Anaerobic microbial metabolism of isoprene

Miriam Kronen

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

School of Civil and Environmental Engineering

Faculty of Engineering

August 2019

Thesis/Dissertation Sheet

Surname/Family Name	:	**Kronen**
Given Name/s	:	**Miriam Magdalena**
Abbreviation for degree as give in the University calendar	:	**PhD**
Faculty	:	**Engineering**
School	:	**Civil and Environmental Engineering**
Thesis Title	:	**Anaerobic microbial metabolism of isoprene**

Abstract

Isoprene (IUPAC: 2-methyl-1,3-butadiene, $CH_2 = C(CH_3)-CH = CH_2$) represents the most abundant biogenic volatile organic compound (BVOC) on Earth which is similar in magnitude to methane sources. It comprises one third of the total global BVOC emission, and influences atmospheric chemistry, leading to increasing global temperatures and raising ozone concentrations. The conjugated diene system of isoprene reacts preferentially with hydroxyl radicals ($^{\bullet}OH$) in the atmosphere resulting in complex secondary organic aerosols that effect climate and human health. The impacts of isoprene on the climate and atmosphere are well studied however its biological and biogeochemical role remains unclear. Little is known about microbiological processes serving as terrestrial sinks for isoprene. Whilst aerobic isoprene degrading bacteria have been identified, there is nothing known about anaerobic, isoprene-metabolizing organisms. Given the environmental abundance and ubiquity of isoprene, it was hypothesized that it is available to anaerobic microorganisms. In this thesis various inocula were examined for anaerobic microbial depletion of isoprene.

Under anaerobic conditions isoprene was reduced stoichiometrically to methylbutene isomers (i.e. 2-methyl-1-butene (>97%), 3-methyl-1-butene (≤2%), 2-methyl-2-butene (≤1%)). The reduction was attributed to a novel hydrogenotrophic *Acetobacterium wieringae* strain that used isoprene as a terminal electron acceptor simultaneously with HCO_3^-. In the presence of isoprene, the strain generated 40% less acetate relative to isoprene free controls but achieved a similar biomass yield. Taken together these findings suggest that isoprene reduction was coupled to energy conservation. Genomic and proteomic analysis identified a five gene operon in the genome of *A. wieringae* strain ISORED-2 that is upregulated in the presence of isoprene. One gene in this operon encodes a nickel-dependent enzyme that contains a binding site for NADH, FAD and 4Fe-4S ferredoxin and is speculated to be an isoprene reductase. Phylogenetic analysis of the putative isoprene reductase revealed that its homologs are mostly spread among Firmicutes, but could also be found in Spirochaetes, Tenericutes, Actinobacteria, Chloroflexi, Bacteroidetes and some Proteobacteria.

Methylbutene isomers were measured in the headspace of wetland samples collected from three different sample sites and results imply that isoprene reduction could be at least one of the responsible metabolisms resulting in methylbutene formation. Since isoprene reduction was found in natural environments like wetlands, it suggests its presence in other anaerobic environments. Furthermore, isoprene had an inhibitory effect on microbial methanogenesis indicating an interconnection between isoprene emission and methane biosynthesis.

FOR OFFICE USE ONLY Date of completion of requirements for Award:

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 the right to archive and to make available my thesis or dissertation in whole or part in the University libraries in all forms of media, now or here after known, subject to the provisions of the Copyright Act 1968. I retain all proprietary rights, such as patent rights. I also retain the right to use in future works (such as articles or books) all or part of this thesis or dissertation. I also authorise University Microfilms to use the 350 word abstract of my thesis in Dissertation Abstract International (this is applicable to doctoral theses only). I have either used no substantial portions of copyright material in my thesis or I have obtained permission to use copyright material; where permission has not been granted I have applied/will apply for a partial restriction of the digital copy of my thesis or dissertation.'

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. No emendation of content has occurred and if there are any minor variations in formatting, they are the result of the conversion to digital format.’

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 student contributed greater than 50% of the content in the publication and is the "primary author", ie. the student was responsible primarily for the planning, execution and preparation of the work for publication
- The student 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*

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

☐ *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 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.

Name	Signature	Date (dd/mm/yy)

Publications derived from this thesis

Miriam Kronen, Matthew Lee*, Zackary L Jones & Michael J Manefield. "Reductive metabolism of the important atmospheric gas isoprene by homoacetogens". *The ISME Journal* (2019) **13**, 1168–1182. (*co-first author). Incorporated in Chapter 2.

Contributor	**Statement of contribution**
Miriam Kronen	Conception and design (20%), analysis and interpretation of data (35%), drafting the article (100%).
Matthew Lee	Conception and design (45%), analysis and interpretation of data (35%), critically revising article so as to contribute to the interpretation (60%).
Zackary L Jones	Conception and design (0%), analysis and interpretation of data (10%), critically revising article so as to contribute to the interpretation (5%).
Michael J Manefield	Conception and design (35%), analysis and interpretation of data (20%), critically revising article so as to contribute to the interpretation (35%).

Other contributions

Dr. Xabier Vázquez-Campos assembled the genomes of the MAGs ISORED-1 and ISORED-2 and helped with bioinformatics analysis in Chapter 3 and 4.

Abstract

Isoprene (IUPAC: 2-methyl-1,3-butadiene, $CH_2 = C(CH_3)\text{-}CH = CH_2$) represents the most abundant biogenic volatile organic compound (BVOC) on Earth which is similar in magnitude to methane sources. It comprises one third of the total global BVOC emission, and influences atmospheric chemistry, leading to increasing global temperatures and raising ozone concentrations. The conjugated diene system of isoprene reacts preferentially with hydroxyl radicals ($^{\bullet}OH$) in the atmosphere resulting in complex secondary organic aerosols that effect climate and human health. The impacts of isoprene on the climate and atmosphere are well studied; however its biological and biogeochemical role remains unclear. Little is known about microbiological processes serving as terrestrial sinks for isoprene. Whilst aerobic isoprene degrading bacteria have been identified, there is nothing known about anaerobic, isoprene-metabolizing organisms. Given the environmental abundance and ubiquity of isoprene, it was hypothesized that it is available to anaerobic microorganisms. In this thesis various inocula were examined for anaerobic microbial depletion of isoprene.

Under anaerobic conditions isoprene was reduced stoichiometrically to methylbutene isomers (i.e. 2-methyl-1-butene (>97%), 3-methyl-1-butene (≤2%), 2-methyl-2-butene (≤1%)). The reduction was attributed to a novel hydrogenotrophic *Acetobacterium wieringae* strain that used isoprene as a terminal electron acceptor simultaneously with HCO_3^-. In the presence of isoprene, the strain generated 40% less acetate relative to isoprene free controls but achieved a similar biomass yield. Taken together these findings suggest that isoprene reduction was coupled to energy conservation. Genomic and proteomic analysis identified a five gene operon in the genome of *A. wieringae* strain ISORED-2 that is upregulated in the presence of isoprene. One gene in this operon encodes a nickel-dependent enzyme that contains a binding site for NADH, FAD and 4Fe-4S ferredoxin and is speculated to be an isoprene reductase. Phylogenetic analysis of the putative isoprene reductase revealed that its homologs are mostly spread among Firmicutes, but could also be found in Spirochaetes, Tenericutes, Actinobacteria, Chloroflexi, Bacteroidetes and some Proteobacteria.

Methylbutene isomers were measured in the headspace of wetland samples collected from three different sample sites and results imply that isoprene reduction could be at least one of the responsible metabolisms resulting in methylbutene formation. Since isoprene reduction was found in natural environments like wetlands, it suggests its presence in other anaerobic environments. Furthermore, isoprene had an inhibitory effect on microbial methanogenesis indicating an interconnection between isoprene emission and methane biosynthesis.

Acknowledgments

I would like to thank Professor Michael Manefield for the opportunity to pursue my PhD in Australia, for funding me and for proving me with an exciting research topic that led to many new discoveries. Mike, I highly appreciate your great enthusiasm for the project and honor your style of guidance in creating space for creativity and self-development.

Then I would like to thank my joint-supervisor Dr Matthew Lee for his constant support throughout my whole PhD, scientifically and personally. Matt, even when times were tough, I could always rely on you for encouragement, scientific direction, life advices and funny conversations in "German" to keep me going. Thank you for everything.

This whole journey would have been not even half as much fun without my dear friend Sophie. Thank you for always being there, in good and in bad times, and for sharing all these moments together taking care of our thousands of babies. I think the best part of my PhD was to do it with you. This is fine.

Sally and Josie, our lunches were often the best part of the day. Thank you so much for these funny and wonderful moments we shared together, for your support in any life situation, in the PhD and the personal world.

I am very thankful to have met Xabier Vázquez-Campos whose amazing bioinformatic skills boosted my PhD project and I cannot thank you enough for the time, thoughts and ideas you have contributed to my PhD project.

Further I would like to thank the rest of Manefield crew especially Önder for his experimental advice and support and meaningful conversations about politics and life. I thank Miaomiao for her constant support and friendship during the whole PhD journey. I thank Miao (Iris) for her help in bioinformatics as well as her life advices and for being a living example of someone doing a second PhD. I thank James Bevington for his great enthusiasm on Mars research and for always coming up with some highly interesting research ideas to discuss that lighten up the day. I also thank Alumni members such as Dr Rob, Emile and Mac for great outdoor adventures and a lot of fun during the first years of my PhD and I thank Zack for his guidance in bioinformatics.

I thank Achim for discussing science together in the most interesting ways.

Many thanks also to Dr Ling Zhong for her assistances in mass spectrometry in BMSF-UNSW and for her guidance and help in proteomic analysis.

I feel very thankful for my friends Kerstin, Alex and Dominik who I can always rely on and who can understand best what it feels like in the PhD world. Having you here with me in Australia is one of the best things in my life and I am so lucky to have you as my friends. Thanks for always being there <3.

Thank you to my friends for sticking with me during this difficult time especially, Alice, Ria, Ben, Danielle, Kathrin, Erika, Alex, Toby, Jack, Lucy, Anna, Vivienne, Kassandra, Elisabetta and Karl.

Thank you to the very special Mali, her love and companionship were the light of every day <3.

A very special thanks to Kilian who, with great perseverance, supported me during the most difficult time of this thesis and whose great intellect and coding skills helped me so many times when I was stuck. Thank you from the bottom of my heart <3.

Thank you to my family for encouraging me in all of my pursuits and for always being there for me <3.

Lastly, I am forever thankful to my parents for giving me the opportunities that have made me who I am. They have always supported me throughout my life and this whole journey would not have been possible if not for them, and I dedicate this thesis to them <3<3.

"Fluctuat nec mergitur"

Table of Contents

List of Figures

List of Tables

Abbreviations

AAI	Average amino acid identity
ANI	Average nucleotide identity
ATP	Adenosine triphosphate
BLAST	Basic local alignment search tool
BVOC	Biogenic volatile organic compound
CCN	Cloud condensation nuclei
CoA	Coenzyme A
COG	Clusters of orthologous groups
DGGE	Denaturing gradient gel electrophoresis
DNA	Deoxyribonucleic acid
DSMZ	Deutsche Sammlung von Mikroorganismen und Zellkulturen
FAD	Flavin adenine dinucleotide
GC-FID	Gas chromatography - Flame ionization detector
GC-PDD	Gas chromatography - Pulsed discharge detector
HPLC	High-performance liquid chromatography
kDA	Kilodalton
LC MS	Liquid chromatography–mass spectrometry
MAG	Metagenome-assembled genome
MEP	Methylerythritol phosphate
MS	Mass spectrometry
MVA	Mevalonic acid
NAD(P)H	Nicotinamide adenine dinucleotide (phosphate)
NCBI	National Center for Biotechnology Information
OTU	Operational taxonomic unit
PCR	Polymerase chain reaction
PPi	Inorganic pyrophosphate
qPCR	Quantitative polymerase chain reaction
RNA	Ribonucleic acid
ROS	Reactive oxygen species
RT-PCR	Reverse transcription polymerase chain reaction
SOA	Secondary organic aerosols
UPLC	Ultra-performance liquid chromatography

Chapter 1 Literature review

1.1 Isoprene emission and sources

Terrestrial ecosystems emit numerous biogenic volatile organic compounds (BVOCs), that affect atmospheric chemistry and its chemical composition (Kesselmeier and Staudt, 1999; Atkinson and Arey, 2003; Laothawornkitkul *et al.*, 2009). BVOCs are produced by different environmental sources (e.g. soil microbes, animals, leaf litter) but most of the total BVOCs are released from foliage (Sharkey and Singsaas, 1995; Fall and Copley, 2000; Greenberg *et al.*, 2012; Guenther *et al.*, 2012). The dominant emission is isoprene (IUPAC: 2-methyl-1,3-butadiene (CH_2=C(CH_3)-CH=CH_2)) an unsaturated hydrocarbon with a molar mass of 68.11 g/mol and a boiling point of 34°C (Kesselmeier and Staudt, 1999; Sanadze, 2004; Sharkey and Monson, 2017; McGenity *et al.*, 2018). An annual global terrestrial isoprene emission of 500-600 Tg y^{-1} (Guenther *et al.*, 2006, 2012) and an oceanic emission of 0.1-1.2 Tg y^{-1} (Alvarez *et al.*, 2009; Shaw *et al.*, 2010) has been estimated, which is similar in magnitude to methane sources at 526-569 Tg y^{-1} (Kirschke *et al.*, 2013). Isoprene represents about one third of the total global BVOC flux (Guenther *et al.*, 2012).

Isoprene emission from plants was first discovered by Sanadze, (1957) who found that isoprene is formed by photosynthesizing mesophyll cells in leaves. Rasmussen and Went (1965) confirmed these results independently and both eventually measured isoprene by using mass spectrometry (Sharkey and Monson, 2017; Sanadze, 2004). Wiedinmyer *et al.* (2004) described enclosure measurements of foliar emissions of isoprene and other BVOCs and compiled this information into an online database (http://bai.acom.ucar.edu/Data/BVOC/). This database documents ~1800 plant species from which BVOC emission has been studied. Isoprene is mainly emitted by woody plants of which genera *Quercus* (oaks) and *Populus* (poplars) (Harley *et al.*, 1999) have the highest rate of emission. Eucalyptus trees also emit significant amounts of isoprene (He *et al.*, 2000; Winters *et al.*, 2009). Besides trees, isoprene is also released from shrubs, grasses, sedges (Guenther *et al.*, 2006, 2012) and ferns (Tingey *et al.*, 1987). Together, terrestrial plants are considered as the main source of (atmospheric) isoprene in the isoprene budget (Sharkey *et al.*, 2008; McGenity *et al.*, 2018).

Apart from plants, many soil bacteria are known to produce isoprene including numerous Gram-positive (e.g., *Bacillus* species) and Gram-negative bacteria (e.g.,

Escherichia coli, *Pseudomonas fluorescens*, various Actinomycetes *Pseudonocardia*, *Saccharomonospora*, *Streptomyces*, *Thermomonospora*) (Kuzma *et al*., 1995; Schöller *et al*., 1997, 2002; Wagner *et al*., 1999; Fall and Copley, 2000; Effmert *et al*., 2012). Isoprene emission was also detected in root-associated fungi (Bäck et al., 2010). In marine environments, photosynthetic microalgae (chlorophytes, coccolithophores, haptophytes, cyanobacteria, nitrogen fixers, diatoms, dinoflagellates, picoeukaryotes) are the dominant source of isoprene (Shaw *et al*., 2010) besides macroalgae and sea weed (Broadgate *et al*., 2004). In contrast to marine atmospheres, freshwater biomes have received little attention for their BVOC contribution. Recently it has been found that freshwater lakes, via microalgae, also emit isoprene to the atmosphere (Steinke *et al*., 2018). In wetlands and subarctic peatlands isoprene is emitted from aquatic plants such as different mosses (Janson and De Serves, 1998; Janson *et al*., 1999; Tiiva *et al*., 2007) or sedge species (Tiiva *et al*., 2007; Ekberg *et al*., 2009).

Moreover, human activity is changing global isoprene emissions. Natural forests are converted to agricultural land due to increasing demands for food crops and biofuels. Cultivation of biofuel crops, especially oil palm, affects isoprene emissions from vegetation (Armstrong, 2013, Ashworth *et al*., 2012). For example, in South-East Asia, oil palm plantations emit five times more isoprene than the native rainforest (Fowler *et al*., 2011). However, when isoprene-emitting forest and grasslands were replaced with food crops, particularly in Brazil and sub-Saharan Africa, isoprene emission was reduced. Depending on the balance in future agricultural cultivation trends, it is estimated that by 2030, the global isoprene emission could increase by up to 1.4%, or decline by up to 1.7% (Armstrong, 2013). Furthermore, isoprene is produced anthropogenically as a by-product in the oil industry for the synthesis of plastics and rubber and as a component of roasted coffee and tobacco smoke (van Ginkel *et al.*, 1987). Isoprene is also the main hydrocarbon found in human breath and many animals produce isoprene (Gelmont *et al.*, 1981; King *et al.*, 2010).

Isoprene is highly reactive due to its carbon-carbon double bonds which react with atmospheric oxidants and has a reaction lifetime of 1.4 h, 0.8 h and 1.3 days when reacting with $^{\bullet}OH$, NO_3 or O_3 respectively (Levy, 1971; Monson and Holland, 2001; Atkinson and Arey, 2003; Pacifico *et al.*, 2009). It has the potential to play a dominant role on the oxidation capacity of the troposphere but its oxidation reaction chain is very complex and knowledge of these steps is still incomplete (Berndt *et al.*, 2019). The

main source for tropospheric OH radicals is generated by photolysis of ozone (O_3) in the presence of water ((1) & (2)) (Levy, 1971; Reeves *et al.*, 2002).

$$(1)\ O_3 + hv \rightarrow O(^1D) + O_2$$

$$(2)\ O(^1D) + H_2O \rightarrow 2\ OH$$

Whilst atmospheric $^{\bullet}OH$ can react with most chemical species, $^{\bullet}OH$ primarily react with carbon monoxide (CO) and hydrocarbons like isoprene (Allodi *et al.*, 2008) to produce peroxy radicals (HO_2 (3) and RO_2 (4)) that can convert NO to NO_2.

$$(3)\ OH + CO \rightarrow HO_2 + CO_2$$

$$(4)\ RH + OH + O_2 \rightarrow RO_2 + H_2O$$

$$(5)\ RO_2 + NO + O_2 \rightarrow NO_2 + R'CHO + HO_2$$

$$(6)\ HO_2 + NO \rightarrow NO_2 + OH$$

$$(7)\ HO_2 + O_3 \rightarrow OH + 2\ O_2$$

$$(8)\ OH + O_3 \rightarrow HO_2 + O_2$$

Reaction (7), (8) and (1) & (2) are fates for O_3 while reaction (5) and (6) produce NO_2 which via photolysis gives ground-state oxygen $O(^3P)$ that reacts with O_2 to form O_3 ((9) & (10)) [adapted from (Levy, 1971; Reeves *et al.*, 2002; Sharkey *et al.*, 2008)].

$$(9)\ NO_2 + h\nu \rightarrow NO + O(^3P)$$

$$(10)\ O(^3P) + O_2 \rightarrow O_3$$

Thus, in areas with high NO levels the reaction of peroxy radicals with NO to form NO_2 (reaction (5) & (6)) exceeds the total amount of O_3 destroyed via photolysis giving net photochemical production of O_3 whereas isoprene oxidation at lower NO concentrations leads to O_3 consumption (Trainer *et al.*, 1987; Wiedinmyer *et al.*, 2006; Sharkey *et al.*, 2008). High nitrogen oxide concentrations occur in air masses that receive input from fossil fuel combustion or fertilizer use. According to the 2011 US national emission inventory (https://www.epa.gov/air-emissions-inventories) 90% of NO_x emitted is of anthropogenic origin (Xu *et al.*, 2015). The resulting high concentrations of tropospheric ozone can trigger a variety of health problems including chest pain, coughing and throat irritation (Galizia and Kinney, 1999). Rapid reactions between $^{\bullet}OH$ and isoprene also result in increased levels of methane in the air as oxidation by $^{\bullet}OH$ is the main atmospheric sink for this long-lived greenhouse gas (Collins *et al.*, 2002; Pike and Young, 2009).

Oxidized isoprene in the atmosphere also forms secondary organic aerosols (SOA) (Krechmer *et al.*, 2015). Aerosols can be colloidal atmospheric solid particles, liquid droplets in air or a gas. By acting as cloud condensation nuclei (CCN), SOAs influence cloud formation and climate (Zhao *et al.*, 2017). They are linked to adverse health hazards (Pope and Dockery, 2006), cause visibility impairment by light scattering and absorbance (Jacob, 1999) and affect the Earth's climate directly (scattering absorbance) and indirectly (nucleating cloud formation) (Engelhart *et al.*, 2011). Aerosols lead to scattering which takes place by reflection, refraction or diffraction of the radiation beam influencing Earth´s albedo as light is reflected back to space. Scattering also limits visibility in the troposphere (Jacob, 1999). For example, the characteristic blue haze over the forests of the Australian Blue Mountains results from scattering of light by microscopic particles composed of isoprenoids (i.e. monoterpenes and isoprene) (Siwko *et al.*, 2007).

Previously it was assumed that the photooxidation of isoprene does not result in SOA (Pandis *et al.*, 1991), however it was later reported that isoprene is a substantial source of SOA (Claeys *et al.*, 2004; Krechmer *et al.*, 2015). Under low-NO, peroxyl radicals (RO_2) react with hydroperoxyl radicals (HO_2) and form isoprene hydroxyhydroperoxides (ISOPOOH, $C_5H_{10}O_3$). Two of the ISOPOOH isomers are atmospherically important (4,3-ISOPOOH and 1,2-ISOPOOH) (Rivera-Rios *et al.*, 2014). Further oxidation of ISOPOOH by $^{\bullet}OH$ leads to the formation of isoprene epoxydiols (IEPOX) which include trans-β-IEPOX, *cis*-β-IEPOX, δ1-IEPOX and δ4-IEPOX of which the β-isomers dominate. The reaction pathways of isoprene result in many products and are estimated to contribute at least 25-50% of the annual global average of SOA production (Liu *et al.*, 2015).

1.2 The role of isoprene

Despite the importance of isoprene in atmospheric chemistry, its biological role remains unclear (Sharkey and Monson, 2017). Curiously not all plants produce isoprene, suggesting that isoprene production serves a non-essential function or one that is carried out in alternative ways in non-emitters (Sharkey, 2013). There are different proposals on the biological function of isoprene. Isoprene emission may be a vent for unwanted energy (Sanadze, 2004) or metabolites (Rosenstiel *et al.*, 2004) or a flowering hormone (Terry *et al.*, 1995). The most widely accepted hypothesis is that isoprene emission by

plants protects them against environmental stress such as heat and therefore enhances leaf thermotolerance (Singsaas *et al.*, 1997; Sharkey *et al.*, 2008; Sharkey and Monson, 2017). Isoprene emission from leaves is light dependent (Sanadze, 2004) and due to its high volatility the residence time in the thylakoid membrane is very short. Therefore isoprene primarily protects plant membranes against transient exposure to high temperatures rather than continuous heat (Singsaas *et al.*, 1999). Isoprene-producing plants use carbon directly from the Calvin Cycle and emit 1-2% of their fixed carbon as isoprene (Sanadze, 2004; Sharkey *et al.*, 2008).

Isoprene is also thought to play an important role in the protection against ozone and other reactive oxygen species (ROS) in plants. It can prevent damage after ozone exposure and loss of photosynthetic capacity by ROS (Loreto and Velikova, 2001). The mechanism how isoprene protects plants from heat or ROS remains to be resolved. It is thought that isoprene reduces heat damage directly rather than acting only through quenching ROS generated from heat stress. The loss of membrane integrity could lead to enhanced levels of malondialdeyhyde, a membrane oxidation product (Velikova and Loreto, 2005; Sharkey *et al.*, 2008; Sharkey and Monson, 2017). Thus it is possible that isoprene protects the membrane integrity by stabilizing the lipids and therefore reduces the likelihood of the membranes to undergo a heat induced phase transition. Isoprene intercalates into small free volumes in the membrane increasing the order of the lipids and thereby preventing fluidity. At lower temperatures isoprene is lost quickly preventing the membranes from freezing. Thus isoprene is well suited to act as a transient modulator of membrane order (Siwko *et al.*, 2007). Although isoprene protects against both heat and ROS, the physiology of isoprene emission appears to be related to protection against heat. It seems improbable that ozone stress was the major selective force during evolution because NO and O_3 only became high enough to damage plants since the onset of the combustion-engine (Jacob, 1999; Lerdau, 2007).

1.3 Isoprene biosynthesis

In plants isoprene is produced as the first end-product of the methylerythritol phosphate (MEP) pathway by the action of an isoprene synthase (IspS; EC 4.2.3.27). IspS catalyzes isoprene formation via elimination of inorganic pyrophosphate (PPi) from dimethylallyl pyrophosphate (DMAPP) [**Figure 1.1**] (Silver and Fall, 1991) .

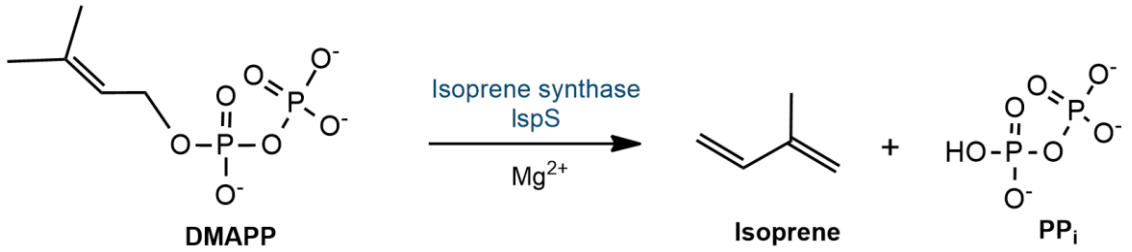

Figure 1.1 The isoprene synthase reaction. The reaction is thought to proceed through the delocalized carbon cation intermediate followed by proton abstraction (adapted from Silver & Fall, 1991).

Dimethylallyl pyrophosphate and isopentenyl diphosphates (IPP) are the precursor molecules of all isoprenoids in the living cell. There are two biosynthetic pathways for the synthesis of DMAPP: the methyl-erythritol-4-phosphate (MEP) pathway and the mevalonic acid (MVA) pathway [**Figure 1.2**].

The MEP pathway is of bacterial origin and present in most Gram-negative bacteria including *Escherichia coli* (Rohmer *et al.*, 1993), the Gram-positive bacteria *Bacillus subtilis* (Wagner *et al.*, 2000) and *Mycobacterium tuberculosis* (Putra *et al.*, 1998), plant plastids (Vranová *et al.*, 2013) and unicellular algae (Schwender *et al.*, 1996). MVA mostly functions in eukaryotes like yeasts and fungi (Disch and Rohmer, 1998), plants (Vranová *et al.*, 2013), some Gram-positive bacteria (*Enterococci*, *Staphylocci* and *Streptococci*) (Wilding *et al.*, 2000) and animals (Lange *et al.*, 2000). However, with the notable exception of some plants and some bacteria (*Streptomyces aeriouvifer*) most organisms only use one pathway (Vranová *et al.*, 2013).

Isoprene production in plants has been well studied but only in 1995 Kuzma *et al.* accidentally discovered isoprene production in prokaryotes. During the attempt to use expression cloning of a plant isoprene synthase gene, control experiments showed iso prene production in *Escherichia coli.* Kuzma discovered more species that emit isoprene but found *Bacillus subtilis* to be the highest isoprene producer (Kuzma *et al.*, 1995; Fall and Copley, 2000). Isoprene formation by various actinomycetes, including *Saccharomonospora*, *Streptomyces* and *Thermomonospora* isolates has also been detected (Schöller *et al.*, 2002). The biosynthetic pathway for isoprene formation in *B. subtilis* has been studied in detail and it appears that isoprene is a product of the MEP pathway [**Figure 1.2**].

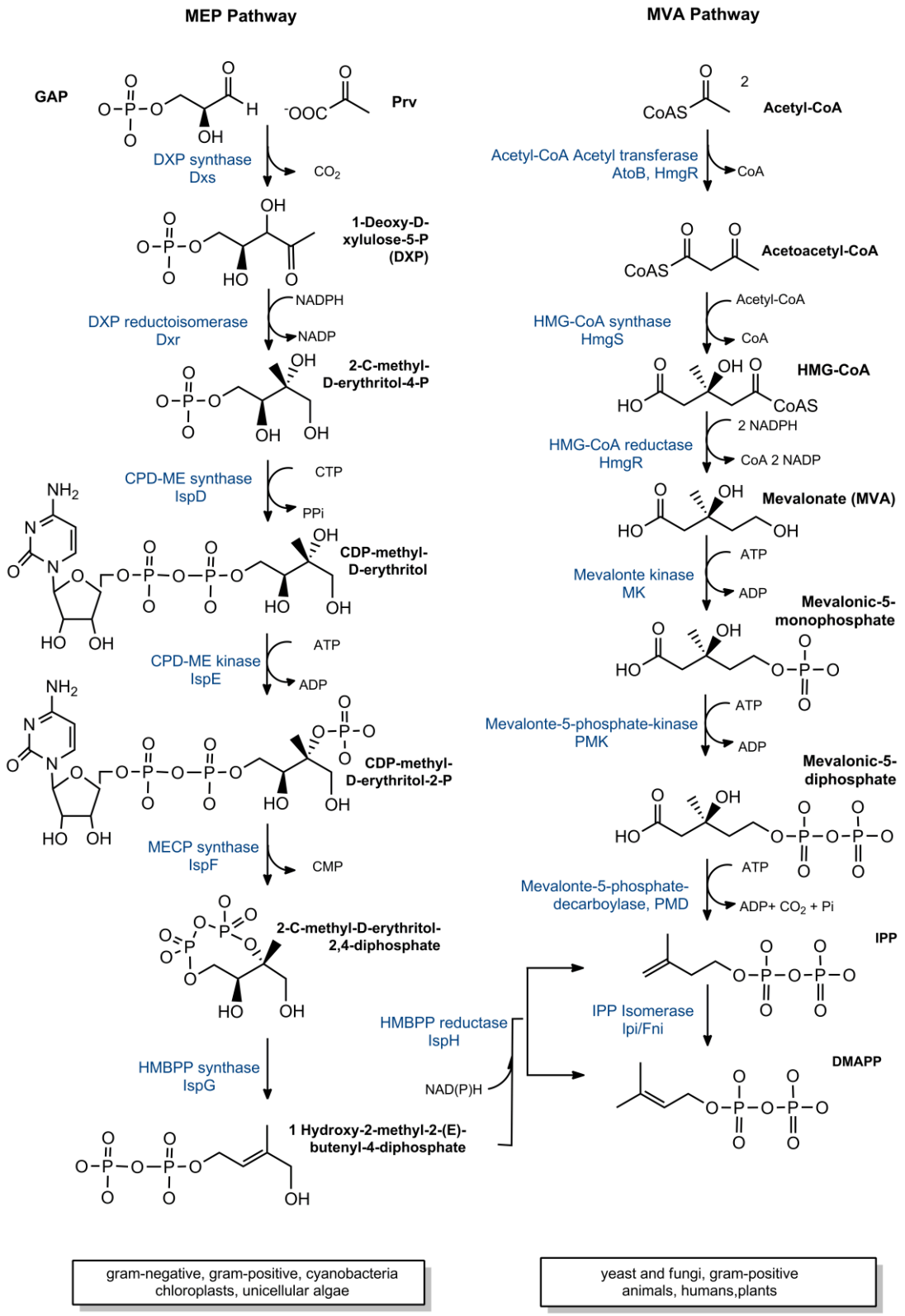

Figure 1.2 Mevalonic acid (MVA) and methylerythritol phosphate (MEP) biosynthetic pathways. Both pathways lead to the formation of isopentenyl diphosphates (IPP) and dimethylally pyrophosphate (DMAPP). DMAPP is the precursor for all isoprenoids in the living cell as well as isoprene [adapted from (Lange *et al.*, 2000; Zurbriggen *et al.*, 2012)].

An enzyme that catalyzes DMAPP dependent formation of isoprene has been detected and partially purified from *B. subtilis*. Thus, isoprene is thought to be formed from DMAPP in *B. subtilis* as in the case of plant chloroplasts by the isoprene synthase (Sivy *et al.,* 2002) [**Figure 1.1**]. A search for a homologous protein of the isoprene synthase from the poplar tree in the genome of *B. subtilis* did not yield a candidate (Julsing *et al.*, 2007) and until now the encoding gene could not be identified. However, plant isoprene synthases have been cloned and characterized from poplar (*Populus alba*, *Populus tremuloide* and *Populus nigra*) (Miller *et al.*, 2001; Sasaki *et al.*, 2005). The plant IspS enzyme has been used upon heterologous expression in microorganisms for commercial isoprene production for renewable fuel development (Zhao *et al.*, 2011). The function of emitted isoprene in bacteria has been postulated as being a signal molecule in the natural environment of the microorganism. Another possible explanation for isoprene emission is the efflux as an overflow metabolite in the bacterial pathway to isoprenoid structures (Fall and Copley, 2000).

1.4 Aerobic isoprene degradation

While the sources of atmospheric isoprene have been well-studied the terrestrial sinks of isoprene remain poorly understood. It is thought that soil, containing isoprene degrading microorganisms, may serve as an important sink of isoprene as shown by (Cleveland and Yavitt, 1997). They estimated the global soil isoprene sink at 20.4 Tg yr^{-1} which is a small fraction (4%) of the estimated annual biogenic emission of isoprene (Guenther *et al.*, 1995; Pegoraro *et al.*, 2005). Ground level isoprene concentrations are relatively low and range between 3-35.8 ppbv (Cleveland and Yavitt, 1998;Wiedinmyer *et al.*, 2005). In most studies on the fate of isoprene, higher isoprene concentration were used including 2-200 ppbv (Gray *et al.*, 2015), 508 ppbv (Cleveland and Yavitt, 1998), 0.581 mM (Alvarez *et al.*, 2009), 2 mM (van Hylckama Vlieg *et al.*, 1998) and 5 mM (Crombie *et al.*, 2015). But even at high concentrations the ability for the microbes to consume isoprene is not satiated suggesting the isoprene consuming microbial community is large relative to the isoprene supply or/and the enzymes are very efficient (Gray *et al.*, 2015).

A range of bacteria have been described that aerobically grow on isoprene as their sole carbon and energy source. These mainly belong to Actinobacteria, including *Arthrobacter* sp. (Cleveland and Yavitt, 1998), *Nocardia* sp. (van Ginkel *et al.*, 1987),

Mycobacterium sp. (Shennan, 2006; Johnston *et al.*, 2017), *Gordonia* (Johnston *et al.*, 2017) and *Leifsonia* (Alvarez *et al.*, 2009) and *Rhodococcus* (Ewers *et al.*, 1990; van Hylckama Vlieg *et al.*, 1998; El Khawand *et al.*, 2016). But isoprene consumption was also demonstrated in other groups such as Alphaproteobacteria (*Xanthobacter* sp.; *Shinella* sp.*; Kaistobacter* sp.; *Caulobacter* sp.) as well as Gammaproteobacteria (*Rhodanobacter*) and Betaproteobacteria (*Variovorax* sp.) and Bacteroidetes (*Dyandobacter*) (Alvarez *et al.*, 2009; Gray *et al.*, 2015; Crombie *et al.*, 2018).

Gray et al. 2015 studied a wide range of microbial taxa related to isoprene degradation in soil and many of the taxa they identified are related to those that have already been associated with metabolism of other hydrocarbons. Fungi were also identified as potential isoprene degrader such as *Zygomycota, Paecilomyces carneus, Penicillium glabrum* and *Eupenicillium pinetorum* which belong to the Trichocomaceae family. Other taxa outside of the *Eurotiomycetes* family that responded positively to increasing isoprene included *Exophiala equine*.

The most detailed characterized isoprene degrading organism is *Rhodococcus* sp. strain AD45 (van Hylckama Vlieg *et al.*, 1997, 1998, 2000). Biochemical analysis of its isoprene metabolism led to the identification of isoprene epoxide (1,2-epoxy-2-methyl-3-butene) and two glutathione adducts 1-hydroxy-2-glutathionyl-2-methyl-3-butene (HGMB) and 2-glutathionyl-2-methyl-3-butenoic acid (GMBA) as the intermediate products of isoprene degradation. Genes encoding the corresponding enzymes catalyzing these reactions were also identified. The metabolism (**Figure 1.3 A**) starts with the oxidation of the methyl-substituted double bond by a monooxygenase (IsoMO) encode by the *isoABCDEF* gene cluster. This reaction yields isoprene epoxide which is a substrate for IsoI, a glutathione *S*-transferase. IsoI catalyzes the glutathione dependent opening of the epoxy ring, resulting in the stable glutathione conjugate HGMB. HGMB is then oxidized to a carboxylate GMBA by a dehydrogenase IsoH. Two additional gene products of the *iso* gene cluster were identified, IsoG which is similar to racemases/CoA-transferases and IsoJ, a second glutathione *S*-tranferase but no role in isoprene degradation could be established (van Hylckama Vlieg *et al.*, 2000).

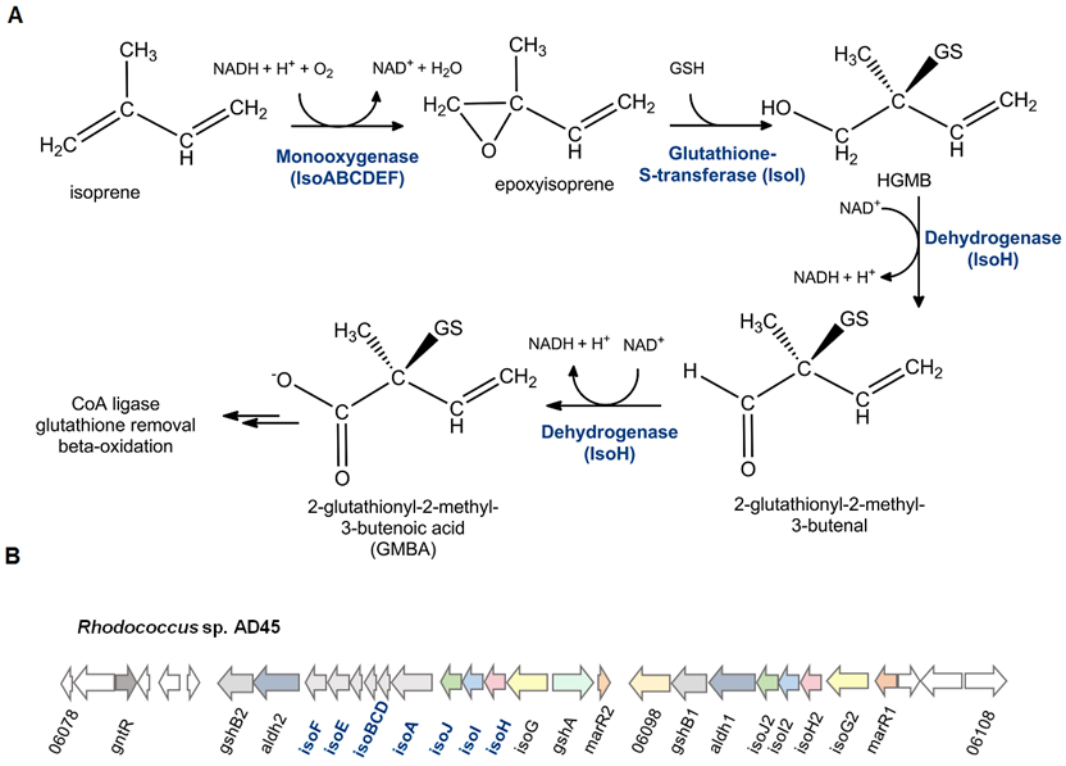

Figure 1.3 Isoprene degradation pathway as predicted by Hylckama Vlieg (A). The region of the *Rhodococcus* sp. AD45 plasmid carrying the isoprene metabolic genes (**B**). Adapted from (van Hylckama Vlieg *et al.*, 2000; Crombie *et al.*, 2015).

Crombie *et al*. (2015) identified additional isoprene metabolic genes in *Rhodococcus* sp. AD45 and published the first complete genome of an isoprene degrading isolate. Transcriptome analysis by RNAseq revealed that transcripts of 22 genes were highly upregulated by isoprene or epoxyisoprene. These 22 genes were found to be concentrated in a small region (**Figure 1.3 B**) on a megaplasmid and included the *isoABCDEFGHIJ* gene cluster as previously described (van Hylckama Vlieg et al., 2000). Furthermore, Crombie *et al*. (2015) identified additional copies of *isoGHIJ* (*isoG2H2J2I2*) approximately 11 kbp upstream the *isoABCDEFGHIJ* cluster. Also an aldehyde dehydrogenase (*aldh1*) and glutathione synthetase (*gshB1*) that are located downstream *isoJ2* and a second copy of these two genes (*aldh2* & *gshB2*) located downstream of the monooxygenase genes (*isoABCDEF*), a glutamate-cysteine ligase (*gshA*) upstream *isoG* followed by a marR-type transcriptional regulator (*marR2*) and a second marR-type regulator *(marR1)* is predicted ahead of *isoG2*. They postulated that *isoGHIJABCDEF* is co-transcribed from an operon with a promoter upstream of *isoG* (Crombie *et al.*, 2015). However, the downstream processes in the aerobic isoprene metabolism are not fully elucidated (McGenity *et al.*, 2018). The Janssen group proposed a pathway where GMBA is converted to a CoA-ester which, after the removal

of the glutathione group, enters central metabolism as acetyl-CoA and propionyl-CoA generated via beta oxidation (van Hylckama Vlieg *et al.*, 2000). Still, genes of the central metabolism that would be required for growth on acetyl-CoA and propionyl-CoA like the glyoxylate cycle genes, the malate synthase and the methylmalonyl-CoA or methylcitrate pathway-encoding genes were not found to be upregulated upon growth of *Rhodococcus* sp. AD45 on isoprene compared to growth on glucose or succinate (Crombie *et al.*, 2015).

Recently, aerobic isoprene degrading organisms were also found in the phyllosphere of isoprene-emitting trees. As previously found in soil, Gram-positive bacteria of the genus *Rhodococcus* were the dominant isoprene degraders on poplar tree leaves but a *Variovorax* strain that could grow on isoprene as the sole carbon and energy source was also isolated. Analysis of its genome confirmed that it contained isoprene metabolic genes with an identical gene organization as in *Rhodococcus* sp. AD45 (Crombie *et al.*, 2018; McGenity *et al.*, 2018).

1.5 Anaerobic isoprene and monoterpene degradation

In contrast to the aerobic isoprene degradation, there are no reports on the anaerobic degradation of isoprene, though one attempt has been reported. Schink (1985) studied the anaerobic degradation of unsaturated hydrocarbons including isoprene by methanogenic enchrichment cultures but no isoprene degradation was obtained.

However, several monoterpenes ($C_{10}H_{16}$), which consist of two linked isoprene (C_5H_8) units were found to be degraded anaerobically but only three pathways are completely known on the genetic and enzymatic level: the ones for camphor (CAM), p-cymene (CYM/CMT), and citronellol/geraniol (ATU/LIU) (Marmulla and Harder, 2014). Pinene, a main component of wood resins, is the most abundant bicyclic monoterpene with the isomers α-pinene and β-pinene. *Pseudomonas rhodesiae* and *P. fluorescens* utilize α-pinene as sole carbon source. First, α-pinene is oxidized to αpinene oxide by a NADH-dependent α-pinene oxygenase (EC 1.14.12.155) and undergoes ring cleavage catalysed by a specific α-pinene oxide lyase (EC 5.5.1.10), forming isonovalal as first product which is isomerized to novalal (Marmulla and Harder, 2014; Best *et al*., 1987; Bicas *et al*., 2008; Linares *et al*., 2009). For pinene, the gene for a key enzyme, the α-pinene oxide lyase (EC 5.5.1.10), is still unknown (Marmulla and Harder, 2014). Many monoterpenes were found to support denitrifying growth of enrichment cultures or pure

strains (Harder and Probian, 1995; Hylemon and Harder, 1998). Like isoprene, monoterpenes are BVOCs emitted by plants (Guenther *et al.*, 1995; Janson *et al.*, 1999; Kesselmeier and Staudt, 1999; Winters *et al.*, 2009) but are also a component of plant resin so monoterpenes reach the surface layers of soils by leaf fall and excreted resins (Marmulla and Harder, 2014). Also plant root systems emit monoterpenes into the rhizosphere (Lackus *et al.*, 2018).

1.6 Research objectives

Isoprene is the single largest BVOC source and a major component of biosphere–atmosphere interactions. The preceding literature review provided an overview about the importance of isoprene in the atmospheric chemistry and its biogenic sources. Despite this, little is known about microbiological processes serving as a terrestrial sink for isoprene. While aerobic isoprene degrading bacteria have been identified, there are no known anaerobic, isoprene-metabolizing organisms.

This thesis explores the anaerobic metabolism of isoprene. We searched for anaerobic microbial depletion of isoprene in different anaerobic environments in Australia for first evidence of isoprene's anaerobic fate. Given the environmental abundance and ubiquity of isoprene and since isoprene is produced by many soil bacteria and fungi, it was hypothesized that it is available to anaerobic microorganisms. To date the physiological role of isoprene in natural environments remains enigmatic. Therefore further insight into the anaerobic microbial isoprene metabolism may also improve the understanding of linkages and feedbacks between atmospheric chemistry and terrestrial microbial activity. Objectives were as follows: 1. Establish isoprene metabolizing enrichment cultures 2. Characterize genomes of anaerobic isoprene metabolizing microorganisms. 3. Identify putative anaerobic isoprene metabolizing genes using comparative proteomics. 4. Discuss the role of anaerobic isoprene metabolism in the biogeochemical cycle.

1.7 Chapter summary

Results of this thesis are organized into four research Chapters. Chapter 2 describes the successful enrichment of an anaerobic isoprene metabolizing culture from waste water treatment sludge. An H_2-consuming homoacetogenic enrichment was shown to utilize isoprene as an electron acceptor in addition to HCO_3^-. The isoprene-reducing community was dominated by *Acetobacterium* spp. and isoprene was shown to be stoichiometrically reduced to three methylbutene isomers (i.e. 2-methyl-1-butene (>97%), 3-methyl-1-butene (≤2%), 2-methyl-2-butene (≤1%)).

Chapter 3 describes the proteogenomic profiling of the isoprene reducing culture which is found to be dominated by a novel *Acetobacterium wieringae* strain. Proteomic analysis identified a five gene operon involved in isoprene reduction.

Chapter 4 explores the phylogeny and potential reaction mechanism of a putative isoprene reductase. The putative isoprene reductase may be a nickel-dependent enzyme that contains a binding site for NADH, FAD and 4Fe-4S ferredoxin. It shares sequence homology with the FAD-dependent oxidoreductase family to which the electron bifurcating NADH-dependent ferredoxin-$NADP^+$-oxidoreductase subunit B also belongs. Homologs of the putative isoprene reductase are widely distributed among anaerobic bacteria especially in the Firmicutes phylum.

Finally, Chapter 5 describes an investigation of isoprene reduction in natural environments and its effect on methanogenesis.

1.8 References

Allodi, M.A., Kirschner, K.N., and Shields, G.C. (2008) Thermodynamics of the hydroxyl radical addition to isoprene. *J. Phys. Chem. A* **112**: 7064–7071.

Alvarez, L.A., Exton, D.A., Timmis, K.N., Suggett, D.J., and McGenity, T.J. (2009) Characterization of marine isoprene-degrading communities. *Environ. Microbiol.* **11**: 3280–3291.

Ashworth, K., Folberth, G., Hewitt, C.N., and Wild, O. (2012) Impacts of near-future cultivation of biofuel feedstocks on atmospheric composition and local air quality. *Atmos. Chem. Phys.* **12**: 919–939.

Atkinson, R. and Arey, J. (2003) Gas-phase tropospheric chemistry of biogenic volatile organic compounds: A review. *Atmos. Environ.* **37**: 197–219.

Armstrong, A. (2013) Isoprene and agriculture. *Nat. Geosci.* **6**: 513–513.

Bäck, J., Aaltonen, H., Hellén, H., Kajos, M.K., Patokoski, J., Taipale, R., et al. (2010) Variable emissions of microbial volatile organic compounds (MVOCs) from root-associated fungi isolated from Scots pine. *Atmos. Environ.* **44**: 3651–3659.

Berndt, T., Hyttinen, N., Herrmann, H., and Hansel, A. (2019) First oxidation products from the reaction of hydroxyl radicals with isoprene for pristine environmental conditions. *Commun. Chem.* **2**: 1–10.

Best, D. J., Floyd, N. C., Magalhaes, A., Burfield, A., and Rhodes, P. M. (1987). Initial enzymatic steps in the degradation of alpha-pinene by *Pseudomonas fluorescens* Ncimb 11671. *Biocatal. Biotransfor.* 1: 147–159

Bicas, J. L., Dionisio, A. P., and Pastore, G. M. (2009). Bio-oxidation of terpenes: an approach for the flavor industry. *Chem. Rev.* **109**: 4518–4531.

Broadgate, W.J., Malin, G., Küpper, F.C., Thompson, A., and Liss, P.S. (2004) Isoprene and other non-methane hydrocarbons from seaweeds: A source of reactive hydrocarbons to the atmosphere. *Mar. Chem.* **88**: 61–73.

Claeys, M., Graham, B., Vas, G., Wang, W., Vermeylen, R., Pashynska, V., et al. (2004) Formation of secondary organic aerosols through photooxidation of isoprene. *Science* **303**: 1173–1176.

Cleveland, C.C. and Yavitt, B. (1997) Consumption of atmospheric isoprene in soil. *Geophys. Res. Lett.* **24**: 2379–2382.

Cleveland, C.C. and Yavitt, J.B. (1998) Microbial consumption of atmospheric isoprene in a temperate forest soil. *Appl. Environ. Microbiol.* **64**: 172–177.

Collins, W.J., Derwent, R.G., Johnson, C.E., and Stevenson, D.S. (2002) The oxidation of organic compounds in the troposphere and their global warming potentials. *Clim. Change* **52**: 453–479.

Crombie, A.T., Khawand, M. El, Rhodius, V.A., Fengler, K.A., Miller, M.C., Whited, G.M., et al. (2015) Regulation of plasmid-encoded isoprene metabolism in *Rhodococcus* , a representative of an important link in the global isoprene cycle. *Environ. Microbiol.* **17**: 3314–3329.

