVKLAEKAGANIFGCVVNTNSNMSLAWNLARTVTFVKACTEAATIPVHVNVGMG VGAVPMSEMPSVDAVSRVDKALVEIGKADGLOIGVGDPFGLEIAHEVACGMGGIRTAGDLVLRM QLSKGMKINDAKKYVAEKLGVSPVDLSDCAIMREIREDLNLGRPMPPDHVAKGMEAKIRIAKAL DIKINSVERFMRMAGLK*
Ga0180325_114324p25	trimethylamine--- corrinoid protein Co-methyltransferase	4656612 - 4658090	Ga0180325_114324 Ga0180325_114325	LSRRNLYAGNSMQEGFGLNVFSQDALDRIHNATLEILWYEGIKIQSEEALEIFHGGGCVVDKKN QKVYIPPHVVEDCIRSTPSTVLLAGRDPKNDIVLDGSRVNFCNFSKGVNVVDPYTGAVRPSTKQ DEANVAILVDALEQYDLLDVAVEARDIDARTANLECYEAMVSNSTKHSTQSPHSFEEAQTLIDM AAAVAGGKDKLRERPITSSTVCPTSPLSIAPETCEPIITYARNRVPLTVLSMAMAGGTSPVTLAGT LVTHNAEVLSGIVLAQLTNKGTPNMYGSSTTIMDLRLASAAIGCPELGMISAAVCKLAQYYNIPS YVAGTOTDSKIGDEQAGHEKTLTMMLAALAGANMIYGLGMVDLGMTLDFGQLVVDNEIAKMV RRTLMGIPVNEETLAVDVIRKVGTGGHFLMEEHTLAHMRTDQSQSKLFDRNTRQNWQAKGAK DLATRATEEARFILENHKPQPLPSSTAQTLREMIEEAERRWGVKK*
Ga0180325_114454s56	glycine/betaine/ sarcosine/D-proline reductase complex selenoprotein B	4793312 - 4794625	Ga0180325_114454 Ga0180325_114455	VSKIKVVHYLNQFFGQMGGEEKADIPPQLREGAVGPGMALNGAMGEGGEIVATVICGDSFFNEN MDAAGGDVVAMIKKYHPDVVIAGPAFNAGRYGVACGAVAKAVSEKLQVPVVSGMYPENPGVD LYRRYAFIVETGNTAAGMRKAIPSMAKLALKLGQGSPVGSPLEEGYIPRGLRVNFFAAERGAKR AVDMLVKKIKGENFETEYPMPVFDRVAPGPPVKDLAKATIALVTSGGIVPKGNPDRIEASSASKY GKYSLAGVRDLTAEAYETAHGGYDPVYANEDADRVLPVDVLRELEDQSVIGKLHEYWYATVGN GTSVANSKKYAAAIAQDLQAAGVDAVILTSTUGTCTRCGATMVKEIESTGVPVVHMCTVAPISL TVGANRIVPTVAIPYPLGNPSLSKEEEKSLRRKLVKKALHALTVPVEGQTVFDQE*
Ga0180325_114673s74	glycine/betaine/ sarcosine reductase complex selenoprotein A	5000370. - 5000816	Ga0180325_114673 Ga0180325_114674	MLSGKKVAVLGDQNGISGQAIEACVKAAGGEVVFSSTECFLUITGGAMDLRNQARVKESAEKYG NRDLIVILGGAEADVCGIAVETVSTGDPSYAGPLSEIPLGLKAYHIFELKDEIPPEVYEEHIGFMET VFPVEDIIKECRAYRSP*

Ga0180325_114675s76	glycine/betaine/ sarcosine/D-proline reductase complex selenoprotein B	5000876 - 5002186	Ga0180325_114675 Ga0180325_114676	MARYQNPVRIVHYLNQFFGGMGGEEMADLKPQVKNEPIGPSKKIQELLGTDYKIVATIMAGDN YFVEHKDLALQKVMAFVKKMNPDIFIAGPCLCMGRYGLACAWICSRVSRELGIPTVLSMSPHNP AFSILPPDIYFVPAGETIRDMKNDVRQFSKMIDKLRQGQRGIIDTDGAFLTGKRGNTLVSANGAT RAVQMLVRILHDEEAGSEIPIAKNNGIRSAPPIADLSTARIALVTEGGIVPLGNPNRLRSARETR WFRYDLPLVEGEMARFQCIHGGFDTNYATIDFNRILPVDTLIGMEKSGGIGRLVPWYFVTCGNL VSVKDSKQMGSEIVRFLIQDQVDGVFLTSTUGTGTRCGATMAKVFESAGIPTVVLTPLVDLAMQ FGAYRIVRGLTVTSPLGDPEVSAEKEKTGRQALVVKALTALTKHITVPMSY*
Ga0180325_114680s81	glycine/betaine/ sarcosine reductase complex selenoprotein A	5005448 - 5005891	Ga0180325_114680 Ga0180325_114681	MLQGKKVAILGDRDGIPGPAIEECVKSAGAEVVFSTTECFVUTAAGAMDLENQARIKDLTEKYG NEDLIVVLGGAEAEASGLAAETVSTGDPTFAGPLSGVSLGLKAYHMFELKEEVDPKVYDEQISM MEMVLDVDEIINEVKTYRG*
Ga0180325_114685s86	glycine/betaine/ sarcosine/D-proline reductase complex selenoprotein B	5009981 - 5011291	Ga0180325_114685 Ga0180325_114686	VENFRVVHYLNQFFAGIGGEDQADMSPAIREGCVGPGLSLDRLWQGQGKIVATVICGDNYFSTN PEATVQELLPLIEAHHPDVVIAGPAFASGRYGLSCGYLCRAVVSKLRIPCVTSMEPENPGVQEAK GLAYIISASDRVARMGRVLPKVARLAAKLACGQSPGAAQEEGYLPRGYRLNTLRTESGAVRAVN MLLAKMNKHSFMTEIPLPQYDQVTAAAPLETGGKIRLALVTESGLVPEGNPDHLESSRATKWL KYPIPGEKYVANKHYSLHGGYDVRFVNADPNRVVPLDAVATLVREGVIGEVCQHYYVTTGMAA PIARATKFGSEIAHDLLQEGVNMVIVTSTUGTGTRCGATLAKEIEREGIPAVIITALPDVARNIGT PRIVQGVAITNPTGDPKKSAEGEVQLRRDLVLRCIRAAHTPIENPTLFQNE*
Ga0180325_114726p27	dimethylamine: corrinoid methyltransferase	5062099 - 5063493	Ga0180325_114726 Ga0180325_114727	LDKFFTRMGDGSAVWMSEEDIRWDLEEGMKDAADRGKIPELTDDEMQQLFEIVTHPLKNVSC ARGNEAVVTFDAGTLKLPVRAGLPMDRMTTILTHERALCSDTMELCTTDYSYKSIKNFVHEEA MSMELAQLNSIIPLFYGAMPNLGQYTKPDGPIDNWSELLPMAKIAEARAAQEEAIEHAVRDMV FIASALYEEGGADGINFDTVGASGDGDVLATLKAAQILKGKYPEIPIEIGMSGEFVLGMHGQLAF DGTRLAGLYPHQQVKVCEKAGATIFGCVINTNSSMSFAWNLARTVTFTKACVEAASIPVHVNAG MGVGGVPLTNTSPTDATSRASKAIIEIGKADGLOIGHGDAYGMAVMHEVTTGMGGIRTAGDLV ARMQMLKGMKLKDAKEYVAGKLGVDLFDLSDSYVMKELREQLDIGTVLDRANASFGMEAKFN IARVLDIPINSVNKFKRKVGMV*

Ga0180325_114729p30	trimethylamine---corrinoid protein Co-methyltransferase	5064364 - 5065839	Ga0180325_114729 Ga0180325_114730	VSRNLHAGFHRIDGFGINVFSRDELYAVHCATLDVLQHVGVRVDSKEAQEIFAGGGATVDPQTD IVKIPPYMVEDALRWAPGTLLLAARDPKKDYILESNRVGFVNFGEGVNIIDPVTRKYRPTNKRD VANAALMSDYLSEMDISYRAVVAQDQPGHVQSLHNAEAIFPNTAKHFFIGADGVKNARKLIKM GAALAGGKDKLKERPLISFNVCPTSLLKLIPECTDVVMEAARAGIPVNIISMAMAGATSPVTLAG TLVTHNAEVLATIVLSQFTCKGAPCIYGSSTTIMDMRYTTAPVGAPELGMISAAVAKMAQYYLL PSFVAGGOADSKVPDAQAAHEKTLTGLLAAQGGANLIYGAGMLELGITFDFAQFVMDNEIFKMI RKAVGGIRVTDADMAVDIIKEIGPGGEFISHAHTFENFRKEQSQSRLIDRTMRDTWLLSGGRDL TERAYEEANHILNTHKVAPLAPGTEGTIRSIVEEAEEEYGIKR*
Ga0180325_114797s98	glycine/betaine/sarcosine reductase complex selenoprotein A	5139489 - 5139932	Ga0180325_114797 Ga0180325_114798	MLEGKKVAILGDRDGIPGPAIEECVKSAGAEVVFSTTECFVUTAAGAMDLENQARIKDLTDKYG NEDLIVVLGGAEAEASGLAAETVSTGDPTFAGPLSGVSLGLKAYHMFELKDEVDPTVYDEQISM MEMVLDVDDIINEVKTYRG*
Ga0180325_114803s	sarcosine reductase complex selenoprotein B	5144611 - 5145921	Ga0180325_114803	MSDKFKVLYYVNQFFGQVGGEDKAGMAPEFRPEKVGPALGFEGLLNKEGEVVGTIICGDNFFNE NKEDALNIILNTVKEAAPDVFVAGPAFNAGRYGVACAEISKAVAERLNIPVVTGMYVENPGLDIC KEIAYVVSTSDSAGGMRKALPAMAAITSKLAKGIEVGSPEEEGYIARGMRKTLFAEKRGSQRAVE MLLARLKGQPFQTEMPMPVFDVVPPAPAIKDLRKATIALCTSGGIVPEGNPDHIQSASAQKWGK YQVGSRDALNAPDFYTTHGGYDPVYANEIPDRVAPLDILKEFEKEGYIGKVYDWFCTTTGTGTA VSKAREFGTEIGAQLKEAGVDGVILTSTUGTCTRCGATMVKEIERYGIPIVHMATITTISESVGAN RIVPTVAIPHPVGNPKLNAEDEHALRRTLVKKALDALATEVTEPTHFE*
Ga0180325_114855s	glycine/betaine/sarcosine/D-proline reductase complex selenoprotein B; pseudogene	5196481 - 5197803	Ga0180325_114855	LPKVRIVYYINQFFGGIGGEEHAGHLFEVKNGPVGPGALLEKLLGNECGIEKTIVCGDNEFNEHE GENIKKILEVIKQVNPDIFVAGPAFTSGRYGLACMKACVAVAENFGIPCVTGIHPENPGAGLQEK QQHVYAVPVGKSAATMKSAIQLFSDLISKIVLNGENGLSVGDGFLSRGIRKNVKVADGAPKRACE MLLNKIKSLPFQSEIAAEQYEVIVPPQPIKDLSKAKIALVTEAALVPEGNPDGIQAARADKWAKY SIKKTDDFVEGSFHSVHGGYDSKWVDKDPDRVLPLDALRYYAQNSEVGQVCEYIYVTCGSMGH VLTMEKIGKEIADQLLKEQVDGVILTATUGTGTRSGATVAKQIEKAGIPAVTCTGLPDVALRVGA NRVYRTEGHFHQPFGDPNKSQEEEMIWRRCQVKKALDSLTQSVDKPTLITFN*
Ga0180325_114860s61	glycine/betaine/sarcosine reductase complex selenoprotein A	5202556 - 5202999	Ga0180325_114860 Ga0180325_114861	MLQGKKVAILGDRDGIPGPAIEECVKSAGAEVVFSTTECFVUTAAGAMDLENQARIKDLTEKYG NEDLIVVLGGAEAEASGLAAETVSTGDPTFAGPLSGVSLGLKAYHMFELKEEVDPKVYDEQISM MEMVLEVDEIINEVKTYRG*