Crombie, A.T., Larke-Mejia, N.L., Emery, H., Dawson, R., Pratscher, J., Murphy, G.P., et al. (2018) Poplar phyllosphere harbors disparate isoprene-degrading bacteria. *Proc. Natl. Acad. Sci. U. S. A.* **115**: 13081–13086.

Disch, A. and Rohmer, M. (1998) On the absence of the glyceraldehyde 3-phosphate/pyruvate pathway for isoprenoid biosynthesis in fungi and yeasts. *FEMS Microbiol. Lett.* **168**: 201–208.

Effmert, U., Kalderás, J., Warnke, R., and Piechulla, B. (2012) Volatile mediated interactions between bacteria and fungi in the soil. *J. Chem. Ecol.* **38**: 665–703.

Ekberg, A., Arneth, A., Hakola, H., Hayward, S., and Holst, T. (2009) Isoprene emission from wetland sedges. *Biogeosciences* **6**: 601–613.

Engelhart, G.J., Moore, R.H., Nenes, A., and Pandis, S.N. (2011) Cloud condensation nuclei activity of isoprene secondary organic aerosol. *J. Geophys. Res.* **116**: 1–11.

Ewers, J., Freier-Schröder, D., and Knackmuss, H.-J. (1990) Selection of trichloroethylene (TCE) degrading bacteria that resist inactivation by TCE. *Arch Microbiol* **154**: 410–413.

Fall, R. and Copley, S.D. (2000) Bacterial sources and sinks of isoprene, a reactive atmospheric hydrocarbon. *Environ. Microbiol.* **2**: 123–130.

Fowler, D., Nemitz, E., Misztal, P., di Marco, C., Skiba, U., Ryder, J., et al. (2011)

Effects of land use on surface-atmosphere exchanges of trace gases and energy in Borneo: Comparing fluxes over oil palm plantations and a rainforest. *Philos. Trans. R. Soc. B Biol. Sci.* **366**: 3196–3209.

Galizia, A. and Kinney, P.L. (1999) Long-term residence in areas of high ozone: Associations with respiratory health in a nationwide sample of nonsmoking young adults. *Environ. Health Perspect.* **107**: 675–679.

Gelmont, D., Stein, R.A., and Mead, J.F. (1981) Isoprene- the main hydrocarbon in human breath. *Biochem. Biophys. Res. Commun.* **99**: 1456–1460.

van Ginkel, C.G., de Jong, E., Tilanus, J.W.R., and de Bont, J.A.M. (1987) Microbial oxidation of isoprene, a biogenic foliage volatile and of 1,3-butadiene, an anthropogenic gas. *FEMS Microbiol. Lett.* **45**: 275–279.

Gray, C.M., Helmig, D., and Fierer, N. (2015) Bacteria and fungi associated with isoprene consumption in soil. *Elem. Sci. Anthr.* **3**: 000053.

Greenberg, J.P., Asensio, D., Turnipseed, A., Guenther, A.B., Karl, T., and Gochis, D. (2012) Contribution of leaf and needle litter to whole ecosystem BVOC fluxes. *Atmos. Environ.* **59**: 302–311.

Guenther, A., Karl, T., Harley, P., Wiedinmyer, C., Palmer, P.I., and Geron, C. (2006) Estimates of global terrestrial isoprene emissions using MEGAN (Model of Emissions of Gases and Aerosols from Nature). *Atmos. Chem. Phys. Discuss.* **6**: 107–173.

Guenther, A., Nicholas, C., Fall, R., Klinger, L., Mckay, W.A., Pierce, T., et al. (1995) A global model of natural volatile organic compound emissions. *J. Geophys. Res.* **100**: 8873–8892.

Guenther, A.B., Jiang, X., Heald, C.L., Sakulyanontvittaya, T., Duhl, T., Emmons, L.K., and Wang, X. (2012) The model of emissions of gases and aerosols from nature version 2.1 (MEGAN2.1): An extended and updated framework for modeling biogenic emissions. *Geosci. Model Dev.* **5**: 1471–1492.

Harder, J. and Probian, C. (1995) Microbial degradation of monoterpenes in the absence of molecular oxygen. *Appl. Environ. Microbiol.* **61**: 3804–3808.

Harley, P.C., Monson, R.K., and Lerdau, M.T. (1999) Ecological and evolutionary

aspects of isoprene emission from plants. *Oecologia* **118**: 109–123.

He, C., Murray, F., and Lyons, T. (2000) Monoterpene and isoprene emissions from 15 Eucalyptus species in Australia. *Atmos. Environ.* **34**: 645–655.

van Hylckama Vlieg, J.E., Kingma, J., van den Wijngaard, A.J., and Janssen, D.B. (1998) A glutathione S-transferase with activity towards cis-1, 2-dichloroepoxyethane is involved in isoprene utilization by *Rhodococcus* sp. strain AD45. *Appl. Environ. Microbiol.* **64**: 2800–2805.

van Hylckama Vlieg, J.E.T., De Koning, W., and Janssen, D.B. (1997) Effect of chlorinated ethene conversion on viability and activity of *Methylosinus trichosporium* OB3b. *Appl. Environ. Microbiol.* **63**: 4961–4964.

van Hylckama Vlieg, J.E.T., Leemhuis, H., Jeffrey, H., Spelberg, L., and Janssen, D.B. (2000) Characterization of the gene cluster involved in isoprene metabolism in *Rhodococcus* sp. strain AD45. *J. Bacteriol.* **187**: 1956–1963.

Hylemon, P.B. and Harder, J. (1998) Biotransformation of monoterpenes, bile acids, and other isoprenoids in anaerobic ecosystems. *FEMS Microbiol. Rev.* **22**: 475–488.

Jacob, D. (1999) Introduction to Atmospheric Chemistry Princeton University Press.

Janson, R., De Serves, C., and Romero, R. (1999) Emission of isoprene and carbonyl compounds from a boreal forest and wetland in Sweden. *Agric. For. Meteorol.* **98–99**: 671–681.

Janson, R. and De Serves, C. (1998) Isoprene emissions from boreal wetlands in Scandinavia. *J. Geophys. Res. Atmos.* **103**: 25513–25517.

Johnston, A., Crombie, A.T., El Khawand, M., Sims, L., Whited, G.M., McGenity, T.J., and Colin Murrell, J. (2017) Identification and characterisation of isoprene-degrading bacteria in an estuarine environment. *Environ. Microbiol.* **19**: 3526–3537.

Julsing, M.K., Rijpkema, M., Woerdenbag, H.J., Quax, W.J., and Kayser, O. (2007) Functional analysis of genes involved in the biosynthesis of isoprene in *Bacillus subtilis*. *Appl. Microbiol. Biotechnol.* **75**: 1377–1384.

Kesselmeier, J. and Staudt, M. (1999) Biogenic volatile organic compounds (VOC): An overview on emission, physiology and ecology. *J. Atmos. Chem.* **33**: 23–88.

El Khawand, M., Crombie, A.T., Johnston, A., Vavlline, D. V, McAuliffe, J.C., Latone, J.A., et al. (2016) Isolation of isoprene degrading bacteria from soils, development of isoA gene probes and identification of the active isoprene-degrading soil community using DNA-stable isotope probing. *Environ. Microbiol.* **18**: 2743–2753.

King, J., Koc, H., Unterkofler, K., Mochalski, P., Kupferthaler, A., Teschl, G., et al. (2010) Physiological modeling of isoprene dynamics in exhaled breath. *J. Theor. Biol.* **267**: 626–637.

Kirschke, S., Bousquet, P., Ciais, P., Saunois, M., Canadell, J.G., Dlugokencky, E.J., et al. (2013) Three decades of global methane sources and sinks. *Nat. Geosci.* **6**: 813–823.

Krechmer, J.E., Coggon, M.M., Massoli, P., Nguyen, T.B., Crounse, J.D., Hu, W., et al. (2015) Formation of low volatility organic compounds and secondary organic aerosol from isoprene hydroxyhydroperoxide low-NO oxidation. *Environ. Sci. Technol.* **49**:10330-10339.

Kuzma, J., Nemecek-Marshall, M., Pollock, W.H., and Fall, R. (1995) Bacteria produce the volatile hydrocarbon isoprene. *Curr. Microbiol.* **30**: 97–103.

Lackus, N.D., Lackner, S., Gershenzon, J., Unsicker, S.B., and Köllner, T.G. (2018) The occurrence and formation of monoterpenes in herbivore-damaged poplar roots. *Sci. Rep.* **8**: 1–13.

Lange, B.M., Rujan, T., Martin, W., and Croteau, R. (2000) Isoprenoid biosynthesis: the evolution of two ancient and distinct pathways across genomes. *Proc. Natl. Acad. Sci. U. S. A.* **97**: 13172–13177.

Laothawornkitkul, J., Taylor, J.E., Paul, N.D., and Hewitt, C.N. (2009) Biogenic volatile organic compounds in the Earth system. *New Phytol.* **183**: 27–51.

Lerdau, M. (2007) A positive feedback with negative consequences. *Science* **316**: 212–213.

Levy, H. (1971) Normal atmosphere: large radical and formaldehyde concentrations

predicted. *Science* **173**: 141–143.

Linares, D., Fontanille, P., and Larroche, C. (2009). Exploration of A-pinene degradation pathway of *Pseudomonas rhodesiae* cip 107491. Application to novalic acid production in a bioreactor. *Food Res. Int.* **42**: 461–469.

Liu, Y., Kuwata, M., Strick, B.F., Geiger, F.M., Thomson, R.J., McKinney, K.A., and Martin, S.T. (2015) Uptake of epoxydiol isomers accounts for half of the particle-phase material produced from isoprene photooxidation via the HO_2 pathway. *Environ. Sci. Technol.* **49**: 250–258.

Loreto, F. and Velikova, V. (2001) Isoprene produced by leaves protects the photosynthetic apparatus against ozone damage , quenches ozone products , and reduces lipid peroxidation of cellular membranes. **127**: 1781–1787.

Marmulla, R. and Harder, J. (2014) Microbial monoterpene transformations-a review. *Front. Microbiol.* **5**: 1–14.

McGenity, T.J., Crombie, A.T., and Murrell, J.C. (2018) Microbial cycling of isoprene, the most abundantly produced biological volatile organic compound on Earth. *ISME J.* **12**: 931–941.

Miller, B., Oschinski, C., and Zimmer, W. (2001) First isolation of an isoprene synthase gene from poplar and successful expression of the gene in *Escherichia coli*. *Planta* **213**: 483–487.

Monson, R.K. and Holland, E.A. (2001) Biospheric trace gas fluxes and their control over tropospheric chemistry. *Annu. Rev. Ecol. Syst.* **32**: 547–576.

Pacifico, F., Harrison, S.P., Jones, C.D., and Sitch, S. (2009) Isoprene emissions and climate. *Atmos. Environ.* **43**: 6121–6135.

Pandis, S.N., Paulson, S.E., Seinfeld, J.H., and Flagan, R.C. (1991) Aerosol formation in the photooxidation of isoprene and β-pinene. *Atmos. Environ. Part A. Gen. Top.* **25**: 997–1008.

Pegoraro, E., Abrell, L., Van Haren, J., Barron-Gafford, G., Grieve, K.A., Malhi, Y., et al. (2005) The effect of elevated atmospheric CO_2 and drought on sources and sinks of isoprene in a temperate and tropical rainforest mesocosm. *Glob. Chang. Biol.* **11**: 1234–1246.

Pike, R.C. and Young, P.J. (2009) How plants can influence tropospheric chemistry: The role of isoprene emissions from the biosphere. *Weather* **64**: 332–336.

Pope, C.A. and Dockery, D.W. (2006) Health effects of fine particulate air pollution : Lines that connect. *J. Air Waste Manag.* **56**: 709–742.

Putra, S.R., Disch, A., Bravo, J.M., and Rohmer, M. (1998) Distribution of mevalonate and glyceraldehyde 3-phosphate/pyruvate routes for isoprenoid biosynthesis in some Gram-negative bacteria and mycobacteria. *FEMS Microbiol. Lett.* **164**: 169–175.

Rasmussen, R.A. and Went, F.W. (1965) Volatile organic material of plant origin in the atmosphere. *Proc. Natl. Acad. Sci.* **53**: 215–220.

Reeves, C.E., Penkett, S. A., Bauguitte, S., Law, K.S., Evans, M.J., Bandy, B.J., et al. (2002) Potential for photochemical ozone formation in the troposphere over the North Atlantic as derived from aircraft observations during ACSOE. *J. Geophys. Res. D Atmos.* **107**: 1–14.

Rivera-Rios, J.C., Nguyen, T.B., Crounse, J.D., Jud, W., Clair, J.M.S., Mikoviny, T., et al. (2014) Conversion of hydroperoxides to carbonyls in field and laboratory instrumentation: Observational bias in diagnosing pristine versus anthropogenically controlled atmospheric chemistry. *Geophys. Res. Lett.* **41**: 8645–8651.

Rohmer, M., Knani, M., Simonin, P., Sutter, B., and Sahm, H. (1993) Isoprenoid biosynthesis in bacteria: a novel pathway for the early steps leading to isopentenyl diphosphate. *Biochem. J.* **295**: 517–524.

Rosenstiel, T.N., Ebbets, A.L., Khatri, W.C., Fall, R., and Monson, R.K. (2004) Induction of poplar leaf nitrate reductase: A test of extrachloroplastic control of isoprene emission rate. *Plant Biol.* **6**: 12–21.

Sanadze, G.A. (2004) Biogenic isoprene (a review). *Russ. J. Plant Physiol.* **51**: 729–741.

Sanadze, G.A. (1957) The nature of gaseous substances emitted by leaves of *Robinia pseudoacacia*. *Soobshch Akad Nauk Gruz SSR* **27**: 747–750.

Sasaki, K., Ohara, K., and Yazaki, K. (2005) Gene expression and characterization of

isoprene synthase from *Populus alba*. *FEBS Lett.* **579**: 2514–2518.

Schink, B. (1985) Inhibition of methanogenesis by ethylene and other unsaturated hydrocarbons. *FEMS Microbiol. Lett.* **31**: 63–68.

Schöller, C.E.G., Gürtler, H., Pedersen, R., Molin, S., and Wilkins, K. (2002) Volatile metabolites from actinomycetes. *J. Agric. Food Chem.* **50**: 2615–2621.

Schwender, J., Seemann, M., Lichtenthaler, H.K., and Rohmer, M. (1996) Biosynthesis of isoprenoids (carotenoids, sterols, prenyl side-chains of chlorophylls and plastoquinone) via a novel pyruvate/glyceraldehyde 3-phosphate non-mevalonate pathway in the green alga *Scenedesmus obliquus*. *Biochem. J.* **316**: 73–80.

Sharkey, T.D. (2013) Is it useful to ask why plants emit isoprene? *Plant, Cell Environ.* **36**: 517–520.

Sharkey, T.D. and Monson, R.K. (2017) Isoprene research - 60 years later, the biology is still enigmatic. *Plant Cell Environ.* **40**: 1671–1678.

Sharkey, T.D. and Singsaas, E.L. (1995) Why plants emit isoprene. *Nature* **374**: 769.

Sharkey, T.D., Wiberley, A.E., and Donohue, A.R. (2008) Isoprene emission from plants: Why and how. *Ann. Bot.* **101**: 5–18.

Shaw, S.L., Gantt, B., and Meskhidze, N. (2010) Production and emissions of marine isoprene and monoterpenes: A review. *Adv. Meteorol.* 1–24.

Shennan, J.L. (2006) Utilisation of C2-C4 gaseous hydrocarbons and isoprene by microorganisms. *J. Chem. Technol. Biotechnol.* **81**: 237–256.

Silver, G.M. and Fall, R. (1991) Enzymatic synthesis of isoprene from dimethylallyl diphosphate in aspen leaf extracts. *Plant Physiol.* **97**: 1588–1591.

Singsaas, E.L., Laporte, M.M., Shi, J.-Z., Monson, R.K., Bowling, D.R., Johnson, K., et al. (1999) Kinetics of leaf temperature fluctuation affect isoprene emission from red oak (*Quercus rubra*) leaves. *Tree Physiol.* **19**: 917–924.

Singsaas, E.L., Lerdau, M., Winter, K., and Sharkey, T.D. (1997) Isoprene increases thermotolerance of isoprene-emitting species. *Plant Physiol.* **115**: 1413–1420.

Sivy, T.L., Shirk, M.C., and Fall, R. (2002) Isoprene synthase activity parallels

fluctuations of isoprene release during growth of *Bacillus subtilis*. *Biochem. Biophys. Res. Commun.* **294**: 71–75.

Siwko, M.E., Marrink, S.J., de Vries, A.H., Kozubek, A., Schoot Uiterkamp, A.J.M., and Mark, A.E. (2007) Does isoprene protect plant membranes from thermal shock? A molecular dynamics study. *Biochim. Biophys. Acta - Biomembr.* **1768**: 198–206.

Steinke, M., Hodapp, B., Subhan, R., Bell, T.G., and Martin-Creuzburg, D. (2018) Flux of the biogenic volatiles isoprene and dimethyl sulfide from an oligotrophic lake. *Sci. Rep.* **8**: 1–10.

Terry, G.M., Stokes, N.J., Hewitt, C.N., and Mansfield, T.A. (1995) Exposure to isoprene promotes flowering in plants. *J. Exp. Bot.* **46**: 1629–1631.

Tingey, D.T., Evans, R.C., Bates, E.H., and Gumpertz, M.L. (1987) Isoprene emissions and photosynthesis in three ferns – The influence of light and temperature. *Physiol. Plant.* **69**: 609–616.

Tiiva, P., Rinnan, R., Faubert, P., Räsänen, J., Holopainen, T., Kyrö, E., and Holopainen, J.K. (2007) Isoprene emission from a subarctic peatland under enhanced UV-B radiation. *New Phytol.* **176**: 346–355.

Trainer, M., Williams, E.J., Parrish, D.D., Buhr, M.P., Allwine, E.J., Westberg, H.H., et al. (1987) Models and observations of the impact of natural hydrocarbons on rural ozone. *Nature* **329**: 705–707.

Velikova, V. and Loreto, F. (2005) On the relationship between isoprene emission and thermotolerance in *Phragmites australis* leaves exposed to high temperatures and during the recovery from a heat stress. *Plant Cell Environ.* **28**: 318–327.

Vranová, E., Coman, D., and Gruissem, W. (2013) Network analysis of the MVA and MEP pathways for isoprenoid synthesis. *Annu. Rev. Plant Biol.* **64**: 665–700.

Wagner, W.P., Nemecek-Marshall, M., and Fall, R. (1999) Three distinct phases of isoprene formation during growth and sporulation of Bacillus subtilis. *J. Bacteriol.* **181**: 4700–4703.

Wagner, W.P., Helmig, D., and Fall, R. (2000) Isoprene biosynthesis in *Bacillus subtilis* via the methylerythritol phosphate pathway. *J. Nat. Prod.* **63**: 37–40.

Wiedinmyer, C., Greenberg, J., Guenther, A., Hopkins, B., Baker, K., Geron, C., et al. (2005) Ozarks Isoprene Experiment (OZIE): Measurements and modeling of the "isoprene volcano." *J. Geophys. Res. D Atmos.* **110**: 1–17.

Wiedinmyer, C., Guenther, A., Harley, P., Hewitt, N., Geron, C., Artaxo, P., and Steinbrecher, R Rasmussen, R. (2004) Global organic emission from vegetation. In, Granier,C. (ed), *Emission of atmospheric trace compounds*. Kluwer Publishing Co, Dordrecht, pp. 115–170.

Wiedinmyer, C., Tie, X., Guenther, A., Neilson, R., and Granier, C. (2006) Future changes in biogenic isoprene emissions: How might they affect regional and global atmospheric chemistry? *Earth Interact.* **10**:.

Wilding, E.I., Brown, J.R., Bryant, A., Chalker, A.F., Holmes, D., Ingraham, K., et al. (2000) Identification, essentiality and evolution of the mevalonate pathway for isopentenyl diphosphate biosynthesis in Gram-positive cocci. *J. Bacteriol.* **182**: 4319–4327.

Winters, A.J., Adams, M.A., Bleby, T.M., Rennenberg, H., Steigner, D., Steinbrecher, R., and Kreuzwieser, J. (2009) Emissions of isoprene, monoterpene and short-chained carbonyl compounds from *Eucalyptus* spp. in southern Australia. *Atmos. Environ.* **43**: 3035–3043.

Xu, L., Guo, H., Boyd, C.M., Bougiatioti, A., Cerully, K.M., Hite, J.R., et al. (2015) Effects of anthropogenic emissions on aerosol formation from isoprene and monoterpenes in the southeastern United States. *Proc. Natl. Acad. Sci.* **112**: 37–42.

Zhao, D.F., Buchholz, A., Tillmann, R., Kleist, E., Wu, C., and Rubach, F. (2017) Environmental conditions regulate the impact of plants on cloud formation. **8**: 1–35.

Zhao, Y., Yang, J., Qin, B., Li, Y., Sun, Y., Su, S., and Xian, M. (2011) Biosynthesis of isoprene in *Escherichia coli* via methylerythritol phosphate (MEP) pathway. *Appl. Microbiol. Biotechnol.* **90**: 1915–1922.

Zurbriggen, A., Kirst, H., and Melis, A. (2012) Isoprene production via the mevalonic acid pathway in *Escherichia coli* (Bacteria). *Bioenergy Res.* **5**: 814–828.

Chapter 2 Reductive metabolism of the important atmospheric gas isoprene by homoacetogens

2.1 Introduction

Ecosystems emit numerous biogenic volatile organic compounds (BVOCs), which affect atmospheric chemistry and therefore the Earth's climate (Kesselmeier and Staudt, 1999; Atkinson and Arey, 2003; Sanadze, 2004; Laothawornkitkul *et al.*, 2009). By mass, between 30 and 50% of the estimated total global BVOC flux is isoprene (2-methyl-1,3-butadiene (CH_2=C(CH_3)-CH=CH_2)), a highly volatile, unsaturated hydrocarbon (Guenther *et al.*, 1995; Arneth *et al.*, 2011). An annual global terrestrial isoprene emission of 500-600 Tg y^{-1} (Guenther *et al.*, 2006, 2012) and an oceanic emission of 0.1-1.2 Tg y^{-1} (Alvarez *et al.*, 2009) has been estimated, which is similar in magnitude to methane sources at 526-569 Tg y^{-1} (Kirschke *et al.*, 2013). The physiological role of isoprene in natural environments remains enigmatic (Sharkey and Monson, 2017). It is mainly emitted by woody plants (Harley *et al.*, 1999; He *et al.*, 2000; Sharkey *et al.*, 2008; Winters *et al.*, 2009), though is also released in the breath of humans and other mammals (Gelmont *et al.*, 1981; King *et al.*, 2010), marine algae (Broadgate *et al.*, 2004) and by Gram-positive (e.g. *Bacillus* species) and Gram-negative bacteria (e.g. *Escherichia coli*, *Pseudomonas fluorescens*, various actinomycetes *Pseudonocardia*, *Saccharomonospora*, *Streptomyces*, *Thermomonospora*) (Kuzma *et al.*, 1995; Fall and Copley, 2000; Effmert *et al.*, 2012).

In the atmosphere isoprene reaction products modulate the oxidizing potential which leads to a prolonged lifetime of greenhouse gases such as methane (Collins *et al.*, 2002; Reeves *et al.*, 2002; Sharkey *et al.*, 2008; Pike and Young, 2009). Additionally, isoprene oxidation products cause secondary organic aerosols (SOA) formation (Krechmer *et al.*, 2015) which affects the climate directly by scattering absorbance and indirectly via cloud nuclei formation (Engelhart *et al.*, 2011; Zhao *et al.*, 2017).

While sources of atmospheric isoprene have been well-studied, terrestrial fates of isoprene remain relatively unexplored. Soils containing isoprene degrading microorganisms could serve as an important sink as shown by Cleveland and Yavitt (Cleveland and Yavitt, 1997). They estimated the global soil isoprene sink at 20.4 Tg yr^{-1} which is 4% of the estimated annual biogenic emission of isoprene (Guenther *et al.*,

1995). Under aerobic conditions, most soil bacteria that have been shown to utilize isoprene as their sole carbon and energy source belong to the phylum Actinobacteria (van Ginkel *et al.*, 1987; Ewers *et al.*, 1990; Cleveland and Yavitt, 1997; van Hylckama Vlieg *et al.*, 1997). The most detailed biochemical characterization of an aerobic isoprene metabolizing organism was conducted on *Rhodococcus* sp. strain AD45 (van Hylckama Vlieg *et al.*, 1997, 1999, 2000; Crombie *et al.*, 2015; El Khawand *et al.*, 2016). These studies identified isoprene epoxide (1,2-epoxy-2-methyl-3-butene) and the two glutathione adducts, 1-hydroxy-2-glutathionyl-2-methyl-3-butene (HGMB) and 2-glutathionyl-2-methyl-3-butenoic acid (GMBA), as intermediates of isoprene oxidation (van Hylckama Vlieg *et al.*, 1997, 1999, 2000). Crombie *et al*. (Crombie *et al.*, 2015) published the whole genome of *Rhodococcus* sp. AD45 and identified additional genes involved in isoprene metabolism, though the complete pathway has not been resolved.

In the present chapter the fate of isoprene in anoxic environments was investigated for the first time. Samples taken from different environments were used to enrich for anaerobic isoprene utilizing microorganisms with the aim of determining the fate of isoprene in anaerobic microbial systems in order to better understand its global degradation pathways. It was shown for the first time that obligate anaerobes also transform isoprene. Moreover, whereas isoprene primarily serves as an electron donor in aerated soils, we provide evidence that it is used as an electron acceptor to support homoacetogenesis.

2.2 Materials and Methods

2.2.1 Chemicals

Isoprene 99%, 3-methyl-1-butene ≥99.0% (GC), 2-methyl-2-butene ≥99.0% (GC), 2-methyl-1-butene ≥99.5% (GC) and ethene 99.9% in steel cylinder were all purchased from Sigma-Aldrich, Castle Hill, Australia. Helium (>99.9999% purity), nitrogen gas (>99.99% purity) and air (instrument grade) were purchased from BOC Gas, Australia. H_2 (>99.99995% purity) was obtained from a H_2 generator (Parker domnick hunter, UK).

2.2.2 Inocula

Sewage sludge samples were obtained from St Marys Sewage Treatment Plant Sydney, Australia and stored anaerobically in the dark at 4°C. Soil samples from Botany Bay, Sydney, Australia were core drilled from 3.6 meters beneath the surface and stored in the dark at room temperature in anaerobic media. Soil samples from 5 cm beneath the surface of Colo River area in Wollemi National Park, Blue Mountains Australia, were stored in the dark at room temperature.

2.2.3 Microbial strains

Acetobacterium species *A. woodii* DSM 1030, *A. malicum* DSM 4132 and *A. wieringae* DSM 1911 were obtained from Deutsche Sammlung von Mikroorganismen und Zellkulturen (DSMZ, Germany).

2.2.4 Culture conditions

Cells were grown in minimal media containing NH_4Cl (1.2 g l^{-1}), $MgCl_2 \cdot 2H_2O$ (0.4 g l^{-1}) and $CaCl_2 \cdot 2H_2O$ (0.1 g l^{-1}). The media was dispensed into culture flasks, flushed with nitrogen for 20 min, crimp-sealed with Teflon faced rubber septa and autoclaved. After autoclaving 1 ml of a trace element solution (HCl (25%), 10 ml; $FeCl_2 \cdot 4H_2O$, 2150 mg; $MnCl_2 \cdot 4H_2O$, 30 mg; $CoCl_2 \cdot 6H_2O$, 50 mg; $CuCl_2 \cdot 2H_2O$, 34 mg; $NiCl_2 \cdot 6H_2O$, 20 mg; $Na_2MoO_2 \cdot 2H_2O$, 30 mg; $ZnSO_4 \cdot 7H_2O$, 24 mg; H_3BO_3, 20 mg; distilled water, 1 l), 1 ml of tungstate/selenate solution (1000x) (Löffler *et al*., 2005), 1 ml vitamin solution (niacin, 100 mg; thiamine hydrochloride, 100 mg; biotin, 40 mg; pyridoxol hydrochlorid, 100 mg; folic acid, 20 mg; riboflavin, 50 mg; lipoic acid, 50 mg; pantothenic acid, 50 mg; Vitamin B12, 7.8 mg; 4-aminobenzoic acid, 50 mg; distilled water, 1 l), and 12.5 ml of 1 M phosphate buffer (prepared from 1 M K_2HPO_4

solution by adjusting pH value to 7.0 with 1M NaH_2PO_4) were added aseptically per liter of media. Vitamins were filter sterilized through a 0.22 micron filter membrane and all other solutions were autoclaved. The medium was reduced by the addition of appropriate volumes of a stock Na_2S solution to a final concentration of 0.025 mM. Cultures were incubated at room temperature in the dark.

2.2.5 Initial enrichment cultures for isoprene reduction

Sludge (50 µl, ~2 x 10^9 cells /ml) was added to sealed culture flasks (120 ml) containing 80 ml anaerobic minimal media. D/L-Lactate was supplied at ~10 mM from a sterile stock solution and isoprene (16 µl) was added as an electron acceptor using a 100 µl glass syringe to a final concentration of 1.3 mM in liquid media. Dilution to extinction series (10^{-1}-10^{-6}) were performed using the same culture condition except in 60 ml flask containing 40 ml anaerobic minimal media.

2.2.6 Growth in H_2 and HCO_3^- containing media

Anaerobic culture flask (120 ml) containing 80 ml minimal media were supplied with 0.5 bar sterile filtered H_2 and 30 mM $NaHCO_3$. Isoprene was added by using a 100 µl glass syringe to a final concentration of 1.3 mM in liquid media. In the case of ethene it was added from a gas stock via a 5 ml and 0.5 ml gas tight syringe. Cultures were inoculated with isoprene reducing enrichment culture (1 ml). In case of fructose it was added from a 1 M anaerobic stock solution to a final concentration of 20 mM.

2.2.7 Standards

Standards were prepared in 120 ml flasks with 80 ml anaerobic minimal medium. Isoprene was added with a glass syringe to prepare standards reaching from 0.4 mM to 5 mM nominal concentrations. 3-methyl-1-butene, 2-methyl-2-butene and 2-methyl-1-butene were added from each stock solution with a glass syringe and combined in one flask as a standard reaching from 0.4 mM to 5 mM nominal concentrations. H_2 standards were prepared at 0.2 to 20 mM nominal concentrations. Ethane, ethene and methane standards were prepared from a 1:1:1 gas mixture (33.33 % gas each) to prepare standards ranging from 0.1 to 12 mM nominal concentrations. All gases were added with different sizes of gas tight syringes. Dimensionless Henry constants for isoprene, ethene, ethane, methylbutenes, H_2 and methane were calculated from Sander (2015).

2.2.8 Isoprene, H_2, HCO_3^- and hydrocarbon analysis

Isoprene, methylbutenes, methane, H_2, CO_2 and hydrocarbon gases were monitored by gas chromatography (GC) using a Shimadzu GC-2010. Isoprene, methylbutenes and methane were analysed by using a GasPro PLOT column (60 m x 0.32 mm, Agilent Technologies) with Helium as a carrier gas (3 ml min^{-1}) and flame ionization detection (FID). The oven temperature was 150°C for 30 sec and was increased by 20°C min^{-1} to a final temperature of 250°C. Gas samples (100 µl) were withdrawn from the flask via a pressure-lockable gas tight syringe and directly injected into the GC.

H_2 was analysed using a HP-PLOT Molesieve column (30 m x 0.32 mm x 0.25 mm, Agilent Technologies, Australia) and pulsed discharged ionization detector (PDD). The carrier gas was Helium (3 ml min^{-1}) and the oven temperature was applied at 50°C for 1.2 min. Gas samples (20 µl) were withdrawn from the flask via a gas lock syringe and directly injected into the Shimadzu GC-2010.

HCO_3^- concentrations were measured by acidification of media (100 µl) with 20 µl 25% HCl in a sealed flask resulting in the release of HCO_3^- as CO_2. CO_2 was measured by GC-PDD using a HP-PLOT Q column (30 m x 0.32 mm, Agilent Technologies, Australia). The carrier gas was Helium (3 ml min^{-1}) and the oven temperature was applied from 50 °C for 1 min to 54.5 °C with a rate of 3.5 °C. Gas samples (40 µl) were withdrawn from the flask via a gas lock syringe and directly injected into a Shimadzu GC-2010.

Formate, acetate, butyrate and propionate were analysed as their ethyl ester derivative by GC-FID using a DB-FFAP column (30 m x 0.32 mm x 0.25 mm, Agilent Technologies) at 40°C for 6 min with helium as the carrier gas (3 ml min^{-1}). Samples (500 µl media) were supplied with 100% ethanol (200 µl) and undiluted sulfuric acid (200 µl) for esterification, sealed immediately and incubated at 60 °C for 45 min. Before injection into the GC, samples were incubated at 80° for 5 min at 500 rpm, 250 µl of headspace sample was withdrawn from the flask via an automatic sampler (Shimadzu AOC-5000 plus) and directly injected into a Shimadzu Plus GC-2010 at 500 µl s^{-1}.

Ethene and ethane amounts were measured by GC-FID using a GS-Q column (30 m x 0.32 mm x 0.25 mm) at 100°C for 2 min with helium as the carrier gas (3 ml min^{-1}). Gas

samples (100 µl) were withdrawn from the flask via a gas lock syringe and directly injected into a Shimadzu GC-2010.

2.2.9 D/L-Lactate analysis

Lactate concentrations were monitored by using the D/L-lactic acid kit from Megazyme following manufacturer's instructions.

2.2.10 DNA extraction and Illumina sequencing

DNA was extracted from 300 µl culture using standard phenol chloroform extraction method. Lysis buffer (Urakawa *et al.*, 2010) was added to the sample and the tube was mechanically agitated in FastPrep Lysis Matrix A tubes (MP Biomedicals). DNA was extracted with sequential phenol (phenol-chloroform-isoamyl alcohol (25:24:1), 7.5 M ammonium acetate, chloroform and isopropanol treatments, precipitated with ethanol using a general protocol, resuspended in 20 µl H_2O and stored at -20°C until further analysis. Regions of 16S rDNA gene were amplified by PCR from extracted DNA with the Q5 high-fidelity DNA polymerase (New England BioLabs) using the universal primers 926F (5`-TCGTCGGCAGCGTCAGATGTGTATAAGAGACAG -[AAA CTYAAAKGAATTGRCGG]-3`) and 1392R (5`GTCTCGTGGGCTCGGAGATGTGTATAAGAGACAG -[ACG GGC GGT GTG TRC-3`) targeting bacteria and archaea (Engelbrektson *et al.*, 2010). The samples were sequenced on an Illumina MiSeq Sequencer (Illumina, USA) using V3 chemistry at the Next Generation Sequencing Facility at Western Sydney University's Hawkesbury Institute for the Environment (Sydney, Australia). 16S rRNA gene amplicon sequences were analysed with QIIME2 (https://qiime2.org) (Caporaso *et al.*, 2010) utilizing the dada2 pipeline (Callahan *et al.*, 2016). Sequencing quality was first visualized with FastQC (www.bioinformatics.babraham.ac.uk) resulting in forwards and reverse reads being trimmed at 290 base pairs and 240 base pairs respectively. Forward and reverse sequences that passed the default quality control were merged and non-overlapping sequences were discarded. Chimeras were analysed and removed via the consensus method within the dada2 pipeline. Remaining sequences had taxonomy assigned with the RDP classifier (Wang *et al.*, 2007) using the Greengenes 13_8 database (DeSantis *et al.*, 2006). Taxa present at less than 2% abundance were removed for clarity.

2.2.11 Cloning

16S rDNA gene fragment cloning was performed on DNA samples from the isoprene reducing culture after 29 days. Part of the *Acetobacterium* 16S rDNA fragment was PCR-amplified with the Aceto572F and Aceto784R (Duhamel and Edwards, 2006) and a constant annealing temperature of 59°C for 34 cycles. All products were cloned into the pCR™2.1-TOPO® vector with TOPO TA Cloning Kit (Invitrogen, Carlsbad, CA) according to manufactures instructions. Plasmid DNA was extracted from overnight *Escherichia coli* (DH5α™-T1®) cultures using the PureYield™ Plasmid Miniprep System (Promega, Fitchburg, WI).

2.2.12 Quantitative real-time PCR

Quantitative real-time PCR targeting *Acetobacterium* was performed on a Biorad real time PCR system by using QuantiTect SYBR green PCR mastermix (Qiagen, Germany) and *Acetobacterium* primers Aceto572F and Aceto784R. The thermocycling program was as follows: initial denaturation for 3 min at 98°C; 39 cycles of [95°C for 20 s, 59 °C annealing for 50 s] and a final melting curve analysis from 60 to 99°C. The standard curve was generated with serial dilutions of a known quantity of 16S rDNA *Acetobacterium* gene-contained in plasmids generated as described above. Five copies of 16S rRNA genes per *Acetobacterium* genome were calculated based on the sequenced genome of *A. woodii* DSM 1030 to convert gene copies to cell numbers (Ding *et al.*, 2013).

2.3 Results

2.3.1 Isoprene transformation in anaerobic microcosms

Isoprene did not serve as an electron donor for the reduction of SO_4^{2-}, NO_3^{2-} and Fe^{3+} after 12 months of observation in any of the tested inocula (data not shown). However, the evolution of isoprene reduction products 2-methyl-1-butene (>97%), 3-methyl-1-butene (≤2%) and 2-methyl-2-butene (≤1%) was observed in activated sludge microcosms, suggesting the compound is reduced by the inocula [Mass spectra: Supplement **Figure S 1**]. To investigate the reduction of isoprene more closely quadruplicate anaerobic microcosms were prepared with lactate (900 µmoles) as the carbon and energy source, isoprene (160 µmoles) as the electron acceptor, and activated sludge as inoculum (1.7×10^7 cells/ml final concentration).

In the first round of enrichment cultures, isoprene depleted at a rate of ~0.3 µmoles d^{-1} after a 50 day lag period, with concomitant production of 15 µmoles methylbutenes after 200 days [**Figure 2.1 AB**]. The predominant isoprene reduction product was 2-methyl-1-butene (98%, 12.3 µmoles), with lesser amounts of 3-methyl-1-butene (2%). Lactate (900 µmoles) was completely consumed after 50 days in the presence or absence of isoprene [**Figure 2.2 A**] and was fermented to acetate, propionate (400 µmoles each) and H_2 with associated biomass generation [**Figure 2.2 BCDE**]. Isoprene reduction occurred subsequent to lactate consumption (after 50 days). Cultures without lactate amendment generated low quantities of fatty acids (<6 µmoles) and H_2 (~3-10 µmoles) presumably derived from biomass supplied in the inoculum, but did not convert isoprene to methylbutenes. Methane production was not detected in the presence of isoprene and lactate, however in the presence of lactate alone methane production commenced after 48 days and increased to 700 µmoles within 100 days with associated depletion of acetate and propionate [**Figure 2.1 C, Figure 2.2 CD**]. Autoclaved and non-inoculated controls did not show any depletion of isoprene or formation of methylbutenes.

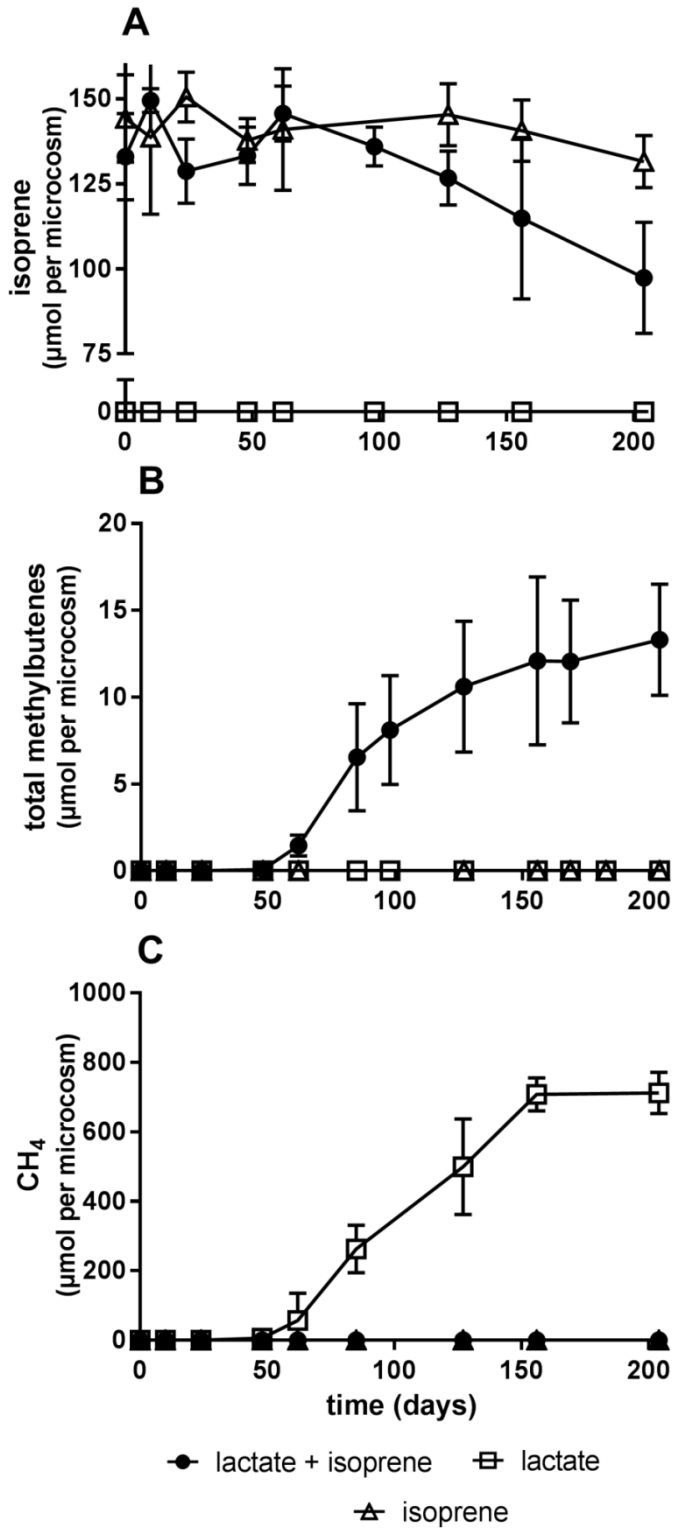

Figure 2.1 Depletion of isoprene (A) and subsequent production of methylbutenes (B) was only observed in microcosms, containing sludge, lactate and isoprene. In control samples supplemented with only lactate or isoprene, no isoprene depletion or methylbutene formation was detected. Methane (**C**) production only occurred in cultures supplied with lactate. Error bars represent one standard deviation (n = 4).

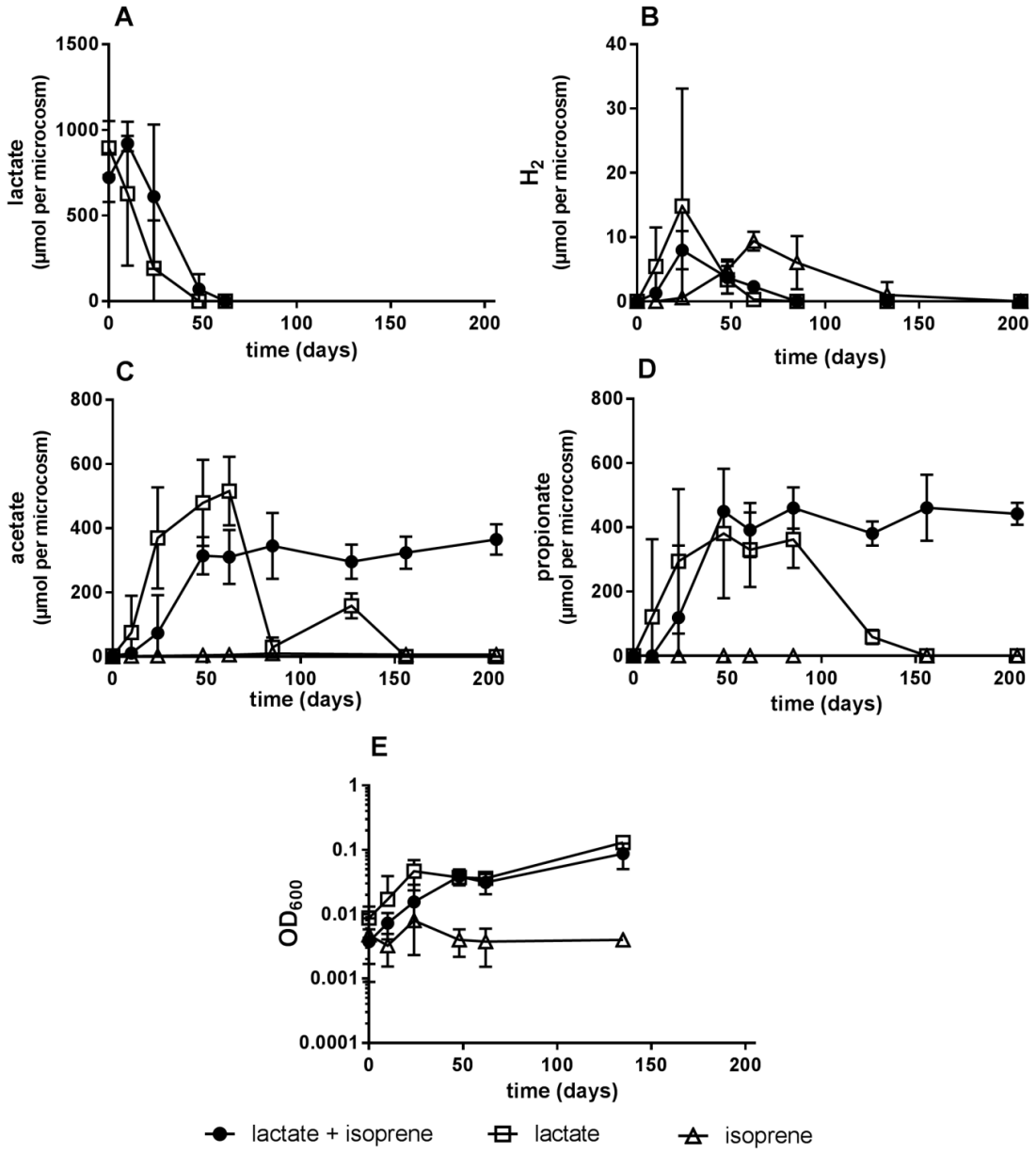

Figure 2.2 Change in quantity of lactate (A), H_2 (B), acetate (C), propionate (D) and biomass formation (E) in anaerobic cultures containing sludge and amended with lactate, isoprene, or both. Acetate and propionate were depleted in cultures with lactate only but concentrations remained stable after day 200 in cultures with isoprene and lactate. Error bars represent one standard deviation (n = 4).

2.3.2 Community analysis of lactate driven anaerobic isoprene transformations

Bacterial and archaeal community analysis was performed using 16S rRNA gene amplicons derived from DNA extracted from the lactate and isoprene fed cultures sampled before and after isoprene consumption [**Figure 2.3**]. In the isoprene fed cultures, *Acetobacterium*, *Geobacter* and in one replicate *Anaeromusa* where found to be the dominant bacterial genera [**Figure 2.3 A**]. Cultures supplied with lactate but without isoprene were dominated by a consortium of methanogenic archaea and bacteria *Geobacter*, *Clostridium*, *Acetobacterium* and *Anaeromusa* [**Figure 2.3 B**].

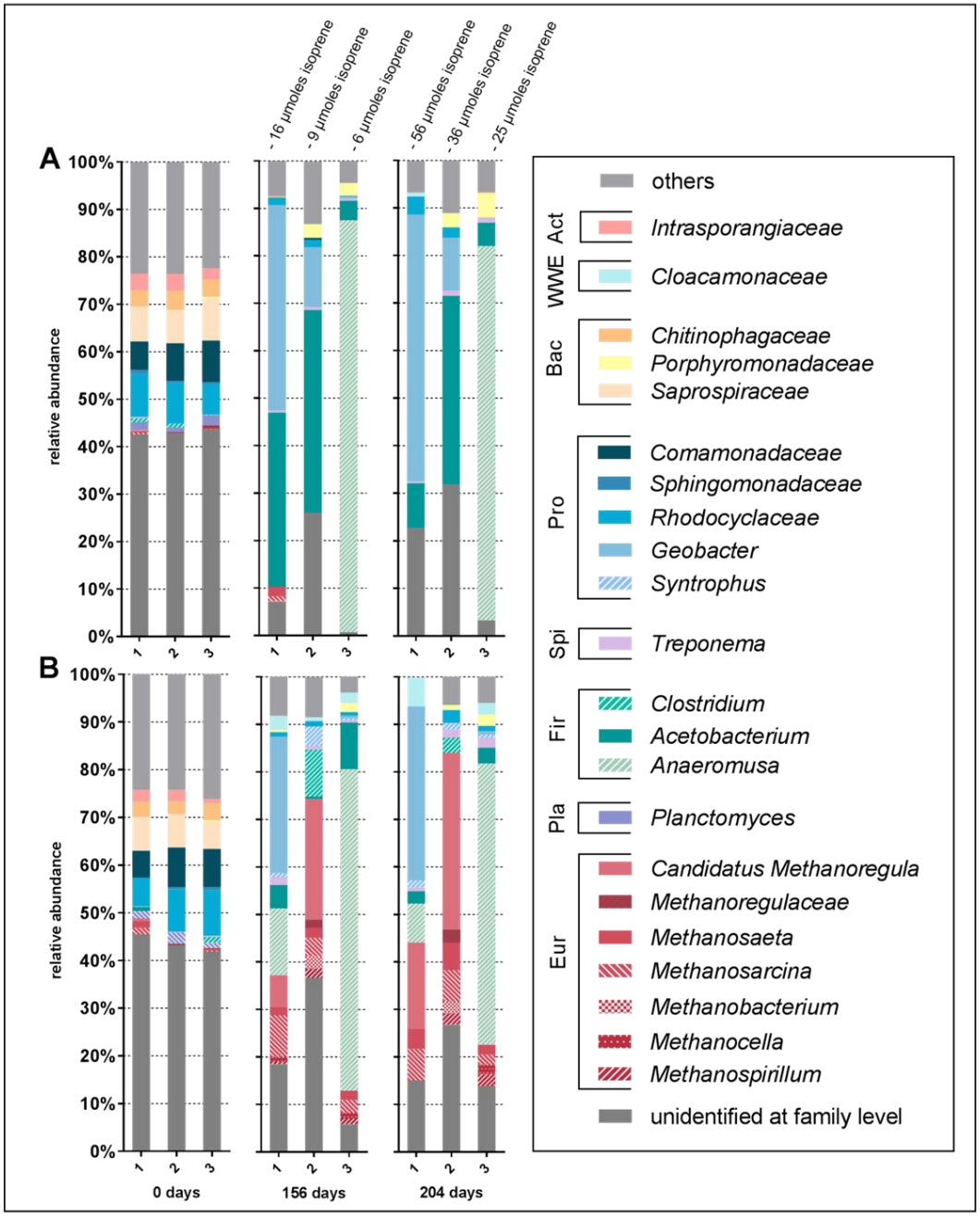

Figure 2.3 Composition of the bacterial and archaeal populations classified at Family level and if possible on Genus level of sludge during growth on lactate + isoprene (A) and lactate only control (B) at time points 0,156 and 204 days. Samples were inoculated with wastewater from St. Mary's treatment plant. Data of three replicates is shown. Other category is the sum of all classifications with less than 2% abundance. Classifications in the legend are clustered according to their phylum (from top to bottom); Actinobacteria, WWE1, Bacteroidetes, Proteobacteria, Spirochaetes, Firmicutes, Planctomycetes, Euryarchaeota. Depleted amounts of isoprene in each replicate are shown at the top. Only three out of four replicates were analysed by Illumina sequencing.

2.3.3 Characterization of H_2 driven isoprene transformations

The lag in isoprene reduction relative to lactate depletion suggested that lactate fermentation products were serving as carbon and energy sources. Given that acetate and propionate concentrations remained stable after day 200 in cultures amended with isoprene [**Figure 2.2 CD]** it was hypothesized that H_2 and HCO_3^- were serving as electron donor and carbon source respectively.

To test this hypothesis a dilution to extinction experiment was set-up with H_2 as the electron donor, HCO_3^- as the carbon source, and isoprene as the electron acceptor using a pooled enrichment culture derived from lactate and isoprene amended cultures as an inoculum. Isoprene reduction was observed in the 10^{-6} dilution with isoprene quantitatively transformed to 2-methyl-1-butene (98%) at a rate of 1.6 µmoles h^{-1}.

Having established the defined conditions for the cultivation of isoprene reducing bacteria (i.e. H_2/HCO_3^-/isoprene), quadruplicate cultures were prepared to characterize H_2/HCO_3^-/isoprene consumption rates and formation of volatile fatty acids (i.e acetate, and formate). Isoprene (130 µmoles) was transformed to methylbutenes (125 µmoles) within 92 h [**Figure 2.4 ABC**]. The predominant methylbutene was 2-methyl-1-butene (97%) with lesser amounts to 3-methyl-1-butene (2%) and 2-methyl-2-butene (≤1%) [**Figure 2.4 BC**]. Depletion of H_2 and HCO_3^- correlated with an increase in acetate. There was no significant difference in H_2 and HCO_3^- depletion or acetate and formate production when comparing isoprene and isoprene free cultures [**Figure 2.5 ABCD**]. When cultures were incubated with isoprene and H_2 or isoprene and HCO_3, neither acetate formation nor isoprene reduction was observed. To determine if methylbutene can be further reduced, H_2 was resupplied. After 8 days no further reduction of methylbutene to methylbutane was observed [Supplement **Figure S 2**].

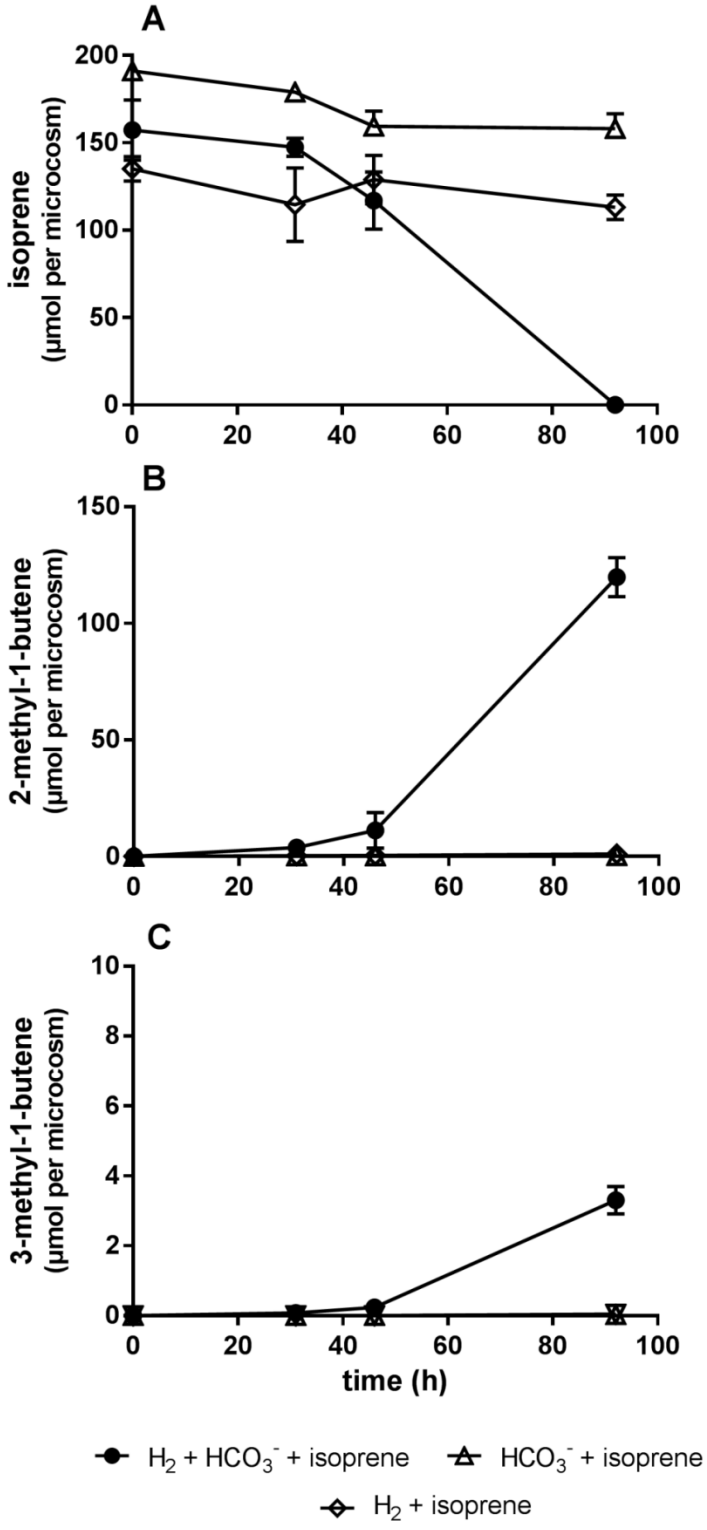

Figure 2.4 Depletion of isoprene (A) in 80 ml anaerobic cultures containing an enriched isoprene reducing culture supplied with H_2 + HCO_3^- + isoprene and reciprocal production of 2-methyl-1-butene (B) and 3-methyl-1-butene (C). In control samples supplemented with only H_2 + isoprene or only HCO_3^- + isoprene no methylbutene formation or isoprene depletion was detected. Error bars represent one standard deviation (n = 4).

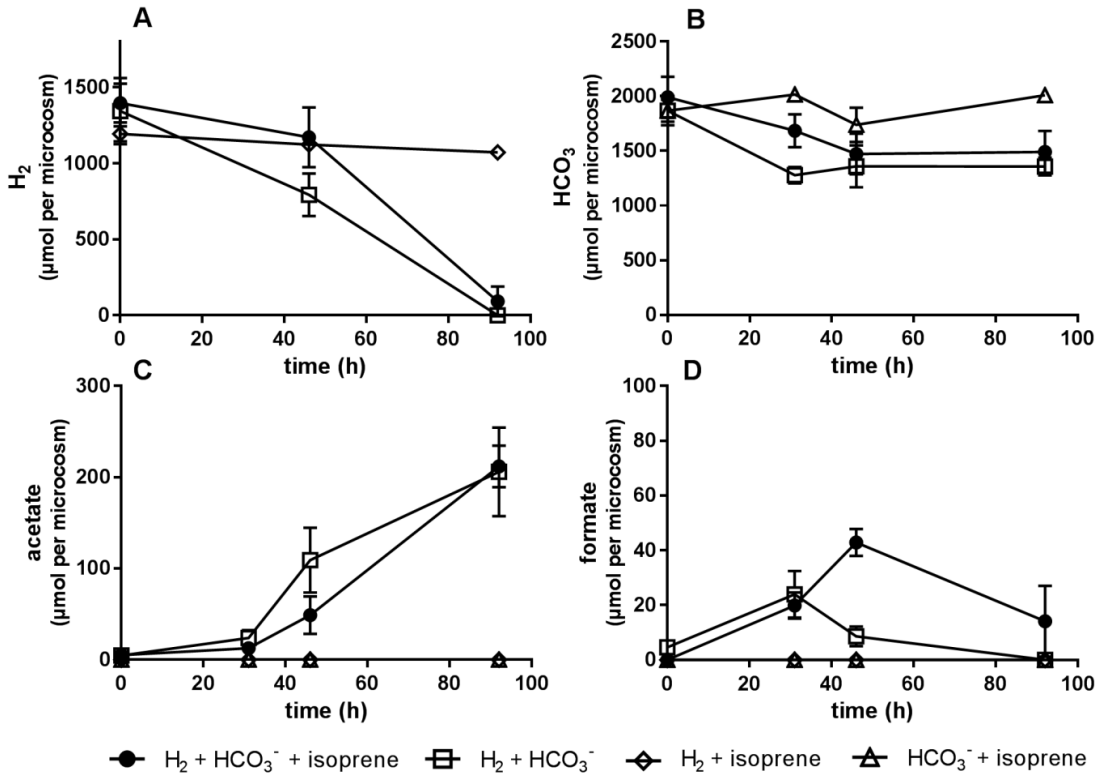

Figure 2.5 Consumption of H_2 (A) and HCO_3^- (B) and formation of acetate (C) and formate (D) anaerobic cultures containing an enriched isoprene reducing culture supplied with and without isoprene. In control samples supplemented with H_2 + isoprene or HCO_3^- + isoprene no acetate or formate formation was detected. Error bars represent one standard deviation (n = 4).

2.3.4 Community analysis of H_2 driven anaerobic isoprene transformations

Illumina sequencing of 16S rRNA gene amplicons from H_2 supplied, isoprene reducing enrichment cultures revealed enrichment of *Acetobacterium* to 92-100% relative abundance [**Figure 2.4 A, Figure 2.6**]. *Comamonadaceae* were also present (2% - 7%). There was no notable difference between bacterial community compositions in the presence vs. absence of isoprene [**Figure 2.6**].

To further probe an isoprene dependent difference in community composition a new set of cultures was monitored over a longer period by resupply of isoprene, H_2 and HCO_3^- when depleted [**Figure 2.7**]. Isoprene was depleted within 3-5 days after each resupply at a similar rate as observed before (1.6 µmoles h^{-1}) [**Figure 2.7 A**]. Methylbutene accumulated up to 800 µmoles respectively [**Figure 2.7 B**]. Illumina sequencing of 16S rRNA amplicons again showed no appreciable difference in microbial community composition when comparing cultures with or without isoprene. Both communities were

once again dominated by *Acetobacterium* spp. [Supplement **Figure S 3**]. Additionally, *Acetobacterium* cell numbers were similar after 29 days (i.e ~4.8 ×10^7 ± 1.3 × 10^7 cells ml^{-1} for cultures with isoprene and 5.6 × 10^7 ± 2.8 × 10^7 cells ml^{-1} for those without) [**Figure 2.7 E**]. Importantly however, there was a significant difference in the amount of acetate produced in cultures with and without isoprene. After day 16, cultures with isoprene produced ~250 µmoles of acetate while those without produced ~400 µmoles [**Figure 2.7 C**].

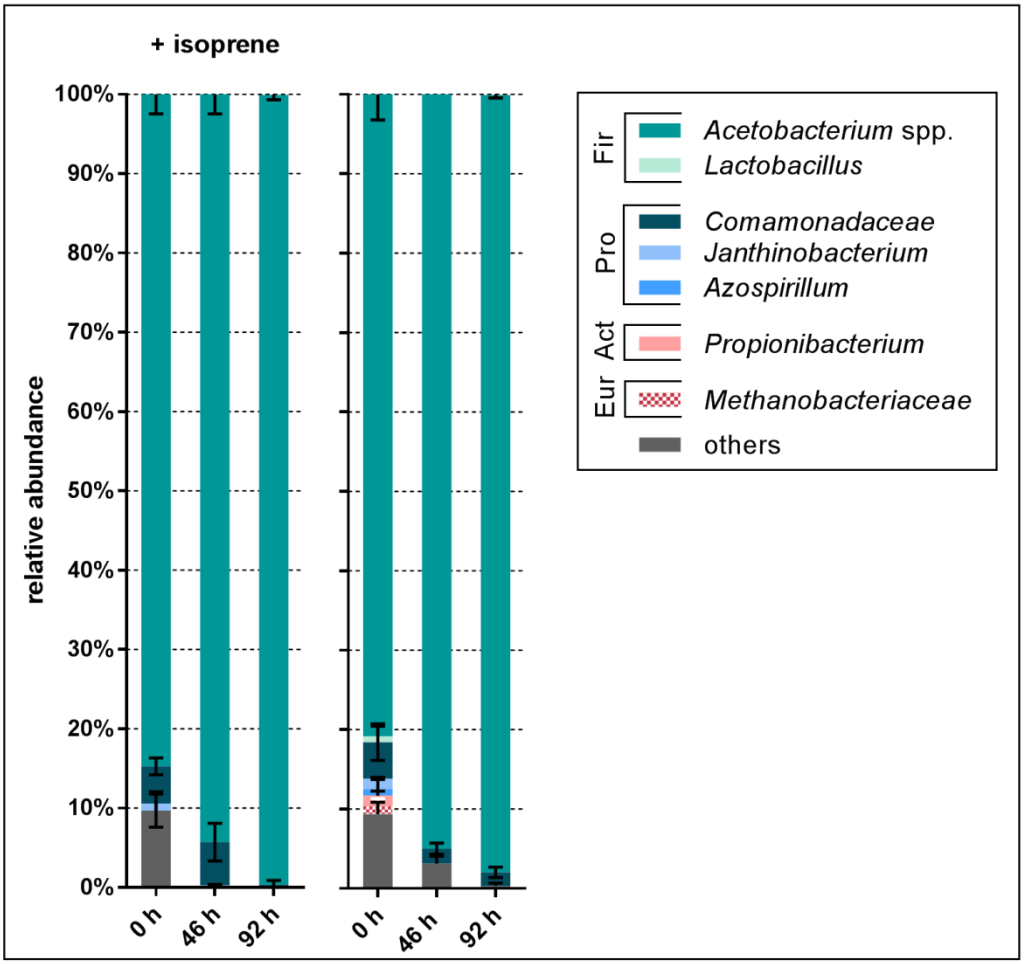

Figure 2.6 Composition of the bacterial populations at Family level and if possible on Genus level of the isoprene reducing culture during growth on H_2 + HCO_3^- + isoprene and on H_2 + HCO_3^- controls at different time points 0, 46 and 96 hours. Other category is the sum of all classifications with less than 2% abundance. Error bars represent one standard deviation (n = 3). Classifications in the legend are clustered according to their phylum (from top to bottom); Firmicutes, Proteobacteria, Actinobacteria, Euryarchaeota.

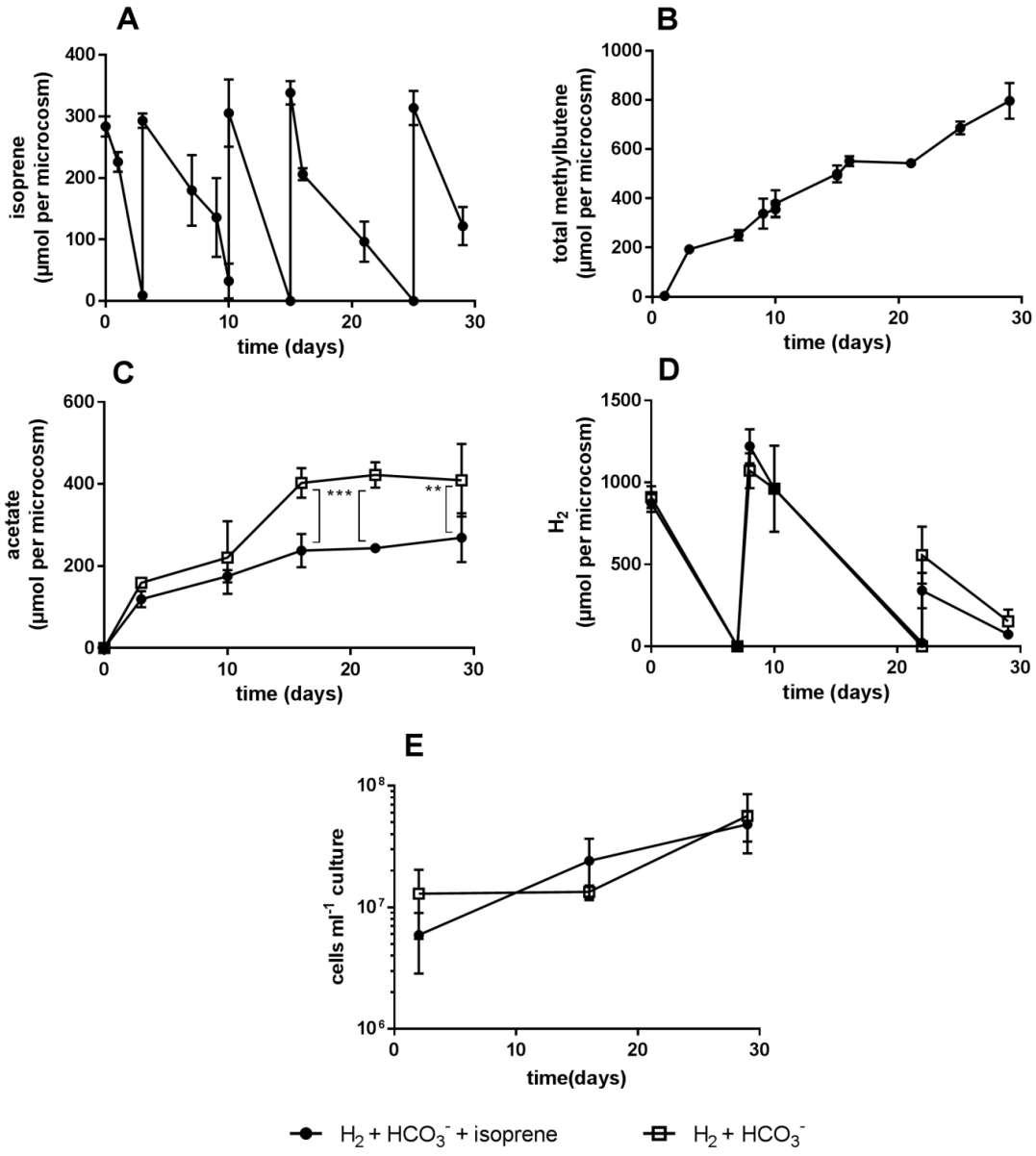

Figure 2.7 Depletion of isoprene (A) and reciprocal production of total methylbutenes (B) in anaerobic cultures containing an enriched isoprene reducing culture amended with H_2 and HCO_3^- with and without isoprene. Cultures amended with isoprene produced 40% less acetate (**C**) while still consuming the same amount of H_2 (**D**). *Acetobacterium* cells ml^{-1} (**E**) calculated from 16S rDNA gene copies ml^{-1} demonstrate equiviant grow in both conditions. *** represent a p-value <0.0001, ** represent a p-value = 0.005 analysed by a 2way ANOVA. Error bars represent one standard deviation (n = 4).

2.3.5 No transformation of isoprene by pure *Acetobacterium* strains

A. woodii DSM 1030, *A. malicum* DSM 4132 and *A. wieringae* DSM 1911 were tested for isoprene reduction on H_2/HCO_3^-/isoprene and fructose/isoprene but showed no isoprene reduction activity. *A. woodii* DSM 1030 was tested because it is the model organism of *Acetobacterium* species and is best studied example for CO_2-alternative electron acceptors (e.g. caffeate) in acetogens. *A. wieringae* DSM 1911 was tested because a clone library was constructed from the 16S rRNA gene that showed that 16 out of 18 clones exhibit highest similarity to *A. wieringae* [see **Figure S3 1**]. *A. malicum* DSM 4132 was tested because 2 out of 18 clones showed highest similarity to *A. malcium* [see **Figure S3 1**].

2.3.6 No transformation of ethene by isoprene reducing enrichment culture

To test the substrate specificity of the isoprene reducing culture, ethene was tested for reduction. Ethene (C_2H_4) contains a single double bond and is another volatile alkene that is emitted by plants.

Quadruplicate anaerobic cultures (80 ml) were prepared with two ethene concentrations (i.e. 10 or 160 µmoles per flask). H_2 was supplied as the electron donor. Samples were inoculated with an active isoprene reducing culture (1 ml, 0.8% v/v). Ethene remained unchanged after 22 days of incubation. Acetogenic growth was not affected by ethene and cells reduced CO_2 to around 600 µmoles of acetate [Supplement **Figure S 4, Figure S 5**].

2.4 Discussion

2.4.1 Isoprene serves as an electron acceptor

Even though isoprene is a highly abundant, energy rich metabolite, nothing is known about its anaerobic metabolism. In this study we tested the utility of isoprene as both an electron donor for inorganic oxide reduction and as an electron acceptor where lactate and molecular H_2 were electron donors.

Anaerobic isoprene oxidation coupled to inorganic oxide reduction (SO_4^{2-}, NO_3^{2-} or Fe^{3+}) could not be demonstrated in any of the tested inocula after one year of incubation. Under standard conditions, anaerobic isoprene oxidation is exergonic, considering the theoretical stoichiometries of isoprene mineralization and SO_4^{2-}, Fe^{3+} or NO_3^- reduction (isoprene energy of formation calculated as 197 kJ/mol (Thauer *et al.*, 1977; Dolfing and Janssen, 1994)).

(11) $2\ C_5H_8 + 2\ H_2O + 7\ SO_4^{2-} \rightarrow 7\ HS^- + 3\ H^+ + 10\ HCO_3^-$ $\Delta G°= -605$ kJ/mol

(12) $C_5H_8 + H_2O + 14\ NO_3^- \rightarrow 5\ HCO_3^- + 5\ H^+ + 14\ NO_2^-$ $\Delta G°= -2044$ kJ/mol

(13) $C_5H_8 + 15\ H_2O + 28\ Fe^{3+} \rightarrow 5\ HCO_3^- + 33\ H^+ + 28\ Fe^{2+}$ $\Delta G°= -2914$ kJ/mol

Evidently the enzyme systems required for the above transformations are either non-existent, extremely rare or inhibited or not induced under the conditions applied.

However, reductive isoprene transformation to 2-methyl-1-butene, 3-methyl-1-butene and 2-methyl-2-butene was observed under methanogenic conditions after two months of incubation. Subsequent enrichment resulted in increased rates of isoprene reduction (from ~0.3 µmoles d^{-1} to 40 µmoles d^{-1}). Isoprene reduction to methylbutene is also thermodynamically favorable considering theoretical stoichiometries with H_2 as electron donor [**Figure 2.8**] (Dolfing and Janssen, 1994; Speight and Lange, 2005).

(14) $H_2 + C_5H_8 \rightarrow C_5H_{10}$ $\Delta G°= -137$ kJ/mol

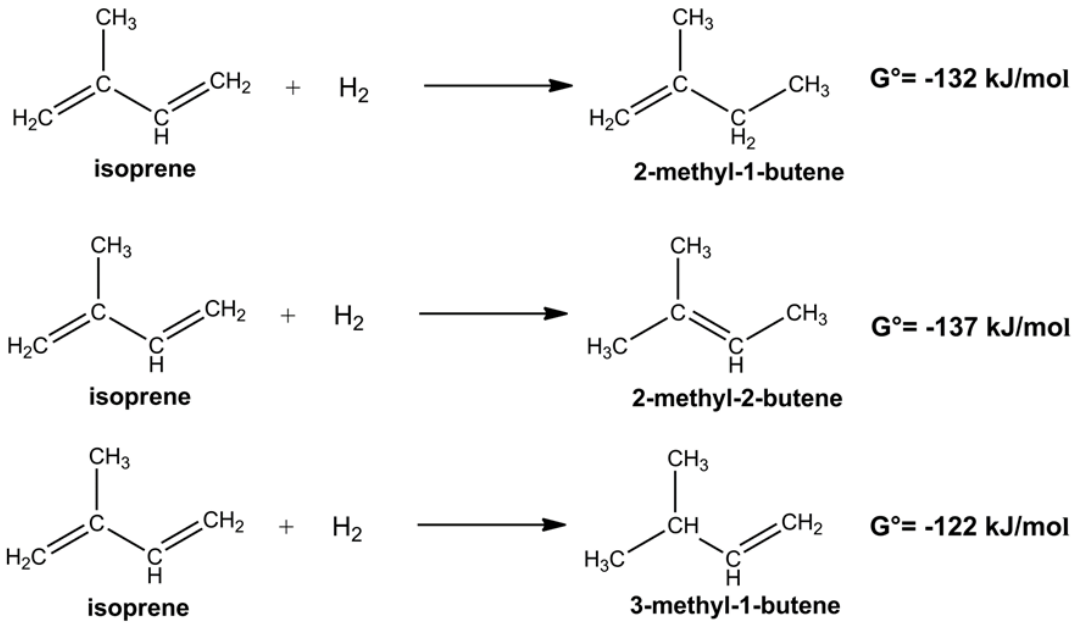

Figure 2.8 Chemical equations and Gibbs free energy of formation of the reduction of isoprene to 3-methyl-1-butene, 2-methyl-1-butene and 2-methyl-2-butene with H_2 as the electron donor. Gibbs free energy of formation of isoprene was estimated by the group contribution method (Dolfing & Janssens, 1994) to be 197 kJ/mol.

2.4.2 *Acetobacterium* spp. reduce isoprene

Illumina sequencing targeting archaeal and bacterial 16S rRNA gene amplicons revealed that acetogenic *Acetobacterium* spp. dominated the isoprene reducing enrichment culture. Acetogens are anaerobic bacteria that use CO_2 as a terminal electron acceptor for energy conservation and carbon fixation utilizing the reductive acetyl coenzyme A (acetyl-CoA) pathway or Wood-Ljungdahl pathway (Ljungdahl and Wood, 1969; Müller, 2003):

(15) $CO_2 + 4\ H_2 \rightarrow CH_3COOH + 2\ H_2O$ $\Delta G^{\circ} = -105$ kJ/mol

They are a phylogenetically and metabolically diverse group using a variety of different electron donors and acceptors (Drake and Küsel, 2013). Reducing equivalents can be generated from oxidation of H_2, C_1-compounds like methanol and formate, sugars, organic acids, and alcohols (Diekert and Wohlfarth, 1994). Besides CO_2, acetogens can also use alternative electron acceptors such as acrylate derivatives (Bache and Pfennig, 1981), fumarate (Dorn *et al.*, 1978), nitrate (Seifritz *et al.*, 1993), chlorethenes, chlorethanes (Terzenbach and Blaut, 1994) and brominated/aromatic compounds (Ding *et al.*, 2013). The most studied example for CO_2-alternative electron acceptors in

acetogens is the reduction of the carbon-carbon double bond in phenylacrylates (e.g. caffeate) by the model organism *Acetobacterium woodii* (Bache and Pfennig, 1981; Dilling *et al.*, 2007; Hess *et al.*, 2013).

As shown in this study, *Acetobacterium* spp. appear also to utilize isoprene as an alternative electron acceptor to CO_2. Isoprene reduction to methylbutenes depended on the presence of H_2 and CO_2 [**Figure 2.4**], common substrates for acetogens, which can otherwise only be metabolized by methanogens under anaerobic conditions. Methanogens were not present in isoprene amended enrichment cultures and no methane formation was observed. In cultures resupplied with H_2, HCO_3^- and isoprene, 40 % less acetate was formed compared to H_2 and HCO_3^- supplied cultures [**Figure 2.7 C**] suggesting that the *Acetobacterium* spp. transferred electrons from H_2 to isoprene at the expense of CO_2 reduction. In cultures without isoprene, 400 µmoles acetate were produced which required oxidation of 1600 µmoles H_2 (Eq.(15)) or 3200 µmoles of electrons. However, cultures with isoprene produced only 240 µmoles of acetate requiring 960 µmoles H_2. In addition 800 µmoles H_2 were required to reduce 800 µmoles of isoprene. Together that totals 1760 µmoles H_2 or 3520 µmoles of electrons transferred in cultures supplied with isoprene. Since similar amounts of electrons were transferred in cultures with (3520 µmoles) and without isoprene (3200 µmoles) and similar cell densities in each condition were observed, it can be concluded that the shortage of acetate in isoprene amended microcosms results from the reduction of isoprene instead of CO_2. Isoprene reduction might therefore also be coupled to energy conservation. Similar results were found by Hansen *et al*. (Hansen *et al.*, 1988) in the *A. woodii* NZva16 strain grown on H_2, CO_2 and caffeate, a key intermediate in lignin synthesis in plants. It was shown that 11 mM caffeate was reduced to hydrocaffeate and ~14-40% less acetate was formed compared to cultures without caffeate.

Considering the $\Delta G°$ values of the isoprene/methylbutene couple (Eq.(14)), isoprene reduction is energetically favored over CO_2 reduction (Eq.(15)). This could lead to preferential use of isoprene over CO_2. However, in the case of caffeate (caffeate/hydrocaffeate), its reduction and acetogenesis from H_2 and CO_2 were either catalyzed simultaneously or CO_2 was preferred over caffeate (Hansen *et al.*, 1988; Dilling *et al.*, 2007). (Tschech and Pfennig, 1984) observed that caffeate was preferentially used over CO_2 as an electron acceptor in the presences of methanol (methanol/caffeate 1:3). In the enrichment culture generated in this study it remains to

be tested whether isoprene reduction also occurs simultaneously catalyzed by a dominating *Acetobacterium* spp. by multiple or if one species only reduces isoprene and not CO_2.

It was unexpected that pure *Acetobacterium* isolates (*A. woodii* DSM 1030, *A. malicum* DSM 4132 and *A. wieringae* DSM 1911) showed no isoprene reduction activity. However, *Acetobacterium* are a particularly tight phylogenetic group with 96-99 % 16S gene sequence similarity between species (Willems and Collins, 1996). It is possible that an uncultured species of *Acetobacterium* is responsible for isoprene reduction, or the gene enabling the reduction was acquired via horizontal gene transfer. These possibilities will be investigated in future studies. Also, *Comamonadaceae* cannot be conclusively excluded from involvement in isoprene reduction as Illumina sequencing data did not reveal the genus of the lineage present. There is a possibility that *Hydrogenophaga* lineages, some of which can oxidize H_2 and fix CO_2 (Willems, 2014), contribute to isoprene reduction (see 3.4.3 for more detail).

Bioenergetics and growth efficiencies of acetogens can be evaluated using the acetate-to-biomass ratios (Daniel *et al.*, 1990). Even though less acetate was formed in isoprene amended cultures, similar cell densities were achieved. This suggests that isoprene reduction is coupled to ATP synthesis as shown for caffeate reduction by *A. woodii* (Hansen *et al.*, 1988). The current model in *A. woodii* suggests that electrons flow from H_2 to NAD^+ and ferredoxin which are reduced by an electron-bifurcating hydrogenase (Hansen *et al.*, 1988). A Na^+-translocating ferredoxin: NAD^+-oxidoreductase then oxidizes ferredoxin (Rnf complex) and generates a Na^+ gradient over the cytoplasmic membrane for ATP synthesis (Imkamp and Müller, 2002; Hess *et al.*, 2013; Schuchmann and Müller, 2014). It is possible that some of these enzymes might be involved in isoprene reduction. But in contrast to phenylacrylates, fumarate and chlorethenes, isoprene is an unsubstituted alkene making nucleophilic attack more difficult.

2.4.3 Hydrogenation of isoprene

In the present study it was shown that isoprene is hydrogenated to three methylbutene isomers with proportions remaining constant throughout the experiments (2-methyl-1-butene 98%, 3-methyl-1-butene 1-2% and 2-methyl-2-butene <1%) i.e. fully reduced alkanes were not detected [Supplement **Figure S 2**]. Hydrogenation of a double bond is

a thermodynamically favorable reaction because it forms a more stable (lower energy) product. The released heat is referred to as the heat of hydrogenation ($\Delta Hh_{298}°$), which reflects the stability of a molecule. Isoprene, due to its conjugated doubled bonds, has a $\Delta Hh_{298}°$ value of -55 kcal mol^{-1} compared to its reduced forms ($\Delta Hh_{298}°$ 3-methyl-1-butene; -29.9 kcal mol^{-1}, 2-methyl-1-butene; -28.24 kcal mol^{-1} and 2-methyl-2-butene; -26.74 kcal mol^{-1} (Prosen and Rossini, 1946). Therefore, reduction of a single double bond requires more activation energy and a different set of enzymes. This could explain why ethene ($H_2C=CH_2$) was not reduced to ethane by the active isoprene reducing culture. A few observations of microbial hydrogenation of ethene to ethane have been made (De Bruin *et al.*, 1992; Koene-Cottaar and Schraa, 1998; Mundle *et al.*, 2011; Elsgaard, 2013) but a pure bacterial culture or responsible enzymes have not been identified. It remains to be tested whether the enzymatic hydrogenation of isoprene to methylbutene occurs directly or via an intermediate hydration product (e.g. 3-methyl-3-buten-1-ol).

2.4.4 The physiological role of isoprene reduction-metabolic strategies of acetogens

Depletion of electron acceptors creates a niche for acetogens and methanogens due to their ability to obtain energy from CO_2 reduction via H_2 oxidation (Avery *et al.*, 2003; Lever, 2012). In these environments acetogens compete with methanogens for H_2.

(16) $HCO_3^- + 4H_2 + H^+ \rightarrow CH_4 + H_2O$ $\Delta G°= -135$ kJ/mol

Thermodynamically, hydrogenotrophic methanogenesis (Eq.(16)) is favoured over acetogenesis (Eq.(15)), therefore acetogens are physiologically less competitive for H_2 when it is a limiting resource. Yet, their ability to use different electron acceptors enhances the *in situ* competitiveness of acetogens (Drake *et al.*, 1997, 2008). Given that $\Delta G°$ values for hydrogenotrophic isoprene reduction (Eq.(14)) and bicarbonate reduction to methane (Eq.(16)) are equivalent, the use of isoprene as an alternative electron acceptor would enable acetogens to compete with methanogens at similar H_2 threshold concentrations.

According to Lever *et al.* (2012) there are two hypotheses for how the wide metabolic spectrum of acetogens enables them to coexist with sulfate reducers and methanogens. Firstly, metabolic versatility leads to niche differentiation with respect to substrate and secondly acetogens can pool energy from a broad range of metabolic reactions (e.g.

simultaneous CO_2/caffeate and now CO_2/isoprene reduction). Being able to utilize a substrate that inhibits a potential competitor could add a third strategy, whereby environments with isoprene present logically favouring acetogenic organisms.

In this study isoprene was observed to inhibit methanogenesis. The inhibitory effect of isoprene on methanogenesis in sediment slurries has been observed previously (Schink, 1985) where isoprene (3.6 mM in liquid phase) partially inhibited methanogenesis. Little is known about isoprene concentrations in anaerobic environments and hence it cannot be concluded yet if isoprene mediated inhibition of methanogenesis is ecologically or biogeochemically relevant. Therefore, it is also not yet clear whether isoprene reducers can consume atmospheric isoprene or if they consume isoprene generated in soil sources such as bacteria (Fall and Copley, 2000). Isoprene concentrations used in this and other studies on isoprene biotransformation are in the order of 10^6 times higher than observed under natural conditions (Ewers *et al.*, 1990; Cleveland and Yavitt, 1998; van Hylckama Vlieg *et al.*, 1998; Alvarez *et al.*, 2009; Crombie *et al.*, 2015; Gray *et al.*, 2015). Regardless, even at high concentrations the observed capacity for microbes to consume isoprene is far from satiated, suggesting the isoprene consuming microbial community is large relative to the isoprene supply and/or enzymes involved are highly efficient (Gray *et al.*, 2015).

2.5 Conclusion

This study explored the anaerobic metabolism of isoprene, which is quantitatively the most abundant volatile hydrocarbon emitted by plants. Isoprene was shown to act as an electron acceptor for homoacetogenic bacteria belonging to the *Acetobacterium* genera and was shown to be reduced to three different methylbutene isomers in an H_2 dependent manner. Isoprene had an inhibitory effect on methanogenesis so there may be a relationship between isoprene emission and methane biosynthesis. The discovery of biohydrogenation of this unsubstituted, unsaturated alkene whose functional group makes up 60 % of all Natural Products (isoprenoids or terpenoids) on Earth (Firn, 2009) warrants further investigation. Future experiments should explore the isoprene reduction mechanism, the enzymes involved and its ecological role in biogenic methane sources. Overall, this research demonstrates that isoprene is capable of being reduced in anaerobic environments, implicating a potential previously undiscovered isoprene sink.

2.6 References

Alvarez, L.A., Exton, D.A., Timmis, K.N., Suggett, D.J., and McGenity, T.J. (2009) Characterization of marine isoprene-degrading communities. *Environ. Microbiol.* **11**: 3280–3291.

Arneth, A., Schurgers, G., Lathiere, J., Duhl, T., Beerling, D.J., Hewitt, C.N., et al. (2011) Global terrestrial isoprene emission models: Sensitivity to variability in climate and vegetation. *Atmos. Chem. Phys.* **11**: 8037–8052.

Atkinson, R. and Arey, J. (2003) Gas-phase tropospheric chemistry of biogenic volatile organic compounds: A review. *Atmos. Environ.* **37**: 197–219.

Avery, G.B., Shannon, R.D., White, J.R., Martens, C.S., and Alperin, M.J. (2003) Controls on methane production in a tidal freshwater estuary and a peatland: Methane production via acetate fermentation and CO_2 reduction. *Biogeochemistry* **62**: 19–37.

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

Broadgate, W.J., Malin, G., Küpper, F.C., Thompson, A., and Liss, P.S. (2004) Isoprene and other non-methane hydrocarbons from seaweeds: A source of reactive hydrocarbons to the atmosphere. *Mar. Chem.* **88**: 61–73.

De Bruin, W.P., Kotterman, M.J.J., Posthumus, M.A., Schraa, G., and Zehnder, A.J.B. (1992) Complete biological reductive transformation of tetrachloroethene to ethane. *Appl. Environ. Microbiol.* **58**: 1996–2000.

Callahan, B.J., McMurdie, P.J., Rosen, M.J., Han, A.W., Johnson, A.J.A., and Holmes, S.P. (2016) DADA2: High-resolution sample inference from Illumina amplicon data. *Nat. Methods* **13**: 581–583.

Caporaso, J.G., Kuczynski, J., Stombaugh, J., Bittinger, K., Bushman, F.D., Costello, E.K., et al. (2010) QIIME allows analysis of high- throughput community sequencing data. *Nat. Methods* **7**: 335–336.

Cleveland, C.C. and Yavitt, B. (1997) Consumption of atmospheric isoprene in soil. *Geophys. Res. Lett.* **24**: 2379–2382.

Cleveland, C.C. and Yavitt, J.B. (1998) Microbial consumption of atmospheric isoprene in a temperate forest soil. **64**: 172–177.

Collins, W.J., Derwent, R.G., Johnson, C.E., and Stevenson, D.S. (2002) The oxidation of organic compounds in the troposphere and their global warming potentials. *Clim. Change* **52**: 453–479.

Crombie, A.T., Khawand, M. El, Rhodius, V.A., Fengler, K.A., Miller, M.C., Whited, G.M., et al. (2015) Regulation of plasmid-encoded isoprene metabolism in *Rhodococcus* , a representative of an important link in the global isoprene cycle. *Environ. Microbiol.* **17**: 3314–3329.

Daniel, S.L., Hsu, T., Dean, S.I., and Drake, H.L. (1990) Characterization of the H2- and CO-dependent chemolithotrophic potentials of the acetogens *Clostridium thermoaceticum* and *Acetogenium kivui*. *J. Bacteriol.* **172**: 4464–4471.

DeSantis, T.Z., Hugenholtz, P., Larsen, N., Rojas, M., Brodie, E.L., Keller, K., et al. (2006) Greengenes, a chimera-checked 16S rRNA gene database and workbench compatible with ARB. *Appl. Environ. Microbiol.* **72**: 5069–5072.

Diekert, G. and Wohlfarth, G. (1994) Metabolism of homoacetogens. *Antonie Van Leeuwenhoek* **66**: 209–221.

Dilling, S., Imkamp, F., Schmidt, S., and Müller, V. (2007) Regulation of caffeate respiration in the acetogenic bacterium *Acetobacterium woodii*. *Appl. Environ. Microbiol.* **73**: 3630–3636.

Ding, C., Chow, W.L., and He, J. (2013) Isolation of *Acetobacterium* sp. strain AG, which reductively debrominates octa- and pentabrominated diphenyl ether technical mixtures. *Appl. Environ. Microbiol.* **79**: 1110–1117.

Dolfing, J. and Janssen, D.B. (1994) Estimation of Gibbs free energies of formation of chlorinated aliphatic compounds. *Biodegradation* **5**: 21–28.

Dorn, M., Andreesen, J.R., and Gottschalk, G. (1978) Fumarate reductase of *Clostridium formicoaceticum*. *Arch. Microbiol.* **119**: 7–11.

Drake, H.L., Daniel, S.L., Küsel, K., Matthies, C., Kuhner, C., and Braus-Stromeyer, S. (1997) Acetogenic bacteria: what are the in situ consequences of their diverse metabolic versatilities? *BioFactors* **6**: 13–24.

Drake, H.L., Gößner, A.S., and Daniel, S.L. (2008) Old acetogens, new light. *Ann. N. Y. Acad. Sci.* **1125**: 100–128.

Drake HL, Küsel K, M.C. (2013) Acetogenic prokaryotes. In, *The prokaryotes: Prokaryotic physiology and biochemistry*. Berlin Heidelberg: Springer, pp. 1–60.

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

Effmert, U., Kalderás, J., Warnke, R., and Piechulla, B. (2012) Volatile mediated interactions between bacteria and fungi in the soil. *J. Chem. Ecol.* **38**: 665–703.

Elsgaard, L. (2013) Reductive transformation and inhibitory effect of ethylene under methanogenic conditions in peat-soil. *Soil Biol. Biochem.* **60**: 19–22.

Engelbrektson, A., Kunin, V., Wrighton, K.C., Zvenigorodsky, N., Chen, F., Ochman, H., and Hugenholtz, P. (2010) Experimental factors affecting PCR-based estimates of microbial species richness and evenness. *ISME J.* **4**: 642–647.

Engelhart, G.J., Moore, R.H., Nenes, A., and Pandis, S.N. (2011) Cloud condensation nuclei activity of isoprene secondary organic aerosol. *J. Geophys. Res.* **116**: 1–11.

Ewers, J., Freier-Schröder, D., and Knackmuss, H.-J. (1990) Selection of trichloroethylene (TCE) degrading bacteria that resist inactivation by TCE. *Arch Microbiol* **154**: 410–413.

Fall, R. and Copley, S.D. (2000) Bacterial sources and sinks of isoprene, a reactive atmospheric hydrocarbon. *Environ. Microbiol.* **2**: 123–130.

Firn, R. (2009) The main classes of NPs—Only a few pathways lead to the majority of NPs. In, *Nature's Chemicals*. Oxford: Oxford University Press, pp. 1–25.

Gelmont, D., Stein, R.A., and Mead, J.F. (1981) Isoprene- the main hydrocarbon in human breath. *Biochem. Biophys. Res. Commun.* **99**: 1456–1460.

van Ginkel, C.G., de Jong, E., Tilanus, J.W.R., and de Bont, J.A.M. (1987) Microbial oxidation of isoprene, a biogenic foliage volatile and of 1,3-butadiene, an anthropogenic gas. *FEMS Microbiol. Lett.* **45**: 275–279.

Gray, C.M., Helmig, D., and Fierer, N. (2015) Bacteria and fungi associated with

isoprene consumption in soil. *Elem. Sci. Anthr.* **3**: 000053.

Guenther, A., Karl, T., Harley, P., Wiedinmyer, C., Palmer, P.I., and Geron, C. (2006) Estimates of global terrestrial isoprene emissions using MEGAN (Model of Emissions of Gases and Aerosols from Nature). *Atmos. Chem. Phys. Discuss.* **6**: 107–173.

Guenther, A., Nicholas, C., Fall, R., Klinger, L., Mckay, W.A., Pierce, T., et al. (1995) A global model of natural volatile organic compound emissions. *J. Geophys. Res.* **100**: 8873–8892.

Guenther, A.B., Jiang, X., Heald, C.L., Sakulyanontvittaya, T., Duhl, T., Emmons, L.K., and Wang, X. (2012) The model of emissions of gases and aerosols from nature version 2.1 (MEGAN2.1): An extended and updated framework for modeling biogenic emissions. *Geosci. Model Dev.* **5**: 1471–1492.

Hansen, B., Bokranz, M., Schönheit, P., and Kröger, A. (1988) ATP formation coupled to caffeate reduction by H_2 in *Acetobacterium woodii* NZva16. *Arch. Microbiol.* **150**: 447–451.

Harley, P.C., Monson, R.K., and Lerdau, M.T. (1999) Ecological and evolutionary aspects of isoprene emission from plants. *Oecologia* **118**: 109–123.

He, C., Murray, F., and Lyons, T. (2000) Monoterpene and isoprene emissions from 15 *Eucalyptus* species in Australia. *Atmos. Environ.* **34**: 645–655.

Hess, V., González, J.M., Parthasarathy, A., Buckel, W., and Müller, V. (2013) Caffeate respiration in the acetogenic bacterium *Acetobacterium woodii*: A coenzyme a loop saves energy for caffeate activation. *Appl. Environ. Microbiol.* **79**: 1942–1947.

van Hylckama Vlieg, J.E., Kingma, J., Kruizinga, W., and Janssen, D.B. (1999) Purification of a glutathione S-transferase and a glutathione conjugate-specific dehydrogenase involved in isoprene metabolism in *Rhodococcus* sp. strain AD45. *J. Bacteriol.* **181**: 2094–101.

van Hylckama Vlieg, J.E., Kingma, J., van den Wijngaard, A.J., and Janssen, D.B. (1998) A glutathione S-transferase with activity towards cis-1, 2-dichloroepoxyethane is involved in isoprene utilization by *Rhodococcus* sp. strain AD45. *Appl. Environ. Microbiol.* **64**: 2800–2805.

van Hylckama Vlieg, J.E.T., De Koning, W., and Janssen, D.B. (1997) Effect of chlorinated ethene conversion on viability and activity of *Methylosinus trichosporium* OB3b. *Appl. Environ. Microbiol.* **63**: 4961–4964.

van Hylckama Vlieg, J.E.T., Leemhuis, H., Jeffrey, H., Spelberg, L., and Janssen, D.B. (2000) Characterization of the gene cluster involved in isoprene metabolism in *Rhodococcus* sp . strain AD45. *J. Bacteriol.* **187**: 1956–1963.

Imkamp, F. and Müller, V. (2002) Chemiosmotic energy conservation with Na^+ as the coupling ion during hydrogen-dependent caffeate reduction by *Acetobacterium woodii*. *J. Bacteriol.* **184**: 1947–1951.

Kesselmeier, J. and Staudt, M. (1999) Biogenic volatile organic compounds (VOC): An overview on emission, physiology and ecology. *J. Atmos. Chem.* **33**: 23–88.

El Khawand, M., Crombie, A.T., Johnston, A., Vavlline, D. V, McAuliffe, J.C., Latone, J.A., et al. (2016) Isolation of isoprene degrading bacteria from soils, development of isoA gene probes and identification of the active isoprene-degrading soil community using DNA-stable isotope probing. *Environ. Microbiol.* **18**: 2743–2753.

King, J., Koc, H., Unterkofler, K., Mochalski, P., Kupferthaler, A., Teschl, G., et al. (2010) Physiological modeling of isoprene dynamics in exhaled breath. *J. Theor. Biol.* **267**: 626–637.

Kirschke, S., Bousquet, P., Ciais, P., Saunois, M., Canadell, J.G., Dlugokencky, E.J., et al. (2013) Three decades of global methane sources and sinks. *Nat. Geosci.* **6**: 813–823.

Koene-Cottaar, F.H.M. and Schraa, G. (1998) Anaerobic reduction of ethene to ethane in an enrichment culture. *FEMS Microbiol. Ecol.* **25**: 251–256.

Krechmer, J.E., Coggon, M.M., Massoli, P., Nguyen, T.B., Crounse, J.D., Hu, W., et al. (2015) Formation of low volatility organic compounds and secondary organic aerosol from isoprene hydroxyhydroperoxide low-NO oxidation. *Environ. Sci. Technol.* 150814105132000.

Kuzma, J., Nemecek-Marshall, M., Pollock, W.H., and Fall, R. (1995) Bacteria produce the volatile hydrocarbon isoprene. *Curr. Microbiol.* **30**: 97–103.

Laothawornkitkul, J., Taylor, J.E., Paul, N.D., and Hewitt, C.N. (2009) Biogenic volatile organic compounds in the Earth system. *New Phytol.* **183**: 27–51.

Lever, M.A. (2012) Acetogenesis in the energy-starved deep biosphere-a paradox? *Front. Microbiol.* **2**: 1–18.

Ljungdahl, L.G. and Wood, H.G. (1969) Total synthesis of acetate from CO_2 by heterotrophic bacteria. *Annu. Rev. Microbiol.* **23**: 515–538.

Löffler, F.E., Sanford, R.A., and Ritalahti, K.M. (2005) Enrichment, cultivation of reductively dechlorinating bacteria. *Methods Enzymol.* **397**: 77–111.

Müller, V. (2003) Energy conservation in acetogenic bacteria. *Appl. Environ. Microbiol.* **69**: 6345–6353.