Ga0180325_114888p89	trimethylamine---corrinoid protein Co-methyltransferase	5232064 - 5233548	Ga0180325_114888 Ga0180325_114889	MRTNSEAAVNLTRGFGLKMFSEDELYAIHLATLQVLERTGIKVESEEALEIFDAGGAKVNKQTH LVKFPAYIVEDAIRSAPSKVVLNARNPVHNVILEGKRVHFTNFGEGIMVIDPFTGAYRKSTKQDC ANAALICDALDEVDVILRAVAAHDVFVPTHALHEMEVCFHNTSKPVFNGGVNARLAEYLFQMG AAVCGGMDKFKERPILSLNVCPTSPLQLTSHCTDAIIKCAEYGVPVNILSMAMAGGSSPVTLGGT LVTHNAEVLSGVILNQLTRKGAPVIYGSSTTMMDLKTTTAPVGAPELGMINAAVAALAQYYLLP SWVAGGOVDSKIADAQAAHETTMSTLLTGLAGANLIYGVGMLELGITFSFEQMVMDNDIIKMV RKVLKGIEISDETLAVDVIDQVGAGGDFLSQEHTIKYMRTEQSRPKILDRQMRYAWKDKGSKDL TAVAHEEAVSLLQNHKPEPLTESVQAELQSIIAEAEAEFAAQMKKK*
Ga0180325_114961p62	trimethylamine---corrinoid protein Co-methyltransferase	5318237 - 5319715	Ga0180325_114961 Ga0180325_114962	VKRNFQVGISHISGFSLNALTADELYAVHCAILEVLQDSGLKVDSREAQAIFEGGGCKVNPKTNIV KIPADVVEDAICSAPSTFLLAGRNPKNDLVVGGKSSAFFNFGEAIYLYDPFTLAYRKSTKEDVGN AALICDAMEEMEICNRGMGADEYPGPIQSIHNADAIFSHTSKHCFIGPNSGYNYKKVVEMAAAI QGGMEALQERPIYSATVCPTSPLQLLPEMSDVVIEAARYGLPVNIVSMAMAGASSPVTLAGTLV TQGAEVLGGIVLNQLTRKGSPVTFGAVSTIMDMRLATAPVGAPELGMLSAAMANLAQYYKIPSF GAGGOSDSKVPDAQAAQEKALSLLTSSMMGVNLVLGIGMLEGGLTYDFAQIVMDCELIRMVRRI VRGITVDDESLAVDVIKQVGAAGEFMSHEHTFRHFRNEHSQSKIADRTIRANWLAKGSKEMVE RAYEEARYILKNHQPDPLPSGVDSAIGRIMEEANEHYGIKQN*
Ga0180325_115204p05	trimethylamine---corrinoid protein Co-methyltransferase	5606151 - 5607593	Ga0180325_115204 Ga0180325_115205	MYQEEVLGYSILREQDLAAIHQGTLDVLAETGLKVFSEKAREIYSGAGCIVDDQNMIVKISPHIVN DAIDSAPGRILLAARDPKHDIILEGKKVVFKNFATGVKVLDPDTLDYRPSTKADLGNIARFCDSL EEVDFFTLAVSAQDVHPKVRDLHEGEVVLNNTAKHFSHDTQSIKSTKRFLEMAATIAGGMDQL RERPIVSLGTCPVSPLVLNAECTDLIIEAAQAGIPMNVLSMGLAGGTTPVTMAGTLVVTNAEVLG GIILAQLVNKGTPMMYGTSNTIMDLIYTTSPVGAPEHAMFSAAVGQLGHYYNIPTDVGGTOCDS KISDIQAGHEKTLTALLPALTGSNILYGMGLLESGIAFCYAQMLVDREFVRMVKKVMQGIAVHK DTLALEVIKAVGAGGNYLMEEHTLKYMRQEQSRAKLIDRRTRKGWEETGGQDMITRARSEAR QILAGYRPMPLDPKVAARLRQIVQEAEDELK*

Ga0180325_115207p08	dimethylamine: corrinoid methyltransferase	5609138 - 5610514	Ga0180325_115207 Ga0180325_115208	MVFTRLGDGTMIEAEISEIRADLEAGTQDAAARAEIAPLSADDLARLCDIVCRPGKVVGVEKGNE IIMTGDSPSIGSVPRGFPVNRIQMLQTYERVCGMDTAEAGFIDYSYKAIKTVASEERSWVEQASH ILTIPLFYGAMPDLGRYSRPDGPVPNWAELLPQGKIAQARAAQEEAIELCVEDLVYVASEMYEG GAQGIDFDTSGAAGDADFYAALKATEILKKKYPEMCIEMGMAGEFVLGMHSELRYDGVRLAGL YPHDQVKLAQKAGVTIFGPVVNNVCNKSFSYNLARAVTFCKACADASGIPVHPNVGLGVGGVTT VEVLPVDVVSRVSVAMAEVAKADGLOVGSGDFCGMVATHTLASGMGGLRTAGDLVARMQMT RGMRIGEAKKYVADQLHVSVKELSDEIIMNEVRDDLNIGRVVSPHGKARGIQAQMNIGKILGIEI NSVKRFLEKID*
Ga0180325_115252s	betaine reductase complex selenoprotein B	5663441 - 5664736	Ga0180325_115252	MKKALLYVNQFFGGVGGEHHTDFEPIIKKGPIGPGLALKGALKGAEITHTVICGDNFMASNQEE ALKRIEGFLSGKEFNLFLAGPAFRAGRYGVNCGEMCKFMYEKYGVMGVTSMHEENPGVEMYRE NPFYILKGSEGAAKMRQDIAAMAAFANKLIAGEEILWASAEGYFPRGIRKEVFVDKTSADRSVD MLLAKLNGQPFETEFKIEVCDIVKPAKAVDIKKAKIGFISTGGLVPKGNPDHLPSGTSTIFIRYDIS GMDSLKPGSYECIHGGSMPDKINANPEVLFPLATLKQLEKEGQIGEVDSYFYSTTGNLTSMKNA TRIGAGIAESFKDNNIDAAILTSTUGSCTRCGATIVKEIERAGIPVAHICNLTAVAQITGSNRVIAG PCLNSPCCDVNLPEEQQKQQLRMIVTSALRALSTDIKKQTIF*
Ga0180325_115257s	glycine/betaine/ sarcosine reductase complex selenoprotein A	5667632 - 5668076	Ga0180325_115257	MLQGKKVAILGDRDGIPGPAIEECVKSAGAEVVFSTTECFVUTAAGAMDLENQARIKDLAEKYG NEDLIVVLGGAEAEASGLAAETVSTGDPTFAGPLSGVSLGLKAYHMFELKEEVDPGVYDEQISM MEMVLEVDEIINEVKTYRG*
Ga0180325_115329p30	trimethylamine---corrinoid protein Co-methyltransferase	5755164 - 5756642	Ga0180325_115329 Ga0180325_115330	VVKRRISSGQAMLGGFGINIFTENDFMAIHYGTLEVLEKTGVFVDNPEAIDLYESGGARVDRSNK KVKIPASLVDECLHSAPKKVLLAGRDAKNDILLEGTRVHFCSFGIGLNVYDPFTGAYRKSTKKDV GDVARLCDYLEDIDMLECTLTPNDVHPNVYNLHILEANLRNTTKPCLSDPDPGLFPWILEMASA VAGGEDKLRERPIISGIVCPQSPMTFHHSCCEGIMQYARHELPMIVLPMAMAGGTSPVTLAGTVI SHNVEVLAGLVLAQIVHKGAPIIYGSSTTMLDLKTATATVGCPELAMLNAALAKMAQFYLLPSW VAGGOADSKILDAQAGQEKMLTALLPALAGANMVFGSGMLESGIALSFGSLVADNENARMIRR VLQGIPVNDITMAVDVIKEMGTNGLYLVNEHTLEHFRAHQSQPVVIDRRIRQRWLDDGARDYA FRAEEYARNILQNHQPAPLPDAVSEKVNAIVEDAEKRLIPKKK*

Ga0180325_115331p32	dimethylamine: corrinoid methyltransferase	5756725 - 5758116	Ga0180325_115331 Ga0180325_115332	MEKIMTFMGDGSRIFMNAAEIRADLESGTADAADRGKIPELTEDEINHLFDIITAKSKVVGVENG KEIVTTTDSGIKIPMQGFVPVDRSTNAQIHERVLCSDTFELSHIDYSFKQIKTILPTEQSVLEYTQ MNMTAPVLYGAMPNLGLYTVPDGPVANWNELLPQGEIQAALAAQEEAVEYAVKDMVFVSGG MFDAGADGIDFDTVGASGDADFLAALKATQILREKYPDYAVEMGMAGEFVLGIHGDLYFGPDR LAGLYPHQQVRVAEKAGVTIFGPVVNTNSLRSFPWNIARVCTFIKACAEAARIPIHPNVGMGVG GIPLTLKVPVELATRAVKAIVEIGKADGLOTGEGDPFSTEAAFELAAGMGGIRTSGDLVARVQLA KALKIDEAKKYVAEKLKLSLRDLTDPVLMEEVRLDLGLGRVQPYPENALGMEAKFNIAELLNIKI KSVENFKKKSRIQL*
Ga0180325_115773p74	trimethylamine---corrinoid protein Co-methyltransferase	6243806 - 6245314	Ga0180325_115773 Ga0180325_115774	LTVRGVHAGSHRMEGIGLNMFTYDELYDIHLATLDVLWNIGCKVESEEAVEIFDGAGCTIDRAN HNVKIPAYLVEDALRSIPKTFRACGRNPENDWLNENNRTGFVNFGEAVQMIDPYTKEIRKPKK KDVDDVTRFCEAMDQIIIFERALGPSEVDPDVSQVHIAESFYNNCTKHAYIGMNRPENMRAVAK MGYVVSGGEDQFRQRPLFSQSTDPVSPLVHSKGATDTLIQAIRCGVPAKINSMGLCGGTTCVNL ASTLVTHNAEMLSMFVLAQLVKKGHPLVYGTSTAIMDLRTTTACVGSPELALFSAACAKMAQF YNIPVWVAGG OTDSKVPDAQSAFEFCMTALLPALAGANMIYGLGLLEGGLTWDYAMLVMQNEMVKQILHCVK GIPVNDEEMALEVIKSVGPGGEFISHEHTFQNFRRLSAPVLLDRHNRDGWKAAGSKDIVQKSYE KAHEILENYKPTPLPENIQKQLKDIVAEAEAETTEIKAKEKEAMRKGK*
Ga0180325_115799p00	dimethylamine: corrinoid methyltransferase	6269722 - 6271107	Ga0180325_115799 Ga0180325_115800	MKKYLTRYGDNYADFISGDQIKADIESGSQDAAERAHIPPLTQDEMDYLYEIIISPQKIVSVEPGN EVVVSFDAGTLKLPVRNGIPMDRLQAILTQERALASDSIELCHVDYSYKPVKAIISEERQTMEQA QLMTTLPLLYGAMPNLGLYTRPDGPVGNWSELLPAGKIAEAREAQEEAIQHAIRDIVFVASQMI ESGADGINLDTVGASGDADVKASLEAVKILKQKYPQIGFETGMAGEFNMGMHGLLEIDGERIAG LYPHKQVKMAEKCGVSIFGMAVNTNSNMSLAWNLARVCTFAKETVKVSTIPVHANVGMGVGG IPLSLIPAADAVSRVDKALIEIAKVDGLOVGAGDTFGMAVTHEATAGMGGIRSAGDLVFRMELA GMKINEAKTFVAEKLQVGVKDLADCAVMKDVREDLNMGTTQSRPNAGVGIQTKFRIAEIMGISI NSVEKFKKEAGI*

Ga0180325_115804p05	trimethylamine---corrinoid protein Co-methyltransferase	6275841 - 6277319	Ga0180325_115804 Ga0180325_115805	VGRNLKAGYHRQDGFGLNMFSDDELYAIHCATLDVLKNSGIRILSKEAQDIYDGGGAIVDRKNNI VKIPPYMVEDAIQSAPSTLLLAGRNPKNDIVLEANRTGFTNFGEGIMIIDPYTKAYRRTTKKDVG DVARVCDALDAIDVHERAVSAQDVPAAVAPLHEVEINLTNTSKHLFQGCGGAKNLRKVVEMAA AVVGGKDKLRERPIYSCITCPVSPLQLVPESTEVIIECARLGVPINILSMALAGGTSAVTLAGTLVT HNAEVLGGIVLNQLTSKGAPVIYGSSTTMMDLKYTTSPVGCPELGMINAAVAKLAQYYLLPSWV AGGOADSKVPDAQAAHEKTITAILPALAGANLIYGLGMLELGMTMDYAQLVMDNEIARMIKQA VGGIDVTDEDLAVDVIKQVGAAGEFVSHEHTFHHFRRVQSTTRLIDRRMREAWLADGAKDFTQ RAYEQAIDILENYKPDPLPAGAAETFRAIIEEAEKEYGVKKK*

Table S3 Genes involved in the quaternary amine metabolic model presented in Figure 3.9.