Mundle, S.O.C., Johnson, T., Lacrampe-Couloume, G., Perez-de-Mora, A., Duhamel, M., Edwards, E.A., et al. (2011) Monitoring biodegradation of ethene and bioremediation of chlorinated ethenes at a contaminated site using compound-specific isotope analysis (CSIA). *Environ. Sci. Technol.* **46**: 1731–1738.

Pike, R.C. and Young, P.J. (2009) How plants can influence tropospheric chemistry: The role of isoprene emissions from the biosphere. *Weather* **64**: 332–336.

Prosen, E.J. and Rossini, F.D. (1946) Heats of formation, hydrogenation, and combustion of the monoolefin hydrocarbons through the hexenes, and of the higher 1-alkenes, in the gaseous state at 25°C. *J. Res. Natl. Bur. Stand. (1934).* **36**: 269–275.

Reeves, C.E., Penkett, S. A., Bauguitte, S., Law, K.S., Evans, M.J., Bandy, B.J., et al. (2002) Potential for photochemical ozone formation in the troposphere over the North Atlantic as derived from aircraft observations during ACSOE. *J. Geophys. Res. D Atmos.* **107**: 1–14.

Sanadze, G.A. (2004) Biogenic isoprene (a review). *Russ. J. Plant Physiol.* **51**: 729–741.

Sander, R. (2015) Compilation of Henry's law constants, version 4.0. *Atmos. Chem. Phys.* **14**: 29615–30521.

Schink, B. (1985) Inhibition of methanogenesis by ethylene and other unsaturated

hydrocarbons. *FEMS Microbiol. Lett.* **31**: 63–68.

Schuchmann, K. and Müller, V. (2014) Autotrophy at the thermodynamic limit of life: A model for energy conservation in acetogenic bacteria. *Nat. Rev. Microbiol.* **12**: 809–821.

Seifritz, C., Daniel, S.L., Gossner, A., and Drake, H.L. (1993) Nitrate as a preferred electron sink for the acetogen *Clostridium thermoaceticum*. *J. Bacteriol.* **175**: 8008–8013.

Sharkey, T.D. and Monson, R.K. (2017) Isoprene research - 60 years later, the biology is still enigmatic. *Plant Cell Environ.* **40**: 1671–1678.

Sharkey, T.D., Wiberley, A.E., and Donohue, A.R. (2008) Isoprene emission from plants: Why and how. *Ann. Bot.* **101**: 5–18.

Speight, J.G. and Lange, N.A. (2005) Lange's handbook of chemistry. New York, N.Y. : McGraw-Hill, 1999, p. 1561.

Terzenbach, D.P. and Blaut, M. (1994) Transformation of tetrachloroethylene to trichloroethylene by homoacetogenic bacteria. *FEMS Microbiol. Lett.* **123**: 213–8.

Thauer, R.K., Jungermann, K., and Decker, K. (1977) Energy conservation in chemotrophic anaerobic bacteria. *Bacteriol. Rev.* **41**: 100–180.

Tschech, A. and Pfennig, N. (1984) Growth yield increase linked to caffeate reduction in *Acetobacterium woodii*. *Arch. Microbiol.* **137**: 163–167.

Urakawa, H., Martens-Habbena, W., and Stahl, D.A. (2010) High abundance of ammonia-oxidizing archaea in coastal waters, determined using a modified DNA extraction method. *Appl. Environ. Microbiol.* **76**: 2129–2135.

Wang, Q., Garrity, G.M., Tiedje, J.M., and Cole, J.R. (2007) Naïve Bayesian classifier for rapid assignment of rRNA sequences into the new bacterial taxonomy. *Appl. Environ. Microbiol.* **73**: 5261–5267.

Willems, A. (2014) The Family *Comamonadaceae*. In, Rosenberg E., DeLong E.F., Lory S., Stackebrandt E.,T.F. (ed), *The Prokaryotes*. Springer, Berlin, Heidelberg, pp. 777–851.

Willems, A. and Collins, M.D. (1996) Phylogenetic relationships of the genera

Acetobacterium and Eubacterium sensu stricto and reclassification of *Eubacterium alactolyticum* as *Pseudoramibacter alactolyticus* gen. nov., comb. nov. *Int. J. Syst. Bacteriol.* **46**: 1083–1087.

Winters, A.J., Adams, M.A., Bleby, T.M., Rennenberg, H., Steigner, D., Steinbrecher, R., and Kreuzwieser, J. (2009) Emissions of isoprene, monoterpene and short-chained carbonyl compounds from *Eucalyptus* spp. in southern Australia. *Atmos. Environ.* **43**: 3035–3043.

Zhao, D.F., Buchholz, A., Tillmann, R., Kleist, E., Wu, C., and Rubach, F. (2017) Environmental conditions regulate the impact of plants on cloud formation. **8**: 1–35.

2.7 Supplementary Material

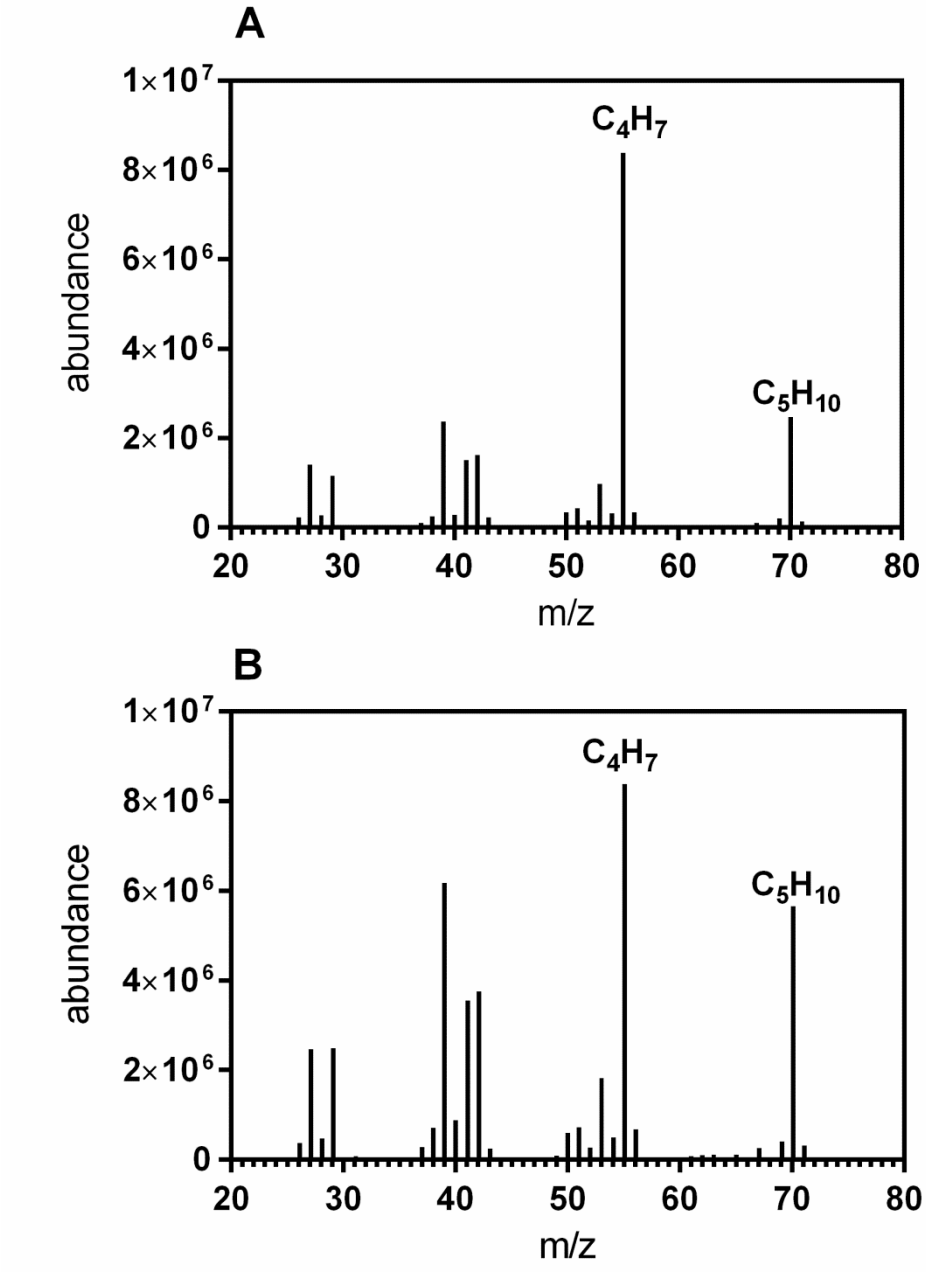

Figure S 1 Massspectra of 2-metyhl-1-butene (A) and 3-methyl-1-butene (B) derived from isoprene reducing cultures.

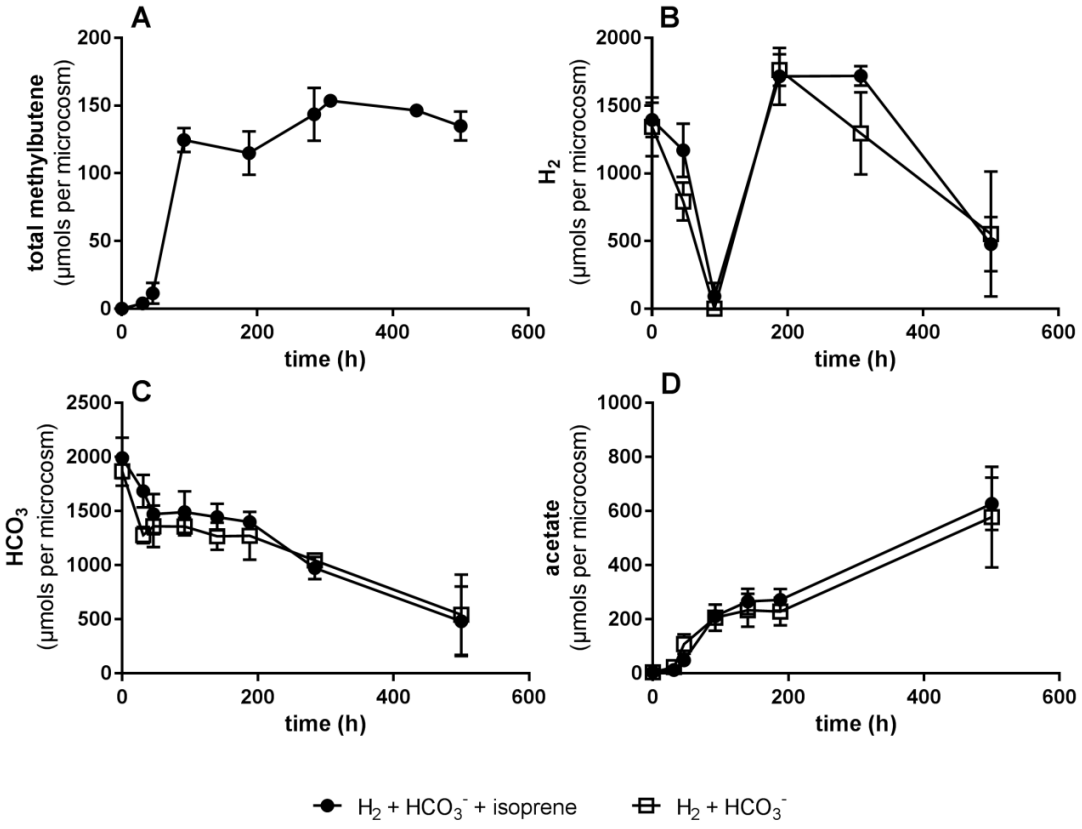

Figure S 2 Methylbutene amounts of highly enriched isoprene reducing culture after the addition of new H_2 after 200 h (A). Consumption of H_2 (**B**) and HCO_3^- (**C**) and acetate formation (**D**) in anaerobic cultures. Error bars represent one standard deviation (n = 4).

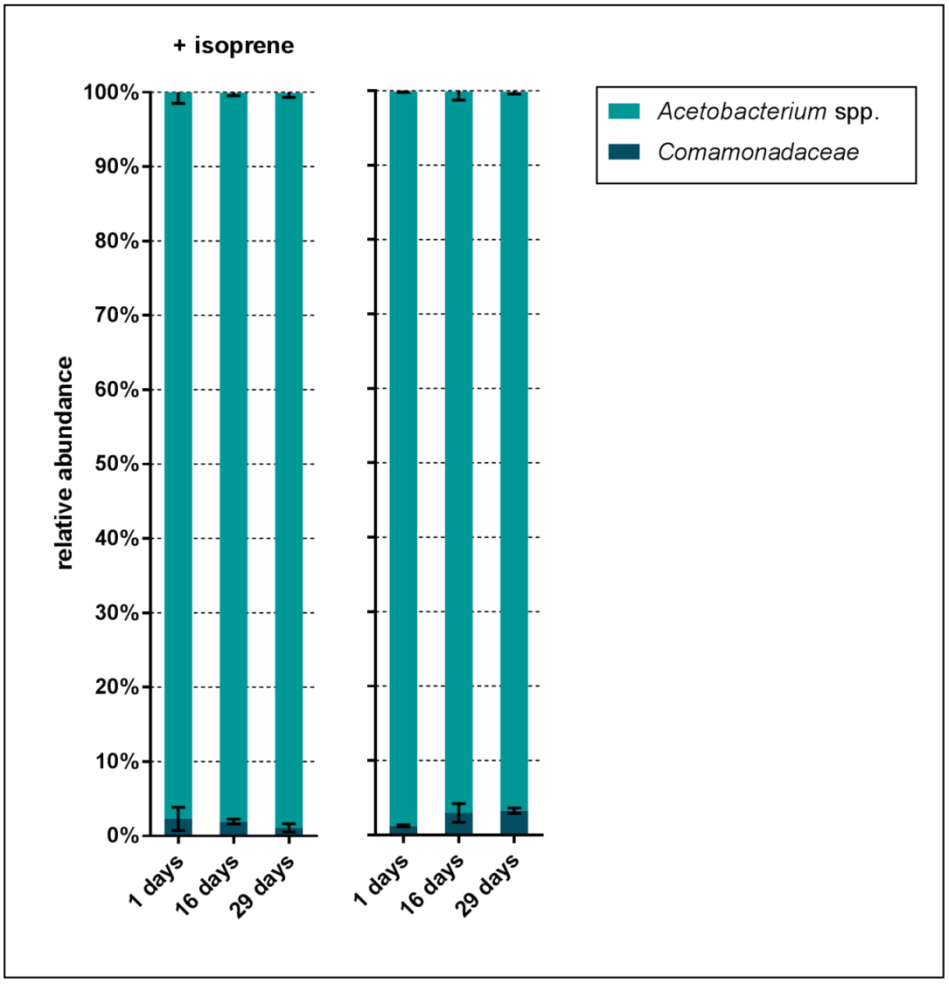

Figure S 3 Composition of the bacterial populations at Family and if possible on Genus level (based on 16S rDNA illumina sequencing) of an isoprene reducing, continuous grown batch culture on H_2 + HCO_3^- + isoprene and a H_2 + HCO_3^- control at different time points 1, 16 and 29 days. In both set ups *Acetobacterium* was found most abundant. *Comamonadaceae* are also present in smaller numbers (0.8 % - 4 %). Error bars represent one standard deviation (n = 3).

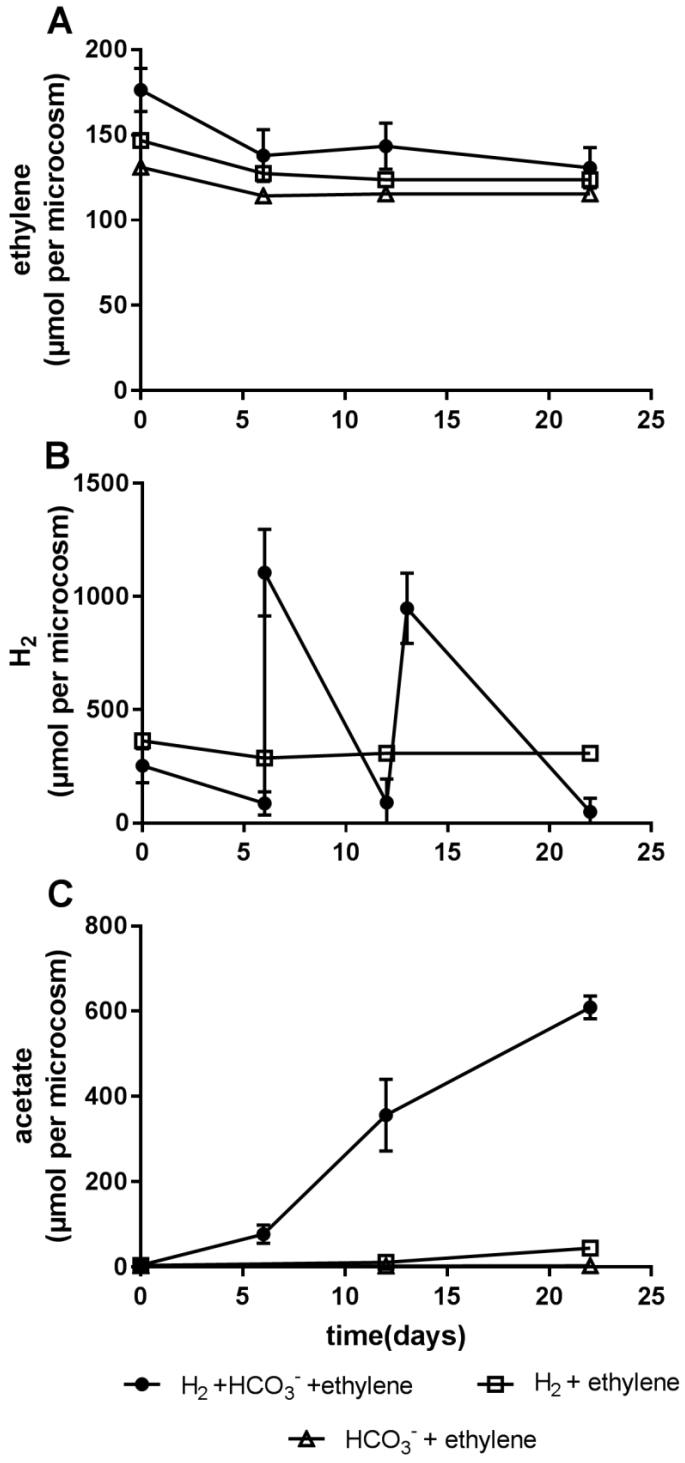

Figure S 4 Incubation of the active isoprene reducing culture with 160 µmoles of ethene, H_2 and HCO_3^-. Ethene concentration stayed stable over time H_2 + HCO_3^- + ethene and the controls (**A**). H_2 was depleted and acetate was formed respectively (**B** &**C**). Error bars represent one standard deviation (n = 4).

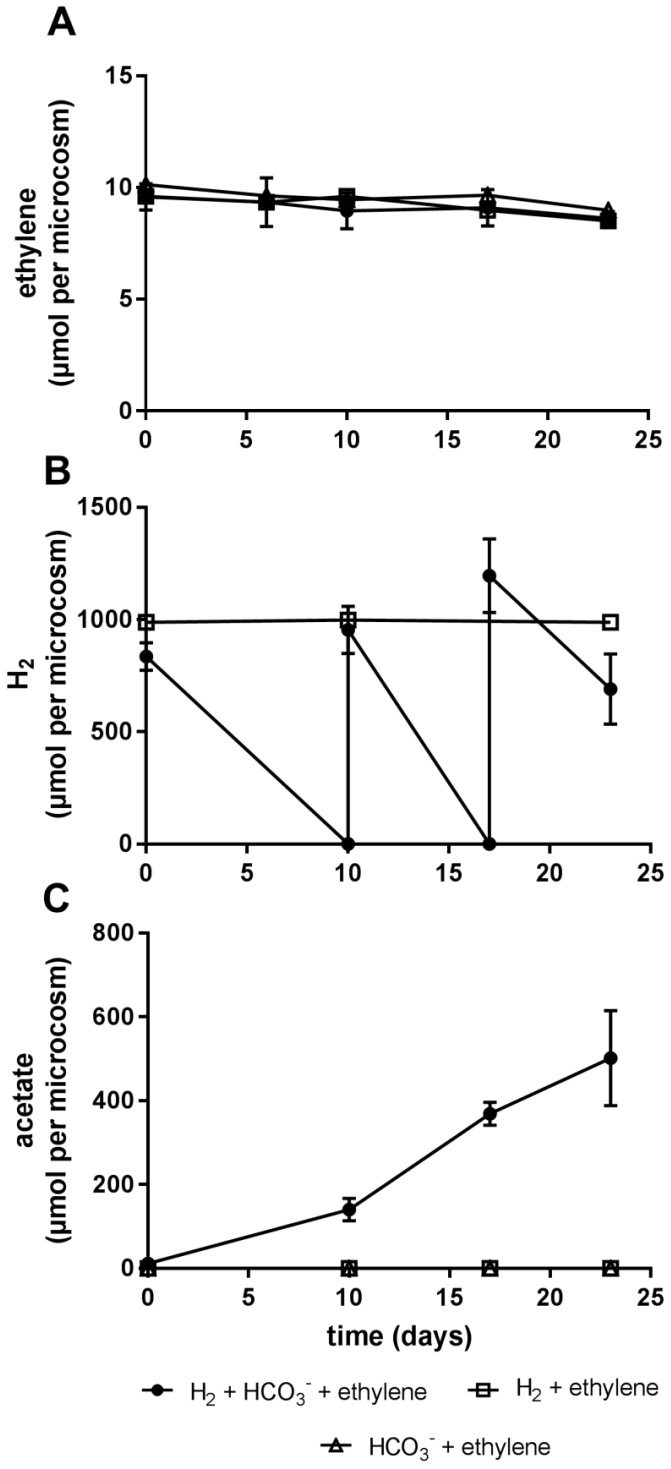

Figure S 5 Incubation of the active isoprene reducing culture with 10 µmoles of ethene, H_2 and HCO_3^-. Ethene concentration stayed stable over time H_2 + HCO_3^- + ethene and the controls (**A**). H_2 was depleted and acetate was formed respectively (**B** &**C**). Error bars represent one standard deviation (n = 4).

Chapter 3 Characterization and proteogenomic profiling of isoprene reducing culture

3.1 Introduction

Isoprene represents the most abundant BVOC on Earth (Kesselmeier and Staudt, 1999; Atkinson and Arey, 2003; Sanadze, 2004; McGenity *et al.*, 2018). Soils and marine environments are isoprene sources but also serve as isoprene sinks harbouring aerobic isoprene degrading organisms (McGenity *et al.*, 2018). The fate of isoprene in anaerobic enrichment cultures was also described recently (Chapter 2; Kronen *et al.*, 2019). In anoxic environments isoprene can serve as an electron acceptor and can be reduced to three methylbutene isomers 2-methyl-1-butene (>97%), 3-methyl-1-butene (≤2%), 2-methyl-2-butene (≤1%). The isoprene reducing culture was dominated by *Acetobacterium* sp., which in the presence of isoprene, formed 40% less acetate, suggesting that isoprene reduction is coupled to energy conservation. *Acetobacterium* sp. utilized isoprene as an alternative electron acceptor to CO_2 and reduction depended on the presence of H_2 and CO_2. *Acetobacterium* spp. and other acetogens usually use CO_2 as a terminal electron acceptor for energy conservation and carbon fixation via the Wood-Ljungdahl pathway (Ljungdahl and Wood, 1969; Müller, 2003)

$$CO_2 + 4\ H_2 \rightarrow CH_3COOH + 2\ H_2O \qquad \Delta G^\circ = -105\ kJ/mol$$

In recent years evidence has gathered that homoacetogens can also reduce alternative electron acceptors to CO_2 such as acrylate derivatives (Bache and Pfennig, 1981), fumarate (Dorn *et al.*, 1978), nitrate (Seifritz *et al.*, 1993), chlorethenes, chlorethanes (Terzenbach and Blaut, 1994), and brominated/aromatic compounds (Ding *et al.*, 2013). Reduction of the carbon–carbon double bond in phenylacrylates (e.g., caffeate) by the model organism *Acetobacterium woodii* (Bache and Pfennig, 1981; Dilling *et al.*, 2007; Hess *et al.*, 2013) is the best studied example for CO_2-alternative electron acceptors in *Acetobacterium* spp. The overall reduction reaction is linked to ATP synthesis by a chemiosmotic mechanism with Na^+ as the coupling ion (Imkamp and Müller, 2002; Dilling *et al.*, 2007; Imkamp *et al.*, 2007). The caffeate reduction operon and corresponding enzymes have been characterised in *A. woodii* (Hess *et al.*, 2011, 2013) and a caffeyl-CoA reductase-Etf complex was found to use flavin-dependent electron bifurcation to drive the endergonic reduction of ferredoxin with NADH as

electron donor by coupling it to the exergonic NADH-dependent reduction of caffeyl-CoA (Bertsch *et al.*, 2013).

The aim of this Chapter was to identify genes involved in isoprene reduction. DNA from the isoprene reducing enrichment culture was sequenced. Comparative metagenomic and metaproteomic analysis was performed on an enriched culture grown with H_2/HCO_3^-/isoprene and compared against H_2/HCO_3^- cells to understand the enzymes and mechanisms involved in isoprene reduction.

3.2 Materials and Methods

3.2.1 Chemicals

Isoprene 99%, 3-methyl-1-butene ≥ 99.0% (GC), 2-methyl-2-butene ≥ 99.0% (GC) and 2-methyl-1- butene ≥ 99.5% (GC) were all purchased from Sigma-Aldrich, Castle Hill, Australia. Helium (>99.9999% purity), nitrogen gas (>99.99% purity), and air (zero grade purity) were purchased from BOC Gas, Australia. H_2 (>99.99995% purity) was obtained from a H_2 generator (Parker domnick hunter, UK).

3.2.2 Growth conditions and media

Unless otherwise stated as described previously in Chapter 2 or Kronen *et al.* (2019)**.**

3.2.3 Standard preparation and isoprene, methylbutene analysis

Unless otherwise stated as described previously in Chapter 2 or Kronen *et al.* (2019)**.**

3.2.4 DNA extraction and Illumina sequencing

DNA was extracted from the highly enriched isoprene reducing cultures anaerobically grown on H_2, HCO_3^- with and without isoprene as described previously in Chapter 2 or Kronen *et al.* (2019). Libraries were prepared for both samples using the Nextera XT DNA Sample Preparation Kit according to manufacturer's protocol (Illumina). Sequencing reactions were carried out using the MiSeq v2 (2 x150 bp) chemistry (Illumina) by the Ramaciotti Centre for Genomics at the University of New South Wales (Sydney, Australia).

3.2.5 Genome assembly and binning

Quality trimming was performed with BBDuk. Filtered reads from both samples were co-assembled with MegaHIT v1.1.3 (Li *et al.*, 2016) and default parameters. Contigs ≥2.5 kbp were binned manually under Anvi'o v5.2.0 (Eren *et al.*, 2015). Protein coding genes of the metagenome and genomes from databases (see below) were predicted with Prodigal v2.6.3 (Hyatt *et al.*, 2010).

3.2.6 Annotation

Metagenome assembled genomes (MAGs) derived from binning were identified and named based on the Genome Taxonomy Database with GTDB-Tk v0.1.3 (Parks *et al.*, 2018). Predicted proteins were annotated with InterProScan v5.25-64 (Jones *et al.*,

2014). Results were then parsed with a custom script, iprs2anvio.sh (https://github.com/xvazquezc/stuff/blob/master/iprs2anvio.sh) and integrated in the Anvi'o workflow. Predicted proteins were also assigned to bacterial orthologous groups using the bactNOG database from eggNOG v4.5.1 (Huerta-Cepas *et al.*, 2016) with EggNOG-mapper v1.0.3-3-g3e22728 (Huerta-Cepas *et al.*, 2016). Genome annotation was performed with a modified version of Prokka v1.13.3 (Seemann, 2014) in which Prodigal generates partial gene calls at the end of contigs in order to minimise differences between Prokka- and Anvi'o-based gene predictions. Prokka-based annotations were imported in PathwayTools v22.0 (Karp *et al.*, 2016).

3.2.7 *Acetobacterium* pangenome

Pangenomic analysis of the genus *Acetobacterium* with eight reference *Acetobacterium* genomes [Supplement **Table S3 3**] was performed with Anvi'o v5.2.0 (Eren *et al.*, 2015) following the standard pangenomics workflow (http://merenlab.org/2016/11/08/pangenomics-v2). Genes were clustered with an MCL inflation value of 6 (van Dongen and Abreu-Goodger, 2012). Gene clusters were grouped based on their presence in all 9 genomes (core), at least 4 out of 9 genomes (soft core) or unique to the organism (singleton). Average Nucleotide Identity (ANI) between *Acetobacterium* genomes was calculated with pyani v0.2.7 (Pritchard *et al.*, 2016). Average Amino Acid Identity was calculated with CompareM v0.0.23 (https://github.com/dparks1134/CompareM).

3.2.8 Cell preparation for metaproteomics

The isoprene reducing enrichment cultures were grown in 12 culture flasks (120 ml) containing anaerobic minimal media (80 ml) supplied with 0.5 bar sterile filtered H_2 and 30 mM $NaHCO_3$. Applicable isoprene (99.9%) was supplied at 1.3 mM (nominal aqueous phase concentration) to six flasks and resupplied every four days. H_2 was resupplied every four days to all 12 flasks and after 10 days of incubation at 30°C, cells were harvested. As cells formed aggregates during growth they were pipetted from the bottom of each flask and transferred to a 2 ml reaction tube, cells of 2 samples were pooled (i.e. from 6 to 3 for each condition) and centrifuged at 10 000 g for 10 min and then stored at -20°C until use.

3.2.9 Protein extraction and digestion

Harvested cells suspended in 100 µl of lysis buffer (Urakawa *et al.*, 2010) were mechanically disrupted in FastPrep Lysis Matrix A tubes (MP Biomedicals) at 30 Hz for 10 min. The crude extract was passed through a 30 kDa Amicon Ultra-0.5 mL Centrifugal Filters and washed 6 times with 200 µl of 50 mM ammonium bicarbonate buffer (pH 6.9). Sample protein concentrations were determined with the Quick Start Bradford Protein Assay following manufacturer's instructions (Bio-Rad Laboratories, Australia), and adjusted to 2 ug ul^{-1} and 10 ul (20 ug) were used for Filter Aided Sample Preparation (FASP) (Wiśniewski, 2016, 2017). Samples were treated with 5 mM dithiotreitol (DTT) at 37°C for 30 min. Protein lysates were then transferred to new 30 kDa Amicon Ultra-0.5 mL Centrifugal Filters and treated following FASP method (https://www.biochem.mpg.de/226356/FASP). Trypsin solution (1 µl of a 200 ng ul^{-1}) was added for digestions. Peptides were eluted in 2 x 20 µl 50 mM ammonium bicarbonate buffer and stored at -20 °C until LC-MS/MS analysis.

3.2.10 LC-MS/MS analysis

Samples analysis was performed at the Bioanalytical Mass Spectrum Facility (BMSF) at the University of New South Wales. Peptides were separated by nano-HPLC using Ultimate 3000 UPLC and autosampler system (Thermo Fisher Scientific) coupled to a LTQ-Orbitrap mass spectrometer (Thermo Fisher Scientific). Chromatographic separation was performed in a fritless nano column (75 µ x ~10 cm) containing C18 media (1.9 µ, 120Å, DrMaisch, Ammerbuch- Entringen Germany). Peptides were injected into the column at a flow rate of 0.2 µl/min and eluted using a gradient from 2% to 45% acetonitrile in 0.1% formic acid over 60 min. High voltage (2000 V) was applied to low volume tee (Upchurch Scientific) and the column tip positioned ~ 0.5 cm from the heated capillary (T=275°C) of an OrbitrapVelos (Thermo Electron, Bremen, Germany) mass spectrometer. Positive ions were generated by electrospray and the Orbitrap operated in data dependent acquisition mode (DDA). A survey scan m/z 350-1750 was acquired in the Orbitrap (Resolution = 30,000 at m/z 400, with an accumulation target value of 1,000,000 ions) with lock mass enabled. Up to the 10 most abundant ions (>4,000 counts) with charge states > +2 were sequentially isolated and fragmented within the linear ion trap using collisional induced dissociation with an activation q = 0.25 and activation time of 30 ms at a target value of 30,000 ions. M/z ratios selected for MS/ MS were dynamically excluded for 30 seconds.

3.2.11 MS data analysis

The raw MS data was processed using the MaxQuant software (version 1.6.2.1) (Cox and Mann, 2008) and searched against a costum data-base of all predicted proteins in the metagenome of the isoprene reducing culture (6517 sequences). Enzyme specificity was set to trypsin/P, cleaving C terminal to lysine and arginine and a maximum number of two missed cleavages was allowed. Carbamidomethylation of cysteine was set as a fixed modification and oxidation of methionines, acetylation of protein N termini were set as variable modifications. The minimum length of a peptide was set to 7 amino acids and a maximum peptide mass was 4600 Da. The minimal score for modified peptides was 40 and the minimal delta score for modified peptides was 6.

Peptide intensities were normalised using MaxLFQ (Cox *et al.*, 2014). Downstream analysis was performed in R v3.5.1 with the package DEP v1.4.0 (Zhang *et al.*, 2018). First, MaxQuant output data was filtered, retaining only proteins detected by at least two unique peptides, and detected in all replicates of the treatments. LFQ intensities were normalised with vsn (Huber *et al.*, 2002) and missing values imputed by left-censored imputation (MinProb function). Differential expression analysis was conducted with Limma (Ritchie *et al.*, 2015). Proteins were considered to be differentially expressed if they had an adjusted (FDR) P value < 0.05.

3.2.12 Cell suspension experiment

Cells were harvested as described for proteomic analysis but in anaerobic chamber under $N_2/H_2/CO_2$ atmosphere. Cells from six 120 ml flasks (80 ml medium) from H_2/HCO_3^-/isoprene and six flasks from H_2/HCO_3^- grown cells were pooled by pipetting them to two separate 6 ml anaerobic glass flasks. Cells were washed in minimal media containing NH_4Cl (1.2 g l^{-1}), $MgCl_2 \cdot 2H_2O$ (0.4g l^{-1}), $CaCl_2 \cdot 2H_2O$ (0.1 g l^{-1}), 12.5 mM phosphate buffer (pH 7.0) and 1 mM titanium citrate used to ensure low redox conditions. OD_{600} was measured spectrophotometrically outside of the anaerobic chamber and the OD_{600} and the volume (1.57 ml) was then adjusted between the two samples. The flasks were crimp-sealed with Teflon faced rubber septa and flushed with N_2 for 30 mins to remove incumbent isoprene, methylbutenes and CO_2.

The headspace was measured for the presence of isoprene and methylbutene before the experiment was started. Then H_2 (0.7 bar) was added to the samples. To start the experiment, HCO_{3-} was added from a 1 M stock solution to a final concentration of 60 mM and isoprene was added from a 9 mM aqueous stock solution to a final

concentration of 1 mM. Cells were incubated at 30℃ at 180 rpm and the headspace was periodically measured for isoprene depletion and methylbutene production. Additionally, liquid samples (0.04 ml) were periodically sampled by a syringe and frozen at -20℃ for acetate analysis.

3.2.13 Acetate analysis

Acetate was analysed as its ethyl ester derivative by GC-FID using a DB-FFAP column (30 m x 0.32 mm × 0.25 mm, Agilent Technologies) at 40 °C for 6 min with helium as the carrier gas (3 ml min^{-1}). Samples (30 µl) were supplied with 100% ethanol (20 µl) and undiluted sulfuric acid (20 µl) for esterification, sealed immediately and incubated at 60 °C for 45 min. Before injection into the GC, samples were incubated at 80 °C for 5 min at 500 rpm, 250 µl of headspace sample was withdrawn from the flask via an automatic sampler (Shimadzu AOC-5000 plus) and directly injected into a Shimadzu Plus GC-2010 at 500 µl s^{-1}.

3.3 Results

3.3.1 Genome recovery and community structure of isoprene reducing culture

Previously, a novel isoprene reducing culture dominated by *Acetobacterium* sp. was enriched. This culture was further enriched via serial dilutions and anaerobic plating. The isolation process was monitored via Illumina sequencing, clone libraries [Supplement **Figure S3 1**] and denaturing gradient gel electrophoresis (DGGE) [Supplement **Figure S3 2**]. When samples attained their highest purity [Supplement **Figure S3 2 E**] the metagenomes of two samples, one grown with H_2 and HCO_3^- and one with H_2, HCO_3^- and isoprene, were sequenced. Two metagenome-assembled genomes (MAGs) were obtained from the isoprene reducing culture. MAG ISORED-1 showed 74% average amino acid identity (AAI) and 79% average nucleotide identity (ANI) to *Comamonas aquatica* CJD and ISORED-2 showed 97% AAI/ANI to *Acetobacterium wieringae* DSM 1911 [Supplement **Table S3 1, Table S3 2**]. *A. wieringae* strain ISORED-2 is the dominant organism based on the proportions of sequencing coverage [**Table 3.1**]. Relative abundance in sequenced samples from the H_2/HCO_3^-/isoprene sample contained 89% *A. wieringae* ISORED-2 and 11% *Comamonas* sp. strain ISORED-1, while the H_2/HCO_3^- sample contained 77.2% and 22.8% respectively. The genome of *A. wieringae* ISORED-2 was predicted to be 99.28% complete with a genome size of 3.81 Mbp and 44% GC content. The genome of *Comamonas* sp. ISORED-1 was also 99.28% complete with a genome size of 3.22 Mbp and GC content of 63%.

Table 3.1 Summary of MAGs derived from isoprene reducing culture.

				Coverage	
MAGS	**Genome size Mbp**	**GC content %**	**Complete-ness %**	**H_2/HCO_3^- (rel. ab. %)**	**H_2/HCO_3^-/ isoprene (rel. ab. %)**
Comamonas sp. ISORED-1	3.22	63.7	99.28	34.4 (22.8)	16.3 (11.1)
Acetobacterium wieringae ISORED-2	3.81	44	99.28	116.6 (77.2)	131.0 (89.0)

3.3.2 *Acetobacterium wieringae* ISORED-2 unique genes

Even though *A. wieringae* DSM 1911 is closely related to *A. wieringae* ISORED-2, *A. wieringae* DSM 1911 was not able to catalyse isoprene reduction and neither were *A. woodii* DSM 1030 or *A. malicum* DSM 4132. Therefore, comparative pangenome analysis with eight publicly available *Acetobacterium* genomes [Supplement **Table S3 3**] was performed to find unique features in the *A. wieringae* ISORED-2 genome. Anvi'o pangenome analysis shows the protein coding sequences from all nine genomes (33035 in total) grouped into 8190 gene clusters (GCs) [**Figure 3.1**]. A set of shared GCs (1492 GC) was clearly shown across the nine genomes (core). The Anvi'o visualization also highlights the protein sets that are unique in each of the *Acetobacterium* genomes [**Figure 3.1**, **Table 3.2**].

Table 3.2 Summary of shared and unique gene cluster within *Acetobacterium* genus. The "core" selection corresponds to the gene clusters that contain genes from all the genomes. The "soft core" selection corresponds to gene clusters that contain genes from at least 7 genomes and the shell from at least 4 genomes.

Name	Gene cluster	Gene calls
Core (9 genomes)	1493	14589
Soft core (≥7 genomes)	532	4568
Shell (≥ 4 genomes)	1111	6490
***A. wieringae* ISORED-2**	**318**	**352**
A. wieringae all	74	243
A. wieringae DSM 1911	213	230
A. sp. MES1	182	209
A. woodii	558	577
A. dehalogenans	371	392
A. bakii	745	777
A. KB-1	244	255
A. sp. UBA6819	603	610
HWG-Firmicutes-4	463	476

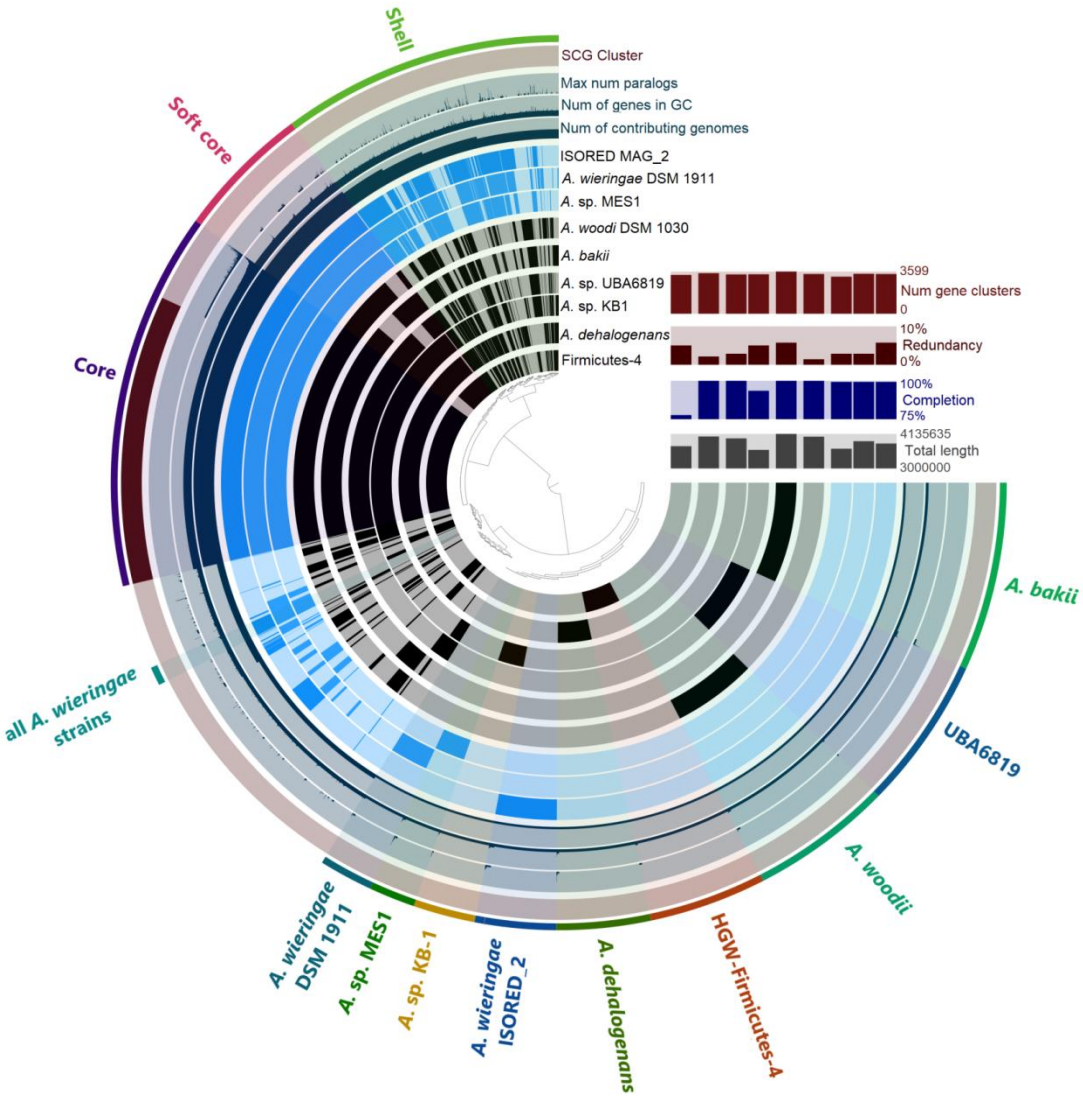

Figure 3.1 Anvi'o pangenome analysis of nine *Acetobacterium* genomes. Each of the 8190 gene clusters contains one or more genes contributed by one or more isolate genomes. The "core" selection corresponds to the gene clusters that contain genes from all the genomes. The "soft core" selection corresponds to gene clusters that contain genes from at least 7 genomes and the shell from at least 4 genomes. "Singletons" selection corresponds to clusters that contain one or multiple genes from a single genome. MAG specific selection marks clusters that contain genes unique to *A. wieringae* ISORED-2 and others. On the right hand side additional data is provided for each isolate. [Supplement **Table S3 3**].

For *A. wieringae* ISORED-2 a total repertoire of 3628 proteins in 3386 gene clusters were identified. Of these gene clusters, 1492 were part of the core genome, 740 were part of the shell and 318 GCs with 352 proteins were unique to the *A. wieringae* ISORED-2 genome amongst *Acetobacterium* spp. [**Table 3.2**]. Of the 352 unique proteins, 46.3% were hypothetical proteins with 31 hypothetical proteins that matched *Acetobacterium* or *Firmicutes* strains, 51 hypothetical proteins matching other taxa and 81 proteins not recognized by BLASTp against the NCBI database. The rest of the proteins (189) could be assigned by COG functions. A total of 91 proteins were phage related such as terminases, phage major capsid protein, site specific recombinase and phage portal protein. Another 98 proteins were related to plasmid mobilisation, bacterial defence system bacteriophage exclusion (BREX) and chemotaxis. One gene cluster was found within the unique proteins that could conceivably catalyse isoprene reduction. The cluster harboured three FAD-dependent oxidoreductases, one 4Fe-4S-ferredoxin and two hydrogenase nickel incorporation proteins HypA. The details of the BLASTp-against-NCBI results of all 352 proteins are presented in the Supplementary Material [Supplement **Table S3 4**].

3.3.3 *De novo* protein synthesis in response to isoprene

To test whether *de novo* protein synthesis is required to induce isoprene reduction, cell suspension experiments were performed. Cells grown with H_2/HCO_3^- or H_2/HCO_3^- /isoprene were harvested and resuspended separately to an OD_{600} of 7.5-7.89 in phosphate buffered minimal media.

When cells suspensions involving H_2/HCO_3^- acquired cells were incubated with H_2/HCO_3^- and isoprene, acetogenesis commenced immediately at 70 nmol min^{-1} for 135 min and then 24 nmol min^{-1} thereafter [**Figure 3.2 A**]. However, a 95 min lag in isoprene reduction was observed, after which isoprene reduction commenced at 1.25 nmol min^{-1}.

In contrast, isoprene reduction and acetogenesis commenced immediately in cell suspensions involving H_2/HCO_3^-/isoprene acquired cells [**Figure 3.2 B**]. Initial isoprene reduction and acetogenic rates were 30 nmols min^{-1} and 37 nmols min^{-1} respectively. Isoprene reduction stopped after 135 mins, at which time the acetogenic rate increased to 77 nmol min^{-1}.

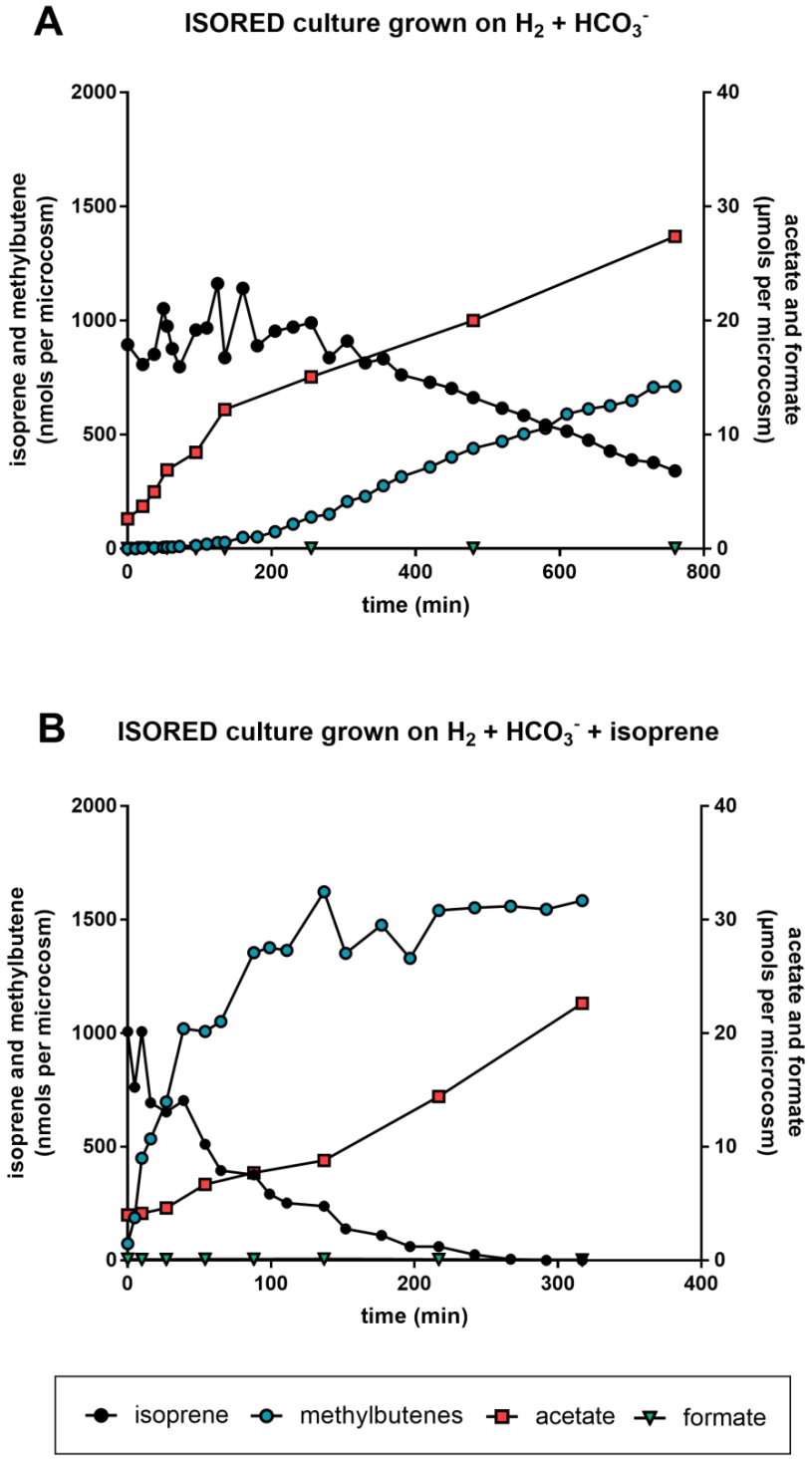

Figure 3.2 Induction of isoprene reduction during acetogenesis from H_2 plus CO_2 by isoprene reducing culture dominated by *A. wieringae* ISORED-2. Cell suspensions of *A. wieringae* ISORED-2 grown on H_2/HCO_3^- without (**A**) and with isoprene (**B**) were incubated under N_2 atmosphere at 30°C in shaking conditions in the presence of 0.5 bar H_2, 40 mM HCO_3^- and 1.3 mM isoprene. At times indicated samples were withdrawn and analysed to determine isoprene, methylbutenes and acetate. Please note that the time scales between A and B are different. Please note that on the left y-axis the unit is nmols per microcosm and on the right y-axis it is µmols per microcosm.