Gene	Protein Product	JGI Gene Locus
Wood-Ljungdahl pathway, carbonyl branch		
	Acetyl-CoA synthase corrinoid activation protein (Ferredoxin)	Ga0180325_111202
acsE	5-methyltetrahydrofolate:corrinoid/iron sulphur protein co-methyltransferase	Ga0180325_111199
acsD	acetyl-CoA synthase corrinoid iron-sulphur protein, small subunit; CO dehydrogenase/acetyl-CoA synthase delta subunit	Ga0180325_111200
acsC	acetyl-CoA synthase corrinoid iron-sulphur protein, large subunit; acetyl-CoA decarbonylase/synthase complex gamma subunit	Ga0180325_111203
acsB	CO-methylating acetyl-CoA synthase; CO dehydrogenase/acetyl-CoA synthase alpha subunit	Ga0180325_111204
cooC	CO dehydrogenase maturation protein	Ga0180325_111201
		Ga0180325_111206
cooT	CO dehydrogenase maturation protein	Ga0180325_115201
		Ga0180325_11684
cooS (acsA)	CO dehydrogenase (acceptor)	Ga0180325_111205
		Ga0180325_11614
		Ga0180325_113455
cooA	CO-responsive transcriptional activator	Ga0180325_113457
cooF	CO dehydrogenase iron-sulphur protein	Ga0180325_113456
Wood-Ljungdahl pathway, methyl branch		
metF	methylenetetrahydrofolate reductase [NAD(P)H]	Ga0180325_111198
fhs	formate-tetrahydrofolate ligase	Ga0180325_114722
fdhF	formate dehyrogenase, alpha subunit	Ga0180325_112878s80
hoxE	formate dehydrogenase gamma subunit (NAD^+-reducing hydrogenase)	Ga0180325_112877
hoxF	formate dehydrogenase beta subunit (NAD^+-reducing hydrogenase)	Ga0180325_112876
folD	methylenetetrahydrofolate dehydrogenase($NADP^+$)/methenyltetrahydrofolate cyclohydrase	Ga0180325_11378
Acetate metabolism		
pduL	phosphate propanoyltransferase	Ga0180325_112350
		Ga0180325_112835
		Ga0180325_115608
ackA	acetate kinase	Ga0180325_111561

acsA	acetate-CoA ligase	Ga0180325_115029
Choline to glycine betaine		
	choline dehydrogenase	Ga0180325_11215
	betaine aldehyde dehydrogenase	Ga0180325_114191
Putative glycine betaine/dimethylglycine methyltransferases		
mtgB / MT_1	glycine betaine / dimethylgylcine methyltransferase	Ga0180325_1148
	glycine betaine / dimethylgylcine methyltransferase	Ga0180325_11427
	glycine betaine / dimethylgylcine methyltransferase	Ga0180325_114740
	glycine betaine / dimethylgylcine methyltransferase	Ga0180325_115476
	glycine betaine / dimethylgylcine methyltransferase	Ga0180325_115483
	glycine betaine / dimethylgylcine methyltransferase	Ga0180325_115497
	glycine betaine / dimethylgylcine methyltransferase	Ga0180325_115521
mtgA / MT_2	5-methyltetrahydrofolate--homocysteine methyltransferase	Ga0180325_11439
	5-methyltetrahydrofolate--homocysteine methyltransferase	Ga0180325_111153
	5-methyltetrahydrofolate--homocysteine methyltransferase	Ga0180325_111236
	5-methyltetrahydrofolate--homocysteine methyltransferase	Ga0180325_111281
	5-methyltetrahydrofolate--homocysteine methyltransferase	Ga0180325_111615
	5-methyltetrahydrofolate--homocysteine methyltransferase	Ga0180325_111616
	5-methyltetrahydrofolate--homocysteine methyltransferase	Ga0180325_111677
	5-methyltetrahydrofolate--homocysteine methyltransferase/trimethylamine corrinoid protein	Ga0180325_111811
	5-methyltetrahydrofolate--homocysteine methyltransferase	Ga0180325_112018
	5-methyltetrahydrofolate--homocysteine methyltransferase	Ga0180325_112642
	5-methyltetrahydrofolate--homocysteine methyltransferase	Ga0180325_112645
	5-methyltetrahydrofolate--homocysteine methyltransferase/trimethylamine corrinoid protein	Ga0180325_113010
	5-methyltetrahydrofolate--homocysteine methyltransferase	Ga0180325_113360
	5-methyltetrahydrofolate--homocysteine methyltransferase	Ga0180325_113423
	5-methyltetrahydrofolate--homocysteine methyltransferase	Ga0180325_113624
	5-methyltetrahydrofolate--homocysteine methyltransferase	Ga0180325_113658
	5-methyltetrahydrofolate--homocysteine methyltransferase	Ga0180325_113813
	5-methyltetrahydrofolate--homocysteine methyltransferase	Ga0180325_113829
	5-methyltetrahydrofolate--homocysteine methyltransferase	Ga0180325_114450
	5-methyltetrahydrofolate--homocysteine methyltransferase	Ga0180325_114734
	5-methyltetrahydrofolate--homocysteine methyltransferase	Ga0180325_114736
	5-methyltetrahydrofolate--homocysteine methyltransferase	Ga0180325_115284
	5-methyltetrahydrofolate--homocysteine methyltransferase	Ga0180325_115354
	5-methyltetrahydrofolate--homocysteine methyltransferase	Ga0180325_115420
	5-methyltetrahydrofolate--homocysteine methyltransferase	Ga0180325_115485
	5-methyltetrahydrofolate--homocysteine methyltransferase	Ga0180325_115487

	5-methyltetrahydrofolate--homocysteine methyltransferase	Ga0180325_115517
	5-methyltetrahydrofolate--homocysteine methyltransferase	Ga0180325_115771
	5-methyltetrahydrofolate--homocysteine methyltransferase	Ga0180325_115806
mtbC / corrinoid protein	Methanogenic corrinoid protein MtbC1	Ga0180325_11500
	Methanogenic corrinoid protein MtbC1	Ga0180325_11502
	Methanogenic corrinoid protein MtbC1	Ga0180325_11517
	Methanogenic corrinoid protein MtbC1	Ga0180325_11521
	methyltransferase cognate corrinoid proteins	Ga0180325_111156
	Methanogenic corrinoid protein MtbC1	Ga0180325_111663
	Methanogenic corrinoid protein MtbC1	Ga0180325_111666
	Methanogenic corrinoid protein MtbC1	Ga0180325_111699
	Methanogenic corrinoid protein MtbC1	Ga0180325_111701
	Methanogenic corrinoid protein MtbC1	Ga0180325_112242
	Methanogenic corrinoid protein MtbC1	Ga0180325_112245
	methyltransferase cognate corrinoid proteins	Ga0180325_112545
	methyltransferase cognate corrinoid proteins	Ga0180325_114062
	Methanogenic corrinoid protein MtbC1	Ga0180325_114649
	Methanogenic corrinoid protein MtbC1	Ga0180325_114771
	Methanogenic corrinoid protein MtbC1	Ga0180325_114772
	methyltransferase cognate corrinoid proteins	Ga0180325_115389
	methyltransferase cognate corrinoid proteins	Ga0180325_115394
	Methanogenic corrinoid protein MtbC1	Ga0180325_115712
Sarcosine reductase cluster		
trxB	thioredoxin reductase (NADPH)	Ga0180325_114795
trxA	thioredoxin 1	Ga0180325_114796
grdA	glycine reductase complex selenoprotein A	Ga0180325_114797s98
grdC	glycine/sarcosine/betaine reductase complex, protein C subunit beta	Ga0180325_114799
grdD	glycine/sarcosine/betaine reductase complex, protein C subunit alpha	Ga0180325_114800
grdG	sarcosine reductase complex, protein B subunit gamma	Ga0180325_114802
grdF	sarcosine reductase complex, protein B subunits alpha and beta	Ga0180325_114803s

Table S4 Methanol-specific methyltransferase genes in the strain DCMF genome.

IMG Locus Tag (Ga0180325_)	Gene Symbol	Product Name	Length (AA)
111151	mtaA	[methyl-Co(III) methanol-specific corrinoid protein]:coenzyme M methyltransferase	340
111155	mtaA	[methyl-Co(III) methanol-specific corrinoid protein]:coenzyme M methyltransferase	338
111671	mtaA	[methyl-Co(III) methanol-specific corrinoid protein]:coenzyme M methyltransferase	349
111674	mtaA	[methyl-Co(III) methanol-specific corrinoid protein]:coenzyme M methyltransferase	349
112541	mtaA	methanol-specific methylcobalamin: coenzyme M methyltransferase	338
112544	mtaA	[methyl-Co(III) methanol-specific corrinoid protein]:coenzyme M methyltransferase	357
112644	mtaB	methanol:corrinoid methyltransferase	458
115395	mtaA	[methyl-Co(III) methanol-specific corrinoid protein]:coenzyme M methyltransferase	345
115702	mtaA	[methyl-Co(III) methanol-specific corrinoid protein]:coenzyme M methyltransferase	381
115703	mtbC	methanol corrinoid protein	279

Table S5 List of significantly differentially expressed proteins between DCM- and glycine betaine-amended cells. Differential expression was assessed via t-test, FDR < 0.01. See Dataset S2 for the full list of proteins identified.

Function	Locus Tag	Product	DCM	glycine betaine	DCM	glycine betaine	Sig?	-Log p-value	Q-value	Differ-ence
	Ga0180325_1112	phosphoribosylaminoimidazolecarboxamide formyltransferase / IMP cyclohydrolase	1.32	0.95	29.87	29.10	+	2.504	0.007	0.76
	Ga0180325_1120	ATP phosphoribosyltransferase regulatory subunit	-0.10	1.74	25.99	31.38	+	2.858	0.002	-5.39
	Ga0180325_1121	imidazoleglycerol-phosphate dehydratase	-1.11	0.96	23.25	29.13	+	2.557	0.005	-5.88
	Ga0180325_1122	glutamine amidotransferase	-1.14	0.52	23.17	27.86	+	3.338	0.002	-4.70
	Ga0180325_1123	1-(5-phosphoribosyl)-5-[(5-phosphoribosylamino)methylideneamino] imidazole-4-carboxamide isomerase	-1.07	0.65	23.37	28.25	+	3.058	0.002	-4.88
	Ga0180325_1124	phosphoribosyl-AMP cyclohydrolase	-1.20	0.86	23.00	28.84	+	4.899	0.000	-5.84
	Ga0180325_1125	phosphoribosyl-ATP pyrophosphatase	-1.19	0.97	23.03	29.17	+	3.626	0.001	-6.14
	Ga0180325_1167	exodeoxyribonuclease-3	0.76	0.96	28.34	29.12	+	2.411	0.008	-0.79
	Ga0180325_1181	homocitrate synthase NifV	-0.72	1.79	24.31	31.51	+	3.601	0.001	-7.20
	Ga0180325_11107	Tripartite-type tricarboxylate transporter, receptor component TctC	-0.51	-0.08	24.88	26.15	+	2.581	0.005	-1.27
	Ga0180325_11218	thiosulfate/3-mercaptopyruvate sulfurtransferase	-1.12	0.58	23.22	28.05	+	3.532	0.002	-4.83
	Ga0180325_11220	D-3-phosphoglycerate dehydrogenase	0.22	2.50	26.86	33.56	+	3.406	0.002	-6.70