3.3.4 Identification of genes involved in isoprene reduction through proteomics

Since *de novo* protein synthesis is required for isoprene reduction, label-free comparative metaproteomics was performed to identify proteins and corresponding genes involved in isoprene metabolism using proteins predicted from the metagenome as reference database.

The enriched isoprene reducing culture was grown with H_2/HCO_3^- and H_2/HCO_3^- /isoprene with three replicates each and the proteome was sequenced via LC-MS/MS and analysed via label free protein quantification (LFQ) in MaxQuant. A total of 1531 proteins were identified: 1279 proteins belonged to *A. wieringae* ISORED-2 and 252 proteins were assigned to *Comamonas* ISORED-1 [**Figure 3.3**].

A two-sided t-test identified 13 proteins of *A. wieringae* ISORED-2 [**Table 3.3**] and 140 proteins of *Comamonas* sp. ISORED-1 [Supplement **Table S3 5**] that significantly differed in expression between cells grown on H_2/HCO_3^-/isoprene vs. H_2/HCO_3^-. Twelve proteins were upregulated in isoprene supplied cells and they all belonged to *A. wieringae* ISORED-2. Of the 141 proteins downregulated in isoprene grown cells, one belonged to *A. wieringae* ISORED-2 and 140 to *Comamonas* sp. ISORED-1 [**Figure 3.3**]. The significantly different logFC in *Comamonas* sp. ISORED-1 protein abundances might be a result of uneven amounts of total *Comamonas* protein in the H_2/HCO_3^- vs. H_2/HCO_3/isoprene samples and not a result of true differential expression. Based on the ratio of the coverage values from metagenome sequencing, *Comamonas* sp. ISORED-1 showed a relative abundance in H_2/HCO_3 culture of 23% which is double that in isoprene-amended cultures (11%) [**Table 3.1**]. Conversely, *A. wieringae* ISORED-2 relative abundances were 89% and 77% in H_2/HCO_3/isoprene and H_2/HCO_3^- grown culture respectively [**Table 3.1**] (i.e. a 1.15x relative difference in protein abundance). Differential expression of the 13 *A. wierinage* ISORED-2 proteins was therefore considered a true reflection of upregulation because if the difference in intensity was an artefact more than only 13 proteins would be affected.

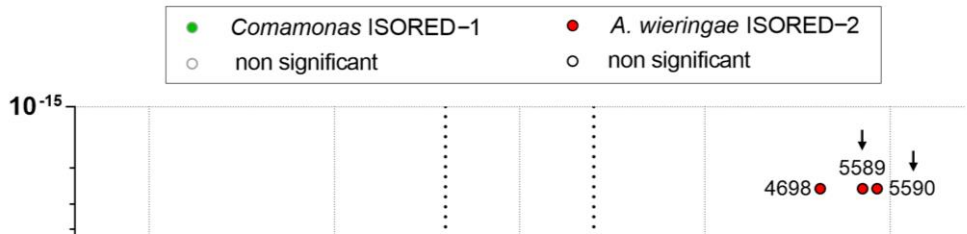

Figure 3.3 Volcano plot of the metaproteomics data analysed by Maxquant comparing the ISORED culture grown on H_2 /HCO_3^-/isoprene vs. H_2/HCO_3^- grown cells. Significant data points are based on a minimum abs(logFC) of 2 and an adjusted p-value of 0.05. Labelling of the significant points is based on the MAGs *Acetobacterium wieringae* ISORED-2 and *Comamonas* sp. ISORED-1. Proteins that are located adjacent to each other in the genome of *A. wieringae* ISORED-2 indicating they belong to the same operon are highlighted by a black arrow.

Interestingly, four out of 13 regulated proteins from *A. wieringae* ISORED-2 are located adjacent to each other in the genome indicating they belong to the same operon (5587, LFC 10.8x; 5590, LFC 9.65x; 5591, LFC 6.63x; 5589 LFC 9.26x) [**Figure 3.3, Table 3.3**]. Since three of these proteins (5587, 5590, 5591) are also unique to *A. wieringae* ISORED-2 amongst *Acetobacterium* it implicates their involvement in isoprene reduction.

Protein 5587 is an FAD-dependent oxidoreductase with the best orthologous group match ENOG4107QZ5. Protein 5591 (ENOG4105WMM) is a HypA homologue, and 5589 (ENOG4107RSS) is a HypB homologue, both involved in inserting Nickel in the catalytic centre of [NiFe]-hydrogenases. Protein 5590 (ENOG4105DQ9) belongs to the 4Fe-4S superfamily (SSF54862) but its sequence does not match any specific family. The other nine differentially expressed proteins in *A. wieringae* ISORED-2 cells grown on isoprene are common to other *Acetobacterium* lineages [**Table 3.3**] and therefore less likely to be involved in isoprene reduction. In addition, they are also not as strongly

upregulated as those in the putative isoprene operon with the exception of protein 4698 [**Figure 3.3**].

Table 3.3 List of proteins that significantly differed in expression between cells grown on H_2/HCO_3^-/isoprene vs. H_2/HCO_3^- in *A. wieringae* ISORED-2. Significant data points are based on a minimum abs(logFC) of 2 and an adjusted p-value of 0.05. Shown are the EggNOG matches with their functional description. Proteins that are unique to *A. wieringae* ISORED-2 amongst *Acetobacterium* spp. are bold. Proteins that are located adjacent to each other are highlighted in grey.

Protein ID	LFC	p-value (adj)	Size AA	EggNOG	Protein/ Enzyme	Function
3185	-3.32	6.34E-05	281	ENOG4108S8M	degv family	unknown
2947	2.46	1.58E-02	149	ENOG41083D8	heat shock protein	stress responds
6170	2.47	1.07E-02	556	ENOG4105CKU	formyltetrahydrofolate synthetase	one-carbon metabolic process
2942	2.48	2.18E-03	461	ENOG4105H0F	uroporphyrinogen decarboxylase (URO-D)	porphyrin-containing compound metabolic process
5235	2.52	2.39E-02	67	ENOG4105VJV	50s ribosomal protein L35	structural constituent of ribosome
5948	2.58	3.12E-02	513	ENOG410ND7V	xylulokinase	carbohydrate metabolic process
1821	2.59	1.14E-03	1025	ENOG4108UGT	methyl-accepting chemotaxis	signal transducer activity
6165	4.56	1.83E-05	167	ENOG4105WDP	acetyltransferase	N-acetyltransferase activity
5591	6.63	2.00E-09	117	ENOG4105WMM	hydrogenase accessory protein HypA	plays a role in a hydrogenase nickel cofactor insertion step
4698	8.11	1.02E-13	327	ENOG41	methyltransf	unknown

				060FG	erase 1 (EC 2.1.1.-)	
5589	9.26	1.02E-13	225	ENOG4107RSS	hydrogenase accessory protein HypB	nickel cation binding, transition metal ion binding hydrolase activity
5590	9.65	1.02E-13	350	ENOG4105DQ9	4Fe-4S ferredoxin	iron-sulfur cluster binding and electron carrier activity
5587	10.8	2.12E-10	901	ENOG4107QZ5	glutamate synthase, FAD-dependent oxidoreductase	oxidoreductase activity; iron-sulfur cluster binding; flavin adenine dinucleotide binding oxidoreductase activity;

3.4 Discussion

3.4.1 Organisms involved in isoprene reduction

This study aimed to investigate organisms, genes and corresponding enzymes implicated in the reduction of isoprene. Previously, an isoprene reducing culture dominated by *Acetobacterium* sp. was enriched, now identified as *Acetobacterium wieringae* ISORED-2. The genome of *A. wieringae* ISORED-2 shows highest sequence similarity to *A. wieringae* DSM 1911 and *A.* sp. MES-1 (Braun and Gottschalk, 1982; Ross *et al.*, 2017) and like other *Acetobacterium* spp. encodes the Wood-Ljungdahl pathway (WLP) for autotrophic growth (Ragsdale and Pierce, 2008), the Na^+-translocating ferredoxin: NAD^+-oxidoreductase (Rnf complex) (Westphal *et al.*, 2018), F_1F_0-ATPase (Fritz and Müller, 2007), the electron transfer flavoproteins (Imkamp *et al.*, 2007) and one hydrogenase with high sequence similarity to the electron-bifurcating [FeFe]-hydrogenase (Schuchmann and Müller, 2012). In addition, it also contains 352 unique genes [**Figure 3.1,** Supplement **Table S3 4**] and since *A. wieringae* DSM 1911 did not exhibit isoprene reducing activity (Kronen *et al.*, 2019), the set of unique genes most likely contains those responsible for isoprene reduction. Even though *A. wieringae* ISORED-2 dominates the isoprene reducing culture (89%), a *Comamonas* sp. which shows highest sequence similarity to *Comamonas aquatica* CJG (Dai *et al.*, 2016), is still present at ~11% of the community in isoprene-grown cells [**Table 3.1**].

3.4.2 Isoprene dependent gene regulation

Analysis of the metaproteome revealed that *A. wieringae* ISORED-2 contributed 83.5% of total identified proteins suggesting high transcriptional and translational activity in *A. wieringae* ISORED-2. Comparative proteomic analysis identified 1279 proteins that belonged to *A. wieringae* ISORED-2 of which only 13 were differentially expressed in response to isoprene [**Figure 3.3, Table 3.3**]. Another 252 proteins were assigned to the *Comamonas* MAG of which 140 were differentially expressed [**Figure 3.3**, Supplement **Table S3 5**].

Of the 13 proteins regulated by isoprene, 12 were upregulated. Four of these were adjacent to each other in a five gene operon of which four genes were unique amongst *Acetobacterium* to *A. wieringae* ISORED-2 [**Table 3.3**, Supplement **Table S3 4**]. This putative five gene operon harbours genes coding for an FAD-dependent oxidoreductase

(5587), three nickel-inserting, hydrogenase maturation factors HypA (5588, 5591), HypB (5589) and one 4Fe-4S ferredoxin (5590) [**Table 3.3**]. Apart from these genes, no other genes responding to isoprene [**Table 3.3**] are unique to this *Acetobacterium* MAG implicating them in isoprene reduction. Regulated genes common to other *Acetobacterium* lineages include the acetyltransferases 6165 (ENOG4105WDP) and 4698 (ENOG41060FG), the latter with 12 copies on the MAG. The chemotaxis protein 1821, also common to the *Acetobacterium* lineages, could be involved in a chemotactic response to isoprene since *Acetobacterium* spp. are capable of a chemotaxis (Müller and Bowien, 1995).

3.4.3 *Comamonas* sp. ISORED-1 metabolism

It is unclear why *Comamonas* sp. ISORED-1 persists in the culture, or if it has any trophic relationship with *A. wieringae* ISORED-2 under the established culturing conditions. Detected *Comamonas* sp. ISORED-1 proteins included proteins for beta-oxidation of fatty acids as well as enzymes involved in the glyoxylate cycle, the citric acid cycle, the 2-methylcitrate cycle and enzyme complexes from the respiratory chain [Supplement **Table S3 5**]. Hence, *Comamonas* sp. ISORED-1 could feed on bacterial necromass so called necromass recycling that has been shown for some Spirochaetes strains co-cultured with a sulphate reducer (Dong *et al.*, 2018). The Spirochaetes were feeding mainly on proteins and carbohydrates derived from dead cells from *Desulfobacterium* N47. It might be possible that *Comamonas* grows on fatty acids obtained from hydrolysis of membranes from dead *Acetobacterium* cells and channels the obtained acetyl-CoA (even and odd fatty acids) and propionyl-CoA (odd fatty acids) in the glyoxylate and 2-methylcitrate cycle respectively (Textor *et al.*, 1997; Muñoz-Elías and McKinney, 2006; Dolan *et al.*, 2018). Further *Comamonas* sp. ISORED-1 expressed the acetyl-CoA ligase, which can activate acetate produced by *A. wieringae* ISORED-2 into the glyoxylate cycle (Dolan and Welch, 2018).

Also, *Comamonas* sp. ISORED-1 involvement in isoprene reduction cannot be excluded. However, in H_2/HCO_3^- grown cells *Comamonas* sp. ISORED-1 made up to 22% of the community compared to H_2/HCO_3^-/isoprene grown cells where *Comamonas* resembled only 11% of the community, which is an indication that these cells benefit from the absence of isoprene rather than depend on it for growth.

3.5 Conclusion

Two organisms were present in the isoprene reducing culture, a *Comamonas* sp. designated MAG ISORED-1 and a novel *Acetobacterium wieringae* strain designated MAG ISORED-2. *A. wieringae* ISORED-2 is the dominant organism the isoprene reducing enrichment cultures and pangenomic analysis revealed that 352 genes were unique amongst *Acetobacterium* spp. to *A. wieringae* ISORED-2. This also includes four adjacently located genes, upregulated upon growth with isoprene identified in comparative proteomic analysis. These genes belong to a putative five gene isoprene operon and encode an FAD-dependent oxidoreductase (5587), three nickel-inserting, hydrogenase maturation factors HypA (5588, 5591), HypB (5589) and one 4Fe-4S ferredoxin (5590). Future experiments will focus on the operon organisation, phylogeny and characterisation of genes in the putative isoprene operon. How *Comamonas* sp. ISORED-1 persists in the culture is still unclear. Proteomic data suggests it may grow on bacterial necromass derived from *A. wieringae* ISORED-2 cell membrane degradation. Its involvement in isoprene reduction can also not be completely excluded.

3.6 References

Atkinson, R. and Arey, J. (2003) Gas-phase tropospheric chemistry of biogenic volatile organic compounds: A review. *Atmos. Environ.* **37**: 197–219.

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

Bertsch, J., Parthasarathy, A., Buckel, W., and Müller, V. (2013) An electron-bifurcating caffeyl-CoA reductase. *J. Biol. Chem.* **288**: 11304–11311.

Braun, M. and Gottschalk, G. (1982) *Acetobacterium wieringae* sp. nov., a new species producing acetic acid from molecular hydrogen and carbon dioxide. *Zbl.Bakt.Hyg.,I.Abt.Orig.C* **3**: 368–376.

Broadgate, W.J., Malin, G., Küpper, F.C., Thompson, A., and Liss, P.S. (2004) Isoprene and other non-methane hydrocarbons from seaweeds: A source of reactive hydrocarbons to the atmosphere. *Mar. Chem.* **88**: 61–73.

Cox, J., Hein, M.Y., Luber, C.A., Paron, I., Nagaraj, N., and 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, J. and Mann, M. (2008) MaxQuant enables high peptide identification rates, individualized p.p.b.-range mass accuracies and proteome-wide protein quantification. *Nat. Biotechnol.* **26**: 1367–1372.

Dai, W., Zhu, Y., Wang, X., Sakenova, N., Yang, Z., Wang, H., et al. (2016) Draft genome sequence of the bacterium *Comamonas aquatica* CJG. *Genome Announc.* **4**: 4–5.

Dani, K.G.S. and Loreto, F. (2017) Trade-off between dimethyl sulfide and isoprene emissions from marine phytoplankton. *Trends Plant Sci.* **22**: 361–372.

Dilling, S., Imkamp, F., Schmidt, S., and Müller, V. (2007) Regulation of caffeate respiration in the acetogenic bacterium *Acetobacterium woodii*. *Appl. Environ. Microbiol.* **73**: 3630–3636.

Ding, C., Chow, W.L., and He, J. (2013) Isolation of *Acetobacterium* sp. strain AG, which reductively debrominates octa- and pentabrominated diphenyl ether technical mixtures. *Appl. Environ. Microbiol.* **79**: 1110–1117.

Dolan, S.K. and Welch, M. (2018) The glyoxylate shunt, 60 years on. *Annu. Rev. Microbiol.* **72**: 309–330.

Dolan, S.K., Wijaya, A., Geddis, S.M., Spring, D.R., Silva-Rocha, R., and Welch, M. (2018) Loving the poison: The methylcitrate cycle and bacterial pathogenesis. *Microbiol. (United Kingdom)* **164**: 251–259.

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.

van Dongen, S. and Abreu-Goodger, C. (2012) Using MCL to extract clusters from networks. In, van Helden,J., Toussaint,A., and Thieffry,D. (eds), *Bacterial Molecular Networks. Methods in Molecular Biology (Methods and Protocols)*. Springer, New York, NY, pp. 281–295.

Dorn, M., Andreesen, J.R., and Gottschalk, G. (1978) Fumarate reductase of Clostridium formicoaceticum. *Arch. Microbiol.* **119**: 7–11.

Eren, A.M., Esen, Ö.C., Quince, C., Vineis, J.H., Morrison, H.G., Sogin, M.L., and Delmont, T.O. (2015) Anvi'o: an advanced analysis and visualization platform for 'omics data. *PeerJ* **3**: e1319.

Fritz, M. and Müller, V. (2007) An intermediate step in the evolution of ATPases - The F_1F_0-ATPase from *Acetobacterium woodii* contains F-type and V-type rotor subunits and is capable of ATP synthesis. *FEBS J.* **274**: 3421–3428.

Gelmont, D., Stein, R.A., and Mead, J.F. (1981) Isoprene- the main hydrocarbon in human breath. *Biochem. Biophys. Res. Commun.* **99**: 1456–1460.

Hackenberg, S.C., Andrews, S.J., Airs, R., Arnold, S.R., Bouman, H.A., Brewin, R.J.W., et al. (2017) Potential controls of isoprene in the surface ocean. *Global Biogeochem. Cycles* **31**: 644–662.

Harley, P.C., Monson, R.K., and Lerdau, M.T. (1999) Ecological and evolutionary

aspects of isoprene emission from plants. *Oecologia* **118**: 109–123.

He, C., Murray, F., and Lyons, T. (2000) Monoterpene and isoprene emissions from 15 *Eucalyptus* species in Australia. *Atmos. Environ.* **34**: 645–655.

Hess, V., González, J.M., Parthasarathy, A., Buckel, W., and Müller, V. (2013) Caffeate respiration in the acetogenic bacterium *Acetobacterium woodii*: A coenzyme a loop saves energy for caffeate activation. *Appl. Environ. Microbiol.* **79**: 1942–1947.

Hess, V., Vitt, S., and Müller, V. (2011) A caffeyl-coenzyme A synthetase initiates caffeate activation prior to caffeate reduction in the acetogenic bacterium *Acetobacterium woodii*. *J. Bacteriol.* **193**: 971–978.

Huber, W., von Heydebreck, A., Sültmann, H., Poustka, A., and Vingron, M. (2002) Variance stabilization applied to microarray data calibration and to the quantification of differential expression. *Bioinformatics* **18**: 96–104.

Huerta-Cepas, J., Szklarczyk, D., Forslund, K., Cook, H., Heller, D., Walter, M.C., 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.

Hyatt, D., Chen, G.L., LoCascio, P.F., Land, M.L., Larimer, F.W., and Hauser, L.J. (2010) Prodigal: Prokaryotic gene recognition and translation initiation site identification. *BMC Bioinformatics* **11**:.

Imkamp, F., Biegel, E., Jayamani, E., Buckel, W., and Müller, V. (2007) Dissection of the caffeate respiratory chain in the acetogen *Acetobacterium woodii*: Identification of an Rnf-type NADH dehydrogenase as a potential coupling site. *J. Bacteriol.* **189**: 8145–8153.

Imkamp, F. and Müller, V. (2002) Chemiosmotic energy conservation with Na^+ as the coupling ion during hydrogen-dependent caffeate reduction by *Acetobacterium woodii*. *J. Bacteriol.* **184**: 1947–1951.

Janson, R. and De Serves, C. (1998) Isoprene emissions from boreal wetlands in Scandinavia. *J. Geophys. Res. Atmos.* **103**: 25513–25517.

Jones, P., Binns, D., Chang, H.Y., Fraser, M., Li, W., McAnulla, C., et al. (2014)

InterProScan 5: Genome-scale protein function classification. *Bioinformatics* **30**: 1236–1240.

Karp, P.D., Latendresse, M., Paley, S.M., Krummenacker, M., Ong, Q.D., Billington, R., et al. (2016) Pathway tools version 19.0 update: Software for pathway/genome informatics and systems biology. *Brief. Bioinform.* **17**: 877–890.

Kesselmeier, J. and Staudt, M. (1999) Biogenic volatile organic compounds (VOC): An overview on emission, physiology and ecology. *J. Atmos. Chem.* **33**: 23–88.

King, J., Koc, H., Unterkofler, K., Mochalski, P., Kupferthaler, A., Teschl, G., et al. (2010) Physiological modeling of isoprene dynamics in exhaled breath. *J. Theor. Biol.* **267**: 626–637.

Kronen, M., Lee, M., Jones, Z.L., and Manefield, M.J. (2019) Reductive metabolism of the important atmospheric gas isoprene by homoacetogens. *ISME J.* **13**: 1168–1182.

Le, S.Q. and Gascuel, O. (2008) An improved general amino acid replacement matrix. *Mol. Biol. Evol.* **25**: 1307–1320.

Li, D., Luo, R., Liu, C.M., Leung, C.M., Ting, H.F., Sadakane, K., et al. (2016) MEGAHIT v1.0: A fast and scalable metagenome assembler driven by advanced methodologies and community practices. *Methods* **102**: 3–11.

Ljungdahl, L.G. and Wood, H.G. (1969) Total synthesis of acetate from CO_2 by heterotrophic bacteria. *Annu. Rev. Microbiol.* **23**: 515–538.

McGenity, T.J., Crombie, A.T., and Murrell, J.C. (2018) Microbial cycling of isoprene, the most abundantly produced biological volatile organic compound on Earth. *ISME J.* **12**: 931–941.

Minh, B.Q., Nguyen, M.A.T., and von Haeseler, A. (2013) Ultrafast approximation for phylogenetic bootstrap. *Mol. Biol. Evol.* **30**: 1188–1195.

Müller, V. (2003) Energy conservation in acetogenic bacteria. *Appl. Environ. Microbiol.* **69**: 6345–6353.

Müller, V. and Bowien, S. (1995) Differential effects of sodium ions on motility in the

homoacetogenic bacteria *Acetobacterium woodii* and *Sporomusa sphaeroides*. *Arch. Microbiol.* **164**: 363–369.

Muñoz-Elías, E.J. and McKinney, J.D. (2006) Carbon metabolism of intracellular bacteria. *Cell. Microbiol.* **8**: 10–22.

Nguyen, L.T., Schmidt, H.A., Von Haeseler, A., and Minh, B.Q. (2015) IQ-TREE: A fast and effective stochastic algorithm for estimating maximum-likelihood phylogenies. *Mol. Biol. Evol.* **32**: 268–274.

Parks, D.H., Chuvochina, M., Waite, D.W., Rinke, C., Skarshewski, A., Chaumeil, P.-A., and Hugenholtz, P. (2018) A standardized bacterial taxonomy based on genome phylogeny substantially revises the tree of life. *Nat. Biotechnol.* **36**: 996–1004.

Pritchard, L., Glover, R.H., Humphris, S., Elphinstone, J.G., and Toth, I.K. (2016) Genomics and taxonomy in diagnostics for food security: Soft-rotting enterobacterial plant pathogens. *Anal. Methods* **8**: 12–24.

Ragsdale, S.W. and Pierce, E. (2008) Acetogenesis and the Wood-Ljungdahl Pathway of CO_2 fixation. *Biochim. Biophys. Acta* **1784**: 1873–1898.

Ritchie, M.E., Phipson, B., Wu, D., Hu, Y., Law, C.W., Shi, W., and Smyth, G.K. (2015) Limma powers differential expression analyses for RNA-sequencing and microarray studies. *Nucleic Acids Res.* **43**: e47.

Ross, D.E., Marshall, C.W., May, H.D., and Norman, R.S. (2017) Metagenome-assembled genome sequences of *Acetobacterium* sp. strain MES1 and *Desulfovibrio* sp. strain MES5 from a cathode-associated acetogenic microbial community. *Genome Announc.* **5**: e00938-17.

Sanadze, G.A. (2004) Biogenic isoprene (a review). *Russ. J. Plant Physiol.* **51**: 729–741.

Schuchmann, K. and Müller, V. (2012) A bacterial electron-bifurcating hydrogenase. *J. Biol. Chem.* **287**: 31165–31171.

Seemann, T. (2014) Prokka: Rapid prokaryotic genome annotation. *Bioinformatics* **30**: 2068–2069.

Seifritz, C., Daniel, S.L., Gossner, A., and Drake, H.L. (1993) Nitrate as a preferred electron sink for the acetogen *Clostridium thermoaceticum*. *J. Bacteriol.* **175**: 8008–8013.

Sharkey, T.D., Wiberley, A.E., and Donohue, A.R. (2008) Isoprene emission from plants: Why and how. *Ann. Bot.* **101**: 5–18.

Terzenbach, D.P. and Blaut, M. (1994) Transformation of tetrachloroethylene to trichloroethylene by homoacetogenic bacteria. *FEMS Microbiol. Lett.* **123**: 213–8.

Textor, S., Wendisch, V.F., De Graaf, A.A., Müller, U., Linder, M.I., Linder, D., and Buckel, W. (1997) Propionate oxidation in *Escherichia coli*: Evidence for operation of a methylcitrate cycle in bacteria. *Arch. Microbiol.* **168**: 428–436.

Urakawa, H., Martens-Habbena, W., and Stahl, D.A. (2010) High abundance of ammonia-oxidizing archaea in coastal waters, determined using a modified DNA extraction method. *Appl. Environ. Microbiol.* **76**: 2129–2135.

Westphal, L., Wiechmann, A., Baker, J., Minton, N.P., and Müller, V. (2018) The Rnf complex is an energy-coupled transhydrogenase essential to reversibly link cellular NADH and ferredoxin pools in the acetogen *Acetobacterium woodii*. *J. Bacteriol.* **200**: 1–13.

Winters, A.J., Adams, M.A., Bleby, T.M., Rennenberg, H., Steigner, D., Steinbrecher, R., and Kreuzwieser, J. (2009) Emissions of isoprene, monoterpene and short-chained carbonyl compounds from *Eucalyptus* spp. in southern Australia. *Atmos. Environ.* **43**: 3035–3043.

Wiśniewski, J.R. (2017) Filter-aided sample preparation: The versatile and efficient method for proteomic analysis. *Methods Enzymol.* **585**: 15–27.

Wiśniewski, J.R. (2016) Quantitative evaluation of filter aided sample preparation (FASP) and multienzyme digestion FASP protocols. *Anal. Chem.* **88**: 5438–5443.

Zhang, X., Smits, A.H., van Tilburg, G.B.A., Ovaa, H., Huber, W., and Vermeulen, M. (2018) Proteome-wide identification of ubiquitin interactions using UbIA-MS. *Nat. Protoc.* **13**: 530–550.

3.7 Supplementary material

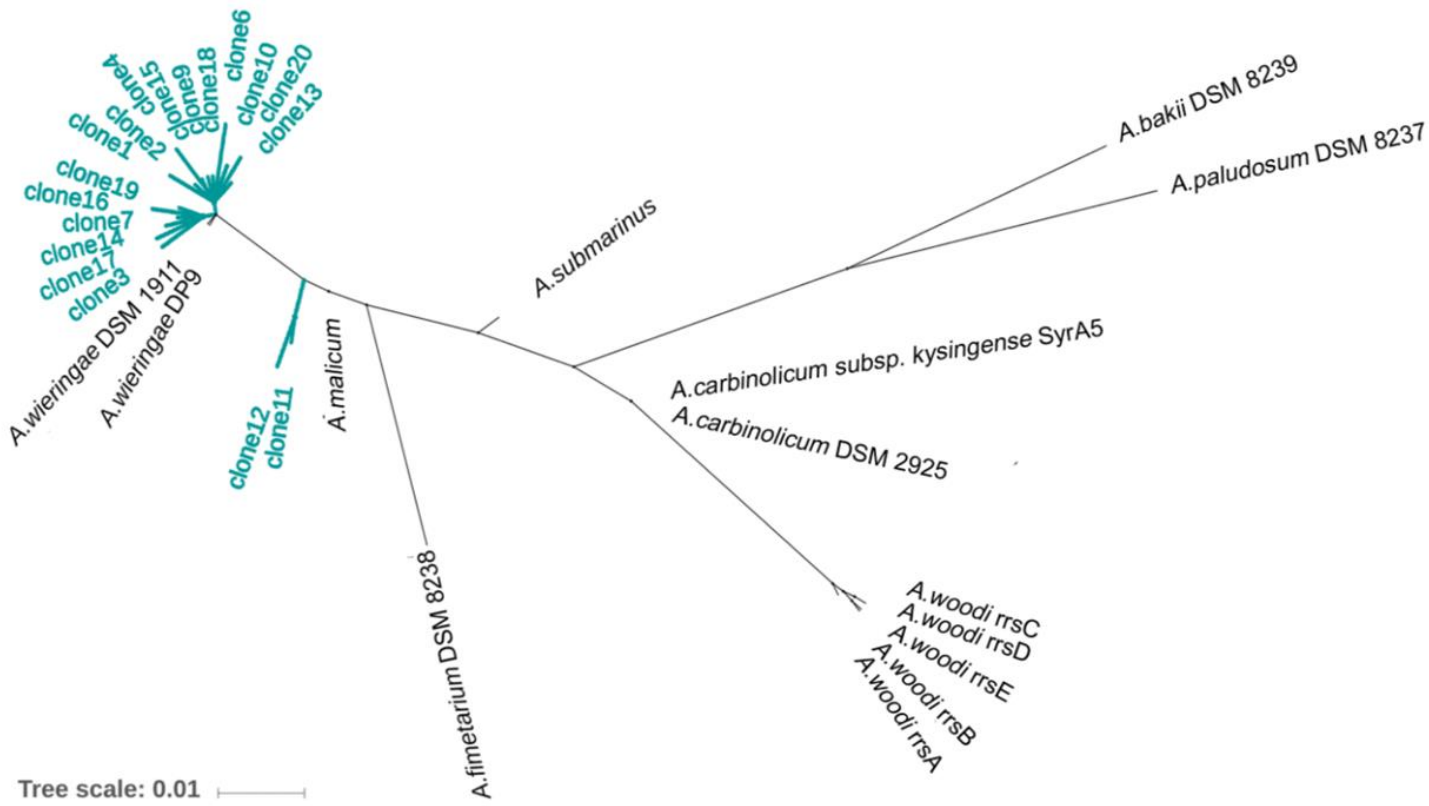

Figure S3 1 Unrooted 16S rRNA gene maximum likelihood phylogenetic tree. The tree relates 19 full length 16S rRNA genes of *Acetobacterium* strains isolated from isoprene enrichment culture to *Acetobacterium* isolates from NCBI. 1000 bootstrap replicates were performed with IQ tree.

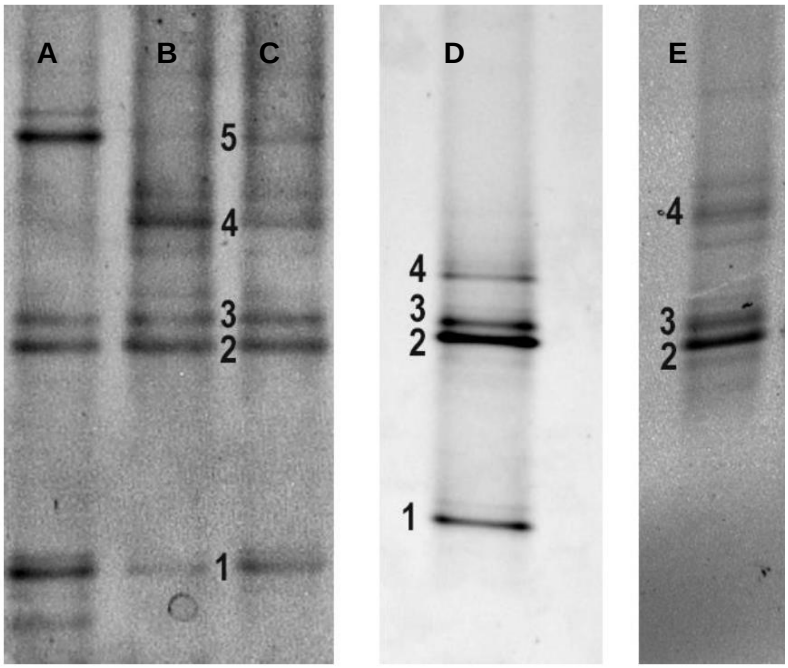

Figure S3 2 DGGE profile of isoprene reducing enrichment culture grown on H_2/HCO_3^- /isoprene over time. 1 *Comamonas* spp., 2-4 *Acetobacterium* spp. 5 *Eubacterium* spp. identified via Sanger sequencing of the PCR product. On the left enrichment after 20 month, three different colonies picked from anaerobic minimal plates (**A-C**), in the middle further dilution to extinction of picked colony after 1 month of incubation (**D**), the right enrichment after further dilution to extinction treatment for another 6 month (**E**). The culture with DGGE profile (**E**) was used for Metagenome sequencing where it seemed like *Comamonas* spp. had disappeared which later turned out not to be the case.

Table S3 1 Average nucleotide identity (ANI) between enriched *A. wieringae* ISORED-2 and close relative genomes available on NCBI (unhighlighted cells; top right) and average amino acid identity (AAI) pairwise comparisons (highlighted cells; bottom left).

	A. woodi	***A. dehalogens***	***A. bakii***	***A. wieringae* DSM 1911**	***A.* MES1**	***A.* sp. UAB6819**	**HGW-Firmicutes-4**	***A.* sp. KB-1**	***A. wieringae* ISORED_2**
A. woodi	100	77.7	75	77.2	77	77.2	76.8	77.4	76.8
A. dehalogens	80.6	100	75.36	83.6	83.3	83.2	87.8	83.1	83.8
A. bakii	74.3	75.3	100	75	75	75.2	75.2	75.5	75
***A. wieringae* DSM 1911**	79.4	88.2	74.5	100	96.9	80.7	82.5	81	97.2
***A.* MES1**	80.1	88.4	74.5	97.4	100	80.7	82.2	80.8	97.1
***A.* sp. UAB6819**	80.4	87.2	74.8	85	85.2	100	82.5	98.3	80.8
HGW-Firmicutes-4	79.1	91	74.8	86.8	86.7	86	100	82.6	82.7
***A.* sp. KB-1**	80.5	87.2	75	85.3	85.3	97.8	86.3	100	80.6
***A. wieringae* ISORED-2**	79.9	88	74.8	97.1	97.2	85	86.3	84.6	100

Table S3 2 Average nucleotide identity (ANI) between enriched *Comamonas* sp. ISORED-1 and close relative genomes available on NCBI (unhighlighted cells; top right) and average amino acid identity (AAI) pairwise comparisons (highlighted cells; bottom left).

	C_aquatica_CJG	C_aquatica_DA1877	C_aquatica_NBRC_14918	C_aquatica_UBA1287	C_kerstersii_121606	C_kerstersii_12322_1	C_kerstersii_202149	C_kerstersii_8943	C_kerstersii_J29	C_kerstersii_UBA11446	C_sp_UBA2122	C_sp_UBA7840	C_terrigena_FDAARGOS_394	C_terrigena_NBRC_13299	C_terrigena_UBA1883	ISORED_MAG-1
C_aquatica_CJG	100.0	97.2	97.0	96.7	81.6	81.8	81.6	81.7	81.7	81.3	81.9	81.9	81.4	81.4	81.3	79.1
C_aquatica_DA1877	98.1	100.0	97.0	96.9	81.1	81.3	81.1	81.1	81.1	81.0	81.6	81.6	81.1	81.1	81.0	78.7
C_aquatica_NBRC_14918	98	98.2	100.0	96.8	81.2	81.3	81.1	81.2	81.2	81.0	81.6	81.4	80.9	81.0	81.0	78.6
C_aquatica_UBA1287	97.8	97.9	97.8	100.0	81.2	81.3	81.2	81.3	81.3	81.1	81.8	81.7	81.1	81.1	81.2	78.7
C_kerstersii_121606	83	82.1	82.2	82.9	100.0	99.3	99.8	99.3	99.3	99.0	80.5	80.4	79.1	79.1	79.3	77.4
C_kerstersii_12322_1	83.4	82.4	82.5	82.9	99.4	100.0	99.3	99.4	99.5	99.2	80.6	80.5	79.4	79.4	79.6	77.4
C_kerstersii_202149	83.02	82	82.1	82.8	99.9	99.46	100.0	99.2	99.3	99.0	80.5	80.4	79.2	79.2	79.3	77.3
C_kerstersii_8943	83.15	82.2	82.1	82.9	99.4	99.5	99.3	100.0	100.0	99.1	80.5	80.5	79.3	79.4	79.3	77.3
C_kerstersii_J29	83.2	82.2	82.2	82.9	99.4	99.5	99.3	100	100.0	99.0	80.6	80.5	79.6	79.6	79.5	77.5
C_kerstersii_UBA11446	82.8	82.16	82.2	83	99.2	99.2	99.2	99	99.1	100.0	80.4	80.4	79.0	79.0	79.2	77.3
C_sp_UBA2122	85.2	84.7	84.7	85.12	83.1	83.4	83	83.4	83.42	83.3	100.0	96.2	78.6	78.6	78.8	76.6
C_sp_UBA7840	85.3	84.5	84.4	85.5	83.12	83.2	83	83	83	83.26	97.5	100.0	78.6	78.7	78.8	76.7
C_terrigena_FDAARGOS_394	77.5	76.8	76.8	77.6	75.8	76	75.7	75.7	75.73	75.7	76.2	76	100.0	99.9	98.4	78.1
C_terrigena_NBRC_13299	77.46	76.8	76.8	77.6	75.8	76	75.74	75.6	75.72	75.7	76.2	76	99.99	100.0	98.4	78.0
C_terrigena_UBA1883	77.32	76.9	76.9	77.6	75.8	76	75.7	75.64	75.4	75.7	76.2	76	98.5	98.5	100.0	78.0
ISORED_MAG-1	74.11	73.62	73.5	74.5	73.07	73.4	73.03	73	73.1	72.9	72.8	73	71.9	71.87	72	100.0

Table S3 3 List of complete *Acetobacterium* genomes used for pangenome analysis with ANVIO.

ID	NCBI Organism Name	NCBI Taxonomy	GTDB Taxonomy
GCA_002237635.1	*Acetobacterium* sp. MES1	d__Bacteria; p__Firmicutes; c__Clostridia; o__Clostridiales; f__Eubacteriaceae; g__Acetobacterium;	d__Bacteria; p__Firmicutes_A; c__Clostridia; o__Eubacteriales; f__Eubacteriaceae; g__Acetobacterium; s__Acetobacterium wieringae
GCA_002452525.1	*Acetobacterium* sp. UBA6819	d__Bacteria; p__Firmicutes; c__Clostridia; o__Clostridiales; f__Eubacteriaceae; g__Acetobacterium;	d__Bacteria; p__Firmicutes_A; c__Clostridia; o__Eubacteriales; f__Eubacteriaceae; g__Acetobacterium;
GCA_002841405.1	Firmicutes bacterium HGW-Firmicutes-4	d__Bacteria; p__Firmicutes; c__; o__; f__; g__; s__	d__Bacteria; p__Firmicutes_A; c__Clostridia; o__Eubacteriales; f__Eubacteriaceae; g__Acetobacterium; s__
GCF_000247605.1	*Acetobacterium woodii* DSM 1030	d__Bacteria; p__Firmicutes; c__Clostridia; o__Clostridiales; f__Eubacteriaceae; g__Acetobacterium; s__Acetobacterium woodii	d__Bacteria; p__Firmicutes_A; c__Clostridia; o__Eubacteriales; f__Eubacteriaceae; g__Acetobacterium; s__Acetobacterium woodii
GCF_000472665.1	*Acetobacterium dehalogenans* DSM 11527	d__Bacteria; p__Firmicutes; c__Clostridia; o__Clostridiales; f__Eubacteriaceae; g__Acetobacterium; s__Acetobacterium dehalogenans	d__Bacteria; p__Firmicutes_A; c__Clostridia; o__Eubacteriales; f__Eubacteriaceae; g__Acetobacterium; s__Acetobacterium dehalogenans
GCF_001263355.1	*Acetobacterium bakii*	d__Bacteria; p__Firmicutes; c__Clostridia; o__Clostridiales; f__Eubacteriaceae; g__Acetobacterium; s__Acetobacterium bakii	d__Bacteria; p__Firmicutes_A; c__Clostridia; o__Eubacteriales; f__Eubacteriaceae; g__Acetobacterium; s__Acetobacterium bakii
GCF_001766835.1	*Acetobacterium wieringae* DSM1911	d__Bacteria; p__Firmicutes; c__Clostridia; o__Clostridiales; f__Eubacteriaceae; g__Acetobacterium; s__Acetobacterium wieringae	d__Bacteria; p__Firmicutes_A; c__Clostridia; o__Eubacteriales; f__Eubacteriaceae; g__Acetobacterium; s__Acetobacterium wieringae
GCA_003260995.1	*Acetobacterium* sp. KB-1	d__Bacteria; p__Firmicutes; c__Clostridia; o__Clostridiales; f__Eubacteriaceae; g__Acetobacterium; s__	NA

Table S3 4 List of 350 unique proteins encoded in the genome of *A. wieringae* ISORED-2. Shown are BLASTp-against-NCBI results of each protein. Proteins are marked in different colours; light grey marks phage-related proteins, dark grey marks proteins with not match on NCBI, yellow marks hypothetical proteins, orange marks proteins with assigned function, dark red marks important gene clusters.

ID	Function	Size AA	Coverage	ID	E value
4	terminase small subunit [Streptococcus iniae]	44	93	71	3.00E-11
5	Phage gp6-like head-tail connector protein [Roseburia faecis]	110	100	51	1.00E-33
6	N4-gp56 family major capsid protein [Streptococcus gordonii]	277	98	47	2.00E-76
7	DUF4355 domain-containing protein [Clostridium sp. OM04-12AA]	135	79	47	1.00E-24
8	Hypothetical protein	64			
9	uncharacterized protein BN631_01216 [Eubacterium dolichum CAG:375]	58	94	38	2.3
10	-	59			
11	hypothetical protein [Faecalicatena contorta]	74	50	46	2.00E-04
12	-	43			
14	hypothetical protein [Acetobacterium wieringae]	88	100	43	3.00E-16
19	-	68			
22	lysine transporter LysE Clostridium puniceum	195	99	70	Mar-92
23	DUF2000 domain-containing protein [Bacillus cereus]	135	100	79	3.00E-77
24	LysR family transcriptional regulator [Lactococcus lactis]	299	98	58	3.00E-120
25	IS4 family transposase [Clostridium uliginosum]	369	98	49	5.00E-123
58	-	35			
59	Hypothetical protein	210			
60	-	75			
61	hypothetical protein [Bacillus pseudomycoides]	435			
62	XRE family transcriptional regulator [Firmicutes bacterium HGW-Firmicutes-5]	144	100	47	2.00E-33
67	-	81			
68	-	77			

69	hypothetical protein [Clostridium beijerinckii]	237			
105	hypothetical protein ACWI_09930 [Acetobacterium wieringae]	83	59	55	9.00E-11
106	-	180			
107	hypothetical protein [Acetobacterium woodii]	178	99	47	1.00E-34
108	hypothetical protein [Acetobacterium woodii]	2231	100	67	0
149	phage major capsid protein [Desulfotomaculum reducens]	247	98	78	6.00E-155
150	-	105			
151	hypothetical protein [Ruminococcaceae bacterium]	120			
157	mtfA protein [Firmicutes bacterium HGW-Firmicutes-12]	230	99	58	4.00E-96
158	hypothetical protein CVU87_09485 [Firmicutes bacterium HGW-Firmicutes-12]	270			
171		84			
172		117			
173		71			
174		68			
175	XRE family transcriptional regulator [bacterium 0.1xD8-82]	251	26	64	5.00E-18
176	kappa-carrageenase precursor [Clostridium aurantibutyricum]	452	56	49	1.00E-62
179	restriction endonuclease subunit R [Syntrophomonas wolfei]	85	100	96	3.00E-51
180	DUF2130 domain-containing protein [Syntrophomonas wolfei]	47	89	100	2.00E-20
181	Metallo-beta-lactamase superfamily protein [Eubacterium aggregans]	274	100	81	5.00E-167
182	HD domain-containing protein [Eubacterium aggregans]	418	99	80	0
232	reverse transcriptase [Geobacillus sp. 1017]	49	93	70	5.00E-16
233	-	112			
234	-	704			
235	-	111			
236	hypothetical protein CVU87_09480 [Firmicutes bacterium HGW-Firmicutes-12]	80	100	75	4.00E-37
237	hypothetical protein CVU87_09485 [Firmicutes bacterium HGW-Firmicutes-12]	1108	100	78	0
256	integrase [Clostridiisalibacter paucivorans]	312	100	82	0

330	PBSX family phage terminase large subunit [Clostridium cadaveris]	188	96	76	1.00E-101
333	-	91			
336	-	34			
368	FAD-dependent oxidoreductase [Acetobacterium wieringae]	257	96	95	8.00E-171
383	-	36			
408	-	62			
647	-	74			
659	-	92			
668	-	196			
669	XRE family transcriptional regulator [Viridibacillus sp. FSL H7-0596]	66	56	43	0.51
670	MULTISPECIES: DNA-binding protein [Caldicellulosiruptor]	66	83	47	5.00E-10
672	pathogenicity island protein [Alkalicoccus saliphilus]	122	90	48	3.00E-30
673	-	92			
674	-	77			
675	-	41			
760	hypothetical protein [Acetobacterium woodii]	99	84	54	1.00E-24
762	hypothetical protein [Acetobacterium wieringae]	291	100	46	2.00E-84
763	hypothetical protein [Acetobacterium bakii]	313	100	41	1.00E-80
764	CpaF family protein [Acetobacterium bakii]	446	96	51	2.00E-149
767	prepilin peptidase [Acetobacterium bakii]	177	71	38	1.00E-20
768	hypothetical protein [Acetobacterium bakii]	110	93	39	4.00E-20
769	hypothetical protein CVU92_08985 [Firmicutes bacterium HGW-Firmicutes-17]	73	95	44	3.00E-13
770	hypothetical protein CVU92_08980 [Firmicutes bacterium HGW-Firmicutes-17]	227	98	44	1.00E-68
771	putative uncharacterized protein [Oscillibacter sp. CAG:241]	80	90	50	2.00E-15
772	hypothetical protein M918_13470 [Clostridium sp. BL8]	109	94	42	3.00E-18
774	hypothetical protein [Pseudobacteroides cellulosolvens]	89	94	30	1.00E-04
775	hypothetical protein [Acetobacterium woodii]	225	96	29	3.00E-25

777	plasmid mobilization relaxosome protein MobC [Acetobacterium wieringae]	176	93	46	7.00E-45
778	type II secretion protein F [Escherichia albertii]	76	73	38	5.7
779	hypothetical protein [Acetobacterium woodii]	439	23	33	2.00E-07
780	hypothetical protein TRSC58_01502 [Trypanosoma rangeli SC58]	140	97	28	5.7
781	type IA DNA topoisomerase [Faecalitalea sp. Marseille-P3755]	688	100	49	0
784	-				
834	hypothetical protein BI182_15400 [Acetobacterium sp. MES1]	454	100	47	4.00E-104
835	SIR2 family protein [Stomatobaculum longum]	777	96	44	0
836	hypothetical protein [Alloprevotella sp. E39]	69	60	40	1.5
842	-	193			
927	transposase mutator type [[Clostridium] sphenoides]	28	96	63	5.00E-04
929	chromosome partitioning protein ParB [Acetobacterium woodii]	305	97	43	2.00E-68
1015	hypothetical protein [Clostridium formicaceticum]	125	92	51	6.00E-35
1016	IS66 family transposase [Peptostreptococcus russellii]	529	94	56	0
1017	hypothetical protein [Candidatus Bathyarchaeota archaeon]	270	84	29	2.00E-18
1019	PKD domain-containing protein [Acetobacterium dehalogenans]	943	50	71	0
1020	reverse transcriptase [[Eubacterium] angustum]	119	97	57	7.00E-38
1021	four helix bundle protein [Andreesenia angusta]	111	100	50	9.00E-39
1022	-	702			
1023	hypothetical protein CVU87_09475 [Firmicutes bacterium HGW-Firmicutes-12]	111	100	48	3.00E-22
1024	hypothetical protein CVU87_09480 [Firmicutes bacterium HGW-Firmicutes-12]	81	100	74	2.00E-35
1025	hypothetical protein CVU87_09485 [Firmicutes bacterium HGW-Firmicutes-12]	1199	100	75	0
1053	NYN domain-containing protein [[Clostridium] lavalense]	221	99	71	3.00E-113
1054	hypothetical protein BAE30_15915 [Acidithiobacillus caldus]	131	69	38	1.00E-09
1055	helix-turn-helix domain-containing protein [Clostridium novyi]	206	64	33	8.00E-11
1056	XRE family transcriptional regulator [Tepidibacillus decaturensis]	67	95	50	2.00E-14
1057	-	46			