	Ga0180325_11221	Uncharacterized conserved protein, DUF1015 family	0.58	1.54	27.84	30.79	+	3.936	0.000	-2.95
	Ga0180325_11222	phosphoserine aminotransferase apoenzyme	1.36	2.66	29.96	34.03	+	4.427	0.000	-4.06
	Ga0180325_11341	CxxC-x17-CxxC domain-containing protein	-1.00	1.20	23.56	29.83	+	3.834	0.001	-6.27
	Ga0180325_11355	aconitase	-0.99	0.88	23.58	28.91	+	2.507	0.007	-5.33
	Ga0180325_11376	DNA binding domain-containing protein, excisionase family	-0.87	0.98	23.89	29.20	+	2.853	0.002	-5.31
Wood-Ljungdahl pathway	Ga0180325_11378	methenyltetrahydrofolate cyclohydrolase /5,10-methylenetetrahydrofolate dehydrogenase (NADP+)	1.36	1.62	29.96	31.02	+	3.646	0.001	-1.06
	Ga0180325_11379	hypothetical protein	-0.99	0.06	23.58	26.56	+	2.828	0.003	-2.98
	Ga0180325_11392	hypothetical protein	-1.07	-0.36	23.36	25.33	+	2.479	0.007	-1.97
Methyltransferase system	Ga0180325_11439	5-methyltetrahydrofolate--homocysteine methyltransferase	1.18	-1.27	29.47	22.74	+	2.498	0.007	6.73
	Ga0180325_11473	Putative zinc-finger	-0.25	-1.21	25.59	22.90	+	2.427	0.007	2.68
	Ga0180325_11513	Multimeric flavodoxin WrbA	-0.74	-0.45	24.26	25.07	+	3.144	0.002	-0.81
	Ga0180325_11576	two component transcriptional regulator, winged helix family	-0.16	-0.49	25.82	24.95	+	3.075	0.002	0.87
	Ga0180325_11578	phosphate ABC transporter substrate-binding protein, PhoT family	-1.14	-0.15	23.16	25.95	+	2.833	0.003	-2.79
	Ga0180325_11589	protein of unknown function (DUF1540)	-1.09	0.26	23.30	27.11	+	2.693	0.004	-3.81
	Ga0180325_11590	Coat F domain-containing protein	-1.07	-0.11	23.36	26.05	+	3.748	0.001	-2.69
	Ga0180325_11591	Coat F domain-containing protein	-1.03	0.08	23.48	26.60	+	3.107	0.002	-3.12
	Ga0180325_11621	aspartate-semialdehyde dehydrogenase	0.31	-0.13	27.10	26.00	+	2.616	0.005	1.11

S-layer	Ga0180325_11624	S-layer homology domain-containing protein	-0.48	0.77	24.97	28.58	+	2.684	0.004	-3.61
	Ga0180325_11625	hypothetical protein	-1.26	0.43	22.84	27.61	+	4.591	0.000	-4.77
	Ga0180325_11673	nif11-like leader peptide domain-containing protein	-0.29	0.30	25.47	27.24	+	2.567	0.005	-1.76
	Ga0180325_11689	DNA-binding regulatory protein, YebC/PmpR family	0.38	-0.44	27.29	25.12	+	3.278	0.001	2.17
	Ga0180325_11701	protein translocase subunit yajC	0.64	1.02	28.01	29.30	+	2.405	0.008	-1.30
	Ga0180325_11702	preprotein translocase subunit SecD	1.64	0.94	30.72	29.07	+	4.489	0.000	1.65
	Ga0180325_11703	protein translocase subunit secF	1.51	0.95	30.36	29.11	+	2.817	0.003	1.25
	Ga0180325_11712	DNA-binding transcriptional regulator, FrmR family	0.48	-0.27	27.56	25.59	+	2.851	0.002	1.97
	Ga0180325_11721	alanyl-tRNA synthetase	1.04	0.75	29.08	28.52	+	2.474	0.007	0.56
	Ga0180325_11774	3-dehydroquinate dehydratase	0.43	-0.22	27.43	25.75	+	2.603	0.005	1.68
Sporulation	Ga0180325_11784	stage III sporulation protein AH	-1.22	-0.25	22.95	25.64	+	2.675	0.004	-2.69
	Ga0180325_11787	Small integral membrane protein (DUF2273)	-0.04	-1.11	26.16	23.17	+	3.153	0.002	2.99
Energy conservation	Ga0180325_11791	NADH-quinone oxidoreductase subunit E	-0.44	0.71	25.08	28.41	+	2.285	0.009	-3.34
Energy conservation	Ga0180325_11792	NADH-quinone oxidoreductase subunit F	1.86	2.19	31.32	32.68	+	4.020	0.000	-1.37
Wood-Ljungdahl pathway	Ga0180325_11793	formate dehydrogenase major subunit	2.27	2.76	32.43	34.31	+	3.933	0.000	-1.88
Sporulation	Ga0180325_11831	two-component system, response regulator, stage 0 sporulation protein A	-0.06	1.07	26.10	29.45	+	3.663	0.001	-3.36
	Ga0180325_11833	TRAP-type C4-dicarboxylate transport system, substrate-binding protein	2.27	-0.60	32.42	24.65	+	4.320	0.000	7.77

	Ga0180325_11836	opine dehydrogenase	0.51	-1.22	27.64	22.87	+	3.776	0.001	4.77
Sporulation	Ga0180325_11851	anti-anti-sigma regulatory factor, SpoIIAA	-1.01	0.54	23.53	27.92	+	5.206	0.000	-4.39
Sporulation	Ga0180325_11852	stage II sporulation protein AB (anti-sigma F factor)	-0.68	0.47	24.42	27.73	+	4.234	0.000	-3.31
	Ga0180325_11863	tryptophan synthase beta chain	1.63	0.80	30.68	28.68	+	3.492	0.002	2.00
	Ga0180325_11885	3-deoxy-D-arabinoheptulosonate-7-phosphate synthase	1.27	1.58	29.72	30.91	+	2.583	0.005	-1.20
	Ga0180325_11890	4-hydroxy-3-methylbut-2-enyl diphosphate reductase	1.95	1.54	31.57	30.81	+	2.637	0.004	0.76
	Ga0180325_11896	GTP-binding protein	0.60	-0.03	27.90	26.29	+	2.981	0.002	1.61
Sporulation	Ga0180325_11899	stage IV sporulation protein A	-0.65	1.54	24.48	30.80	+	2.826	0.003	-6.31
	Ga0180325_11903	aspartate kinase	1.19	0.68	29.50	28.33	+	2.732	0.004	1.17
	Ga0180325_11964	hypothetical protein	0.17	0.61	26.74	28.13	+	3.295	0.001	-1.40
	Ga0180325_11990	hypothetical protein	-1.11	0.00	23.25	26.38	+	2.611	0.005	-3.13
	Ga0180325_11999	ribosome-binding factor A	0.59	-0.02	27.86	26.30	+	2.695	0.004	1.56
	Ga0180325_111029	recombination protein RecA	1.32	0.49	29.85	27.78	+	3.396	0.002	2.07
	Ga0180325_111031	metal dependent phosphohydrolase	1.88	1.55	31.38	30.82	+	2.455	0.007	0.56
	Ga0180325_111042	inosine-5'-monophosphate dehydrogenase	0.84	0.48	28.55	27.75	+	2.269	0.010	0.80
	Ga0180325_111060	hypothetical protein	1.07	0.08	29.19	26.61	+	4.614	0.000	2.57
	Ga0180325_111124	selenium metabolism protein YedF	0.77	1.78	28.35	31.48	+	2.658	0.004	-3.13
	Ga0180325_111125	cysteine desulfurase family protein	-0.09	0.64	26.02	28.22	+	2.615	0.005	-2.20
	Ga0180325_111131	D-3-phosphoglycerate dehydrogenase	2.14	0.60	32.08	28.10	+	3.393	0.002	3.98

	Ga0180325_111133	succinate dehydrogenase / fumarate reductase iron-sulfur subunit	-0.15	-0.80	25.86	24.08	+	3.464	0.002	1.78
	Ga0180325_111136	Endonuclease IV	-0.12	-0.84	25.95	23.96	+	3.063	0.002	1.99
Methyltransferase system	Ga0180325_111152	trimethylamine---corrinoid protein Co-methyltransferase	-1.64	2.93	21.81	34.78	+	3.658	0.001	-12.98
Methyltransferase system	Ga0180325_111154	trimethylamine---corrinoid protein Co-methyltransferase	-1.24	2.95	22.90	34.84	+	4.694	0.000	-11.94
Methyltransferase system	Ga0180325_111156	methyltransferase cognate corrinoid proteins	-1.15	0.93	23.13	29.04	+	3.933	0.000	-5.91
Methyltransferase system	Ga0180325_111158	Uroporphyrinogen decarboxylase (URO-D)	-1.03	0.06	23.45	26.56	+	4.146	0.000	-3.10
	Ga0180325_111159	Benzoyl-CoA reductase/2-hydroxyglutaryl-CoA dehydratase subunit, BcrC/BadD/HgdB	-1.34	0.72	22.63	28.45	+	3.595	0.001	-5.82
	Ga0180325_111166	GTP cyclohydrolase I	0.38	-0.05	27.29	26.23	+	3.455	0.002	1.06
	Ga0180325_111174	malate dehydrogenase (oxaloacetate-decarboxylating)	0.48	-0.92	27.58	23.73	+	2.235	0.010	3.85
	Ga0180325_111179	hypothetical protein	0.52	-0.32	27.67	25.46	+	2.900	0.002	2.21
	Ga0180325_111180	propionyl-CoA carboxylase carboxyltransferase subunit	1.66	0.70	30.77	28.38	+	4.111	0.000	2.39
	Ga0180325_111183	methylmalonyl-CoA mutase, N-terminal domain	1.92	0.66	31.49	28.27	+	5.306	0.000	3.21
	Ga0180325_111195	aminotransferase	-1.48	-0.84	22.25	23.97	+	2.579	0.005	-1.71
	Ga0180325_111196	DNA-binding transcriptional regulator, Lrp family	-1.17	0.64	23.09	28.21	+	3.215	0.002	-5.12
Wood-Ljungdahl pathway	Ga0180325_111198	methylenetetrahydrofolate reductase (NADPH)	2.13	2.79	32.07	34.39	+	4.293	0.000	-2.33
Wood-Ljungdahl pathway	Ga0180325_111201	CO dehydrogenase maturation factor	1.69	0.91	30.87	28.98	+	3.590	0.001	1.89

Wood-Ljungdahl pathway	Ga0180325_111203	CO-methylating acetyl-CoA synthase corrinoid iron-sulfur protein large subunit precursor	3.45	2.88	35.66	34.66	+	3.074	0.002	1.00
Wood-Ljungdahl pathway	Ga0180325_111205	Ni-dependent carbon monoxide dehydrogenase precursor	2.71	2.30	33.64	33.00	+	2.879	0.002	0.64
	Ga0180325_111207	EDD domain protein, DegV family	-0.16	-1.39	25.82	22.39	+	3.266	0.001	3.43
	Ga0180325_111211	tryptophan synthase, alpha chain	0.07	-1.30	26.46	22.63	+	2.982	0.002	3.83
	Ga0180325_111212	tryptophan synthase, beta chain	0.06	-1.06	26.44	23.34	+	3.426	0.002	3.10
	Ga0180325_111214	indole-3-glycerol phosphate synthase	-0.23	-1.17	25.65	23.00	+	2.501	0.007	2.65
DCM-associated cluster	Ga0180325_111230	Kef-type K+ transport system, membrane component KefB	0.07	-1.08	26.46	23.28	+	2.272	0.010	3.19
DCM-associated cluster; methyltransferase system	Ga0180325_111231	Uroporphyrinogen decarboxylase (URO-D)	1.48	-1.56	30.29	21.89	+	4.779	0.000	8.40
DCM-associated cluster; methyltransferase system	Ga0180325_111232	tetrahydromethanopterin S-methyltransferase subunit H	2.77	0.25	33.79	27.10	+	3.917	0.000	6.70
DCM-associated cluster; methyltransferase system	Ga0180325_111233	methyltransferase, MtaA/CmuA family	3.08	1.10	34.63	29.55	+	4.053	0.000	5.08
DCM-associated cluster	Ga0180325_111234	two component transcriptional regulator, AraC family	0.53	-1.29	27.72	22.67	+	4.906	0.000	5.05
DCM-associated cluster; methyltransferase system	Ga0180325_111235	uroporphyrinogen decarboxylase	3.43	1.54	35.59	30.79	+	4.838	0.000	4.80
DCM-associated cluster; methyltransferase system	Ga0180325_111236	5-methyltetrahydrofolate--homocysteine methyltransferase	2.67	1.35	33.53	30.24	+	4.766	0.000	3.29