1058	DNA-binding protein [Intestinimonas butyriciproducens]	60	93	52	1.00E-13
1059	hypothetical protein [Clostridiales bacterium Marseille-P5551]	679	44	46	3.00E-72
1060	-	101			
1061	-	68			
1135	terminase small subunit [Streptococcus iniae]	44	93	71	3.00E-11
1136	phage gp6-like head-tail connector family protein [Clostridioides difficile]	110	100	50	2.00E-32
1137	Uncharacterised protein [uncultured Ruminococcus sp.]	277	99	70	1.00E-143
1138	DUF4355 domain-containing protein [Clostridium sp. OM04-12AA]	138	78	47	5.00E-25
1139	uncharacterized protein BN631_01216 [Eubacterium dolichum CAG:375]	72	62	40	4.10E+00
1140	xanthine dehydrogenase family protein molybdopterin-binding subunit [Jiangella alba]	60	73	41	5.20E-01
1141	hypothetical protein [Faecalicatena contorta]	76	48	43	3.00E-04
1142	-	43			
1171	hypothetical protein SAMN02744040_00124 [Tepidibacter thalassicus DSM 15285]	47	89	60	4.00E-07
1172	XkdX family protein [Planococcus faecalis]	43	95	78	7.00E-14
1173	aspartate aminotransferase family protein [Aquimixticola sp. Bin25-node5]	61	72	36	8.30E-01
1174	fibronectin type III domain-containing protein [Oscillibacter sp. ER4]	102	100	34	5.00E-06
1252	hypothetical protein C7955_101286 [Eubacterium limosum]	57	64	38	1.20E+00
1253	PTS sugar transporter subunit IIA [Aerococcus urinae]	72	70	32	6.50E+00
1255	hypothetical protein [Geobacillus sp. 1017]	304	56	33	1.00E-17
1259	-	42			
1264	sigma-70 family RNA polymerase sigma factor [Clostridium vincentii]	131	100	36	2.00E-21
1265	HNH endonuclease [Faecalibacterium prausnitzii]	159	66	63	3.00E-38
1266	RNA polymerase subunit sigma-70 [Clostridium merdae]	125	89	71	9.00E-51
1267	terminase large subunit [Desulfotomaculum reducens]	556	100	72	0.00E+00
1268	phage portal protein [Clostridium aceticum	389	98	61	1.00E-171
1269	phage related protease [[Clostridium] sordellii]	297	95	52	9.00E-86
1270	phage major capsid protein [Clostridium amylolyticum]	186	90	63	3.00E-66

1296	DNA-binding response regulator [Acetobacterium sp. MES1]	49	100	100	2.00E-26
1864	methyltransferase [Acetobacterium sp. MES1]	384	97	46	2.00E-116
2043	Pimeloyl-ACP methyl ester carboxylesterase [Seinonella peptonophila]	319	95	58	5.00E-128
2044	MerR family transcriptional regulator [[Clostridium] clostridioforme]	279	99	38	2.00E-55
2306	permease-like protein [Trypanosoma rangeli SC58]	76	60	35	9.60E-01
2508	preprotein translocase subunit YajC [Porphyromonas sp. COT-239 OH1446]	66	62	41	5.80E+00
3171	hypothetical protein CAPSK01_002458 [Candidatus Accumulibacter sp. SK-01]	80	43	49	3.70E+00
3591	magnesium-translocating P-type ATPase [Clostridium minihomine]	881	100	80	0.00E+00
3592	-	100			
3594	zinc ribbon domain-containing protein [Proteiniclasticum ruminis]	119	100	99	1.00E-80
3595	zinc ribbon domain-containing protein [Eubacterium aggregans]	442	100	95	0.00E+00
3596	hypothetical protein [Oxobacter pfennigii]	153	83	87	5.00E-64
3597	zinc ribbon domain-containing protein [Proteiniborus ethanoligenes]	375	100	39	2.00E-90
3598	hypothetical protein [Proteiniborus ethanoligenes]	517	96	48	1.00E-150
3599	DUF4352 domain-containing protein [Hathewaya proteolytica]	186	99	79	9.00E-100
3600	hypothetical protein [Eubacterium aggregans]	687	100	89	0.00E+00
3601	hypothetical protein [Eubacterium aggregans]	169	100	97	9.00E-115
3603	hypothetical protein [Syntrophomonas wolfei]	298	43	92	3.00E-85
3606	cell division protein MraZ [Bacillus sp. OV166]	85	100	87	8.00E-44
3607	aminoglycoside nucleotidyltransferase [Aerococcus sp. SJQ22]	164	100	99	1.00E-115
3610	Msr family ABC-F type ribosomal protection protein [Exiguobacterium sp. S3-2]	487	100	97	0.00E+00
3660	**peptide ABC transporter substrate-binding protein [Desulfosporosinus orientis]**	521	100	88	0.00E+00
3661	**ABC transporter permease [Lactonifactor longoviformis]**	315	100	69	5.00E-166
3662	**glutathione ABC transporter permease GsiD [Firmicutes bacterium HGW-Firmicutes-17]**	306	91	90	0.00E+00
3663	**dipeptide ABC transporter ATP-binding protein DppD [Firmicutes bacterium HGW-Firmicutes-17]**	330	99	74	0.00E+00
3664	**ABC transporter ATP-binding protein [Firmicutes bacterium HGW-Firmicutes-17]**	256	100	84	2.00E-160

3665	**ABC transporter ATP-binding protein [Firmicutes bacterium HGW-Firmicutes-17]**	107	97	84	3.00E-49
3972	-	56			
3973	-	57			
3974	hypothetical protein [Acetobacterium woodii]	85	62	45	
3976	-	120			
3977	-	26			
4202	hypothetical protein [Clostridium perfringens]	133	92	35	3.00E-18
4231	-	46			
4323	**methyl-accepting chemotaxis protein, partial [Firmicutes bacterium HGW-Firmicutes-17]**	63	100	98	1.00E-31
4442	**sensor histidine kinase [Domibacillus enclensis]**	744	79	54	0
4443	**histidine kinase [Firmicutes bacterium HGW-Firmicutes-12]**	116	79	53	6.00E-23
4444	**two-component system response regulator [Firmicutes bacterium HGW-Firmicutes-1]**	354	96	57	1.00E-144
4445	-	338			
4449	hypothetical protein [Pseudoramibacter alactolyticus]	1008	95	42	0.00E+00
4450	phosphoadenosine phosphosulfate reductase [Selenomonas ruminantium]	808	96	51	0.00E+00
4451	hypothetical protein [Desulfosporosinus lacus]	1399	100	77	0.00E+00
4452	BREX-4 system phosphatase PglZ [Desulfosporosinus lacus]	267	97	64	7.00E-117
4489	hypothetical protein [Acetobacterium bakii]	71	100	66	1.00E-09
4501	-	29			
4613	PBSX family phage terminase large subunit [Clostridium cadaveris]	188	96	75	3.00E-101
4616		278	38	36	5.00E-07
4617	zinc-ribbon domain-containing protein [Acetobacterium woodii]	241	69	33	8.00E-10
4618	-	787			
4619	toxin-antitoxin system HicB family antitoxin [Schwartzia succinivorans]	59	91	50	3.00E-13
4621	DUF1643 domain-containing protein [Ruminococcus bromii]	202	87	48	8.00E-48
4622	-	42			
4623	hypothetical protein [Parasporobacterium paucivorans]	103			

4624	hypothetical protein [Proteiniclasticum ruminis]	120			
4625	recombinase family protein [Clostridium tertium]	553	92	34	5.00E-93
4626	-	53			
4627	DNA primase [Marasmitruncus massiliensis]	512	80	37	3.00E-73
4628	von Willebrand factor D and EGF domain-containing protein isoform X2 [Larimichthys crocea]	175	44	33	7.5
4629	-	199			
4630	-	168			
4631	DNA polymerase I [Clostridium butyricum DORA_1]	373	95	35	2.00E-75
4632	LysR family transcriptional regulator [Pseudoalteromonas rubra]	424	12	38	9.4
4633	hypothetical protein [Kluyvera ascorbata]	37			
4655	hypothetical protein [Bacillus cereus]	538	100	47	5.00E-160
4656	DNA polymerase III subunit gamma/tau [Brochothrix thermosphacta]	285	98	46	3.00E-70
4657	zinc-ribbon domain-containing protein [Aerococcus urinaeequi]	599	85	31	1.00E-57
4658	ATP-binding protein [Staphylococcus sp. EZ-P03]	544	94	40	8.00E-120
4659	-	136			
4660	hypothetical protein [Megasphaera genomosp. type_1]	290	100	64	4.00E-127
4661	AAA family ATPase [Megasphaera genomosp. type_1]	676	77	77	0.00E+00
4662	hypothetical protein [Megasphaera genomosp. type_1]	502	100	64	0.00E+00
4664	GTP-binding protein [Acetobacterium sp. KB-1]	380	94	43	8.00E-97
4759	hypothetical protein [Acetobacterium sp. KB-1]	1207	75	64	0.00E+00
4764	hypothetical protein CVU92_00875 [Firmicutes bacterium HGW-Firmicutes-17]	53	60	81	5.00E-07
4773	cell wall-binding protein [Desulfitobacterium hafniense]	187	86	36	2.00E-12
4776	serine/threonine-protein phosphatase [Algoriphagus aquaeductus]	59	69	37	2.40E+00
4777	-	89			
4778	-	105			
4779	-	77			
4780	-	56			

4781	phage/plasmid primase P4 family domain-containing protein [Clostridium sp. ASF356]	792	55	55	0
4782	ICEBs1 excisionase [Massilimaliae massiliensis]	74	82	59	2.00E-17
4783	-	46			
4784	toxin-antitoxin system antitoxin component Xre family [Clostridium sp. CAG:58]	66	96	55	5.00E-15
4785	-	255			
4786	site-specific recombinase, phage integrase family [Clostridium sp. MSTE9]	412	87	60	1.00E-160
4893	DUF4430 domain-containing protein [Paenibacillus gorillae]	281	98	29	2.00E-13
4895	-	78			
4902	S-layer homology domain-containing protein [Anoxybacillus gonensis]	883	75	32	4.00E-58
4907	-	240			
4908	DNA primase catalytic core domain protein [Thermoanaerobacterales bacterium 50_218]	845	97	29	3.00E-85
4909	bifunctional 3'-5' exonuclease/DNA polymerase [Eisenbergiella massiliensis]	793	88	31	6.00E-86
4910	-	50			
5012	hypothetical protein CVU87_09485 [Firmicutes bacterium HGW-Firmicutes-12]	270	99	81	4.00E-146
5013	mtfA protein [Firmicutes bacterium HGW-Firmicutes-12]	232	89	62	3.00E-94
5014	phage tail tape measure protein [Firmicutes bacterium HGW-Firmicutes-17]	982	26	49	2.00E-59
5015	hypothetical protein [Bacillus loiseleuriae]	95	82	35	5.00E-07
5019	putative phage head-tail adaptor [Flavonifractor plautii]	120	85	34	2.00E-14
5020	hypothetical protein [Eubacterium limosum]	105	91	60	4.00E-34
5021	BREX-4 system phosphatase PglZ [Desulfosporosinus lacus]	551	93	73	0.00E+00
5022	BREX system Lon protease-like protein BrxL [Desulfosporosinus lacus]	481	100	86	0.00E+00
5023	hypothetical protein [Monoglobus pectinilyticus]	65	100	72	2.00E-23
5024	cysteine desulfurase [Desulfosporosinus lacus]	374	99	67	0.00E+00
5025	hypothetical protein CVU92_06020 [Firmicutes bacterium HGW-Firmicutes-17]	76	96	92	8.00E-41
5027	O-acetyl-ADP-ribose deacetylase [Catenibacterium mitsuokai]	245	99	45	3.00E-65
5028	hypothetical protein CVU92_06440 [Firmicutes bacterium HGW-Firmicutes-17]	441	100	82	0.00E+00
5043	-	184			

5303	**methyl-accepting chemotaxis protein [Acetobacterium wieringae]**	461	100	99	0.00E+00
5305	**methyl-accepting chemotaxis protein [Acetobacterium wieringae]**	211	100	99	1.00E-133
5306	**methyl-accepting chemotaxis protein [Acetobacterium wieringae]**	112	100	98	3.00E-63
5364	hypothetical protein CVU92_00210 [Firmicutes bacterium HGW-Firmicutes-17]	451	100	91	0.00E+00
5365	hypothetical protein CVU92_00205 [Firmicutes bacterium HGW-Firmicutes-17]	173	94	87	7.00E-98
5366	glycosyltransferase [Roseburia inulinivorans]	265	96	59	4.00E-112
5367	GtrA family protein [Firmicutes bacterium HGW-Firmicutes-17]	137	100	93	2.00E-88
5368	glycosyltransferase [Firmicutes bacterium HGW-Firmicutes-17]	325	99	92	0.00E+00
5402	prepilin-type N-terminal cleavage/methylation domain-containing protein [Thiocapsa marina]	265	98	33	3.00E-37
5404	type II secretion system protein [Acetobacterium woodii]	142	95	53	2.00E-35
5405	hypothetical protein CVU99_09255 [Firmicutes bacterium HGW-Firmicutes-4]	722	95	38	2.00E-113
5411	prepilin-type N-terminal cleavage/methylation domain-containing protein [Acetobacterium sp. KB-1]	189	98	59	5.00E-55
5414	prepilin-type N-terminal cleavage/methylation domain-containing protein [Acetobacterium wieringae]	209	100	57	1.00E-52
5498	-	31			
5501	PBSX family phage terminase large subunit [Clostridium cadaveris]	415	96	79	0.00E+00
5502	terminase small subunit [Streptococcus intermedius]	142	94	59	3.00E-46
5503	Phage gp6-like head-tail connector protein [Roseburia faecis]	110	100	53	1.00E-37
5504	hypothetical protein [Massilimaliae massiliensis]	277	99	69	7.00E-141
5505	DUF4355 domain-containing protein [Clostridium sp. OM04-12AA]	135	79	47	2.00E-24
5506	-	64			
5507	-	64			
5508	hypothetical protein [Faecalicatena contorta]	76	48	46	2.00E-04
5509	-	60			
5511	-	61			
5512	methyl-accepting chemotaxis protein [Acetobacterium woodii]	921	66	38	3.00E-139
5534	hypothetical protein [Clostridium butyricum]	225	100	63	4.00E-90

5579	**LysR family transcriptional regulator [Acetobacterium wieringae]**	311	89	31	8.00E-42
5580	**FAD-dependent oxidoreductase [Acetobacterium wieringae]**	661	99	47	0.00E+00
5583	**hypothetical protein A2Y89_05740 [Chloroflexi bacterium RBG_13_51_18]**	164	79	47	3.00E-28
5585	**PucR family transcriptional regulator [Sporomusa silvacetica]**	519	94	30	5.00E-59
5586	**FAD-dependent oxidoreductase [Lachnoclostridium phytofermentans]**	329	89	30	1.00E-34
5587	**MULTISPECIES: FAD-dependent oxidoreductase [Coprobacillus]**	901	99	56	0.00E+00
5588	**hydrogenase nickel incorporation protein HypA [Firmicutes bacterium HGW-Firmicutes-20]**	118	94	52	8.00E-39
5590	**4Fe-4S ferredoxin [Deltaproteobacteria bacterium HGW-Deltaproteobacteria-1]**	350	97	42	3E-91
5591	**hydrogenase maturation nickel metallochaperone HypA, partial [Deltaproteobacteria bacterium HGW-Deltaproteobacteria-5]**	117	98	52	2.00E-40
5592	**PucR family transcriptional regulator [Ruminococcus sp. OF02-6]**	511	99	33	9.00E-70
5601	recombinase [Ruminiclostridium papyrosolvens C7]	493	99	63	0.00E+00
5730	hypothetical protein CVU92_01240 [Firmicutes bacterium HGW-Firmicutes-17]	201	97	35	4.00E-34
5841	DNA primase [Flavonifractor plautii]	563	74	43	9.00E-118
5843	hypothetical protein [Lysinibacillus sphaericus]	85	77	41	1.00E-05
5853	XRE family transcriptional regulator [Ruminococcus sp. AF14-10]	83	89	45	1.00E-14
5857	-	87			
5858	-	83			
5859	hypothetical protein [Acetobacterium dehalogenans]	83	97	47	4.00E-17
5919	clostridial hydrophobic W [Acetobacterium wieringae]	558	89	49	4.00E-100
5923	hypothetical protein [Acetobacterium bakii]	529	99	48	1.00E-150
5924	hypothetical protein A8709_13975 [Paenibacillus pectinilyticus]	67	94	60	7.00E-13
5981	DUF4062 domain-containing protein [Clostridium botulinum]	141	85	69	2.00E-51
5982	DUF4062 domain-containing protein [Clostridium botulinum]	304	89	39	4.00E-51
5983	-	355			
6215	-	79			
6274	site-specific integrase [Eubacterium sp. AM05-23]	405	94	51	4.00E-131

6275	hypothetical protein [Acetobacterium bakii]	245	79	39	5.00E-50
6276	putative repressor LexA [Eubacterium sp. CAG:156]	215	100	47	9.00E-53
6277	XRE family transcriptional regulator [Clostridia bacterium UC5.1-1D1]	70	92	66	2.00E-26
6278	hypothetical protein [Brevibacillus borstelensis]	213	100	64	2.00E-91
6279	MULTISPECIES: XRE family transcriptional regulator [Clostridiales]	130	59	42	3.00E-09
6280	-	55			
6281	DNA-binding protein [Butyricicoccus sp. 1XD8-22]	59	55	45	2.00E-02
6282	-	186			
6283	-	57			
6284	methyl-accepting chemotaxis protein [Jeongeupia sp. USM3]	116	50	36	0.77
6285	-	57			
6286	hypothetical protein [Acetobacterium dehalogenans]	67	98	44	2.00E-11
6287	XRE family transcriptional regulator [Lactobacillus reuteri]	116	50	46	1.00E-05
6288	-	74			
6289	AP2 domain-containing protein [Listeria monocytogenes]	162	94	48	2.00E-40
6290	NAD(P)/FAD-dependent oxidoreductase [Skermanella stibiiresistens]	60	45	56	1.70E+00
6291	PPE family protein [Anaerovirgula multivorans]	284	98	62	5.00E-122
6292	-	77			
6293	-	81			
6294	hypothetical protein [Lactobacillus parabuchneri]	73	82	28	1.40E-02
6295	hypothetical protein [Clostridiales bacterium oral taxon 876]	154	54	51	6.00E-18
6296	DUF2800 domain-containing protein [Thermoanaerobacterales bacterium SK-G1]	385	98	60	2.00E-163
6297	DUF2815 family protein [Thermoanaerobacterales bacterium SK-G1]	273	98	51	2.00E-77
6298	XRE family transcriptional regulator [Clostridiales bacterium DRI-13]	660	100	66	0.00E+00
6300	VRR-NUC domain-containing protein [Clostridium pasteurianum]	95	97	68	2.00E-36
6302	nucleotide exchange factor GrpE [Candidatus Marinimicrobia bacterium]	164	56	32	1.70E-01
6303	hypothetical protein CVU99_02560 [Firmicutes bacterium HGW-Firmicutes-4]	162	98	3.20E+01	4.00E-11

6304	class I SAM-dependent methyltransferase [Staphylococcus pseudintermedius]	155	97	74	6.00E-78
6305	hypothetical protein [Intestinimonas butyriciproducens]	141	65	75	3.00E-41
6306	hypothetical protein [Acetobacterium dehalogenans]	55	100	41	2.00E-03
6307	hypothetical protein CVU92_04025 [Firmicutes bacterium HGW-Firmicutes-17]	171	97	81	2.00E-97
6308	terminase small subunit [Clostridium botulinum]	173	100	55	2.00E-61
6309	terminase [[Clostridium] lavalense]	460	99	72	0.00E+00
6310	phage portal protein [Firmicutes bacterium HGW-Firmicutes-17]	504	96	86	0.00E+00
6311	phage putative head morphogenesis protein, SPP1 gp7 family [Sporomusa acidovorans]	570	99	43	2.00E-151
6312	-	62			
6313	-	109			
6314	FAD-dependent oxidoreductase [Paenibacillus sanguinis]	86	60	40	4.30E+00
6315	MULTISPECIES: hypothetical protein [Enterobacteriaceae]	88	100	54	1.00E-24
6316	-	42			
6317	phage minor structural protein GP20 [Desulfitobacterium dehalogenans]	203	84	44	9.00E-29
6318	hypothetical protein CVU92_02260 [Firmicutes bacterium HGW-Firmicutes-17]	117	100	92	3.00E-66
6319	capsid protein [Eisenbergiella massiliensis]	341	97	32	4.00E-46
6320	phage gp6-like head-tail connector family protein [Clostridioides difficile]	111	100	56	1.00E-41
6321	putative phage protein [Clostridioides difficile P41]	122	91	45	5.00E-23
6322	HK97 gp10 family phage protein [Lysinibacillus xylanilyticus]	137	100	61	4.00E-59
6323	phage protein, partial [[Clostridium] sordellii 8483]	138	96	30	4.00E-09
6324	Putative phage protein [Clostridioides difficile T15]	61	80	47	8.00E-09
6325	phage tail protein [Firmicutes bacterium HGW-Firmicutes-17]	440	100	84	0.00E+00
6326	Phage-like element PBSX protein xkdM [uncultured Clostridium sp.]	166	90	65	1.00E-71
6327	XkdN-like protein [Flavonifractor plautii]	141	95	39	7.00E-23
6328	phage tail tape measure protein [Neglecta sp. Marseille-P3890]	954	35	47	2.00E-92
6329	LysM peptidoglycan-binding domain-containing protein [Clostridium puniceum]	230	97	45	1.00E-54
6330	hydrolase [Firmicutes bacterium HGW-Firmicutes-17]	334	100	94	0.00E+00

6331	DUF2577 domain-containing protein [Lachnospiraceae bacterium oral taxon 082]	111	100	41	2.00E-17
6332	DUF2634 domain-containing protein [Anaeromassilibacillus sp. An200]	127	89	49	4.00E-29
6333	phage tail protein [Firmicutes bacterium HGW-Firmicutes-17]	355	100	94	0.00E+00
6334	hypothetical protein CVU92_02340 [Firmicutes bacterium HGW-Firmicutes-17]	180	100	86	5.00E-112
6335	hypothetical protein DBX41_04985 [Clostridiales bacterium]	226	85	45	9.00E-35
6338	XkdX family protein [Ruminiclostridium papyrosolvens]	42	100	71	4.00E-11
6342	hypothetical protein [Hungatella hathewayi]	347	96	36	1.00E-67
6343	-	48			
6349	-	373			

Table S3 5 List of proteins assigned to *Comamonas* ISORED-1 identified in proteomics analysis. Significant data points (true) are based on a minimum abs(logFC) of 2 and an adjusted p-value of 0.05. Shown are the EggNOG matches with their functional description. Blue marks enzymes involved in the glyoxylate cycle, green marks enzymes from the citric acid cycle, grey marks enzymes from the 2-methylcitrate cycle, orange marks enzymes involved in the beta-oxidation of fatty acids, purple marks enzyme complexes from the respiratory chain, red marks the ATP synthase and yellow marks enzymes involved in respiration of nitrite.

Protein ID	LFC	P value (adj.)		EggNOG	Protein/Enzyme	Classification
161	-5.58	0.000357	TRUE	ENOG4105CP9	extracellular ligand-binding receptor	Amino acid transport
3838	-3.35	0.44	FALSE	ENOG41061MQ	amino acid ABC transporter	Amino acid transport and metabolism
4014	-1.68	0.159	FALSE	ENOG4105E99	stereospecific condensation of phosphoenolpyruvate (PEP) and D-erythrose-4-phosphate (E4P) giving rise to 3-deoxy-D- arabino-heptulosonate-7-phosphate (DAHP) (By similarity)	Amino acid transport and metabolism
4283	-2.75	0.333	FALSE	ENOG4107R37	glutamate dehydrogenase	Amino acid transport and metabolism
4946	-1.76	0.232	FALSE	ENOG4105CG0	the beta subunit is responsible for the synthesis of L- tryptophan from indole and L-serine (By similarity)	Amino acid transport and metabolism
297	-2.94	0.0035	TRUE	ENOG4105CDH	citrulline--aspartate ligase	Amino acid transport and metabolism
572	-3.86	0.000109	TRUE	ENOG4105CFH	aspartokinase	Amino acid transport and metabolism
718	-3.96	0.000199	TRUE	ENOG4108IJ0	aspartate ammonia-lyase	Amino acid transport and metabolism
1111	-0.221	0.878	FALSE	ENOG410B7PU	acetylornithine aminotransferase	Amino acid transport and metabolism
1128	-2.82	0.0144	TRUE	ENOG4105CDP	catalyzes the condensation of (S)-aspartate-beta- semialdehyde (S)-ASA and pyruvate to 4-hydroxy- tetrahydrodipicolinate (HTPA) (By similarity)	Amino acid transport and metabolism
1439	-3.93	0.0278	TRUE	ENOG4106ADB	ABC transporter substrate-binding protein	Amino acid transport and metabolism
1642	-3.5	0.0996	FALSE	ENOG4105E89	ABC transporter	Amino acid transport and metabolism
1656	-1.61	0.187	FALSE	ENOG4105DGA	diaminopimelate dehydrogenase	Amino acid transport and metabolism
2102	-1.7	0.67	FALSE	ENOG4105F85	ABC transporter substrate-binding protein	Amino acid transport and metabolism

2110	-2.89	0.0087	TRUE	ENOG4105C0C	catalyzes the oxidation of 3-carboxy-2-hydroxy-4- methylpentanoate (3-isopropylmalate) to 3-carboxy-4-methyl-2- oxopentanoate. The product decarboxylates to 4-methyl-2 oxopentanoate (By similarity)	Amino acid transport and metabolism
2114	-3.09	0.131	FALSE	ENOG4107QJH	catalyzes the NADPH-dependent formation of L-aspartate- semialdehyde (L-ASA) by the reductive dephosphorylation of L- aspartyl-4-phosphate (By similarity)	Amino acid transport and metabolism
2149	-4.2	4.39E-09	TRUE	ENOG4107TA6	extracellular ligand-binding receptor	Amino acid transport and metabolism
2432	-2.47	0.00168	TRUE	ENOG4108IQK	o-acetylhomoserine	Amino acid transport and metabolism
3774	-0.543	0.832	FALSE	ENOG410B7PX	aminotransferase	Amino acid transport and metabolism
3775	-3.52	0.000247	TRUE	ENOG4105D6E	homoserine dehydrogenase	Amino acid transport and metabolism
3907	-2.53	0.356	FALSE	ENOG41061MQ	amino acid ABC transporter	Amino acid transport and metabolism
4922	-2.88	0.00146	TRUE	ENOG4105C6M	alpha-keto-beta-hydroxylacyl reductoisomerase	Amino acid transport and metabolism
4923	-3.43	0.0132	TRUE	ENOG4108ZP8	acetolactate synthase small subunit	Amino acid transport and metabolism
4924	-3.42	0.000129	TRUE	ENOG4105C7K	acetolactate synthase	Amino acid transport and metabolism
4981	-2.14	0.307	FALSE	ENOG410B7BM	dihydroxy-acid dehydratase	Amino acid transport and metabolism
4095	-3.06	0.000517	TRUE	ENOG4105CGF	aromatic amino acid aminotransferase	Aromatic amino acid aminotransferase
1501	-2.66	0.147	FALSE	ENOG4105C70	catalyzes the reversible conversion of 2- phosphoglycerate into phosphoenolpyruvate. It is essential for the degradation of carbohydrates via glycolysis (By similarity)	Carbohydrate transport and metabolism
720	-3.45	0.000000244	TRUE	ENOG4105C17	glyceraldehyde-3-phosphate dehydrogenase	Carbohydrate transport and metabolism
721	-1.79	0.413	FALSE	ENOG4105CV1	transketolase (EC 2.2.1.1)	Carbohydrate transport and metabolism
1086	-2.36	0.143	FALSE	ENOG4105CH5	extracellular solute-binding protein, family 7	Carbohydrate transport and metabolism
1087	-3.82	0.0000237	TRUE	ENOG4105CH5	extracellular solute-binding protein, family 7	Carbohydrate transport and metabolism
1520	-2.68	0.342	FALSE	ENOG410B993	TRAP dicarboxylate transporter, DctP subunit	Carbohydrate transport and metabolism
2175	-1.78	0.18	FALSE	ENOG41069A1	TRAP dicarboxylate transporter, dctp subunit	Carbohydrate transport and metabolism

2776	-3.19	0.0000563	TRUE	ENOG4108HRN	catalyzes the phosphorylation of pyruvate to phosphoenolpyruvate (By similarity)	Carbohydrate transport and metabolism
3325	-1.79	0.352	FALSE	ENOG410B7W8	fructose-bisphosphate aldolase	Carbohydrate transport and metabolism
3301	-1.82	0.176	FALSE	ENOG4107QMS	site-determining protein	Cell cycle control, cell division, chromosome partitioning
3989	-2.32	0.223	FALSE	ENOG4105DFF	rod shape-determining protein mreb	Cell cycle control, cell division, chromosome partitioning
1468	-2.69	0.451	FALSE	ENOG4107RWK	flagellar hook-associated protein	Cell motility
115	-1.97	0.246	FALSE	ENOG4105I03	flagellin	Cell motility
116	-2.73	0.0484	TRUE	ENOG4105I03	flagellin	cell motility
1467	-2.62	0.539	FALSE	ENOG4105DTJ	flagellar hook-associated protein	Cell motility
2821	-0.634	0.856	FALSE	ENOG4105I03	flagellin	Cell motility
5272	-2.25	0.489	FALSE	ENOG4105DTS	flagellar hook-associated	Cell motility
1483	-2.92	0.0795	FALSE	ENOG4105CWU	biogenesis protein	Cell motility, Intracellular trafficking, secretion, and vesicular transport
2115	-1.26	0.332	FALSE	ENOG41088DR	inherit from COG: Pilus assembly protein	Cell motility, Intracellular trafficking, secretion, and vesicular transport
972	-3.18	0.0000757	TRUE	ENOG4108V3W	chemotaxis protein, CheW	cell motility, Signal transduction mechanisms
213	-1.06	0.512	FALSE	ENOG4105YRI	part of the outer membrane protein assembly complex, which is involved in assembly and insertion of beta-barrel proteins into the outer membrane (By similarity)	Cell wall/membrane/envelope biogenesis
912	-3.24	0.0521	FALSE	ENOG4106FUM	outer membrane efflux protein	Cell wall/membrane/envelope biogenesis
997	-2.78	0.014	TRUE	ENOG4108RFI	molecular chaperone that interacts specifically with outer membrane proteins, thus maintaining the solubility of early folding intermediates during passage through the periplasm (By similarity)	cell wall/membrane/envelope biogenesis
1129	-2.74	0.708	FALSE	ENOG4105GJM	(LipO)protein	Cell wall/membrane/envelope biogenesis

1555	-2.52	0.0581	FALSE	ENOG4105UYK	outer membrane protein	Cell wall/membrane/envelope biogenesis
2431	-2.44	0.00404	TRUE	ENOG4105EIF	outer membrane protein W	Cell wall/membrane/envelope biogenesis
2989	-2.93	0.016	TRUE	ENOG4105C3I	NAD-dependent epimerase dehydratase	Cell wall/membrane/envelope biogenesis
3431	-2.03	0.203	FALSE	ENOG410B8YI	DegT DnrJ EryC1 StrS aminotransferase	Cell wall/membrane/envelope biogenesis
4055	-2.19	0.326	FALSE	ENOG4108Z9W	outer membrane porin	Cell wall/membrane/envelope biogenesis
5649	-3.6	0.0000159	TRUE	ENOG4108ZG0	Ompa motb domain protein	Cell wall/membrane/envelope biogenesis
5658	-2.69	0.0203	TRUE	ENOG4105KCK	peptidoglycan-associated lipoprotein	Cell wall/membrane/envelope biogenesis
889	-2.47	0.307	FALSE	ENOG4105PCN	Swib mdm2 domain-containing protein	Chromatin structure and dynamics
203	-2.98	0.0000264	TRUE	ENOG4105C3E	may play a key role in the regulation of the intracellular concentration of adenosylhomocysteine (By similarity)	coenzyme transport and metabolism
2752	-0.703	0.784	FALSE	ENOG410B7IK	oxygen-independent coproporphyrinogen III oxidase	Coenzyme transport and metabolism
853	-1.65	0.298	FALSE	ENOG4107QM5	**aconitate hydratase**	Energy production and conversion
82	-2.82	0.0038	TRUE	ENOG4107QP3	**malate synthase**	Energy production and conversion
314	-3.21	0.00000176	TRUE	ENOG4105CH8	**succinyl-CoA ligase ADP-forming subunit alpha**	Energy production and conversion
315	-3.09	0.00247	TRUE	ENOG4105CMV	**succinyl-CoA synthetase subunit beta**	Energy production and conversion
548	-2.92	0.0000749	TRUE	ENOG4107Z6K	**F(1)F(0) ATP synthase produces ATP from ADP in the presence of a proton or sodium gradient.**	Energy production and conversion
549	-3.73	0.00111	TRUE	ENOG4107S15	**F(1)F(0) ATP synthase produces ATP from ADP in the presence of a proton or sodium gradient.**	Energy production and conversion

550	-3.19	0.000363	TRUE	ENOG4105CDG	**produces ATP from ADP in the presence of a proton gradient across the membrane. The alpha chain is a regulatory subunit (By similarity)**	Energy production and conversion
551	-4.86	1.94E-11	TRUE	ENOG4105J80	**produces ATP from ADP in the presence of a proton gradient across the membrane. The gamma chain is believed to be important in regulating ATPase activity and the flow of protons through the CF(0) complex.**	Energy production and conversion
552	-3.39	0.000119	TRUE	ENOG4105C4J	**produces ATP from ADP in the presence of a proton gradient across the membrane. The catalytic sites are hosted primarily by the beta subunits (By similarity)**	Energy production and conversion
553	-2.2	0.0216	TRUE	ENOG4105KNM	**produces ATP from ADP in the presence of a proton gradient across the membrane (By similarity)**	Energy production and conversion
855	-1.97	0.0878	FALSE	ENOG410B90Y	**2-methylcitrate synthase catalyzes the conversion of oxaloacetate and propionyl-CoA to 2-methylcitrate and CoA in the second step of the 2-methylcitric acid cycle**	Energy production and conversion
868	-4.62	0.0894	FALSE	ENOG4105D6P	**C-type cytochrome. Part of the cbb3-type cytochrome c oxidase complex (By similarity)**	Energy production and conversion
870	-2.83	0.0000216	TRUE	NOG4105DUJ	**cytochrome C oxidase**	Energy production and conversion
871	-3.23	0.108	FALSE	ENOG4105EUH	**cytochrome C oxidase, cbb3-type, subunit i**	Energy production and conversion
887	-2.8	0.00297	TRUE	ENOG4108I9X	**isocitrate lyase**	Energy production and conversion
935	-2.93	0.0131	TRUE	ENOG4105C7P	**2-oxoglutarate dehydrogenase, E1**	Energy production and conversion
936	-3.66	0.000148	TRUE	ENOG4105C7S	dehydrogenase	Energy production and conversion
937	-2.55	0.0222	TRUE	ENOG4107QN2	dihydrolipoyl dehydrogenase	Energy production and conversion
986	-2.46	0.611	FALSE	ENOG4107GFK	**fumarate hydratase**	Energy production and conversion
1163	-2.48	0.226	FALSE	ENOG4107QSN	dihydrolipoamide acetyltransferase	Energy production and conversion
1164	-2.93	0.0237	TRUE	ENOG4105DAQ	component of the pyruvate dehydrogenase (PDH) complex, that catalyzes the overall conversion of pyruvate to acetyl-CoA and CO_2 (By similarity)	energy production and conversion
1198	-2.82	0.0000774	TRUE	ENOG4105C00	**succinate dehydrogenase (flavoprotein subunit)**	Energy production and conversion
1199	-2.29	0.398	FALSE	ENOG4105E33	**succinate dehydrogenase**	Energy production and conversion
1201	-2.92	0.000128	TRUE	ENOG4105BZN	**citrate synthase**	Energy production and conversion
1387	-2.42	0.0635	FALSE	ENOG4105EIX	cytochrome c1	Energy production and

						conversion
1388	-3.77	0.0973	FALSE	ENOG4105DJ5	component of the ubiquinol-cytochrome c reductase complex (complex III or cytochrome b-c1 complex), which is a respiratory chain that generates an electrochemical potential coupled to ATP synthesis (By similarity)	Energy production and conversion
1389	-2.86	0.0695	FALSE	ENOG4108R5F	component of the ubiquinol-cytochrome c reductase complex (complex III or cytochrome b-c1 complex), which is a respiratory chain that generates an electrochemical potential coupled to ATP synthesis (By similarity)	Energy production and conversion
2380	-2.63	0.000562	TRUE	ENOG4106RE7	cytochrome	Energy production and conversion
2905	-2.09	0.0996	FALSE	ENOG4105CQV	NADH-quinone oxidoreductase subunit D (NuoD) is part of the connecting domain of complex I of the respiratory chain that couples the transfer of electrons from NADH to quinone with the translocation of protons across the membrane	Energy production and conversion
2908	-2.66	0.0952	FALSE	ENOG4107R0R	NADH dehydrogenase subunit g	Energy production and conversion
2910	-1.88	0.05	FALSE	ENOG4105P3U	NDH-1 shuttles electrons from NADH, via FMN and iron- sulfur (Fe-S) centers, to quinones in the respiratory chain. The immediate electron acceptor for the enzyme in this species is believed to be ubiquinone. Couples the redox reaction to proton translocation (for every two electrons transferred, four hydrogen ions are translocated across the cytoplasmic membrane), and thus conserves the redox energy in a proton gradient (By similarity)	Energy production and conversion
3080	-1.68	0.192	FALSE	ENOG410B7N8	CoA-binding domain protein	Energy production and conversion
3150	-2.85	0.643	FALSE	ENOG410BHYC	aldehyde dehydrogenase	Energy production and conversion
3160	-2.94	0.00101	TRUE	ENOG4107QIJ	**bifunctional aconitate hydratase 2/2-methylisocitrate dehydratase**	Energy production and conversion
3163	-3.96	0.00599	TRUE	ENOG4105D9Z	**catalyzes the reversible oxidation of malate to oxaloacetate (By similarity)**	Energy production and conversion
3406	-1.23	0.367	FALSE	ENOG4105F64	nitrite reductase	Energy production and conversion
3519	-2.68	0.0104	TRUE	ENOG4107QTB	sulfide dehydrogenase	Energy production and conversion
3724	-3.54	0.0128	TRUE	ENOG4105F0N	pyrophosphate phospho-hydrolase	Energy production and conversion
3735	-2.39	0.0017	TRUE	ENOG4105C6K	**malic enzyme**	Energy production and conversion
4111	-3.07	0.0259	TRUE	ENOG4105BZJ	electron transfer flavoprotein	Energy production and conversion
4112	-4.09	0.0122	TRUE	ENOG4105C10	electron transfer flavoprotein	Energy production and

						conversion
4170	-2.37	0.357	FALSE	ENOG4105EQJ	nitrous-oxide reductase is part of a bacterial respiratory system which is activated under anaerobic conditions in the presence of nitrate or nitrous oxide (By similarity) NosZ	Energy production and conversion
989	-1.88	0.186	FALSE	ENOG4105CU6	carbamoyl-phosphate synthetase ammonia chain	Nucleotide transport and metabolism
2877	-5.28	1.4E-10	TRUE	ENOG4105VNB	carrier of the growing fatty acid chain in fatty acid biosynthesis (By similarity)	fatty acid biosynthetic process
1107	-2.47	0.18	FALSE	ENOG4105SHT	UspA domain-containing protein	Function unknown
1479	-0.934	0.838	FALSE	ENOG4107N7Z	anti-sigma-28 factor, FlgM	Function unknown
4320	-2.78	0.424	FALSE	ENOG4105WCI	Flp Fap pilin-like protein	Function unknown
4973	-2.73	0.109	FALSE	ENOG4108BG8	domain of unknown function (DUF349)	Function unknown
4975	-4.37	0.0000225	TRUE	ENOG4105C65	catalyzes the reversible interconversion of serine and glycine with tetrahydrofolate (THF) serving as the one-carbon carrier.	hydroxymethyl-, formyl- and related transferase activity
2351	-2.09	0.091	FALSE	ENOG4108RMU	ferritin dps family protein	Inorganic ion transport and metabolism
565	-3.01	0.0559	FALSE	ENOG4105CH6	catalase	Inorganic ion transport and metabolism
2154	-2.3	0.109	FALSE	ENOG4105D05	receptor	Inorganic ion transport and metabolism
2156	-1.24	0.32	FALSE	ENOG4106BSB	receptor	Inorganic ion transport and metabolism
2171	-3.9	0.000649	TRUE	ENOG4107RP9	extracellular solute-binding protein, family 3	Inorganic ion transport and metabolism
2423	-3.19	0.0171	TRUE	ENOG4105CA0	peroxidase	Inorganic ion transport and metabolism
3847	-2.2	0.0108	TRUE	ENOG4108ESX	receptor	Inorganic ion transport and metabolism
3935	-1.38	0.472	FALSE	ENOG4105DSQ	nitric oxide reductase NorB	Inorganic ion transport and metabolism
4959	-3.47	0.0035	TRUE	ENOG4105D01	(LipO)protein	Inorganic ion transport and metabolism
1395	-3.26	0.00222	TRUE	ENOG4105WKS	part of the twin-arginine translocation (Tat) system that transports large folded proteins containing a characteristic twin-arginine motif in their signal peptide across membranes. TatA could form the protein-conducting channel of the Tat system (By similarity)	Intracellular trafficking, secretion, and vesicular transport
3571	-0.967	0.621	FALSE	ENOG4108EYE	type IV pilus secretin PilQ	Intracellular trafficking,

						secretion, and vesicular transport
4015	-3.91	0.000000299	TRUE	ENOG4105IX2	one of the proteins required for the normal export of preproteins out of the cell cytoplasm. It is a molecular chaperone that binds to a subset of precursor proteins, maintaining them in a translocation-competent state. It also specifically binds to its receptor SecA (By similarity)	Intracellular trafficking, secretion, and vesicular transport
4042	-2.89	0.00111	TRUE	ENOG4105CI6	part of the Sec protein translocase complex. Interacts with the SecYEG preprotein conducting channel.	Intracellular trafficking, secretion, and vesicular transport
4166	-4.25	0.00434	TRUE	ENOG4105E9K	**isocitrate dehydrogenase**	Isocitrate dehydrogenase activity
208	-2	0.0153	FALSE	ENOG4105CSJ	enoyl- acyl-carrier-protein reductase NADH	Lipid transport and metabolism
1527	-0.251	0.882	FALSE	ENOG4105CMX	polyhydroxyalkanoate depolymerase, intracellular	Lipid transport and metabolism
2851	-1.32	0.622	FALSE	ENOG4105KM4	acetyl-CoA carboxylase, biotin carboxyl carrier protein	Lipid transport and metabolism
3960	-1.16	0.711	FALSE	ENOG4105D1N	poly(r)-hydroxyalkanoic acid synthase, class	Lipid transport and metabolism
209	-3.01	0.0195	TRUE	ENOG4105DYT	**3-hydroxybutyryl-CoA dehydrogenase**	Lipid transport and metabolism
2878	-2.76	0.00456	TRUE	ENOG4106JX5	beta-ketoacyl synthase, C-terminal domain	Lipid transport and metabolism
3005	-2.24	0.0156	TRUE	ENOG4106DKE	acyl-Coa dehydrogenase	Lipid transport and metabolism
3007	-2.89	0.00259	TRUE	ENOG4107R0H	**3-hydroxyacyl-CoA dehydrogenase**	Lipid transport and metabolism
3008	-3.72	0.00188	TRUE	ENOG4105CHU	**acetyl-CoA acetyltransferase**	Lipid transport and metabolism
3154	-3.1	0.00544	TRUE	ENOG4108IQF	**catalyzes the conversion of acetate into acetyl-CoA (AcCoA), an essential intermediate at the junction of anabolic and catabolic pathways. acetate--CoA ligase**	Lipid transport and metabolism
3415	-2.76	0.0115	TRUE	ENOG4108RKI	**bifunctional beta-hydroxydecanoyl-[acyl-carrier-protein] dehydratase FabA catalyzes the dehydration of beta-hydroxyacyl-ACP to trans-2-acyl-ACP and the isomerization of trans-2-acyl-ACP to cis-3-acyl-ACP, possibly in the same active site**	Lipid transport and metabolism
3562	-3.35	0.00335	TRUE	ENOG4105CHU	**acetyl-coa acetyltransferase**	Lipid transport and metabolism
3766	-2.27	0.0669	FALSE	ENOG4105D1N	poly(r)-hydroxyalkanoic acid synthase, class	Lipid transport and

						metabolism
3891	-3.06	0.00313	TRUE	ENOG4105E39	**enoyl-CoA hydratase**	Lipid transport and metabolism
3903	-2.68	0.295	FALSE	ENOG410B734	**acetyl-CoA acetyltransferase**	Lipid transport and metabolism
3958	-3.28	0.00178	TRUE	ENOG4105C2D	reductase	Lipid transport and metabolism
3959	-3.6	0.0461	TRUE	ENOG4105CHU	**acetyl-coa acetyltransferase yfcY**	Lipid transport and metabolism
3967	-1.54	0.225	FALSE	ENOG4107QP7	CoA-transferase subunit A	Lipid transport and metabolism
4113	-1.14	0.534	FALSE	ENOG410B8M3	**acyl-CoA dehydrogenase**	Lipid transport and metabolism
4937	-2.52	0.398	FALSE	ENOG410B7RZ	methylmalonate-semialdehyde dehydrogenase	Lipid transport and metabolism
4983	-3.78	0.0271	TRUE	ENOG4108IQF	catalyzes the conversion of acetate into acetyl-CoA (AcCoA), an essential intermediate at the junction of anabolic and catabolic pathways.	Lipid transport and metabolism
2424	-2.86	0.000801	TRUE	ENOG4105CEY	**acyl-CoA synthetase**	Lipid transport and metabolism, Secondary metabolites biosynthesis, transport, and catabolism
217	-1.6	0.087	FALSE	ENOG4105C8V	SAICAR synthetase	Nucleotide transport and metabolism
2402	-2.63	0.138	FALSE	ENOG4108UGX	major role in the synthesis of nucleoside triphosphates other than ATP. The ATP gamma phosphate is transferred to the NDP beta phosphate via a ping-pong mechanism, using a phosphorylated active-site intermediate (By similarity)	Nucleotide transport and metabolism
1004	-3	0.0471	TRUE	ENOG4105C41	catalyzes the reversible phosphorylation of UMP to UDP (By similarity)	Nucleotide transport and metabolism
1224	-3.3	0.00314	TRUE	ENOG4105CP4	catalyzes the conversion of inosine 5'-phosphate (IMP) to xanthosine 5'-phosphate (XMP), the first committed and rate- limiting step in the de novo synthesis of guanine nucleotides, and therefore plays an important role in the regulation of cell growth (By similarity)	Nucleotide transport and metabolism
1899	-2.93	0.0000761	TRUE	ENOG4105C5T	phosphoribosyl pyrophosphate synthase	Nucleotide transport and metabolism
2295	-2.47	0.0379	TRUE	ENOG4107QTF	adenylosuccinate lyase	Nucleotide transport and metabolism
3338	-3.54	0.0517	FALSE	ENOG4105CC8	catalyzes the reversible transfer of the terminal phosphate group between ATP and AMP. Plays an important role in cellular energy homeostasis and in adenine nucleotide metabolism (By similarity)	Nucleotide transport and metabolism

3901	-2.81	0.00238	TRUE	ENOG4105DC1	bifunctional purine biosynthesis protein purh	Nucleotide transport and metabolism
4061	-0.596	0.809	FALSE	ENOG410B94P	dihydroorotase EC 3.5.2.3	Nucleotide transport and metabolism
998	-3.31	0.00245	TRUE	ENOG4105E1Z	cell wall/membrane/envelope biogenesis	Part of the outer membrane protein assembly complex, which is involved in assembly and insertion of beta-barrel proteins into the outer membrane
579	-3.41	0.00331	TRUE	ENOG4107XR1	PPIases accelerate the folding of proteins. It catalyzes the cis-trans isomerization of proline imidic peptide bonds in oligopeptides (By similarity)	Post-translational modification, protein turnover, and chaperones
1099	-3.79	0.00745	TRUE	ENOG4108JUU	redoxin	Post-translational modification, protein turnover, and chaperones
1493	-3.11	0.00182	TRUE	ENOG4105C2Z	ATP-dependent CLP protease ATP-binding subunit	Post-translational modification, protein turnover, and chaperones
2392	-4.03	0.0000288	TRUE	ENOG4105K5Y	binds to Cpn60 in the presence of Mg-ATP and suppresses the ATPase activity of the latter (By similarity)	Post-translational modification, protein turnover, and chaperones
2393	-2.81	0.00165	TRUE	ENOG4105CJ9	prevents misfolding and promotes the refolding and proper assembly of unfolded polypeptides generated under stress conditions (By similarity)	Post-translational modification, protein turnover, and chaperones
2414	-1.9	0.22	FALSE	ENOG4105EPW	hflc protein	Post-translational modification, protein turnover, and chaperones
2438	-0.624	0.795	FALSE	ENOG4105C3K	FeS assembly ATPase SufC	Post-translational modification, protein turnover, and chaperones
3848	-3.3	0.096	FALSE	ENOG4105CJX	molecular chaperone. has ATPase activity (By similarity)	Post-translational modification, protein turnover, and chaperones
4068	-3.62	0.000177	TRUE	ENOG4105CFG	acts as a chaperone (By similarity)	Post-translational modification, protein turnover, and chaperones
4069	-1.97	0.516	FALSE	ENOG4105K90	participates actively in the response to hyperosmotic and heat shock by preventing the aggregation of stress-denatured proteins, in association with DnaK and GrpE. It is the	Post-translational modification, protein

					nucleotide exchange factor for DnaK and may function as a thermosensor. Unfolded proteins bind initially to DnaJ	turnover, and chaperones
4157	-3.63	0.0000482	TRUE	ENOG4105C3H	acts as a processive, ATP-dependent zinc metallopeptidase for both cytoplasmic and membrane proteins.	Post-translational modification, protein turnover, and chaperones
4162	-2.77	0.0185	TRUE	ENOG4107TF3	peptidylprolyl cis-trans isomerase	Post-translational modification, protein turnover, and chaperones
2208	-5.91	2.46E-08	TRUE	ENOG4105DEA	involved in protein export. Acts as a chaperone by maintaining the newly synthesized protein in an open conformation	Protein peptidyl-prolyl isomerization
2828	-1.67	0.182	FALSE	ENOG41060M0	Crispr-associated protein, cse4 family	Replication, recombination and repair
5646	-1.58	0.0423	FALSE	ENOG4105C68	can catalyze the hydrolysis of ATP in the presence of single-stranded DNA, the ATP-dependent uptake of single-stranded DNA by duplex DNA, and the ATP-dependent hybridization of homologous single-stranded DNAs. It interacts with LexA causing its activation and leading to its autocatalytic cleavage (By similarity)	Replication, recombination and repair
3351	-4.31	0.00196	TRUE	ENOG4105D27	DNA-dependent RNA polymerase catalyzes the transcription of DNA into RNA using the four ribonucleoside triphosphates as substrates (By similarity)	RNA polymerase activity
1609	-0.777	0.783	FALSE	ENOG4107SA5	trap dicarboxylate transporter, dctm subunit	Secondary metabolites biosynthesis, transport, and catabolism
894	-1.98	0.0798	FALSE	ENOG4105CI3	gtp-binding protein typa	Signal transduction mechanisms
973	-2.86	0.0496	TRUE	ENOG4105CBS	Chea signal transduction histidine kinase	signal transduction mechanisms
2815	-1.92	0.372	FALSE	ENOG4107EIU	Universal stress protein family	Signal transduction mechanisms
3921	0.248	0.874	FALSE	ENOG4107QSJ	CheA signal transduction histidine kinase	Signal transduction mechanisms
4005	-1.94	0.181	FALSE	ENOG410B8HA	methyl-accepting chemotaxis sensory transducer	Signal transduction mechanisms
4040	-2.8	0.0124	TRUE	ENOG4105D4E	response regulator receiver modulated chew protein	Signal transduction mechanisms
3346	-3.93	0.152	FALSE	ENOG4108UIK	this protein binds directly to 23S ribosomal RNA (By similarity)	This protein binds directly to 23S ribosomal RNA (By similarity)
2762	-1.83	0.0263	FALSE	ENOG4105QE8	transcriptional regulator	Transcription
888	-2.94	0.00032	TRUE	ENOG4108YYX	globally modulates RNA abundance by binding to RNase E (Rne) and regulating its endonucleolytic activity	Transcription