DCM-associated cluster	Ga0180325_111237	Histidine kinase	0.88	-1.22	28.65	22.87	+	5.090	0.000	5.78
	Ga0180325_111249	malate dehydrogenase (oxaloacetate-decarboxylating)	1.03	-0.84	29.08	23.97	+	2.919	0.002	5.11
	Ga0180325_111252	succinate dehydrogenase subunit B	0.39	-1.19	27.33	22.96	+	3.206	0.002	4.37
	Ga0180325_111253	succinate dehydrogenase subunit A	1.49	-0.03	30.31	26.28	+	4.735	0.000	4.03
	Ga0180325_111255	2-oxoglutarate ferredoxin oxidoreductase, gamma subunit	1.96	-0.12	31.60	26.02	+	3.143	0.002	5.59
	Ga0180325_111256	2-oxoglutarate ferredoxin oxidoreductase, beta subunit	2.07	-0.16	31.89	25.90	+	2.881	0.002	5.99
	Ga0180325_111257	2-oxoglutarate ferredoxin oxidoreductase subunit alpha	2.21	0.48	32.28	27.75	+	4.227	0.000	4.53
	Ga0180325_111258	2-oxoglutarate ferredoxin oxidoreductase subunit delta	1.56	-0.90	30.51	23.79	+	6.118	0.000	6.71
Methyltransferase system	Ga0180325_111271p72	dimethylamine:corrinoid methyltransferase	1.61	-0.03	30.65	26.29	+	3.995	0.000	4.36
Methyltransferase system	Ga0180325_111273	trimethylamine---corrinoid protein Co-methyltransferase	2.15	0.18	32.11	26.89	+	4.448	0.000	5.22
Methyltransferase system	Ga0180325_111274	trimethylamine corrinoid protein	2.47	0.35	32.99	27.39	+	4.037	0.000	5.60
Methyltransferase system	Ga0180325_111278p79	trimethylamine---corrinoid protein Co-methyltransferase	2.28	-0.59	32.47	24.67	+	4.507	0.000	7.79
	Ga0180325_111326	porphobilinogen synthase	1.27	0.46	29.71	27.69	+	3.072	0.002	2.03
	Ga0180325_111327	hydroxymethylbilane synthase	1.12	0.63	29.30	28.19	+	3.173	0.002	1.11
	Ga0180325_111360	Glycosyltransferase involved in cell wall bisynthesis	-1.16	-0.45	23.12	25.08	+	2.723	0.004	-1.97
	Ga0180325_111361	UDP-N-acetyl-D-mannosaminuronic acid dehydrogenase	-0.20	0.83	25.73	28.75	+	2.782	0.004	-3.02
	Ga0180325_111378	hypothetical protein	0.96	0.52	28.88	27.87	+	2.957	0.002	1.01
	Ga0180325_111407	ribosome recycling factor	1.93	1.53	31.50	30.76	+	2.306	0.009	0.74

Chemotaxis	Ga0180325_111424	methyl-accepting chemotaxis protein	0.15	0.56	26.68	27.99	+	3.866	0.000	-1.30
	Ga0180325_111487	RNA-binding protein (KH domain)	1.23	0.54	29.62	27.94	+	2.509	0.006	1.68
	Ga0180325_111494	methylthioadenosine phosphorylase	0.39	-0.07	27.32	26.18	+	4.108	0.000	1.14
	Ga0180325_111550	3-oxoacyl-[acyl-carrier-protein] reductase	0.66	0.27	28.06	27.15	+	3.695	0.001	0.91
	Ga0180325_111554	phosphate:acyl-[acyl carrier protein] acyltransferase	-0.10	-0.65	26.00	24.50	+	2.295	0.009	1.50
	Ga0180325_111558	hypothetical protein	-0.10	-0.76	26.00	24.18	+	2.749	0.004	1.83
	Ga0180325_111559	LSU ribosomal protein L32P	-0.19	0.54	25.74	27.93	+	2.653	0.004	-2.19
Acetate metabolism	Ga0180325_111561	acetate kinase	1.52	0.87	30.40	28.87	+	2.407	0.008	1.53
Hydrogenase	Ga0180325_111565	NAD(P)-dependent iron-only hydrogenase catalytic subunit	-0.07	-1.26	26.07	22.75	+	3.597	0.001	3.32
Hydrogenase	Ga0180325_111566	NADH-quinone oxidoreductase subunit F	-0.26	-1.36	25.55	22.47	+	2.344	0.009	3.08
	Ga0180325_111583	Small, acid-soluble spore protein, alpha/beta type	-0.66	0.63	24.46	28.19	+	2.297	0.009	-3.73
	Ga0180325_111605	peptide deformylase	-1.38	0.14	22.51	26.78	+	3.302	0.001	-4.28
	Ga0180325_111639	Copper amine oxidase N-terminal domain-containing protein	-1.46	-0.80	22.29	24.06	+	3.157	0.002	-1.77
	Ga0180325_111642	pyruvate ferredoxin oxidoreductase, gamma subunit	0.31	-0.42	27.11	25.16	+	2.572	0.005	1.95
	Ga0180325_111695	Phage-related replication protein YjqB, UPF0714/DUF867 family	-1.05	0.44	23.40	27.63	+	3.431	0.002	-4.23
Chemotaxis	Ga0180325_111723	Methyl-accepting chemotaxis protein	0.58	0.13	27.84	26.74	+	3.007	0.002	1.10
	Ga0180325_111731	REP element-mobilizing transposase RayT	-0.49	-0.69	24.94	24.39	+	2.931	0.002	0.55
	Ga0180325_111733	hypothetical protein	0.96	0.51	28.89	27.84	+	2.814	0.003	1.05

	Ga0180325_111736	aldehyde oxidoreductase	2.44	1.60	32.90	30.96	+	4.444	0.000	1.94
	Ga0180325_111763	methionine adenosyltransferase	0.87	0.59	28.62	28.07	+	2.518	0.006	0.55
	Ga0180325_111808	Adenine deaminase	0.48	-0.47	27.58	25.02	+	2.697	0.004	2.56
Methyltransferase system	Ga0180325_111809	tetrahydromethanopterin S-methyltransferase subunit H	0.91	-1.58	28.73	21.83	+	3.772	0.001	6.91
Methyltransferase system	Ga0180325_111810	monomethylamine:corrinoid methyltransferase	2.18	-1.96	32.20	20.74	+	5.777	0.000	11.46
Methyltransferase system	Ga0180325_111811	5-methyltetrahydrofolate--homocysteine methyltransferase/trimethylamine corrinoid protein	1.33	-1.32	29.88	22.59	+	5.529	0.000	7.29
	Ga0180325_111812	putative hydroxymethylpyrimidine transporter CytX	1.79	-1.10	31.12	23.21	+	3.171	0.002	7.91
	Ga0180325_111829	transcriptional regulator, DeoR family	1.48	0.74	30.30	28.49	+	3.452	0.002	1.81
	Ga0180325_111832	amino acid/amide ABC transporter substrate-binding protein, HAAT family	2.36	3.30	32.69	35.85	+	4.711	0.000	-3.16
	Ga0180325_111833	urea ABC transporter ATP-binding protein	0.64	1.31	28.01	30.14	+	3.326	0.001	-2.13
QA transporters	Ga0180325_111834	amino acid/amide ABC transporter ATP-binding protein 1, HAAT family	-0.92	0.66	23.77	28.27	+	2.656	0.004	-4.50
	Ga0180325_111835	aspartate carbamoyltransferase	1.94	2.40	31.54	33.28	+	3.252	0.001	-1.74
	Ga0180325_111838	carbon-monoxide dehydrogenase medium subunit	1.37	1.48	29.98	30.62	+	2.342	0.009	-0.64
	Ga0180325_111841	xanthine dehydrogenase molybdenum-binding subunit	1.54	1.79	30.44	31.51	+	2.381	0.008	-1.07
	Ga0180325_111847	molybdenum cofactor synthesis domain-containing protein	0.04	0.93	26.37	29.04	+	3.134	0.002	-2.68
	Ga0180325_111849	selenium-dependent molybdenum hydroxylase system protein, YqeB family	-0.26	0.05	25.56	26.53	+	2.709	0.004	-0.97

	Ga0180325_111854	Acyl-CoA reductase	1.21	-0.91	29.56	23.76	+	2.284	0.009	5.79
	Ga0180325_111876	molybdate transport system substrate-binding protein	1.42	0.26	30.13	27.11	+	3.260	0.001	3.02
	Ga0180325_111877	molybdate transport system regulatory protein	0.19	-0.16	26.77	25.92	+	3.135	0.002	0.85
	Ga0180325_111896	Nitroreductase	0.71	1.54	28.19	30.80	+	2.662	0.004	-2.61
	Ga0180325_111910	pyrroline-5-carboxylate reductase	0.35	-1.22	27.21	22.86	+	3.167	0.002	4.35
	Ga0180325_111929	succinyl-CoA synthetase alpha subunit	1.15	0.56	29.38	27.98	+	2.315	0.009	1.40
	Ga0180325_111990	peptide/nickel transport system substrate-binding protein	0.16	0.71	26.69	28.41	+	3.111	0.002	-1.72
	Ga0180325_111994	sarcosine oxidase subunit beta	1.10	-1.31	29.25	22.61	+	3.250	0.001	6.64
	Ga0180325_111996	Fe-S-cluster-containing hydrogenase component 2	1.22	-1.22	29.58	22.87	+	4.242	0.000	6.70
	Ga0180325_111997	sarcosine oxidase subunit alpha	0.56	-1.34	27.79	22.51	+	3.858	0.000	5.28
	Ga0180325_112103	Methyl-viologen-reducing hydrogenase, delta subunit	-1.07	0.15	23.35	26.80	+	2.457	0.007	-3.45
	Ga0180325_112105	Pyridine nucleotide-disulphide oxidoreductase	1.00	1.18	28.99	29.78	+	2.771	0.004	-0.79
	Ga0180325_112106	heterodisulfide reductase subunit B	0.71	0.98	28.19	29.20	+	2.644	0.004	-1.01
	Ga0180325_112107	heterodisulfide reductase subunit C	0.59	1.04	27.86	29.35	+	2.598	0.005	-1.49
	Ga0180325_112137	tyrosyl-tRNA synthetase	1.68	0.78	30.83	28.63	+	3.569	0.001	2.20
	Ga0180325_112148	hypothetical protein	-0.91	0.81	23.78	28.70	+	3.152	0.002	-4.91
	Ga0180325_112178	urocanate hydratase	1.56	-1.30	30.50	22.65	+	3.088	0.002	7.85
	Ga0180325_112181	imidazolonepropionase	0.59	-1.33	27.86	22.56	+	2.483	0.007	5.30
	Ga0180325_112184	4Fe-4S dicluster domain-containing protein	0.04	-1.15	26.37	23.06	+	2.811	0.003	3.31

Chemotaxis	Ga0180325_112197	methyl-accepting chemotaxis sensory transducer with Cache sensor	-0.03	-0.67	26.19	24.45	+	2.344	0.009	1.74
	Ga0180325_112262	Multimeric flavodoxin WrbA	0.80	-0.37	28.44	25.30	+	3.675	0.001	3.15
	Ga0180325_112271	glutamine--fructose-6-phosphate transaminase	0.67	0.24	28.09	27.07	+	2.380	0.008	1.02
	Ga0180325_112278	YbbR domain-containing protein	-1.17	-0.07	23.09	26.16	+	2.340	0.009	-3.07
	Ga0180325_112287	2-isopropylmalate synthase	0.03	0.59	26.35	28.06	+	2.653	0.004	-1.72
	Ga0180325_112294	hypothetical protein	-1.35	0.23	22.60	27.04	+	3.987	0.000	-4.43
	Ga0180325_112295	Coat F domain-containing protein	-1.54	0.27	22.09	27.15	+	3.444	0.002	-5.06
Methyltransferase system	Ga0180325_112298	Trimethylamine:corrinoid methyltransferase	-1.36	-0.83	22.58	23.99	+	3.388	0.002	-1.42
	Ga0180325_112304	peptide/nickel transport system substrate-binding protein	1.70	1.04	30.88	29.36	+	2.320	0.009	1.52
	Ga0180325_112340	hypothetical protein	-1.04	-0.01	23.44	26.35	+	2.604	0.005	-2.91
Ethanolamine BMC	Ga0180325_112352	BMC domain-containing protein	-0.31	-1.19	25.41	22.97	+	2.289	0.010	2.44
	Ga0180325_112379	membrane-bound serine protease (ClpP class)	1.04	0.60	29.09	28.10	+	2.236	0.010	0.99
	Ga0180325_112514	aminomethyltransferase	0.65	1.95	28.03	31.97	+	2.531	0.005	-3.94
	Ga0180325_112557	cell division topological specificity factor MinE	-1.18	-1.70	23.06	21.49	+	2.949	0.002	1.57
	Ga0180325_112577	trigger factor	1.87	1.55	31.35	30.84	+	2.383	0.008	0.51
	Ga0180325_112593	LAO/AO transport system kinase	-0.27	-1.43	25.52	22.26	+	2.576	0.005	3.26
Sporulation	Ga0180325_112624	Sporulation and spore germination	-0.82	0.34	24.04	27.35	+	2.400	0.008	-3.31
	Ga0180325_112626	phenylacetate-CoA ligase	0.48	0.03	27.56	26.45	+	2.458	0.007	1.12