898	-3.68	0.00362	TRUE	ENOG4105CHV	transcription elongation factor NusA	Transcription
3135	-2.71	0.00488	TRUE	ENOG4105C4P	facilitates transcription termination by a mechanism that involves Rho binding to the nascent RNA, activation of Rho's RNA-dependent ATPase activity, and release of the mRNA from the DNA template (By similarity)	Transcription
3345	-4.13	0.00309	TRUE	ENOG4105E5V	participates in transcription elongation, termination and antitermination (By similarity)	Transcription
3350	-3.45	0.0000237	TRUE	ENOG4107QIH	DNA-dependent RNA polymerase catalyzes the transcription of DNA into RNA using the four ribonucleoside triphosphates as substrates (By similarity)	Transcription
3370	-2.71	0.00345	TRUE	ENOG4105CTF	DNA-dependent RNA polymerase catalyzes the transcription of DNA into RNA using the four ribonucleoside triphosphates as substrates (By similarity)	Transcription
3950	-2.5	0.00186	TRUE	ENOG4105CDY	transcriptional regulatory protein	Transcription
4054	-2.95	0.00745	TRUE	ENOG4105C1I	transcriptional regulator GntR family	Transcription
4145	-0.945	0.423	FALSE	ENOG4108ZZ2	nucleotide-binding protein	Transcription
4933	-1.41	0.415	FALSE	ENOG410B7ZC	sigma factors are initiation factors that promote the attachment of RNA polymerase to specific initiation sites and are then released (By similarity)	Transcription
3800	0.66	0.782	FALSE	ENOG4105CWR	in eubacteria ppGpp (guanosine 3'-diphosphate 5-' diphosphate) is a mediator of the stringent response that coordinates a variety of cellular activities in response to changes in nutritional abundance (By similarity)	Transcription, Signal transduction mechanisms
883	-2.76	0.408	FALSE	ENOG4108YZX	binds directly to 23S ribosomal RNA and is necessary for the in vitro assembly process of the 50S ribosomal subunit. It is not involved in the protein synthesizing functions of that subunit (By similarity)	Translation, ribosomal structure and biogenesis
2418	-0.98	0.715	FALSE	ENOG4105CIM	catalyzes the attachment of alanine to tRNA(Ala) in a two-step reaction alanine is first activated by ATP to form Ala- AMP and then transferred to the acceptor end of tRNA(Ala). Also edits incorrectly charged Ser-tRNA(Ala) and Gly-tRNA(Ala) via its editing domain (By similarity)	Translation, ribosomal structure and biogenesis
3073	-3.1	0.153	FALSE	ENOG4108UJD	30S ribosomal protein S9	Translation, ribosomal structure and biogenesis
3357	-3.39	0.136	FALSE	ENOG4105KAR	one of the proteins that surrounds the polypeptide exit tunnel on the outside of the subunit (By similarity)	Translation, ribosomal structure and biogenesis
189	-3.46	0.000000219	TRUE	ENOG4105KDC	50s ribosomal protein l28	Translation, ribosomal structure and biogenesis
525	-3.8	0.00121	TRUE	ENOG4108Z10	involved in the binding of tRNA to the ribosomes (By similarity)	Translation, ribosomal structure and biogenesis
526	-3.19	0.000076	TRUE	ENOG4105EEE	one of the primary rRNA binding proteins, it binds directly near the 3'-end of the 23S rRNA, where it nucleates assembly of the 50S subunit (By similarity)	Translation, ribosomal structure and biogenesis
527	-3.09	0.00439	TRUE	ENOG4106U5A	one of the primary rRNA binding proteins, this protein initially binds near the 5'-end of the 23S rRNA.	Translation, ribosomal structure and biogenesis
529	-2.29	0.129	FALSE	ENOG4105CFD	one of the primary rRNA binding proteins. Required for association of the 30S and 50S	Translation, ribosomal

					subunits to form the 70S ribosome, for tRNA binding and peptide bond formation. It has been suggested to have peptidyltransferase activity	structure and biogenesis
531	-4.17	0.0000236	TRUE	ENOG4105KAP	the globular domain of the protein is located near the polypeptide exit tunnel on the outside of the subunit, while an extended beta-hairpin is found that lines the wall of the exit tunnel in the center of the 70S ribosome (By similarity)	Translation, ribosomal structure and biogenesis
532	-3.68	0.00000688	TRUE	ENOG4105CKE	binds the lower part of the 30S subunit head, binds mRNA in the 70S ribosome, positioning it for translation	Translation, ribosomal structure and biogenesis
533	-2.47	0.0811	FALSE	ENOG4108R70	binds 23S rRNA and is also seen to make contacts with the A and possibly P site tRNAs (By similarity)	Translation, ribosomal structure and biogenesis
534	-3.27	0.0915	FALSE	ENOG41084PQ	50s ribosomal protein l29	Translation, ribosomal structure and biogenesis
535	-2.07	0.11	FALSE	ENOG4105K87	one of the primary rRNA binding proteins, it binds specifically to the 5'-end of 16S ribosomal	Translation, ribosomal structure and biogenesis
886	-0.677	0.842	FALSE	ENOG4105C22	threonyL-tRNA synthetase	Translation, ribosomal structure and biogenesis
897	-2.3	0.00624	TRUE	ENOG4107QHU	one of the essential components for the initiation of protein synthesis. Protects formylmethionyl-tRNA from spontaneous hydrolysis and promotes its binding to the 30S ribosomal subunits. Also involved in the hydrolysis of GTP during the formation of the 70S ribosomal complex (By similarity)	Translation, ribosomal structure and biogenesis
1005	-4.14	0.000392	TRUE	ENOG4105CU7	associates with the EF-Tu.GDP complex and induces the exchange of GDP to GTP. It remains bound to the aminoacyl-tRNA.EF- Tu.GTP complex up to the GTP hydrolysis stage on the ribosome (By similarity)	Translation, ribosomal structure and biogenesis
1006	-3.36	0.0065	TRUE	ENOG4105CE9	30S ribosomal protein S2	Translation, ribosomal structure and biogenesis
1075	-4.86	4.83E-10	TRUE	ENOG4105VCC	30S ribosomal protein S21	Translation, ribosomal structure and biogenesis
1561	-2.95	0.000178	TRUE	ENOG4105CAV	thus facilitating recognition of the initiation point. It is needed to translate mRNA with a short Shine-Dalgarno (SD) purine-rich sequence (By similarity)	Translation, ribosomal structure and biogenesis
1637	-2.52	0.00949	TRUE	ENOG4105CA4	amino acids such as threonine, to avoid such errors, it has a posttransfer editing activity that hydrolyzes mischarged Thr-tRNA(Val) in a tRNA-dependent manner (By similarity)	Translation, ribosomal structure and biogenesis
1900	-2.97	0.0131	TRUE	ENOG4105KW6	this is one of the proteins that binds to the 5S RNA in the ribosome where it forms part of the central protuberance (By similarity)	Translation, ribosomal structure and biogenesis
2136	-1.52	0.0598	FALSE	ENOG4105CRK	lysyL-tRNA synthetase	Translation, ribosomal structure and biogenesis
2504	-3.19	0.00136	TRUE	ENOG4108UKE	interacts with and stabilizes bases of the 16S rRNA that are involved in tRNA selection in the A site and with the mRNA backbone.	Translation, ribosomal structure and biogenesis
2505	-3.42	0.000192	TRUE	ENOG4108UHY	one of the primary rRNA binding proteins, it binds directly to 16S rRNA where it nucleates assembly of the head domain of the 30S subunit.	Translation, ribosomal structure and biogenesis

2506	-3.2	0.0000712	TRUE	ENOG4105CEJ	catalyzes the GTP-dependent ribosomal translocation step during translation elongation.	Translation, ribosomal structure and biogenesis
2892	-2.6	0.058	FALSE	ENOG4105C07	amino acids such as valine, to avoid such errors it has two additional distinct tRNA(Ile)-dependent editing activities. One activity is designated as 'pretransfer' editing and involves the hydrolysis of activated Val-AMP. The other activity is designated 'posttransfer' editing and involves deacylation of mischarged Val-tRNA(Ile) (By similarity)	Translation, ribosomal structure and biogenesis
2899	-3.53	0.000157	TRUE	ENOG4105C62	involved in mRNA degradation, hydrolyses single-stranded polyribonucleotides processively in the 3'- to 5'-direction (By similarity)	Translation, ribosomal structure and biogenesis
3026	-3.23	0.00775	TRUE	ENOG4105C9M	aspartyl-trna synthetase	Translation, ribosomal structure and biogenesis
3057	-2.28	0.339	FALSE	ENOG4108YY1	this protein is located at the 30S-50S ribosomal subunit interface and may play a role in the structure and function of the aminoacyl-tRNA binding site (By similarity)	Translation, ribosomal structure and biogenesis
3072	-2.79	0.0405	TRUE	ENOG4108UM5	this protein is one of the early assembly proteins of the 50S ribosomal subunit, although it is not seen to bind rRNA by itself. It is important during the early stages of 50S assembly (By similarity)	Translation, ribosomal structure and biogenesis
3137	-3.54	0.00188	TRUE	ENOG4105NRA	50s ribosomal protein l31 type b	Translation, ribosomal structure and biogenesis
3347	-3.75	0.00000901	TRUE	ENOG4105C64	binds directly to 23S rRNA. The L1 stalk is quite mobile in the ribosome, and is involved in E site tRNA release	Translation, ribosomal structure and biogenesis
3348	-2.73	0.439	FALSE	ENOG4108VZM	50s ribosomal protein L10	Translation, ribosomal structure and biogenesis
3349	-2.69	0.0196	TRUE	ENOG4105KBC	seems to be the binding site for several of the factors involved in protein synthesis and appears to be essential for accurate translation (By similarity)	Translation, ribosomal structure and biogenesis
3356	-2.99	0.332	FALSE	ENOG4108UNN	binds to 23S rRNA. Forms part of two intersubunit bridges in the 70S ribosome (By similarity)	Translation, ribosomal structure and biogenesis
3358	-3.43	0.00128	TRUE	ENOG4105CW6	this is 1 of the proteins that binds and probably mediates the attachment of the 5S RNA into the large ribosomal subunit, where it forms part of the central protuberance.	Translation, ribosomal structure and biogenesis
3360	-2.17	0.274	FALSE	ENOG4108UJY	One of the primary rRNA binding proteins, it binds directly to 16S rRNA central domain where it helps coordinate assembly of the platform of the 30S subunit (By similarity)	Translation, ribosomal structure and biogenesis
3361	-3.93	2.13E-08	TRUE	ENOG4108R5J	this protein binds to the 23S rRNA, and is important in its secondary structure. It is located near the subunit interface in the base of the L7 L12 stalk, and near the tRNA binding site of the peptidyltransferase center (By similarity)	Translation, ribosomal structure and biogenesis
3362	-2.43	0.00365	TRUE	ENOG4105K4C	this is one of the proteins that binds and probably mediates the attachment of the 5S RNA into the large ribosomal subunit, where it forms part of the central protuberance (By similarity)	Translation, ribosomal structure and biogenesis
3363	-2.78	0.0158	TRUE	ENOG4108RA9	located at the back of the 30S subunit body where it stabilizes the conformation of the head with respect to the body (By similarity)	Translation, ribosomal structure and biogenesis

3365	-2.92	0.000185	TRUE	ENOG4108UZ0	binds to the 23S rRNA (By similarity)	Translation, ribosomal structure and biogenesis
3369	-3.04	0.00157	TRUE	ENOG4105G6W	one of the primary rRNA binding proteins, it binds directly to 16S rRNA where it nucleates assembly of the body of the 30S subunit (By similarity)	Translation, ribosomal structure and biogenesis
3371	-3.59	0.0246	TRUE	ENOG4108ZT0	50S ribosomal protein l17	Translation, ribosomal structure and biogenesis
3815	-2.32	0.141	FALSE	ENOG4105CGR	catalyzes the attachment of serine to tRNA(Ser). Is also able to aminoacylate tRNA(Sec) with serine, to form the misacylated tRNA L-seryl-tRNA(Sec), which will be further converted into selenocysteinyl-tRNA(Sec) (By similarity)	Translation, ribosomal structure and biogenesis
3992	-2.25	0.0187	TRUE	ENOG4105CHT	allows the formation of correctly charged Asn-tRNA(Asn) or Gln-tRNA(Gln) through the transamidation of misacylated Asp- tRNA(Asn) or Glu-tRNA(Gln) in organisms which lack either or both of asparaginyl-tRNA or glutaminyl-tRNA synthetases.	Translation, ribosomal structure and biogenesis
4070	-2.37	0.232	FALSE	ENOG4108ZDX	binds together with S18 to 16S ribosomal RNA (By similarity)	Translation, ribosomal structure and biogenesis
4073	-3.51	1.11E-08	TRUE	ENOG4105K9Q	binds to the 23S rRNA (By similarity)	Translation, ribosomal structure and biogenesis
502	-1.92	0.132	FALSE	ENOG4106Y56	unknown	Unknown
1120	-2.93	0.000438	TRUE	ENOG41061E8	histone family protein nucleoid-structuring protein h-ns	Unknown
3835	-3.53	0.00822	TRUE	ENOG41068H2	domain of unknown function (DUF1840)	Unknown
717	-2.73	0.0182	TRUE	ENOG4105XSC	ribosomal subunit Interface protein	Unknown function
730	-3.42	0.0765	FALSE	ENOG4108YYC	ycei family	Unknown function
848	-2.02	0.0474	TRUE	ENOG4105EBP	FAD-dependent pyridine nucleotide-disulfide oxidoreductase	Unknown function
865	-3.68	0.0131	TRUE	ENOG4105UFG	bacterial protein of unknown function (DUF883)	Unknown function
1937	-2.86	0.0132	TRUE	ENOG4107RKQ	methyl-accepting chemotaxis sensory transducer	Unknown function
2428	-5.52	0.0000242	TRUE	ENOG4108WFY	phasin family	Unknown function
3276	-2.29	0.0123	TRUE	ENOG4108EMM	sodium solute transporter superfamily	Unknown function
3757	-4.09	0.067	FALSE	ENOG4106020	unknown function	Unknown function
4006	-2.09	0.0105	TRUE	ENOG4107SSQ	methyl-accepting chemotaxis sensory transducer	Unknown function
5659	-2.13	0.015	TRUE	ENOG4105UFD	tol-pal system protein YbgF	Unknown function

Chapter 4 Potential role of enzymes encoded in putative isoprene operon

4.1 Introduction

Previously, isoprene was shown to serve as an electron acceptor for the homoacetogen *Acetobacterium wieringae* ISORED-2. The isoprene hydrogenation reaction is particularly interesting not only because isoprene might be an important electron sink in anaerobic environments but also because of the non-activated C=C bond that is reduced. Generally, hydrogenation of a C=C bond is a thermodynamically favourable reaction because it forms a more stable, less reactive product. Biocatalytic hydrogenation of C=C bonds conjugated to an electron-withdrawing group (EWG) (activated C=C bonds) has been described for a variety of substrates e.g. α,β-unsaturated aldehydes, ketones, carboxylic acids and derivatives (Bougioukou and Stewart, 2012; Winkler *et al.*, 2018). Enzymes that can reduce substituent activated C=C bonds are the enoate reductases [E.C.1.3.1.31] (Rohdich *et al.*, 2001; Joo *et al.*, 2017), the Old Yellow Enzymes (OYEs; EC 1.6.99.1) (Toogood *et al.*, 2011; Gao *et al.*, 2012; Garzón-Posse *et al.*, 2018; Winkler *et al.*, 2018) and flavin/deazaflavin oxidoreductases (FDOR superfamily) (Taylor *et al.*, 2010; Greening *et al.*, 2017). In the OYE family, substrate reduction (e.g. enals, enones, α,β-dicarboxylic acids, imides, nitroalkenes) is catalyzed through a nucleophilic addition of a hydride derived from the flavin cofactor ($FMNH_2$) to the C=C double bond via a Michael reaction (Stuermer *et al.*, 2007; Brenna *et al.*, 2012).

$F_{420}H_2$-dependent reductases of the FDOR superfamily promiscuously hydrogenate C=C bonds in diverse organic compounds such as monocycles (cyclohexenone, a dihydropyran, pyrones) and complex quinone, coumarin, and arylmethane compounds (Greening *et al.*, 2017). By using the unique redox cofactor F_{420} (Greening *et al.*, 2016; Ney *et al.*, 2017) with a low standard redox potential ($E^{\circ\prime}$ = -340 mV) and obligate two-electron chemistry it can reduce compounds otherwise recalcitrant to activation (Walsh, 1986; Greening *et al.*, 2016; Greening *et al.*, 2017).

However, biocatalytic hydrogenation of isoprene to methylbutene is an enzymatic challenge since non-activated C=C bonds are difficult targets for a nucleophilic attack. Mechanisms enabling these reactions usually involve hydration of an adjacent carbon atom before reduction occurs (Kishino *et al.*, 2013; Sakurama *et al.*, 2014; Hirata *et al.*, 2015). Four genes that are located adjacent to one another in a putative five gene operon

were found to be upregulated upon growth on isoprene in *A. wieringae* ISORED-2. The five genes in the putative operon encode a putative FAD-dependent oxidoreductase (5587), three nickel-inserting, hydrogenase maturation factors HypA (5588, 5591) and HypB (5589) and one 4Fe-4S ferredoxin (5590).

In Chapter 4, the phylogeny and characteristics of proteins encoded in the putative isoprene operon are discussed. These results have both environmental relevance in the context of distribution of isoprene electron sinks throughout the environment, and industrial importance given that chemoselective hydrogenation of non-activated conjugated and isolated C=C bonds presents a challenge in organic chemistry (Zagozda and Plenkiewicz, 2006; Toogood and Scrutton, 2014; Garzón-Posse *et al.*, 2018)**.**

4.2 Materials and Methods

4.2.1 Chemicals

Isoprene 99%, 3-methyl-1-butene ≥ 99.0% (GC), 2-methyl-2-butene ≥ 99.0% (GC) and 2-methyl-1- butene ≥ 99.5% (GC) all purchased from Sigma-Aldrich, Castle Hill, Australia. Helium (>99.9999% purity), nitrogen gas (>99.99% purity), and air (zero grade purity) were purchased from BOC Gas, Australia. H_2 (>99.99995% purity) was obtained from a H_2 generator (Parker domnick hunter, UK).

4.2.2 Growth conditions and media

Unless otherwise stated as described previously in Chapter 2 or Kronen *et al.* (2019).

4.2.3 Standard preparation and isoprene, methylbutene analysis

Unless otherwise stated as described previously in Chapter 2 or Kronen *et al.* (2019).

4.2.4 DNA sequence analysis

Promoter prediction was performed using the BPROM program and operon prediction was performed with FGENESB (Solovyev and Salamov, 2010) which are available through the Softberry website (www.softberry.com).

4.2.5 Cell preparation for reverse transcription PCR

The previously described isoprene reducing enrichment culture was grown on H_2/HCO_3^- /isoprene in six anaerobic culture flask (120 ml) containing 80 ml minimal media was supplied with 0.7 bar sterile filtered H_2 and 30 mM $NaHCO_3$. Neat isoprene (1.3 mM nominal aqueous concentration) was added to flasks through the septum. H_2 and isoprene were resupplied every second day. After 4 days of incubation at room temperature cells were harvested and used to for RNA extraction.

4.2.6 RNA extraction

Cell aggregates residing on the flask bottom were pipetted into a 2 ml reaction tube. Cells from 3 flasks were pooled and centrifuged at 10 000 g for 10 min and then stored at -20°C until use. The cells were disrupted in lysis buffer (400 µl) (Urakawa *et al.*, 2010) with mechanical agitation (30Hz for 10 min) in FastPrep Lysis Matrix A tubes (MP Biomedicals). RNA was extracted with sequential (phenol-chloroform-isoamyl alcohol (25:24:1) (pH 4.5), 3 M sodium acetate (pH 5.2) and chloroform treatments,

precipitated with isopropanol and GlycoBlue™ Coprecipitant (Thermo Fisher Scientific, Australia) using a general protocol, resuspended in 35 µl H_2O and stored at -20°C until further analysis. Contaminating chromosomal DNA in RNA samples was digested with RNase-free DNase (Qiagen) I and then cleaned on a Spin Column from PureLink® RNA Mini Kit (ThermoFisher Scientific, Australia). This step was repeated 3 times. RNA concentration in samples was quantified with Qubit™ RNA HS Assay Kit (Thermo Fisher Scientific, Australia). RNA samples were stored at -80°C until use.

4.2.7 Reverse transcription PCR

First strand complementary DNA was synthesized from 100 ng DNase I-treated total RNA using random hexamer primers from the Thermo Scientific™ RevertAid™ First Strand cDNA Synthesis Kit (Thermo Fisher Scientific) following manufacturer's instruction. In the negative control the M-MuLV reverse transcriptase was replaced with water. Synthesized cDNAs was used as templates in PCR with the Q5 high-fidelity DNA polymerase (New England BioLabs) using the intergenic region primers (**Table 4.1, Table 4.2**). Chromosomal DNA was used as a template in the same PCR as a positive control.

Table 4.1 Primer sequences used for RT-PCR to amplify the intergenic regions of the isoprene operon.

Primer	**Nucleotide Sequence 5'-3'**	**TM (°C)**
5587 Reverse 1	GTT CAA CGC ATT AAT GTG CGT	58.4
5588 Reverse 2	GCCGACTTCGACTTCGATT	58
5589 Reverse 3	TCT GCT TCT AAG TCG TTA TTC TCG	58
5590 Reverse 4	CCTTAAACGCTCAAACACATCT	55
5591 Reverse 5	TCA GCA TGT AAC AAG ACA CTTTC	55
5586 Forward 1	GGTTCAGGCAATCAGGCGAA	60
5587 Forward 2	ACCACACGATGTAAATTTGATGC	58
5588 Forward 3	AAC GTG TCC GAA ATG CGA A	58
5589 Forward 4	GGATT AAC GAT CTG TCT GAA TGG A	56
5590 Forward 5	TAC CTG CTG AAA CGA TGT TTG ATT	56

Table 4.2 Annealing temperatures of primer pairs used to amplify the intergenic regions of the isoprene operon and their predicted product sizes.

Reverse Primer	Forward Primer	Annealing Temperature (°C)	Product size (bp)	Intergenic region
5587 ORF Reverse Primer 1	5586 Forward 1	55	291	a
5588 ORF Reverse Primer 2	5587 Forward 2	54	226	b
5589 ORF Reverse Primer 3	5588 Forward 3	55	233	c
5590 ORF Reverse Primer 4	5589 Forward 4	54	148	d
5591 ORF Reverse Primer 5	5590 Forward 5	54	134	e

4.2.8 Phylogenetic analysis of oxidoreductase 5587

Phylogenetic trees were constructed based on all protein sequences from the EggNOG database v 4.5.1 (Huerta-Cepas *et al.*, 2016) matching the orthologous group of 5587 (ENOG4107QZ5), and its archaeal homolog (arCOG01292). Additional ENOG4107QZ5 sequences from other *Acetobacterium* genomes, and characterized enzymes from the literature, were also included in the dataset. Lastly, an additional 1000 top hit sequences to 5587 retrieved from NCBI (blastp search on 10 Dec 2018) were clustered with CD-HIT v4.6 (-s 0.8 -c 0.8) and added to the dataset (Li and Godzik, 2006). All sequences were aligned using MAFFT v7.313 (mafft-linsi) (Katoh and Standley, 2013). Alignment was manually trimmed to restrict the phylogeny to the core/conserved region of the proteins, equivalent to positions 158-779 in 5587. In addition, gap-rich columns were removed from the manually-trimmed alignment with BMGE v1.12 (-m BLOSUM30 -g 0.9 -h 1) (Criscuolo and Gribaldo, 2010). The phylogenetic protein tree was constructed with IQ-TREE v1.6.7 (Nguyen *et al.*, 2015) using a LG+I+G4 model (Le and Gascuel, 2008) and 10000 ultrafast bootstrap replicates (Minh *et al.*, 2013). Trees were visualized using iTol interactive tree of life https://itol.embl.de/tree/4918010318257311557452600 (https://itol.embl.de/).

4.3 Results

4.3.1 Putative isoprene reducing operon

The genes of the isoprene-induced proteins (5587-5591) are organised in one gene cluster. The five genes are in the same orientation, intergenic regions range between 7-71 nucleotides and operon prediction analysis (FGENESB) results suggested the five genes are organized in an operon [**Figure 4.1 A**]. Moreover, bands of expected sizes were obtained in reverse-transcription PCR for sets of primers flanking the individual intergenic regions of adjacent genes in the putative isoprene reducing operon (no bands were present in negative controls) [**Figure 4.1 B**].

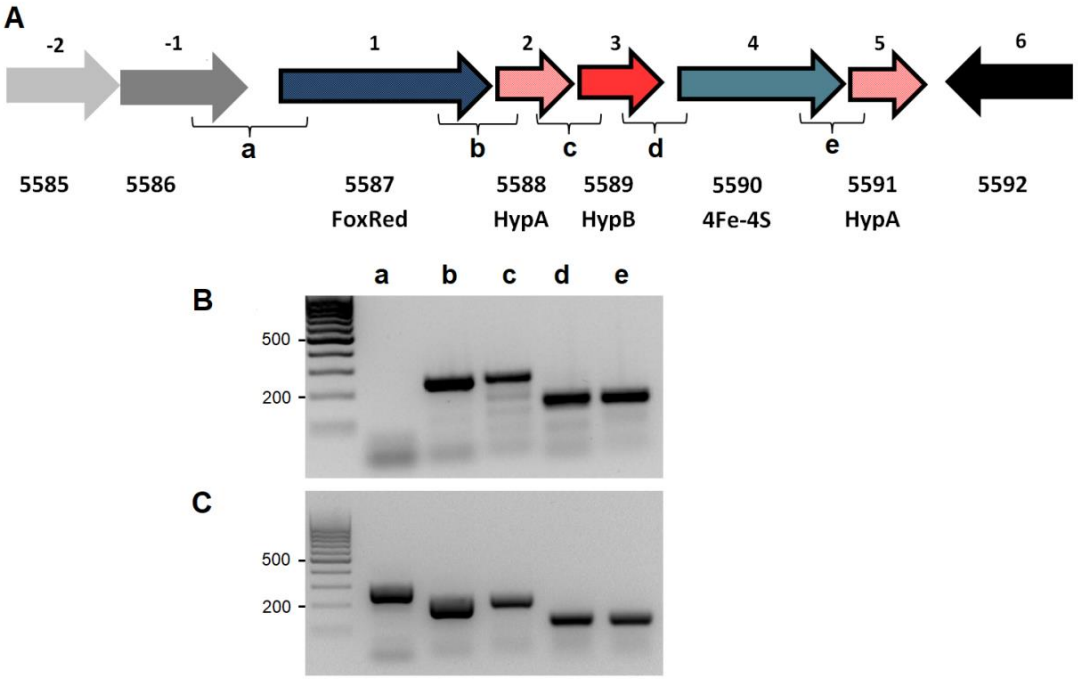

Figure 4.1 Intergenic regions of adjacent genes in the putative isoprene reducing operon were checked by Rt-PCR. (**A**) This five gene operon harbours genes coding for an oxidoreductase (5587), three nickel-inserting, hydrogenase maturation factors HypA (5588, 5591), HypB (5589) and one 4Fe-4S ferredoxin (5590), diagrammatic representation of the operon with open reading frames (ORF) 1 – 5 and their intersections indicated with a– e; (**B**) amplicons with primers connecting intersections of the neighbouring ORFs on cDNA; and (**C**) amplicons with primers connecting intersections of the neighbouring ORFs on chromosomal DNA positive control. No bands appeared in negative controls which employed cDNA amplified without reverse transcriptase in reactions.

Promoter-prediction analysis (BPROM) revealed a potential transcription start site around 44 bp upstream of the open reading frame (ORF) 1 (5587) start codon (sequence of -10 box 'TGTTATAAT' and sequence of -35 box 'ATGTCA') [Supplement **Figure S4 1**]. Two potential transcription-factor (RNA polymerase sigma factor rpoD17 and

ihf) binding sites were predicted at 58 bp and 38 bp upstream of the ORF 1 start codon. Although a strong candidate, this transcription start site has not been verified by 5'RACE or equivalent. Additionally, 4 transcription factor (TF) binding sites could be identified 13 bp upstream of the 5587 start codon (using FIMO with the CollectTF database, output filtered by p-value (<=0.0001), q-values (<=0.05), keeping only matches in the forward sequence). They are highly suggestive binding sites for a Fur (ferric uptake regulator) or NikR (nickel uptake regulator) type of TF, belonging to COG0735 and COG0864 respectively.

4.3.2 Taxanomic distribution and phylogeny of FAD-dependent oxidoreductase 5587 homologues

The putative FAD-dependent oxidoreductase 5587 is the least broadly distributed protein of the five proteins encoded by the putative isoprene operon. The other proteins, 4Fe-4S ferredoxin and HypAAB, are widely distributed and the later are usually found associated with the rest of the hyp genes *hypCDEF*. We therefore searched for 5587 homologs within the EggNOG and NCBI databases to determine the distribution of the putative isoprene operon in other organisms. According to the EggNOG database analysis 5587, which is 900 amino acids (AA) long, is predicted to belong to the bacterial orthologous group ENOG4107QZ5 (141 proteins, 99 species) and archaeal orthologous group arCOG01292 (88 proteins, 55 species). The ENOG4107QZ5/arCOG01292 orthologous group primarily contain proteins of ~500-600 AA long whereas the NCBI BLAST search primarily gave results of full-length homologs ~900 AA [**Figure 4.2 A**]. All sequences, even though they differ in length, share a homologous domain of 500 AA that aligns to position 300-800 AA in 5587.

The phylogenetic analysis of 5587 and its homologs revealed nine distinct clades (**Figure 4.2 B**). Analysis of the genomic context was performed so that some sequences could be grouped into main lineages located within the clades [**Figure 4.2 AB**]: (1) Glutamate synthase β-subunit (GltD) homologs, (2) Archaea-specific homologs, (3) NADH-dependent reduced ferredoxin:$NADP^+$-oxidoreductase large subunit (NfnB) homologs, (4) Dissimilatory sulfite reductase subunit L (DsrL) homologs, (5) NADPH-dependent oxidoreductase subunit C (NsoC) homologs and 2-Oxoglutarate:$NADP^+$-oxidoreductase subunit α (KGOR α), (7) large homologs in Proteobacteria and (9) 5587-like proteins [Supplement **Table S4 1**].

The 5587-like proteins are defined as the most closely related to 5587, which aligned to the entire 900 AA sequence of 5587 with only a few exceptions [**Figure 4.2 AB,** Supplement **Table S4 1**]. Within the 5587-like proteins that show highest amino acid sequence identity (53-56%) to 5587 belong to *Anaerocolumna jejuensis* DSM 15929 (Lachnospiraceae, Firmicutes), Firmicutes bacterium HGW-Firmicutes-20 (Firmicutes), Erysipelotrichaceae bacterium NK3D112 (Erysipelotrichaceae, Firmicutes) and *Coprobacillus* sp. (Erysipelotrichaceae, Firmicutes) [**Figure 4.2 A I,** Supplement **Table S4 1**]. Two of the five closest matches are only partial proteins and are missing the first 300 AA.

Mainly distributed among Firmicutes strains, 5587-like proteins are clustered together in evolutionary group 9 (e.g. *Clostridiaceae*, *Lachnospiraceae*, *Ruminococcaceae*, *Erysipelotrichaceae*, *Eubacteriaceae*, *Defluviitaleaceae*, *Peptostreptococcaceae* and *Veillonellaceae*)**.** A few can also be found in Spirochaetes (*Spirochaetaceae*), Tenericutes (e.g. *Haloplasmataceae*, *Anaeroplasmataceae*), Actinobacteria (e.g. *Coriobacteriaceae*, *Atopobiaceae*) as well as Chloroflexi (*Anaerolineaceae*) and Bacteroidetes (*Bacteroidaceae*). Furthermore, some 5587-like proteins are also encoded by species belonging to Proteobacteria e.g. *Vibrionaceae*, *Shewanellaceae* and *Moritellaceae*. Most species containing 5587-like proteins are strict anaerobes, although genomes of certain facultative microbes such as *Shewanella sedimini* and *Vibrio maritimus* encode homologs.

Other *Acetobacterium* strains also contain 5587-like proteins such as *A. dehalogenans* DSM 11527, *A.* sp. UBA6819 and *A.* sp. KB-1 that share 45% sequence identity with 5587 however, the homologs are relatively distant from 5587 in the tree despite belonging to the same evolutionary group [**Figure 4.2 A II**]. The 5587-like proteins of the other *Acetobacterium* spp. cluster together in the phylogenetic tree but are separate to 5587 from *A. wieringae* ISORED-2 [**Figure 4.2 A I & II**]. *A. wieringae* ISORED-2 has another member of the ENOG4107QZ5 orthologous group in its own genome, 4696 which is clustered with the glutamate synthase β-subunit (GltD) homologs in clade 1 [**Figure 4.2 A III**].

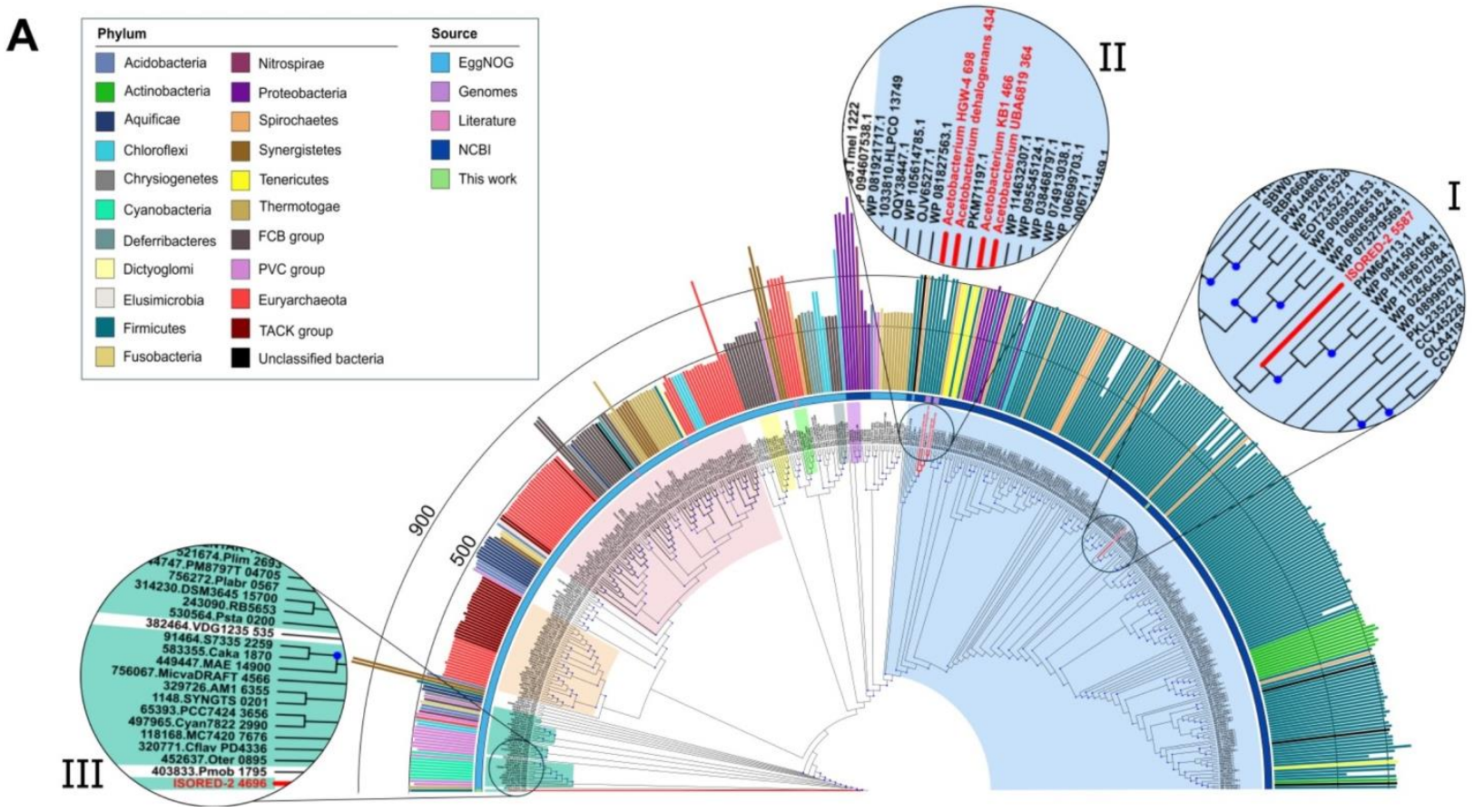

Figure 4.2A Unrooted maximum likelihood phylogenetic tree of protein 5587 and its homologs. (**A**) Phylogeny of amino acid sequences of 5587 and its homologs drawn as an unrooted circular tree. Highlighted are the positions of protein 5587 (I), 5587-like proteins in other *Acetobacterium* strains (II) and another 5587 homolog from *A.wieringae* ISORED-2 4696 (III). The outer circle shows bars representing the protein length coloured according to phylum. Scales are drawn at 900 and 500 amino acids. The following circle indicates the sources of the sequences and the inner circle the sequences position within the clades in the unrooted tree (**B**). See Supplement **Table S4 1** for sequences identity information. See also interactive iTol link https://itol.embl.de/tree/491801031825731155745260O

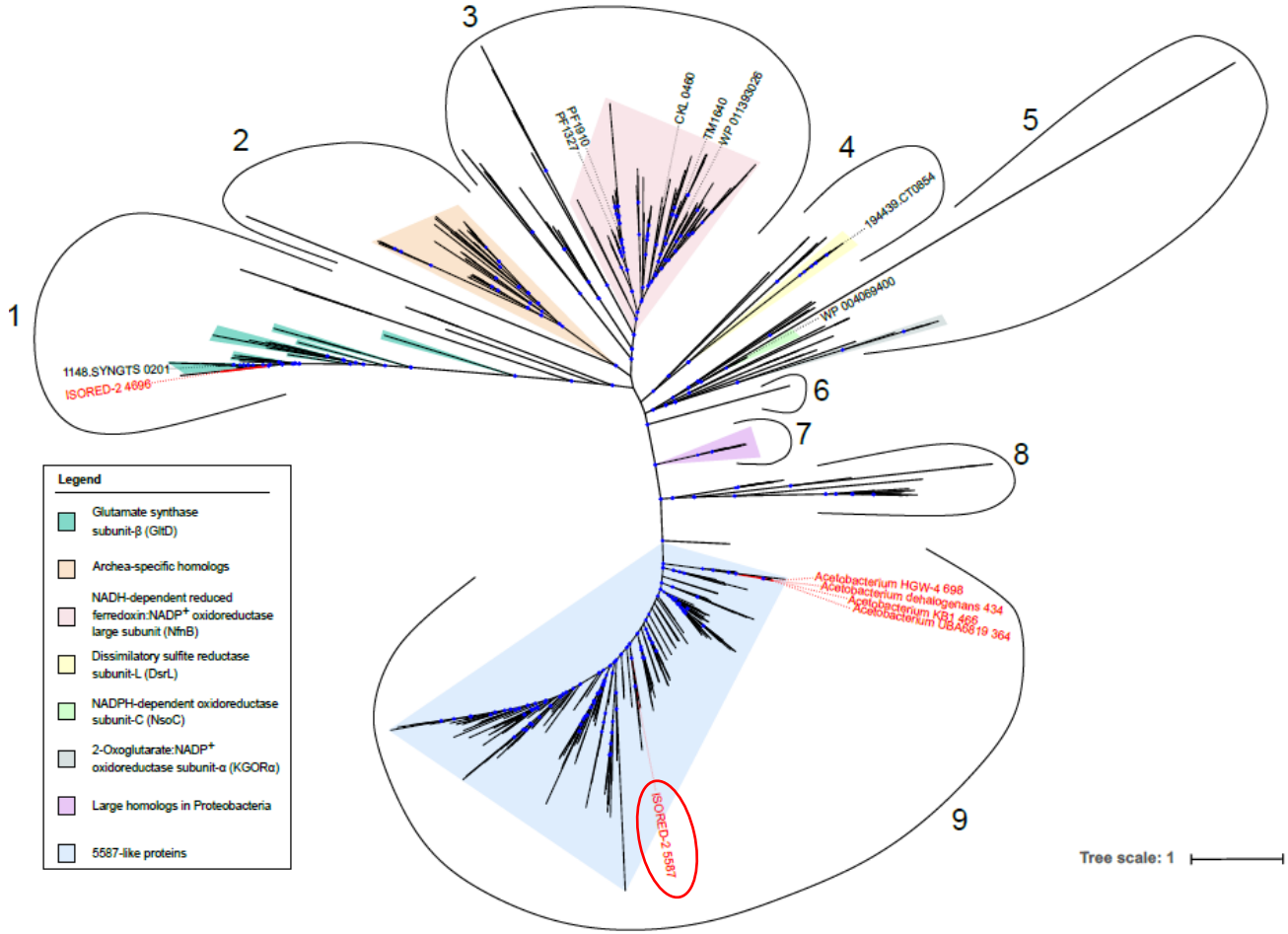

Figure 4.2B Unrooted maximum likelihood phylogenetic tree of protein 5587 and its homologs. The phylogenetic analysis of 5587 and its homologs revealed nine distinct clades containing redox enzymes involved in different reactions. Bootstrap values as a percentage of 1000 replicates are shown as blue circles only values above 90% were considered. Proteins labelled in black indicate characterized proteins in this tree. See Supplement **Table S4 1** for sequences identity information.

Interestingly, the families *Desulfobacteraceae* and *Syntrophobacteraceae* contain a longer version (~1300-1400 AA) of 5587 [evolutionary group 7, **Figure 4.2AB**].

The mostly short ~500 AA sequence homolog EggNOG ENOG4107QZ5/arCOG01292 members are found clustered together on the left part of the tree [**Figure 4.2A**] and are present in many different phyla such as *Chlorobi*, *Bacteroidetes*, *Synergistetes*, *Fusobacteria*, *Thermotogae*, *Deferribacteres*, *Chloroflexi*, *Chlamydiae*, *Acidobacteria*, *Spirochaetes*, *Gemmatimonadetes*. The archaeal arCOG01292 members appear to be restricted to the methanogenic order Methanomicrobia and the thermophilic Thermococci, Thermoplasmata and Thermoprotei. But it is important to note here that the restriction is also due to what sequences are used in the EggNOG database.

4.3.3 Protein domains of putative isoprene reductase

The best characterized and also crystallized proteins in the orthologous group of protein 5587 are the β-subunits of NADH-dependent ferredoxin-$NADP^+$-oxidoreductases (Nfn). The Nfn β-subunits (NfnB) are clustered together in evolutionary group 3 and the characterized NfnB homologs are indicated [**Figure 4.2 AB,** evolutionary group 3, characterized proteins are indicated with ID number; all branches in bold are also NfnB homologs **Figure 4.2A** (Poudel *et al.*, 2018; Liang *et al.*, 2019). See also https://itol.embl.de/tree/491801031825731155745260 0 for detail]. When protein domains between protein 5587 and characterized NfnB homologs from *T. maritima* (TM_1640), *C. kluyveri* (CKL_0460) and *P. furiosus* (PF1327 SUDHA_PYRFU) are compared, they show a very similar domain organisation [**Figure 4.3,** Supplement **Table S4 2]**.

Protein 5587 should therefore also contain an FAD, and NAD(P)H binding site as well as two [4Fe-4S]-clusters (alpha helix ferredoxin) [**Figure 4.3 A**]. In addition, 5587 has an N-terminal extension of 300 amino acids which cannot be assigned to any known protein function with one 4Fe-4S ferredoxin-type iron-sulfur binding domain located at position 289-300 AA. The C-terminal contains a 120 AA extension with an additional 4Fe-4S ferredoxin-type iron-sulfur binding site [**Figure 4.3 A**]. These could be potential binding sites for ferredoxin encoded in the isoprene operon (5590). Interestingly, the ferredoxin 5590 in the isoprene operon aligns to the first 300 amino acids of 5587. This could be an indication of gene fusion/fission which occurred during evolution between the "GltD" subunit and a shorter Fe-S containing protein such as the ferredoxin 5590 (ENOG4105DQ9). Gene fusion is common within the GltD subunits as it was observed

before in the case of Nfn, where NfnA and NfnB are encoded by two open reading frames, but in *C. autoethanogenum*, their homolog is encoded by a single gene (Wang *et al.*, 2010; Liang *et al.*, 2019).

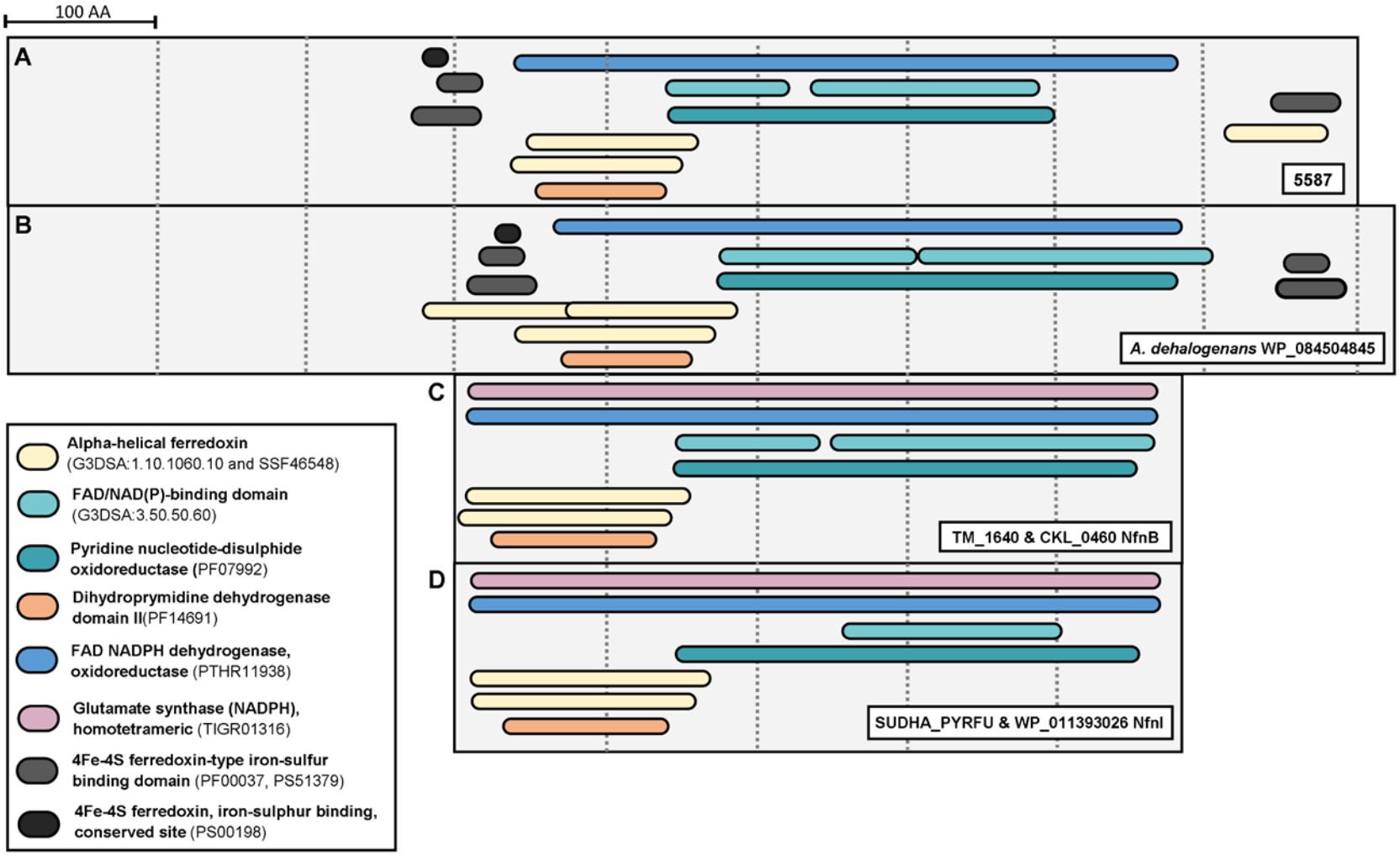

Figure 4.3 Protein domains of 5587, the 5587-like protein of *Acetobacterium dehalogenans* and the characterized large subunits of Nfn. Results of the InterProScan analysis are shown for each protein [detailed information in Supplement **Table S4 2**].

4.3.4 Genome environment of 5587-like proteins

Proteins encoded in the gene neighbourhood of 5587-like proteins were analysed. Around 60% of the gene sequences encoding 5587-like proteins are also associated with *hypA* and *hypB*, 21% are located next to only a *hypA* homolog, 3% only next to a *hypB* homolog and 16% without either [Supplement **Table S4 1**]. In some genomes *hypA* and *hypB* are adjacent to one another with *hypA* preceding *hypB*, in others *hypB* is located upstream of the 5587 homolog, some have a duplicate of *hypB* and a few have a gene located between *hypA* and *hypB* [**Figure 4.4**, **Table S4 1**].