Methanol metabolism; methyltransferase system	Ga0180325_112644	methanol:corrinoid methyltransferase	2.00	0.38	31.71	27.45	+	2.631	0.004	4.25
Methanol metabolism; methyltransferase system	Ga0180325_112645	5-methyltetrahydrofolate--homocysteine methyltransferase	1.56	-0.06	30.51	26.20	+	2.289	0.009	4.31
Chemotaxis	Ga0180325_112762	methyl-accepting chemotaxis sensory transducer with Cache sensor	1.12	1.93	29.30	31.92	+	5.512	0.000	-2.62
Flagellar	Ga0180325_112764	flagellar motor switch protein FliM	-0.35	0.15	25.31	26.80	+	3.418	0.002	-1.49
Chemotaxis	Ga0180325_112765	two-component system, chemotaxis family, response regulator CheY	-0.34	0.68	25.35	28.33	+	2.572	0.005	-2.98
	Ga0180325_112812	Uncharacterized conserved protein	-0.70	-0.35	24.36	25.38	+	2.372	0.008	-1.01
Flagellar	Ga0180325_112818	flagellin	2.40	3.10	32.79	35.29	+	2.321	0.010	-2.50
Chemotaxis	Ga0180325_112826	two-component system, chemotaxis family, sensor kinase CheA	1.77	2.07	31.06	32.32	+	4.609	0.000	-1.26
Acetate metabolism	Ga0180325_112835	putative phosphotransacetylase	2.10	1.45	31.97	30.55	+	3.506	0.002	1.42
	Ga0180325_112849	lactate permease	0.78	0.28	28.38	27.19	+	2.318	0.009	1.19
	Ga0180325_112892	delta-1-pyrroline-5-carboxylate dehydrogenase	0.85	-1.14	28.58	23.09	+	4.411	0.000	5.49
	Ga0180325_112909	Rubrerythrin	0.35	0.54	27.22	27.94	+	3.612	0.001	-0.72
	Ga0180325_112910	Rubrerythrin	0.80	1.12	28.45	29.60	+	3.106	0.002	-1.15
	Ga0180325_112914	ketopantoate hydroxymethyltransferase	-0.18	1.65	25.77	31.12	+	2.964	0.002	-5.35
	Ga0180325_112932	chromate reductase	0.62	1.00	27.94	29.25	+	2.252	0.009	-1.31

	Ga0180325_112988	Acyl-coenzyme A thioesterase PaaI, contains HGG motif	0.15	-0.14	26.67	25.97	+	2.277	0.009	0.70
	Ga0180325_112994	glyceraldehyde-3-phosphate dehydrogenase (NAD+)	2.25	1.48	32.39	30.63	+	3.379	0.002	1.76
	Ga0180325_113014	basic membrane protein A	-1.54	-0.13	22.08	25.99	+	2.906	0.002	-3.91
	Ga0180325_113064	thiamine biosynthesis lipoprotein	-0.12	-0.76	25.95	24.18	+	3.433	0.002	1.77
Energy conservation	Ga0180325_113065	electron transport complex, RnfABCDGE type, B subunit	1.79	1.20	31.14	29.82	+	2.599	0.005	1.31
Energy conservation	Ga0180325_113068	electron transport complex protein RnfG	2.26	1.49	32.39	30.65	+	3.490	0.002	1.75
Energy conservation	Ga0180325_113070	electron transport complex protein RnfC	1.67	1.14	30.80	29.66	+	2.854	0.002	1.15
	Ga0180325_113077	bifunctional phosphoglucose/phosphomannose isomerase	0.27	-0.08	27.00	26.16	+	2.328	0.010	0.85
	Ga0180325_113088	cell division protein FtsX	-0.93	-1.61	23.73	21.74	+	2.986	0.002	1.99
QA transporters	Ga0180325_113093	amino acid/amide ABC transporter ATP-binding protein 2, HAAT family	-0.65	0.97	24.49	29.17	+	2.628	0.004	-4.67
QA transporters	Ga0180325_113094	amino acid/amide ABC transporter ATP-binding protein 1, HAAT family	-0.86	0.76	23.93	28.56	+	2.353	0.009	-4.63
	Ga0180325_113095	branched-chain amino acid transport system permease protein	-1.05	0.16	23.42	26.84	+	2.620	0.004	-3.42
	Ga0180325_113097	amino acid/amide ABC transporter substrate-binding protein, HAAT family	1.55	3.36	30.49	36.04	+	3.129	0.002	-5.56
	Ga0180325_113098	proline racemase	-0.50	1.74	24.91	31.38	+	3.458	0.002	-6.47
	Ga0180325_113110	phosphoglucomutase, alpha-D-glucose phosphate-specific	0.54	0.66	27.73	28.27	+	2.455	0.007	-0.54

	Ga0180325_113118	hypothetical protein	0.83	0.16	28.51	26.82	+	2.660	0.004	1.69
	Ga0180325_113128	hydrophobic/amphiphilic exporter-1, HAE1 family	0.78	0.28	28.37	27.17	+	3.156	0.002	1.20
S-layer	Ga0180325_113131	S-layer homology domain-containing protein	1.11	1.47	29.29	30.60	+	2.502	0.007	-1.31
Sporulation	Ga0180325_113146	putative DeoR family transcriptional regulator, stage III sporulation protein D	-0.55	0.54	24.76	27.93	+	2.236	0.010	-3.16
Energy conservation	Ga0180325_113152	ATP synthase F1 subcomplex epsilon subunit	1.91	2.13	31.45	32.49	+	2.366	0.008	-1.04
Energy conservation	Ga0180325_113153	ATP synthase F1 subcomplex beta subunit	2.45	2.71	32.92	34.18	+	2.460	0.007	-1.25
Energy conservation	Ga0180325_113154	ATP synthase F1 subcomplex gamma subunit	1.70	1.94	30.88	31.94	+	2.547	0.005	-1.07
Energy conservation	Ga0180325_113156	ATP synthase F1 subcomplex delta subunit	1.66	2.11	30.77	32.43	+	3.141	0.002	-1.65
Energy conservation	Ga0180325_113157	ATP synthase F0 subcomplex B subunit	2.52	2.68	33.13	34.08	+	3.615	0.001	-0.96
	Ga0180325_113183	transaldolase	0.80	0.23	28.44	27.04	+	4.679	0.000	1.40
	Ga0180325_113188	Lipid II:glycine glycyltransferase (Peptidoglycan interpeptide bridge formation enzyme)	0.49	0.00	27.61	26.38	+	3.421	0.002	1.23
Type IV pili	Ga0180325_113213	pilus assembly protein CpaF	-0.97	-1.36	23.64	22.46	+	2.308	0.009	1.18
	Ga0180325_113261	cold shock protein (beta-ribbon, CspA family)	-1.36	0.21	22.56	26.99	+	3.185	0.002	-4.43
	Ga0180325_113273	protein of unknown function (DUF4342)	-0.06	1.16	26.11	29.71	+	3.627	0.001	-3.61
	Ga0180325_113287	malate dehydrogenase (oxaloacetate-decarboxylating)	1.03	-0.84	29.08	23.97	+	2.919	0.002	5.11
	Ga0180325_113290	succinate dehydrogenase subunit B	0.39	-1.19	27.33	22.96	+	3.206	0.002	4.37
	Ga0180325_113291	succinate dehydrogenase subunit A	1.49	-0.03	30.31	26.28	+	4.735	0.000	4.03

Chemotaxis	Ga0180325_113341	Methyl-accepting chemotaxis protein (MCP) signalling domain-containing protein	-0.21	0.68	25.69	28.33	+	2.644	0.004	-2.64
	Ga0180325_113428	hypothetical protein	0.19	1.01	26.78	29.26	+	2.924	0.002	-2.48
	Ga0180325_113435	aconitase	1.50	-1.59	30.35	21.81	+	4.876	0.000	8.53
	Ga0180325_113438	citrate lyase subunit beta / citryl-CoA lyase	0.53	-1.41	27.70	22.31	+	4.066	0.000	5.40
	Ga0180325_113448	Murein DD-endopeptidase MepM and murein hydrolase activator NlpD, contain LysM domain	-0.85	-1.68	23.94	21.56	+	3.217	0.002	2.38
	Ga0180325_113470	L-ascorbate metabolism protein UlaG, beta-lactamase superfamily	0.61	0.89	27.91	28.94	+	3.108	0.002	-1.03
	Ga0180325_113487	single-strand DNA-binding protein	0.79	0.39	28.41	27.49	+	3.479	0.002	0.92
	Ga0180325_113536	ornithine carbamoyltransferase	1.86	0.48	31.31	27.75	+	4.269	0.000	3.57
	Ga0180325_113552	CoA-substrate-specific enzyme activase, putative	-0.58	-0.73	24.70	24.27	+	2.332	0.010	0.43
	Ga0180325_113555	acetolactate synthase, small subunit	-1.21	0.22	22.98	27.01	+	3.626	0.001	-4.03
	Ga0180325_113560	arginine deiminase	0.66	-1.13	28.05	23.12	+	4.617	0.000	4.93
	Ga0180325_113564	transporter, SSS family	-0.76	-1.29	24.20	22.68	+	2.486	0.007	1.53
	Ga0180325_113579	formylmethanofuran dehydrogenase, subunit E	0.46	0.04	27.52	26.50	+	2.246	0.010	1.03
	Ga0180325_113580	oligopeptidase F. Metallo peptidase. MEROPS family M03B	1.22	0.70	29.57	28.39	+	3.117	0.002	1.18
	Ga0180325_113627	GTP-binding protein	0.12	-0.27	26.59	25.59	+	2.347	0.009	1.00
	Ga0180325_113638	Nitroreductase	0.99	1.64	28.94	31.10	+	3.481	0.002	-2.16
	Ga0180325_113648	phosphoenolpyruvate carboxykinase (ATP)	1.79	0.83	31.13	28.76	+	3.471	0.002	2.37

	Ga0180325_113652	amino acid ABC transporter substrate-binding protein, PAAT family	1.05	1.96	29.13	32.02	+	5.571	0.000	-2.90
	Ga0180325_113660	Sugar phosphate permease	0.54	-1.63	27.73	21.69	+	3.502	0.002	6.04
Methyltransferase system	Ga0180325_113661	trimethylamine---corrinoid protein Co-methyltransferase	1.79	-1.75	31.13	21.36	+	4.266	0.000	9.78
QA transporters	Ga0180325_113679	amino acid/amide ABC transporter ATP-binding protein 2, HAAT family	-1.32	-0.16	22.67	25.93	+	3.065	0.002	-3.26
	Ga0180325_113683	branched-chain amino acid transport system substrate-binding protein	2.25	2.37	32.37	33.19	+	2.528	0.005	-0.82
	Ga0180325_113711	DNA gyrase subunit A	1.00	0.54	28.97	27.92	+	2.318	0.009	1.05
Pyruvate metabolism	Ga0180325_113745	pyruvate-ferredoxin/flavodoxin oxidoreductase	0.92	1.34	28.77	30.24	+	2.235	0.010	-1.47
	Ga0180325_113780	dTMP kinase	-0.15	-0.71	25.84	24.34	+	2.692	0.004	1.50
Sporulation	Ga0180325_113784	Cell fate regulator YaaT, PSP1 superfamily (controls sporulation, competence, biofilm development)	0.29	-0.11	27.06	26.06	+	3.002	0.002	1.01
Methyltransferase system	Ga0180325_113813	5-methyltetrahydrofolate--homocysteine methyltransferase	-0.39	1.98	25.21	32.06	+	3.286	0.001	-6.85
	Ga0180325_113824	solute:Na+ symporter, SSS family	1.65	-0.89	30.76	23.81	+	3.952	0.000	6.95
	Ga0180325_113825	hypothetical protein	1.24	-1.40	29.64	22.34	+	4.129	0.000	7.30
	Ga0180325_113826	hypothetical protein	-0.34	-1.23	25.33	22.83	+	2.389	0.008	2.50
Methyltransferase system	Ga0180325_113827	trimethylamine---corrinoid protein Co-methyltransferase	1.93	-1.23	31.51	22.83	+	2.766	0.004	8.68
	Ga0180325_113835	ATP-dependent Clp protease ATP-binding subunit ClpB	0.61	0.31	27.92	27.27	+	3.474	0.002	0.66
	Ga0180325_113859	NAD(P) transhydrogenase subunit beta	0.43	-0.26	27.44	25.62	+	2.285	0.009	1.81

	Ga0180325_113873	hypothetical protein	-0.13	-0.49	25.92	24.98	+	3.011	0.002	0.94
	Ga0180325_113966	2-isopropylmalate synthase	0.71	-0.25	28.21	25.66	+	3.915	0.000	2.55
QA transporters	Ga0180325_114038	amino acid/amide ABC transporter substrate-binding protein, HAAT family	1.09	2.19	29.22	32.68	+	4.563	0.000	-3.45
	Ga0180325_114073	TIR domain-containing protein	0.20	0.88	26.80	28.91	+	3.270	0.001	-2.11
	Ga0180325_114074	MTH538 TIR-like domain (DUF1863)	0.32	1.12	27.13	29.59	+	3.737	0.001	-2.46
	Ga0180325_114083	DNA sulfur modification protein DndD	1.36	1.47	29.95	30.61	+	3.336	0.001	-0.65
	Ga0180325_114095	Helix-turn-helix	0.22	0.55	26.86	27.96	+	3.246	0.002	-1.10
	Ga0180325_114144	acetylornithine deacetylase	-1.13	-0.02	23.20	26.32	+	4.660	0.000	-3.12
	Ga0180325_114161	anthranilate phosphoribosyltransferase	0.87	0.30	28.63	27.23	+	2.611	0.005	1.40
	Ga0180325_114165	Tetratricopeptide repeat-containing protein	0.55	-0.24	27.77	25.70	+	2.547	0.005	2.07
	Ga0180325_114173	Membrane protease subunit, stomatin/prohibitin family, contains C-terminal Zn-ribbon domain	0.28	0.44	27.03	27.65	+	2.711	0.004	-0.62
Choline metabolism	Ga0180325_114191	aldehyde dehydrogenase (acceptor)	-1.04	-0.04	23.44	26.27	+	3.743	0.001	-2.83
	Ga0180325_114194	PH domain-containing protein	0.04	0.86	26.36	28.86	+	2.501	0.007	-2.50
	Ga0180325_114200	formate dehydrogenase major subunit	0.75	-1.56	28.30	21.88	+	4.196	0.000	6.41
	Ga0180325_114215	Aldo/keto reductase	0.82	1.30	28.50	30.11	+	2.562	0.005	-1.61
	Ga0180325_114229	HSP20 family protein	-1.21	-0.66	22.97	24.48	+	3.095	0.002	-1.50
	Ga0180325_114273	response regulator receiver modulated diguanylate cyclase	-0.94	0.17	23.71	26.87	+	2.913	0.002	-3.16
	Ga0180325_114292	trk system potassium uptake protein TrkA	-0.13	-0.43	25.92	25.13	+	3.103	0.002	0.79

	Ga0180325_114300	hypothetical protein	-1.12	-0.55	23.21	24.79	+	2.595	0.004	-1.59
	Ga0180325_114316	lysine:proton symporter, AAT family	-0.93	-0.45	23.74	25.07	+	2.439	0.007	-1.33
	Ga0180325_114380	L-glutamine synthetase	1.03	1.98	29.06	32.07	+	2.992	0.002	-3.01
	Ga0180325_114381	glutamate synthase (NADPH/NADH) large chain	0.95	1.99	28.85	32.09	+	3.655	0.001	-3.24
	Ga0180325_114382	glutamate synthase (NADH) small subunit	0.45	1.17	27.49	29.75	+	2.814	0.003	-2.26
Sporulation	Ga0180325_114392	stage V sporulation protein G	0.37	1.28	27.27	30.06	+	2.883	0.002	-2.79
	Ga0180325_114399	peptidyl-prolyl cis-trans isomerase C/foldase protein PrsA	1.76	1.77	31.04	31.47	+	2.671	0.004	-0.43
Sporulation	Ga0180325_114400	AbrB family transcriptional regulator, stage V sporulation protein T	-1.21	0.19	22.98	26.92	+	3.346	0.002	-3.94
	Ga0180325_114401	tetrapyrrole methylase family protein / MazG family protein	1.04	0.38	29.10	27.46	+	4.181	0.000	1.64
	Ga0180325_114418	peptidyl-prolyl cis-trans isomerase B (cyclophilin B)	-1.25	-0.15	22.86	25.94	+	3.323	0.001	-3.08
	Ga0180325_114419	hypothetical protein	-1.26	0.84	22.84	28.79	+	3.186	0.002	-5.95
Chemotaxis	Ga0180325_114424	methyl-accepting chemotaxis sensory transducer with Cache sensor	-0.15	0.33	25.85	27.32	+	2.776	0.004	-1.47
	Ga0180325_114443	TRAP-type C4-dicarboxylate transport system, substrate-binding protein	-1.69	0.00	21.67	26.36	+	2.650	0.004	-4.69
Methyltransferase system	Ga0180325_114447	trimethylamine---corrinoid protein Co-methyltransferase	-1.85	-0.99	21.23	23.53	+	2.592	0.004	-2.30
Methyltransferase system	Ga0180325_114450	5-methyltetrahydrofolate--homocysteine methyltransferase	-1.28	-0.16	22.77	25.92	+	3.228	0.002	-3.15
	Ga0180325_114457	hypothetical protein	-1.13	0.25	23.20	27.10	+	2.317	0.009	-3.90

Methyltransferase system	Ga0180325_114458	trimethylamine---corrinoid protein Co-methyltransferase	-0.47	1.75	24.99	31.41	+	3.621	0.001	-6.43
Methyltransferase system	Ga0180325_114459	trimethylamine---corrinoid protein Co-methyltransferase	-0.57	1.65	24.70	31.13	+	2.788	0.004	-6.42
	Ga0180325_114460	lipoyl(octanoyl) transferase	-1.02	0.73	23.49	28.47	+	2.837	0.002	-4.98
	Ga0180325_114461	lipoic acid synthetase	-1.26	-0.14	22.84	25.98	+	2.589	0.005	-3.14
	Ga0180325_114463	glycine cleavage system H protein	-1.07	0.98	23.35	29.19	+	3.533	0.002	-5.84
	Ga0180325_114464	glycine dehydrogenase (decarboxylating) alpha subunit	-0.20	1.88	25.72	31.77	+	4.129	0.000	-6.06
	Ga0180325_114465	glycine dehydrogenase (decarboxylating) beta subunit	-0.52	1.91	24.86	31.88	+	3.473	0.002	-7.02
	Ga0180325_114466	dihydrolipoamide dehydrogenase	-1.06	0.64	23.39	28.22	+	3.605	0.001	-4.84
	Ga0180325_114485	pantothenate synthetase	-0.41	0.84	25.14	28.79	+	2.768	0.004	-3.65
	Ga0180325_114486	L-aspartate 1-decarboxylase	-1.12	0.26	23.21	27.11	+	2.519	0.005	-3.90
	Ga0180325_114506	lysyl-tRNA synthetase, class II	1.95	1.37	31.58	30.32	+	4.273	0.000	1.26
	Ga0180325_114515	protein arginine kinase activator	-0.03	0.20	26.17	26.96	+	2.723	0.004	-0.78
	Ga0180325_114535	Multimeric flavodoxin WrbA	1.39	-0.13	30.05	26.00	+	4.287	0.000	4.05
	Ga0180325_114559	transcription antitermination protein nusG	1.01	0.27	29.01	27.14	+	2.524	0.005	1.87
	Ga0180325_114576	SSU ribosomal protein S19P	0.39	0.85	27.31	28.82	+	3.497	0.002	-1.51
	Ga0180325_114579	LSU ribosomal protein L16P	-1.31	-0.05	22.71	26.24	+	3.559	0.002	-3.54
	Ga0180325_114586	LSU ribosomal protein L6P	0.95	0.42	28.84	27.57	+	2.377	0.008	1.27
	Ga0180325_114589	LSU ribosomal protein L30P	-0.15	0.39	25.86	27.50	+	2.259	0.010	-1.64
	Ga0180325_114605	LSU ribosomal protein L13P	0.50	0.75	27.62	28.53	+	2.663	0.004	-0.91

	Ga0180325_114618	Tripartite-type tricarboxylate transporter, receptor component TctC	-1.01	0.67	23.52	28.30	+	3.432	0.002	-4.77
	Ga0180325_114628	dihydroorotase/allantoinase	0.41	0.81	27.38	28.70	+	3.365	0.002	-1.32
	Ga0180325_114657	Tripartite-type tricarboxylate transporter, receptor component TctC	-1.35	-0.65	22.58	24.52	+	2.629	0.004	-1.93
	Ga0180325_114665	Tripartite-type tricarboxylate transporter, receptor component TctC	1.19	1.81	29.51	31.58	+	4.510	0.000	-2.07
	Ga0180325_114693	RHS repeat-associated core domain-containing protein	-0.86	-1.35	23.92	22.48	+	2.356	0.009	1.44
	Ga0180325_114711	glutamate N-acetyltransferase	0.24	-0.15	26.92	25.93	+	2.297	0.009	0.99
	Ga0180325_114715	argininosuccinate synthase	1.35	0.71	29.94	28.43	+	2.711	0.004	1.51
Wood-Ljungdahl pathway	Ga0180325_114722	Formate-tetrahydrofolate ligase	2.78	3.13	33.81	35.37	+	3.420	0.002	-1.57
GB metabolism; methyltransferase system	Ga0180325_114734	5-methyltetrahydrofolate--homocysteine methyltransferase	-1.04	1.69	23.45	31.23	+	2.961	0.002	-7.79
GB metabolism; methyltransferase system	Ga0180325_114735	trimethylamine---corrinoid protein Co-methyltransferase	0.61	3.05	27.93	35.13	+	5.768	0.000	-7.20
GB metabolism; methyltransferase system	Ga0180325_114736	5-methyltetrahydrofolate--homocysteine methyltransferase	1.04	3.19	29.09	35.54	+	4.820	0.000	-6.46
GB metabolism; methyltransferase system	Ga0180325_114737	trimethylamine---corrinoid protein Co-methyltransferase	-0.19	2.94	25.76	34.82	+	3.483	0.002	-9.06
GB metabolism	Ga0180325_114738	N-methylhydantoinase A/oxoprolinase/acetone carboxylase, beta subunit	-1.22	1.38	22.96	30.34	+	3.755	0.001	-7.39

GB metabolism; methyltransferase system	Ga0180325_114739	5-methyltetrahydrofolate--homocysteine methyltransferase	1.31	2.19	29.81	32.67	+	3.620	0.001	-2.85
GB metabolism; methyltransferase system	Ga0180325_114740	trimethylamine---corrinoid protein Co-methyltransferase	-0.55	1.77	24.76	31.47	+	3.442	0.002	-6.71
GB metabolism; QA transporters	Ga0180325_114741	glycine betaine transporter	-0.87	0.51	23.91	27.85	+	3.354	0.002	-3.94
GB metabolism; QA transporters	Ga0180325_114742	glycine betaine transporter	-0.17	0.66	25.79	28.28	+	3.972	0.000	-2.49
	Ga0180325_114776	sec-independent protein translocase protein TatA	0.76	-0.03	28.33	26.29	+	2.431	0.007	2.03
	Ga0180325_114794	transcriptional regulator, GntR family	0.88	1.42	28.66	30.45	+	3.337	0.001	-1.79
Sarcosine reductase complex	Ga0180325_114795	thioredoxin reductase (NADPH)	-0.92	1.40	23.75	30.39	+	3.016	0.002	-6.63
Sarcosine reductase complex	Ga0180325_114796	thioredoxin 1	-0.95	2.08	23.69	32.34	+	4.223	0.000	-8.66
Sarcosine reductase complex	Ga0180325_114799	betaine reductase	-1.46	1.72	22.31	31.33	+	3.629	0.001	-9.02
Sarcosine reductase complex	Ga0180325_114800	Fatty acid/phospholipid biosynthesis enzyme	-1.21	1.43	22.97	30.48	+	4.890	0.000	-7.51
Sarcosine reductase complex	Ga0180325_114802	glycine reductase	-1.31	1.93	22.71	31.92	+	3.653	0.001	-9.21
Sarcosine reductase complex	Ga0180325_114803	glycine reductase	-0.72	2.26	24.30	32.87	+	4.352	0.000	-8.57

	Ga0180325_114833	glycyl-tRNA synthetase beta chain	1.29	0.65	29.77	28.24	+	2.730	0.004	1.53
	Ga0180325_114840	glutamate dehydrogenase (NADP+)	2.96	0.64	34.31	28.22	+	4.654	0.000	6.09
	Ga0180325_114842	glutamate dehydrogenase (NADP)	1.61	-0.13	30.63	25.99	+	3.256	0.001	4.64
Sarcosine reductase complex	Ga0180325_114859	thioredoxin 1	-0.95	2.08	23.69	32.34	+	4.223	0.000	-8.66
	Ga0180325_114937	CRISPR-associated autoregulator, Cst2 family	0.40	1.01	27.35	29.28	+	4.137	0.000	-1.94
	Ga0180325_114938	CRISPR-associated protein, Cas5 family	-0.87	-1.55	23.90	21.92	+	2.239	0.010	1.97
	Ga0180325_114976	osmoprotectant transport system substrate-binding protein	-1.02	0.52	23.50	27.88	+	4.274	0.000	-4.37
Chemotaxis	Ga0180325_114987	methyl-accepting chemotaxis sensory transducer	0.10	-0.31	26.53	25.48	+	2.933	0.002	1.05
	Ga0180325_115009	pyruvate carboxylase	-1.11	1.54	23.25	30.80	+	5.262	0.000	-7.55
	Ga0180325_115066	Predicted Fe-Mo cluster-binding protein, NifX family	-0.44	0.92	25.06	29.01	+	3.913	0.000	-3.95
	Ga0180325_115067	Pyruvate/2-oxoglutarate dehydrogenase complex, dihydrolipoamide dehydrogenase (E3) component	1.11	1.62	29.29	31.04	+	4.359	0.000	-1.75
S-layer	Ga0180325_115133	S-layer homology domain-containing protein	1.64	0.33	30.72	27.33	+	3.440	0.002	3.39
	Ga0180325_115134	protein of unknown function (DUF4430)	0.70	-0.11	28.16	26.05	+	4.058	0.000	2.10
	Ga0180325_115136	energy-coupling factor transport system ATP-binding protein	0.69	-1.31	28.13	22.60	+	3.033	0.002	5.53
	Ga0180325_115137	Prenyltransferase and squalene oxidase repeat-containing protein	1.28	0.55	29.73	27.94	+	2.391	0.008	1.79
	Ga0180325_115143	Xylose isomerase-like TIM barrel	0.35	-0.97	27.22	23.60	+	2.425	0.007	3.62

	Ga0180325_115155	peptide/nickel transport system substrate-binding protein	0.77	1.15	28.35	29.67	+	2.254	0.010	-1.32
QA transporters	Ga0180325_115163	peptide/nickel transport system substrate-binding protein	1.34	1.93	29.91	31.91	+	2.358	0.009	-2.00
	Ga0180325_115166	Ubiquinone/menaquinone biosynthesis C-methylase UbiE	-0.38	0.66	25.24	28.28	+	2.752	0.004	-3.04
	Ga0180325_115191	peptide/nickel transport system substrate-binding protein	0.97	1.55	28.91	30.83	+	3.085	0.002	-1.92
	Ga0180325_115198	formylmethanofuran dehydrogenase subunit E	1.87	0.93	31.35	29.04	+	4.005	0.000	2.30
Methyltransferase system	Ga0180325_115284	5-methyltetrahydrofolate--homocysteine methyltransferase	-1.25	0.27	22.86	27.16	+	3.570	0.001	-4.30
	Ga0180325_115323	Protein of unknown function (DUF1638)	1.51	1.92	30.38	31.90	+	3.987	0.000	-1.52
	Ga0180325_115385	Benzoyl-CoA reductase/2-hydroxyglutaryl-CoA dehydratase subunit, BcrC/BadD/HgdB	-1.03	0.19	23.46	26.93	+	2.486	0.007	-3.47
	Ga0180325_115386	fructokinase/ribokinase	-0.97	0.07	23.63	26.58	+	2.789	0.004	-2.96
Methyltransferase system	Ga0180325_115389	methyltransferase cognate corrinoid proteins	-1.13	0.76	23.19	28.56	+	3.497	0.002	-5.37
Methyltransferase system	Ga0180325_115390	uroporphyrinogen decarboxylase	-0.95	0.47	23.68	27.73	+	4.399	0.000	-4.05
Methyltransferase system	Ga0180325_115393	uroporphyrinogen decarboxylase	-1.12	1.54	23.22	30.80	+	4.716	0.000	-7.57
Methyltransferase system	Ga0180325_115394	methyltransferase cognate corrinoid proteins	-0.59	2.11	24.67	32.45	+	3.012	0.002	-7.78
	Ga0180325_115396	Uncharacterized 2Fe-2 and 4Fe-4S clusters-containing protein, contains DUF4445 domain	-1.03	0.72	23.47	28.44	+	4.022	0.000	-4.97
	Ga0180325_115397	histidinol-phosphate aminotransferase	2.22	0.33	32.31	27.31	+	5.315	0.000	4.99
	Ga0180325_115398	indolepyruvate ferredoxin oxidoreductase alpha subunit	1.47	-0.73	30.25	24.27	+	4.555	0.000	5.98

	Ga0180325_115399	indolepyruvate ferredoxin oxidoreductase beta subunit	1.34	-0.93	29.91	23.70	+	3.738	0.001	6.21
	Ga0180325_115400	acetyltransferase	1.14	-0.99	29.36	23.54	+	3.679	0.001	5.82
Chemotaxis	Ga0180325_115401	methyl-accepting chemotaxis sensory transducer with Cache sensor	1.12	0.58	29.31	28.05	+	2.907	0.002	1.26
	Ga0180325_115405	peptide/nickel transport system permease protein	0.06	-1.24	26.43	22.82	+	2.283	0.009	3.61
	Ga0180325_115410	predicted aconitase subunit 2	-0.02	-1.29	26.22	22.66	+	3.451	0.002	3.56
Methyltransferase system	Ga0180325_115414	trimethylamine---corrinoid protein Co-methyltransferase	1.48	-1.49	30.29	22.10	+	3.170	0.002	8.20
	Ga0180325_115416	indolepyruvate ferredoxin oxidoreductase alpha subunit	1.69	-1.27	30.87	22.74	+	3.549	0.002	8.13
	Ga0180325_115417	indolepyruvate ferredoxin oxidoreductase beta subunit	0.98	-1.38	28.93	22.40	+	3.627	0.001	6.53
	Ga0180325_115418	acetyltransferase	1.19	-1.43	29.49	22.26	+	4.824	0.000	7.23
	Ga0180325_115419	histidinol-phosphate aminotransferase	1.80	-1.56	31.16	21.90	+	3.542	0.002	9.26
Methyltransferase system	Ga0180325_115420	5-methyltetrahydrofolate--homocysteine methyltransferase	1.18	-1.27	29.47	22.74	+	2.498	0.007	6.73
	Ga0180325_115421	Acetoacetate decarboxylase	0.96	-1.02	28.88	23.44	+	2.416	0.008	5.44
	Ga0180325_115423	peptide/nickel transport system permease protein	0.06	-1.24	26.43	22.82	+	2.283	0.009	3.61
Methyltransferase system	Ga0180325_115430	trimethylamine---corrinoid protein Co-methyltransferase	0.54	-0.75	27.73	24.21	+	2.473	0.007	3.52
Methyltransferase system	Ga0180325_115483	trimethylamine:corrinoid methyltransferase	1.44	2.17	30.19	32.62	+	3.616	0.001	-2.44
Methyltransferase system	Ga0180325_115486	5-methyltetrahydrofolate--homocysteine methyltransferase	1.31	2.19	29.81	32.67	+	3.620	0.001	-2.85
	Ga0180325_115511	branched-chain amino acid transport system substrate-binding protein	-1.13	-0.03	23.19	26.28	+	2.449	0.007	-3.08

	Ga0180325_115530	peptide/nickel transport system substrate-binding protein	0.04	1.18	26.37	29.77	+	5.241	0.000	-3.40
	Ga0180325_115572	hypothetical protein	-1.32	0.44	22.68	27.63	+	4.756	0.000	-4.95
	Ga0180325_115573	hypothetical protein	-0.03	1.09	26.19	29.50	+	3.474	0.002	-3.31
	Ga0180325_115578	hypothetical protein	-1.48	0.23	22.23	27.03	+	2.894	0.002	-4.79
	Ga0180325_115645	O-acetylhomoserine sulfhydrylase	0.98	0.48	28.93	27.75	+	2.779	0.004	1.18
Energy conservation	Ga0180325_115678	NADH dehydrogenase subunit B	0.12	1.29	26.59	30.07	+	3.353	0.002	-3.48
Energy conservation	Ga0180325_115679	NADH dehydrogenase subunit C	0.30	1.14	27.09	29.65	+	3.360	0.002	-2.56
Energy conservation	Ga0180325_115680	NADH dehydrogenase subunit D	1.23	2.06	29.61	32.31	+	2.969	0.002	-2.70
Energy conservation	Ga0180325_115681	NADH dehydrogenase subunit H	-0.02	0.78	26.22	28.62	+	2.352	0.009	-2.40
Energy conservation	Ga0180325_115682	NADH dehydrogenase subunit I	0.05	0.94	26.41	29.08	+	4.216	0.000	-2.67
Energy conservation	Ga0180325_115686	NADH dehydrogenase subunit M	-0.16	1.03	25.82	29.34	+	4.724	0.000	-3.52
	Ga0180325_115737	Amino acid transporter	-1.06	0.55	23.37	27.97	+	4.418	0.000	-4.59
Corrinoid biosynthesis	Ga0180325_115764	uroporphyrinogen III methyltransferase / synthase	0.74	0.28	28.27	27.19	+	2.778	0.004	1.07
Corrinoid biosynthesis	Ga0180325_115782	adenosylcobinamide kinase /adenosylcobinamide-phosphate guanylyltransferase	0.51	0.04	27.64	26.48	+	2.998	0.002	1.16
Oxidative stress	Ga0180325_115824	Flavorubredoxin	1.15	0.37	29.38	27.43	+	4.446	0.000	1.95
	Ga0180325_115882	hydrophobic/amphiphilic exporter-1, HAE1 family	1.21	0.52	29.56	27.87	+	2.760	0.004	1.68
	Ga0180325_115932	hypothetical protein	-0.34	1.37	25.35	30.31	+	2.775	0.004	-4.96
	Ga0180325_115935	4-aminobutyrate aminotransferase apoenzyme	1.08	-1.52	29.21	22.00	+	3.016	0.002	7.22

Table S6 Classification of Illumina 16S rRNA gene amplicon sequencing samples in relation to the amount of substrate consumed.

Classification	Substrate	Time (d)	Replicate
Start	DCM	0	Inoculum
		0	A, B
	Glycine betaine	0	A, B, C
	Choline	0	A, B, C
	Methanol	0	A, B, C
Pre	DCM	14, 21	A, B
	Glycine betaine	7	A, B
	Choline	7	A, B, C
		11	A
	Methanol	14	A, C
		19	A
Early	DCM	25	A, B
	Glycine betaine	7	C
		11	A, B
	Choline	11	B, C
		15	A
	Methanol	23	A
Mid	DCM	29	A, B
	Glycine betaine	11	C
		15	A, B
	Choline	15	B, C
		21	A
	Methanol	14	B
		23	C
		26	A
Late	DCM	35	A, B
	Glycine betaine	21	A, B, C
	Choline	21	B, C
		25	A, B
		28	A
	Methanol	19	B
		26	C
		30	A, C
Post	Glycine betaine	28	A, B, C
	Choline	25	C
		28	B, C
	Methanol	26	B

Appendix A

Permission from the Dharawal Language Program to use of the placename Warabiya as the species name for the novel bacterium, strain DCMF.

EASTERN ZONE GUJAGA ABORIGINAL CORPORATION

Professor Michael J Manefield
School of Chemical Engineering
School of Civil and Environmental Engineering
University of New South Wales

Email: manefield@unsw.edu.au

Re: *Formamonas warabiya*

Dear Professor Manefield,

In regard to your request to use a placename in the binomial name for a newly discovered bacterium, the Dharawal Language Program has identified the local name for the geographical area where the bacteria were isolated.

Warabiya (wa-ra-bee-ya) has been recorded as the approximate area between Bunnerong Creek near the paper mill on Port Botany Road and the suburb of Botany. Following our selection criteria for accepting words in the language revitalisation process, *warabiya* has been accepted as the place name for the geographical location described above.

Eastern Zone Gujaga Aboriginal Corporation (**Gujaga**) holds Indigenous cultural intellectual copyright for Dharawal language and cultural knowledge on behalf of the Dharawal community. Gujaga, through the Dharawal Language Program has been leading language and culture activities in the La Perouse Aboriginal Community for 18 years and is the only community-based language reclamation program in the coastal Sydney area.

Gujaga supports the use of the placename, Warabiya, to be used as the species name for the research thesis, in publications and related promotional material associated with the identified research project. Gujaga grants a non-exclusive licence to Professor Michael Manefield to use the information for the agreed purpose only.

The Dharawal Language Program should be identified as the provider of the Dharawal language and cultural information on all materials relating to this resource. This will ensure cultural affirmation amongst stakeholders and the Dharawal community.

Hope this helps and good luck with your future works.

Dr Shane Ingrey
Language, Culture and Research

PO BOX 102 MATRAVILLE 2036
1 ELAROO AVE LA PEROUSE 2036
EMAIL: admin@gujaga.org.au

PHONE: (02) 9661-6097
FAX: (02) 9694-1239
ABN: 28 023 435 902
ICN: 365