None of the 5587-like proteins have a ferredoxin from the orthologous group ENOG4105DQ9 located adjacent to it apart from 5587 itself. The 5587-like proteins of the other *Acetobacterium* strains *A*. sp. KB-1 and *A*. *dehalogenans* are, in addition to *hypA* and *hypB*, flanked by another ferredoxin oxidoreductase (ENOG41061TH) and a helix turn helix domain containing protein (ENOG4107MUM) [**Figure 4.4**]. This gene cluster pattern can also be found in other organisms e.g. *Oceanispirochaeta* sp. M1, *Clostridium* sp. CL-2, *Vallitalea* sp. S15 [Supplement **Figure S5,** Supplement **Table S4 1**].

Genes encoding 5587-like proteins can also be found by themselves or in association with other genes encoding for example the GNAT family N-acetyltransferase, [FeFe]-hydrogenase, MATE family efflux transporter, benzoyl-CoA reductase/2-hydroxyglutaryl-CoA dehydratase subunit BcrC/BadD/HgdB, enoate reductase, methyltransferase, SAM-dependent methyltransferase domain-containing protein, HAMP domain-containing histidine kinase, fumarate hydratase, CO dehydrogenase/acetyl-CoA synthase bet [Supplement **Table S4 1**]. The long version of 5587 homologs [evolutionary group 7 **Figure 4.2 AB**] from *Desulfobacteraceae* and *Syntrophobacteraceae* are located in a very different gene environment, next to hydrogenase iron-sulfur subunit, heterodisulfide reductase (*hdrA* or *hdrB*) and ferredoxin from a different orthologous family (ENOG4105QAX) [**Figure 4.4,** Supplement **Table S4 1]**.

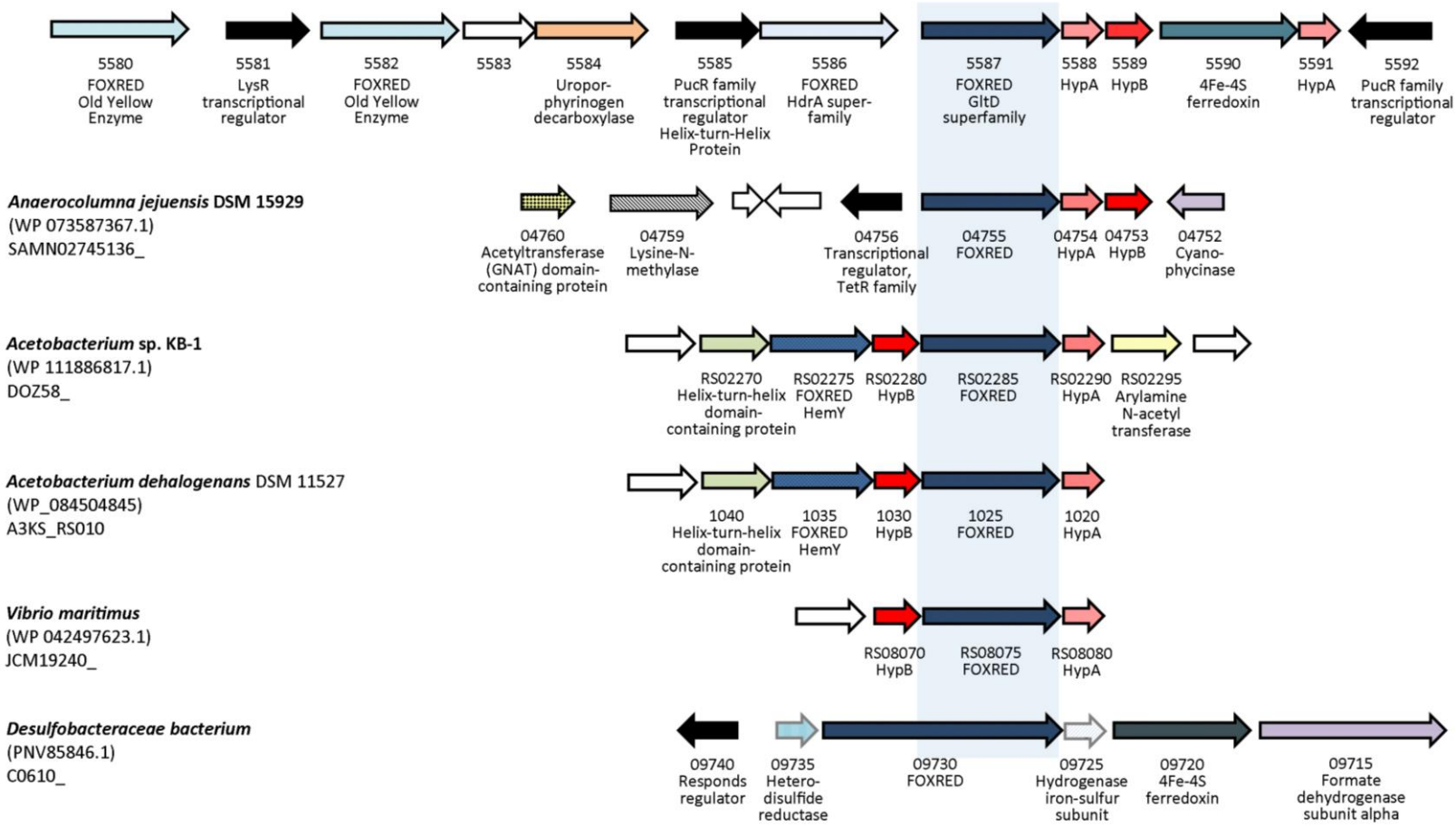

Figure 4.4 Gene environment of putative isoprene reductase 5587 in *A.wieringae* ISORED-2 and selected 5587-like proteins. The *Anaerocolumna jejuensis* DSM 15929 5587-like protein is most closely related to 5587. The gene environment of 5587-like proteins in other *Acetobacterium* strains is different to *A. wieringae* ISORED-2 though this gene arrangement can be found in multiple other organisms. *Desulfobacteraceae* and *Syntrophobacteraceae* contain a longer version (~1300-1400 AA) of 5587 which is not located next to *hypA* or *hypB*. The *Desulfobacteracaea* bacterium is shown as an example how these gene arrangments are organized. Blank arrows indicated proteins of unknown function

4.4 Discussion

4.4.1 Putative isoprene reductase is related to subunit B of glutamate synthase

Proteomics identified genes organized in an operon that encodes the putative isoprene reductase/dehydrogenase 5587 [**Figure 4.1 A**]. Like many other redox proteins 5587 is annotated as a FAD-dependent oxidoreductase or as the small subunit of the glutamate synthase (GltS) (Andersson and Roger, 2002). GltS is an iron-sulfur flavoprotein involved in the assimilation of ammonia and is widely distributed among all organisms (Curti *et al.*, 1996; Vanoni and Curti, 1999). It occurs in three forms that can be categorized based on the electron donor used for the conversion of *L*-glutamine plus 2-oxoglutarate to *L*-glutamate. The three donors are 1) NADPH (NADPH-GltS, EC 1.4.1.13), 2) NADH (NADH-GltS, EC 1.4.1.14) or 3) reduced ferredoxin (Fd-GltS, EC 1.4.7.1).

The bacterial NADPH dependent form of the enzyme is composed of a large 150 kDa α-subunit (GltB) and a small 50 kDa β-subunit (GltD) which together form the active holoenzyme (Vanoni and Curti, 1999). However, the β-subunit itself is an independent oxidoreductase with binding sites for $NAD(P)^+$, FAD and contains two N-terminal [4Fe–4S] clusters (Vanoni *et al.*, 1996; Vanoni and Curti, 1999; Petoukhov *et al.*, 2003; Vanoni and Curti, 2008). Therefore the β-subunit should be recognised as a novel family of Fe/S-containing FAD-dependent NAD(P)H oxidoreductases unless found adjacent to a putative *gltB* gene on the genome or confirmed biochemically (Vanoni and Curti, 1999; Stutz and Reid, 2004).

There are only 30 sequences within the ENOG4107QZ5 orthologous group [**Figure 4.2 AB** evolutionary group 1] that are actual glutamate synthase β-subunits (ie. located adjacent of *gltB*). This group contains GltD protein sequences homologous to the small subunit of cyanobacterial NADH-GltS (Temple *et al.*, 1998). NADH–GltS from cyanobacteria are composed of two different subunits: a 160 kDa large subunit and a small 60 kDa subunit (Okuhara *et al.*, 1999). Five of the GltD homologs in the tree belong to cyanobacterial strains including *Synechocystis* where *gltB* is located far away from *gltD* (148.SYNGTS 0201) (Navarro *et al.*, 1995) [**Figure 4.2 B**]. Protein 4696 in *A. wieringae* ISORED-2, is also a true GltD subunit that is located adjacent to *gltB* which is homologous to the cyanobacterial NADH dependent GltB [**Figure 4.2 A III**].

Whilst protein 5587 is not a true GltD subunit it does contain the homologous domain shared between all proteins in the phylogenetic tree that resembles the β-subunit of glutamate synthases aka the FAD-dependent NAD(P)H oxidoreductase functional protein domain, henceforth referred to here as "GltD". The protein sequences in the phylogenetic tree [**Figure 4.2 AB**] are either homologs of the complete "GltD" or they contain a "GltD" domain integrated into a larger protein such as 5587. Characterized examples and their homologs in the phylogenetic tree [**Figure 4.2 AB**] include the large subunit NfnB of NADH-dependent ferredoxin-$NADP^+$-oxidoreductase (Nfn) system (evolutionary group 3) (Wang *et al.*, 2010; Demmer *et al.*, 2015; Lubner *et al.*, 2017; Nguyen *et al.*, 2017), the DrsL subunit of dissimilatory sulphite reductase (Dsr) (evolutionary group 4) (Dahl *et al.*, 2005; Lübbe *et al.*, 2006; Dahl and Friedrich, 2008; Holkenbrink *et al.*, 2011; Florentino *et al.*, 2019), the NsoC subunit from the (nsoABCD) NADPH-dependent oxidoreductase (evolutionary group 5) (Tóth *et al.*, 2008) and the α-subunit of the 2-oxoglutarate:ferredoxin oxidoreductase (KGOR) (evolutionary group 5) (Ebenau-Jehle *et al.*, 2003).

All these enzymes are involved in oxidoreductive reactions clearly indicating the wide distribution of the "GltD" domain and its ability to transfer reducing equivalents to different protein domains or acceptors enabling its participation in diverse redox reactions. Vanoni and Curti proposed that, during the course of evolution, bacteria have recruited the "GltD" domain for glutamate synthase activity from a much larger GltD-like family of proteins with primarily flavin-dependent oxidoreductase functionality. Protein 5587 could be one example of these larger GltD-like family proteins as well as the large 5587 homologs found in Proteobacteria [evolutionary group 7, **Figure 4.2**]. The idea of a "construction-kit principle" has also become increasingly recognized where a small number of redox modules occur grouped in different arrangements, generating diverse proteins with new metabolic features (Baymann *et al.*, 2003, 2018; Grein *et al.*, 2013).

4.4.2 Potential characteristics of the putative isoprene reductase

The best characterized and crystallized proteins in the orthologous group of 5587 are the β-subunits of NADH-dependent ferredoxin-$NADP^+$-oxidoreductases [**Figure 4.2 AB**, evolutionary group 3]. The enzyme was first identified in *Clostridium kluyveri* (Wang *et al.*, 2010) and its homologs were later characterized in other anaerobes e.g. *Moorella thermoacetica* (Huang *et al.*, 2012), *Thermotoga maritima* (Demmer *et al.*, 2015) and

Pyrococcus furiosus (Lubner *et al.*, 2017). Nfn catalyses the following reaction (Wang *et al.*, 2010).

(17) 2 Fd_{red}^- + NADH + 2 $NADP^+$ + H^+ ⇌ 2 Fd_{ox} + NAD^+ + 2 NADPH $\Delta G°' = -20$ kJ/mol

Nfn is one of 12 known flavin-based electron-bifurcating enzymes using free energy generated by an exergonic oxidation–reduction reaction to drive an endergonic reaction (for detailed reviews see Buckel and Thauer, 2013, 2018a, 2018b; Poudel *et al.*, 2018). The electron-bifurcating enzyme complexes are grouped into four categories: (I) electron-transferring flavoproteins (EtfAB) or EtfAB homolog complexes that contain an FAD in each subunit; (II) NfnB harbouring FAD; (III) complexes with NuoF homologs (HydB, HytB, or HylB) that bind FMN; and, (IV) complexes containing heterodisulfide reductase HdrA with bound FAD (Table 1 in Buckel and Thauer, 2018b).

Nfn possesses the simplest structure among known electron-bifurcating enzymes (Buckel and Thauer, 2018b) since it comprises only two subunits, NfnA (32.6 kDa) and NfnB (49.8 kDa) and it is therefore used as a model for studying the electron bifurcation mechanism. Solved crystal structures from *T. maritima* (TM_1640) and *P. furiosus* (PF1327) revealed that NfnB contains two [4Fe-4S]-cluster as well as binding sites for NADPH and FAD with FAD being the site of electron bifurcation (Liang *et al.*, 2019). When protein domains between 5587 and NfnB from *T. maritima* (TM_1640), *C. kluyveri* (CKL_0460) and *P. furiosus* (PF1327 SUDHA_PYRFU) are compared, they show a very similar domain organisation [**Figure 4.3,** Supplement **Table S4 2]**.

Protein 5587 should therefore also contain an FAD, and NAD(P)H binding site as well as two [4Fe-4S]-clusters (alpha helix ferredoxin) [**Figure 4.3 A**].

4.4.3 Bifurcating enzymes in the *Acetobacterium* genus

Bifurcating enzymes are involved in diverse metabolic pathways e.g. CO_2 fixation, H_2 production, butyric acid formation, dissimilatory sulphate reduction, acetogenesis and methanogenesis (Buckel and Thauer, 2018b). The latter two are regarded as ancient biological processes which use flavin-based bifurcation as an ancient, more primitive energy conserving mechanism in energy-limited environments (Nitschke and Russell, 2012; Martin *et al.*, 2014). Therefore it is not surprising that two out of the 12 known flavin-based bifurcating enzymes can be found in *Acetobacterium* sp. that live on the

thermodynamic limit of life: the bifurcating [FeFe]-hydrogenase (Schuchmann and Müller, 2012), and the caffeyl- CoA reductase (Bertsch *et al.,* 2013).

The tetrameric [FeFe]-hydrogenase (HydABCD) characterized in *A. woodii*, catalyses the endergonic reduction of ferredoxin with H_2 as electron donor by coupling it to the exergonic electron transfer from H_2 to NAD^+ via electron bifurcation (Schuchmann and Müller, 2012):

(18) $2\ H_2 + Fd_{ox} + NAD^+ \rightleftharpoons Fd_{red} + NADH + 3\ H^+$

The Fd_{red} generated is crucial in two scenarios. Firstly, it serves as the reductant for the reduction of CO_2 to CO in the Wood-Ljungdahl pathway (Ragsdale and Kumar, 1996; Ragsdale and Pierce, 2008). Secondly, it fuels the Na^+-translocating ferredoxin: NAD^+-oxidoreductase (Rnf complex) and thus ATP synthesis (Müller *et al.*, 2008; Biegel *et al.*, 2011; Westphal *et al.*, 2018). The Rnf complex oxidizes Fd_{red} to reduce NAD^+ to NADH and establishes a Na^+ gradient over the cytoplasmic membrane with the negative free energy generated from this reaction. The Na^+ gradient then drives ATP synthesis via a Na^+ coupled F_1F_0-ATP synthase (Fritz and Müller, 2007; Fritz *et al.*, 2008). Derived NADH is funnelled via the WLP to carbon dioxide producing acetyl-CoA for biomass production and acetate:

(19) $4\ H_2 + 2\ CO_2 \rightarrow$ acetate $+ 2\ H_2O$ $\qquad \Delta G^{\circ\prime} = -105$ kJ/mol

In the case of caffeate reduction the generated NADH from the [FeFe]-hydrogenase and the Rnf complex is not only channelled to enzymes in the WLP but also to the bifurcating caffeyl-CoA reductase (Tschech and Pfennig, 1984; Imkamp *et al.*, 2007; Hess *et al.*, 2013). The caffeyl-CoA reductase-Etf complex of *A. woodii* couples the endergonic ferredoxin reduction with NADH as reductant to the exergonic reduction of caffeyl-CoA. The reduced ferredoxin is re-oxidized at the Rnf complex again generating a Na^+ gradient (Bertsch *et al.*, 2013). In the case of isoprene a similar scenario could be envisaged. The Rnf complex, the [FeFe]-hydrogenase and the ATP synthase where highly expresses in *A. wierinage* ISORED-2 in cells grown on H_2 /HCO_3^- and H_2/HCO_3^- /isoprene [Supplement **Table S4 3**]. In a bifurcating scenario, NADH derived from the [FeFe]-hydrogenase and Rnf complex would act as the reductant for the exergonic reduction of isoprene to methylbutene which is coupled to the endergonic reduction of ferredoxin [**Figure 4.5**].

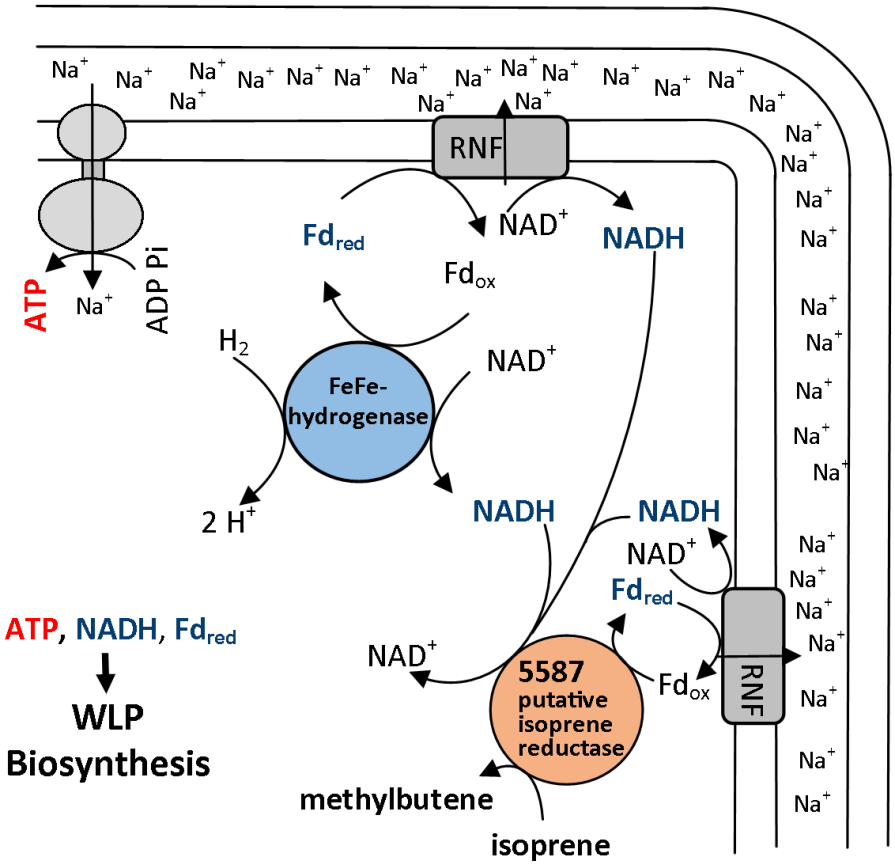

Figure 4.5 Possible electron transfer pathway from H_2 to isoprene catalysed by the bifurcating [FeFe]-hydrogenase and 5587 in *Acetobacterium wieringae* ISORED-2. All enzyme complexes were highly expressed in H_2/HCO_3^-/isoprene cultures [Supplement **Table S4 3**]. Protein 5587 could be a bifurcating enzyme that catalyses the oxidation of NAD(P)H and the simultaneous exergonic reduction of isoprene and endergonic reduction of ferredoxin. Adapted from (Bertsch *et al.*, 2013) for hydrogen-dependent caffeate respiration in *Acetobacterium woodii*.

Requirements for bifurcation are at least one FAD containing subunit, ferredoxin dependency, a low redox potential $2e^-$ donor and a high potential $2e^-$ acceptor (Demmer *et al.*, 2015). Protein 5587 has a conserved FAD binding site and most likely binds ferredoxin or might even contain a ferredoxin like domain [**Figure 4.3**]. The low redox potential $2e^-$ donor would be NAD(P)H and isoprene the high potential $2e^-$ acceptor. Calculating the redox potential of isoprene considering the calculated ΔG_f of -137 kJ/mol (Kronen *et al.*, 2019):

$$E^0 = (-137000\ J/mol) \div (-2\ mol \times 96484\ J/mol\ V) = 709\ mV$$

$$E^{0\prime} = 0.709\ mV + (0.059 \times \log(10^{-7})) = \mathbf{+\ 296\ mV}$$

However, since experimental evidence for protein 5587 is lacking one can only speculate until the enzyme has been purified and characterized. Buckel and Thauer (2018b) indicate that homology to bifurcating enzymes is not sufficient to presume

electron bifurcation in an enzyme. The question is also whether the other enzymes that contain the "GltD" subunit in the phylogenetic tree [**Figure 4.2 AB**] are electron bifurcating enzymes, e.g. the NADPH-dependent oxidoreductase (NsoABCD) in thermophilic Archaea involved in sulphur reduction (Tóth *et al.*, 2008), the dissimilatory sulphite reductase in green and purple sulphur bacteria (Dahl and Friedrich, 2008) or even the NADH-dependent glutamate synthase (Navarro *et al.*, 1995). Potential bifurcating electron transfer pathways have been proposed for the glutamate synthase (Vanoni and Curti, 2008). Still, experimental evidence is missing in all cases except for Nfn. Interestingly, the long versions of protein 5587 present in Proteobacteria [evolutionary group 7 **Figure 4.2**], are located adjacent to homologs of the heterodisulfide reductase subunit A or B (*hdrA* or *hdrB*) which belong to the electron-bifurcating enzyme group IV (see table 1 for detail in Buckel and Thauer, 2018b). This might be another indication that electron bifurcation could be a mechanism involved in hydrogenation of isoprene. If isoprene reduction was a linear process, reduction would have to be couple to ATP synthesis, maybe also through a gradient establishment, since the reduction of isoprene conserves energy (Kronen *et al.*, 2019).

4.4.4 Evidence that 5587 might be a nickel-dependent enzyme

The isoprene operon contains three *hyp* genes (**hy**drogenase **p**leiotropic) which encode proteins that are usually responsible for the acquisition and insertion of nickel during the maturation process of [NiFe]-hydrogenases (Jacobi *et al.*, 1992; Olson *et al.*, 2001; Hube *et al.*, 2002; Maier *et al.*, 2007; Peters *et al.*, 2015). Biosynthesis and maturation of the [NiFe]-hydrogenase active site is a complex multistep process that involves other accessory Hyp proteins (HypCDEF) in addition to HypA and HypB (Watanabe *et al.*, 2012). However, the genome of *A. wieringae* ISORED-2 does not harbour *hypCDEF* or any genes encoding a hydrogenase close to the *hypAB* genes. Therefore, it is more likely that the HypAB proteins incorporate nickel into the putative isoprene reductase which might therefore be a nickel-dependent enzyme. It has been shown before that HypA and HypB are involved in the nickel-dependent maturation of other nickel-dependent enzymes besides NiFe hydrogenases e.g. urease in *H. pylori* (Olson *et al.*, 2001). HypAB seem to be important for the function of the putative isoprene reductase since both or at least one of them were found to be associated with 5587 homologs in 85% of the 5587-like proteins [Supplement **Table S4 1**]. Like all HypA proteins in other organism including *H. pylori*, both HypA proteins (5588, 5591) in *A. wieringae*

ISORED-2 contain conserved binding properties to Ni^{2+} (backbone amides of residues His2, Glu3) and Zn^{2+} (four conserved cysteines (2 x CxxC)) (Xia *et al*., 2009; Kolkowska *et al*., 2015). Proteins 5587 and 5590 also contain potential CxxC metal binding motifs. In protein 5587, two potential CxxC motifs are located adjacent to each other at the C-terminus and one is located closer to the N-terminus. Protein 5590 contains two CxxC motifs close to each other at the C-terminus. If these CxxC are involved in Ni^{2+} binding needs to be experimentally investigated in the future.

Further evidence that protein 5587 might be a nickel-dependent enzyme comes from the promotor region analysis of the isoprene operon. It was demonstrated that in *A. wieringae* ISORED-2 *de novo* synthesis of proteins is required to induce the ability to reduce isoprene, though induction cannot be differentiated from de-repression in our assays. Four transcription factor (TF) binding sites could be identified 13 bp upstream of the 5587 gene start codon [Supplement **Figure S4 1**]. They are highly suggestive of binding sites for a ferric uptake regulator (Fur) or Ni(II)-dependent transcriptional regulator (NikR) type of TF (van Vliet *et al.*, 2002; Muller *et al.*, 2011). In both cases, they can control repression and activation of gene transcription by coupling metal ion binding with a change in their DNA binding affinity and/or specificity, thus translating the amount of a specific metal ion into a change in gene transcription (Ernst *et al.*, 2005; Pich and Merrell, 2013; Musiani *et al.*, 2015).

4.4.5 Isoprene reduction ability in other bacteria

The 5587-like proteins are widely distributed among anaerobic bacteria [**Figure 4.2,** evolutionary group 9] but the operon organisation found in *A.wieringae* ISORED-2 was not found in any other bacterial strains. Even the other evaluated *Acetobacterium* spp. contain 5587-like proteins that show only 45% amino acid sequence identity to 5587 and they are found in different gene cluster organisations. It remains to be experimentally determined whether 5587-like proteins are able to catalyse the reduction of isoprene.

4.5 Conclusion

Homoacetogenic bacteria like *Acetobacterium* spp. live at the thermodynamic limit of life. They therefore have evolved with different metabolic strategies to remain competitive in energy constrained environments. In this Chapter, evidence for a putative isoprene reductase located in a five-gene operon in the genome of *A. wieringae* ISORED-2 is considered. Evidence that the putative isoprene reductase may be a nickel-dependent enzyme that contains a binding site for NADH, FAD and 4Fe-4S ferredoxin is presented. It shares sequence homology with the FAD-dependent oxidoreductase family to which the electron bifurcating NADH-dependent ferredoxin-$NADP^+$-oxidoreductase subunit B also belongs. Proteins homologous to the putative isoprene reductase are widely distributed among anaerobic bacteria especially in the Firmicutes phylum. However the operon arrangement found in *A. wieringae* ISORED-2 could not be found in any other organism on the NCBI database. Other *Acetobacterium* strains contain homologous proteins but in a different gene arrangement. Biochemical testing is required to assign isoprene reducing functionality to the putative isoprene reductase, explore the isoprene reduction mechanism and to verify if electron bifurcation is involved in the reduction process. The putative isoprene reductase/hydrogenase is of particular high interest because of its reduction of a non-activated C=C bond. To the best of our knowledge, no enzyme that is able to perform a direct reduction of a non-activated C=C bond has been isolated or characterized so far, making the proteins/protein family involved in the isoprene reduction process highly interesting. Hydration of isoprene is unlikely since a hydratase and multiple enzymes involved in this kind of reduction processes were not upregulated in the presence of isoprene.

4.6 References

Andersson, J.O. and Roger, A.J. (2002) Evolutionary analyses of the small subunit of glutamate synthase: Gene order conservation, gene fusions, and prokaryote-to-eukaryote lateral gene transfers. *Eukaryot. Cell* **1**: 304–310.

Baymann, F., Lebrun, E., Brugna, M., Schoepp-Cothenet, B., Giudici-Orticoni, M.T., and Nitschke, W. (2003) The redox protein construction kit: Pre-last universal common ancestor evolution of energy-conserving enzymes. *Philos. Trans. R. Soc. B Biol. Sci.* **358**: 267–274.

Baymann, F., Schoepp-Cothenet, B., Duval, S., Guiral, M., Brugna, M., Baffert, C., et al. (2018) On the natural history of flavin-based electron bifurcation. *Front. Microbiol.* **9**: 1357.

Bertsch, J., Parthasarathy, A., Buckel, W., and Müller, V. (2013) An electron-bifurcating caffeyl-CoA reductase. *J. Biol. Chem.* **288**: 11304–11311.

Biegel, E., Schmidt, S., González, J.M., and Müller, V. (2011) Biochemistry, evolution and physiological function of the Rnf complex, a novel ion-motive electron transport complex in prokaryotes. *Cell. Mol. Life Sci.* **68**: 613–634.

Bougioukou, D.J. and Stewart, J.D. (2012) Reduction of C=C double bonds. *Enzym. Catal. Org. Synth. Third Ed.* **2**: 1111–1163.

Brenna, E., Gatti, F.G., Monti, D., Parmeggiani, F., and Sacchetti, A. (2012) Cascade coupling of ene reductases with alcohol dehydrogenases: Enantioselective reduction of prochiral unsaturated aldehydes. *ChemCatChem* **4**: 653–659.

Buckel, W. and Thauer, R.K. (2013) Energy conservation via electron bifurcating ferredoxin reduction and proton/Na^+ translocating ferredoxin oxidation. *Biochim. Biophys. Acta - Bioenerg.* **1827**: 94–113.

Buckel, W. and Thauer, R.K. (2018a) Flavin-based electron bifurcation, a new mechanism of biological energy coupling. *Chem. Rev.* **118**: 3862–3886.

Buckel, W. and Thauer, R.K. (2018b) Flavin-based electron bifurcation, ferredoxin, flavodoxin, and anaerobic respiration with protons (Ech) or NAD^+ (Rnf) as electron acceptors: A historical review. *Front. Microbiol.* **9**: 401.

Criscuolo, A. and Gribaldo, S. (2010) BMGE (Block Mapping and Gathering with Entropy): A new software for selection of phylogenetic informative regions from multiple sequence alignments. *BMC Evol. Biol.* **10**: 210.

Curti, B., Vanoni, M.A., Verzotti, E., and Zanetti, G. (1996) Glutamate synthase: A complex iron-sulfur flavoprotein. *Biochem. Soc. Trans.* **24**: 95–99.

Dahl, C., Engels, S., Pott-sperling, A.S., Schulte, A., Sander, J., Lu, Y., et al. (2005) Novel genes of the *dsr* gene cluster and evidence for close interaction Dsr proteins during sulfur oxidation in the phototrophic sulfur bacterium *Allochromatium vinosum*. *J.Bacteriol.* **187**: 1392–1404.

Dahl, C. and Friedrich, C.G. (2008) Microbial sulfur metabolism Eckey,C. (ed) Springer-Verlag Berlin Heidelberg New York.

Demmer, J.K., Huang, H., Wang, S., Demmer, U., Thauer, R.K., and Ermler, U. (2015) Insights into flavin-based electron bifurcation via the NADH-dependent reduced ferredoxin:NADP oxidoreductase structure. *J. Biol. Chem.* **290**: 21985–21995.

Ebenau-Jehle, C., Boll, M., and Fuchs, G. (2003) 2-Oxoglutarate:$NADP^+$ oxidoreductase in *Azoarcus evansii*: Properties and function in electron transfer reactions in aromatic ring reduction. *J. Bacteriol.* **185**: 6119–6129.

Ernst, F.D., Kuipers, E.J., Heijens, A., Sarwari, R., Stoof, J., Penn, C.W., et al. (2005) The nickel-responsive regulator NikR controls activation and repression of gene transcription in *Helicobacter pylori*. *Infect. Immun.* **73**: 7252–7258.

Florentino, A.P., Pereira, I.A.C., Boeren, S., van den Born, M., Stams, A.J.M., and Sánchez-Andrea, I. (2019) Insight into the sulfur metabolism of *Desulfurella amilsii* by differential proteomics. *Environ. Microbiol.* **21**: 209–225.

Fritz, M., Klyszejko, A.L., Morgner, N., Vonck, J., Brutschy, B., Muller, D.J., et al. (2008) An intermediate step in the evolution of ATPases - A hybrid F_0-V_0 rotor in a bacterial Na^+ F_1F_0 ATP synthase. *FEBS J.* **275**: 1999–2007.

Fritz, M. and Müller, V. (2007) An intermediate step in the evolution of ATPases - The F_1F_0-ATPase from *Acetobacterium woodii* contains F-type and V-type rotor subunits and is capable of ATP synthesis. *FEBS J.* **274**: 3421–3428.

Gao, X., Ren, J., Wu, Q., and Zhu, D. (2012) Biochemical characterization and substrate profiling of a new NADH-dependent enoate reductase from *Lactobacillus casei*. *Enzyme Microb. Technol.* **51**: 26–34.

Garzón-Posse, F., Becerra-Figueroa, L., Hernández-Arias, J., and Gamba-Sánchez, D. (2018) Whole cells as biocatalysts in organic transformations. *Molecules* **23**:.

Grein, F., Ramos, A.R., Venceslau, S.S., and Pereira, I.A.C. (2013) Unifying concepts in anaerobic respiration: Insights from dissimilatory sulfur metabolism. *Biochim. Biophys. Acta - Bioenerg.* **1827**: 145–160.

Greening, C., Ahmed, F.H., Mohamed, A.E., Lee, B.M., Pandey, G., Warden, A.C., et al. (2016) Physiology, biochemistry, and applications of F_{420} - and Fo-dependent redox reactions. **80**: 451–493.

Greening, C., Jirapanjawat, T., Afroze, S., Ney, B., Scott, C., Pandey, G., et al. (2017) Mycobacterial $F_{420}H_2$-dependent reductases promiscuously reduce diverse compounds through a common mechanism. *Front. Microbiol.* **8**: 1–10.

Hess, V., González, J.M., Parthasarathy, A., Buckel, W., and Müller, V. (2013) Caffeate respiration in the acetogenic bacterium *Acetobacterium woodii*: A coenzyme a loop saves energy for caffeate activation. *Appl. Environ. Microbiol.* **79**: 1942–1947.

Hirata, A., Kishino, S., Park, S.-B., Takeuchi, M., Kitamura, N., and Ogawa, J. (2015) A novel unsaturated fatty acid hydratase toward C_{16} to C_{22} fatty acids from *Lactobacillus acidophilus*. *J. Lipid Res.* **56**: 1340–1350.

Holkenbrink, C., Barbas, S.O., Mellerup, A., Otaki, H., and Frigaard, N.U. (2011) Sulfur globule oxidation in green sulfur bacteria is dependent on the dissimilatory sulfite reductase system. *Microbiology* **157**: 1229–1239.

Huang, H., Wang, S., Moll, J., and Thauer, R.K. (2012) Electron bifurcation involved in the energy metabolism of the acetogenic bacterium *Moorella thermoacetica* growing on glucose or H_2 plus CO_2. *J. Bacteriol.* **194**: 3689–3699.

Hube, M., Blokesch, M., and Böck, A. (2002) Network of hydrogenase maturation in *Escherichia coli:* Role of accessory proteins HypA and HybF. *J. Bacteriol.* **184**: 3879–3885.

Huerta-Cepas, J., Szklarczyk, D., Forslund, K., Cook, H., Heller, D., Walter, M.C., 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.

Imkamp, F., Biegel, E., Jayamani, E., Buckel, W., and Müller, V. (2007) Dissection of the caffeate respiratory chain in the acetogen *Acetobacterium woodii*: Identification of an Rnf-type NADH dehydrogenase as a potential coupling site. *J. Bacteriol.* **189**: 8145–8153.

Jacobi, A., Rossmann, R., and Böck, A. (1992) The hyp operon gene products are required for the maturation of catalytically active hydrogenase isoenzymes in *Escherichia coli*. *Arch. Microbiol.* **158**: 444–451.

Joo, J.C., Khusnutdinova, A.N., Flick, R., Kim, T., Bornscheuer, U.T., Yakunin, A.F., and Mahadevan, R. (2017) Alkene hydrogenation activity of enoate reductases for an environmentally benign biosynthesis of adipic acid. *Chem. Sci.* **8**: 1406–1413.

Katoh, K. and Standley, D.M. (2013) MAFFT multiple sequence alignment software version 7: Improvements in performance and usability. *Mol. Biol. Evol.* **30**: 772–780.

Kishino, S., Takeuchi, M., Park, S.-B., Hirata, A., Kitamura, N., Kunisawa, J., et al. (2013) Polyunsaturated fatty acid saturation by gut lactic acid bacteria affecting host lipid composition. *Proc. Natl. Acad. Sci.* **110**: 17808–17813.

Kolkowska, P., Krzywoszynska, K., Potocki, S., Chetana, P.R., Spodzieja, M., Rodziewicz-Motowidlo, S., and Kozlowski, H. (2015) Specificity of the $Zn^{2+,}$ Cd^{2+} and Ni^{2+} ion binding sites in the loop domain of the HypA protein. *Dalt. Trans.* **44**: 9887–9900.

Kronen, M., Lee, M., Jones, Z.L., and Manefield, M.J. (2019) Reductive metabolism of the important atmospheric gas isoprene by homoacetogens. *ISME J.* **13**: 1168–1182.

Le, S.Q. and Gascuel, O. (2008) An improved general amino acid replacement matrix. *Mol. Biol. Evol.* **25**: 1307–1320.

Li, W. and Godzik, A. (2006) Cd-hit: A fast program for clustering and comparing large sets of protein or nucleotide sequences. *Bioinformatics* **22**: 1658–1659.

Liang, J., Huang, H., and Wang, S. (2019) Distribution, evolution, catalytic mechanism, and physiological functions of the flavin-based electron-bifurcating NADH-dependent reduced ferredoxin: $NADP^+$ oxidoreductase. *Front. Microbiol.* **10**: 1–12.

Lübbe, Y.J., Youn, H.S., Timkovich, R., and Dahl, C. (2006) Siro(haem)amide in *Allochromatium vinosum* and relevance of DsrL and DsrN, a homolog of cobyrinic acid a,c-diamide synthase, for sulphur oxidation. *FEMS Microbiol. Lett.* **261**: 194–202.

Lubner, C.E., Jennings, D.P., Mulder, D.W., Schut, G.J., Zadvornyy, O.A., Hoben, J.P., et al. (2017) Mechanistic insights into energy conservation by flavin-based electron bifurcation. *Nat. Chem. Biol.* **13**: 655–659.

Maier, R.J., Benoit, S.L., and Seshadri, S. (2007) Nickel-binding and accessory proteins facilitating Ni-enzyme maturation in *Helicobacter pylori*. *BioMetals* **20**: 655–664.

Martin, W.F., Sousa, F.L., and Lane, N. (2014) Energy at life's origin. *Science* **344**: 1092–1093.

Minh, B.Q., Nguyen, M.A.T., and von Haeseler, A. (2013) Ultrafast approximation for phylogenetic bootstrap. *Mol. Biol. Evol.* **30**: 1188–1195.

Muller, C., Bahlawane, C., Aubert, S., Delay, C.M., Schauer, K., Michaud-Soret, I., and De Reuse, H. (2011) Hierarchical regulation of the NikR-mediated nickel response in *Helicobacter pylori*. *Nucleic Acids Res.* **39**: 7564–7575.

Müller, V., Imkamp, F., Biegel, E., Schmidt, S., and Dilling, S. (2008) Discovery of a ferredoxin:NAD^+-oxidoreductase (Rnf) in *Acetobacterium woodii*: A novel potential coupling site in acetogens. *Ann. N. Y. Acad. Sci.* **1125**: 137–146.

Musiani, F., Zambelli, B., Bazzani, M., Mazzei, L., and Ciurli, S. (2015) Nickel-responsive transcriptional regulators. *Metallomics* **7**: 1305–1318.

Navarro, F., Candau, P., and Florencio, F.J. (1995) Existence of two ferredoxin-glutamate synthases in the cyanobacterium *Synechocystis* sp. PCC 6803. Isolation and insertional inactivation of gltB and gltS genes. *Plant Mol. Biol.* **27**: 753–767.

Nguyen, D.M.N., Schut, G.J., Zadvornyy, O.A., Tokmina-Lukaszewska, M., Poudel, S., Lipscomb, G.L., et al. (2017) Two functionally distinct $NADP^+$-dependent ferredoxin oxidoreductases maintain the primary redox balance of *Pyrococcus furiosus*. *J. Biol. Chem.* **292**: 14603–14616.

Nguyen, L.T., Schmidt, H.A., Von Haeseler, A., and Minh, B.Q. (2015) IQ-TREE: A fast and effective stochastic algorithm for estimating maximum-likelihood phylogenies. *Mol. Biol. Evol.* **32**: 268–274.

Nitschke, W. and Russell, M.J. (2012) Redox bifurcations: Mechanisms and importance to life now, and at its origin. *Bioessays* **34**: 106–109.

Ney, B., Ahmed, F.H., Carere, C.R., Biswas, A., Warden, A.C., Morales, S.E., et al. (2017) The methanogenic redox cofactor F_{420} is widely synthesized by aerobic soil bacteria. *ISME J.* **11**: 125–137.

Okuhara, H., Matsumura, T., Fujita, Y., and Hase, T. (1999) Cloning and inactivation of genes encoding ferredoxin- and NADH-dependent glutamate synthases in the Cyanobacterium *Plectonema boryanum* . Imbalances in nitrogen and carbon assimilations caused by deficiency of the ferredoxin-dependent enzyme. *Plant Physiol.* **20**: 33–42.

Olson, J.W., Mehta, N.S., and Maier, R.J. (2001) Requirement of nickel metabolism proteins HypA and HypB for full activity of both hydrogenase and urease in *Helicobacter pylori*. *Mol. Microbiol.* **39**: 176–182.

Peters, J.W., Schut, G.J., Boyd, E.S., Mulder, D.W., Shepard, E.M., Broderick, J.B., et al. (2015) [FeFe]- and [NiFe]-hydrogenase diversity, mechanism, and maturation. *Biochim. Biophys. Acta - Mol. Cell Res.* **1853**: 1350–1369.

Petoukhov, M. V., Svergun, D.I., Konarev, P. V., Ravasio, S., van den Heuvel, R.H.H., Curti, B., and Vanoni, M.A. (2003) Quaternary structure of *Azospirillum brasilense* NADPH-dependent glutamate synthase in solution as revealed by synchrotron radiation x-ray scattering. *J. Biol. Chem.* **278**: 29933–29939.

Pich, O.Q. and Merrell, D.S. (2013) The ferric uptake regulator of *Helicobacter pylori*: a critical player in the battle for iron and colonization of the stomach. **8**: 725–738.

Poudel, S., Dunham, E.C., Lindsay, M.R., Amenabar, M.J., Fones, E.M., Colman, D.R., and Boyd, E.S. (2018) Origin and evolution of flavin-based electron bifurcating enzymes. *Front. Microbiol.* **9**: 1–26.

Ragsdale, S.W. and Kumar, M. (1996) Nickel-containing carbon monoxide dehydrogenase/acetyl-CoA synthase. *Chem. Rev.* **96**: 2515–2540.

Ragsdale, S.W. and Pierce, E. (2008) Acetogenesis and the Wood-Ljungdahl Pathway of CO_2 fixation. *Biochim. Biophys. Acta* **1784**: 1873–1898.

Rohdich, F., Wiese, A., Feicht, R., Simon, H., and Bacher, A. (2001) Enoate reductases of Clostridia. Cloning, sequencing, and expression. *J. Biol. Chem.* **276**: 5779–5787.

Sakurama, H., Kishino, S., Mihara, K., Ando, A., Kita, K., Takahashi, S., et al. (2014) Biohydrogenation of C_{20} polyunsaturated fatty acids by anaerobic bacteria. *J. Lipid Res.* **55**: 1855–1863.

Schuchmann, K. and Müller, V. (2012) A bacterial electron-bifurcating hydrogenase. *J. Biol. Chem.* **287**: 31165–31171.

Solovyev, V. and Salamov, A. (2010) Automatic annotation of bacterial community sequences and application to infections diagnostic. *Metagenomics its Appl. Agric. Biomed. Environ. Stud.* 61–78.

Stuermer, R., Hauer, B., Hall, M., and Faber, K. (2007) Asymmetric bioreduction of activated C=C bonds using enoate reductases from the old yellow enzyme family. *Curr. Opin. Chem. Biol.* **11**: 203–213.

Stutz, H.E. and Reid, S.J. (2004) GltX from *Clostridium saccharobutylicum* NCP262: Glutamate synthase or oxidoreductase? *Biochim. Biophys. Acta - Gene Struct. Expr.* **1676**: 71–82.

Taylor, M.C., Jackson, C.J., Tattersall, D.B., French, N., Peat, T.S., Newman, J., et al. (2010) Identification and characterization of two families of F_{420} H_2-dependent reductases from Mycobacteria that catalyse aflatoxin degradation. *Mol. Microbiol.* **78**: 561–575.

Temple, S.J., P, V.C., and Stephen, G.J. (1998) Glutamate synthase and nitrogen

assimilation. *Trends Plant Sci.* **3**: 1360–1385.

Toogood, H.S., Mansell, D., Gardiner, J., and Scrutton, N. (2011) Reduction: enantioselective bioreduction of carbon-carbon double bonds. In, *Comprehensive Chirality*. Oxford: Elsevier Science, p. 216.

Toogood, H.S. and Scrutton, N.S. (2014) New developments in 'ene'-reductase catalysed biological hydrogenations. *Curr. Opin. Chem. Biol.* **19**: 107–115.

Tóth, A., Takács, M., Groma, G., Rákhely, G., and Kovács, K.L. (2008) A novel NADPH-dependent oxidoreductase with a unique domain structure in the hyperthermophilic Archaeon, *Thermococcus litoralis*. *FEMS Microbiol. Lett.* **282**: 8–14.

Tschech, A. and Pfennig, N. (1984) Growth yield increase linked to caffeate reduction in *Acetobacterium woodii*. *Arch. Microbiol.* **137**: 163–167.

Urakawa, H., Martens-Habbena, W., and Stahl, D.A. (2010) High abundance of ammonia-oxidizing archaea in coastal waters, determined using a modified DNA extraction method. *Appl. Environ. Microbiol.* **76**: 2129–2135.

Vanoni, M.A. and Curti, B. (1999) Glutamate synthase: A complex iron-sulfur flavoprotein. *Cell. Mol. Life Sci.* **55**: 617–638.

Vanoni, M.A. and Curti, B. (2008) Structure-function studies of glutamate synthases: A class of self-regulated iron-sulfur flavoenzymes essential for nitrogen assimilation. *IUBMB Life* **60**: 287–300.

Vanoni, M.A., Verzotti, E., Zanetti, G., and Curti, B. (1996) Properties of the recombinant β subunit of glutamate synthase. *Eur. J. Biochem.* **236**: 937–946.

van Vliet, A.H.M., Stoof, J., Vlasblom, R., Wainwright, S.A., Hughes, N.J., Kelly, D.J., et al. (2002) The role of the ferric uptake regulator (Fur) in regulation of *Helicobacter pylori* iron uptake. *Helicobacter* **7**: 237–244.

Walsh, C. (1986). Naturally occurring 5-deazaflavin coenzymes: biological redox roles. *Acc. Chem. Res.* **19**, 216–221.

Wang, S., Huang, H., Moll, J., and Thauer, R.K. (2010) $NADP^+$ reduction with reduced

ferredoxin and $NADP^+$ reduction with NADH are coupled via an electron-bifurcating enzyme complex in *Clostridium kluyveri*. *J. Bacteriol.* **192**: 5115–5123.

Watanabe, S., Sasaki, D., Tominaga, T., and Miki, K. (2012) Structural basis of [NiFe] hydrogenase maturation by Hyp proteins. *Biol. Chem.* **393**: 1089–1100.

Westphal, L., Wiechmann, A., Baker, J., Minton, N.P., and Müller, V. (2018) The Rnf complex is an energy-coupled transhydrogenase essential to reversibly link cellular NADH and ferredoxin pools in the acetogen *Acetobacterium woodii*. *J. Bacteriol.* **200**: 1–13.

Winkler, C.K., Faber, K., and Hall, M. (2018) Biocatalytic reduction of activated C=C-bonds and beyond: emerging trends. *Curr. Opin. Chem. Biol.* **43**: 97–105.

Xia, W., Li, H., Sze, K.H., and Sun, H. (2009) Structure of a nickel chaperone, HypA, from *Helicobacter pylori* reveals two distinct metal binding sites. *J. Am. Chem. Soc.* **131**: 10031–10040.

Zagozda, M. and Plenkiewicz, J. (2006) Enantioselective reduction of α,β-unsaturated ketones by *Geotrichum candidum*, *Mortierella isabellina* and *Rhodotorula rubra* yeast. *Tetrahedron Asymmetry* **17**: 1958–1962.

4.7 Supplementary material

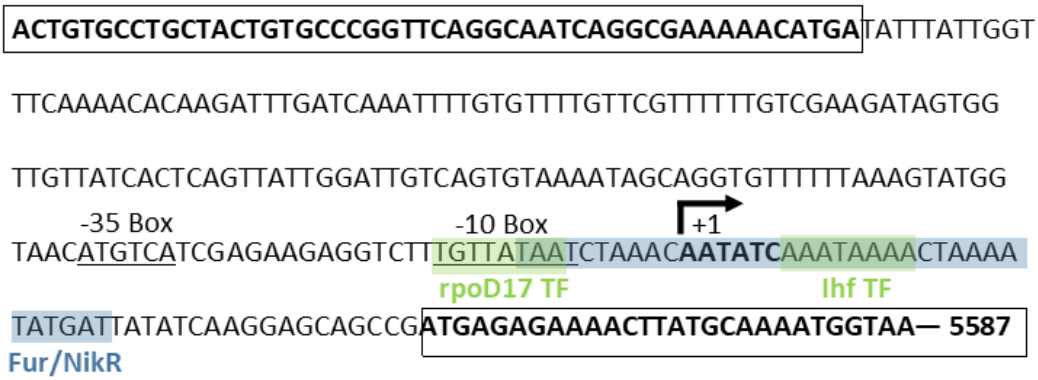

Figure S4 1 Promoter region of isoprene operon. Prediction analysis (BPROM) revealed a potential transcription start site around 44 bp upstream of the open reading frame (ORF) (5587) start codon. Two potential transcription-factor (RNA polymerase sigma factor rpoD17 and ihf) binding sites were predicted at 58 bp and 38 bp upstream of the ORF 1 start codon. Additionally, 4 transcription factor (TF) binding sites could be identified 13 bp upstream of the 5587 start codon. They are highly suggestive binding sites for a Fur (ferric uptake regulator) or NikR (nickel uptake regulator) type of TF.

Table S4 1 Sequence identity information of 5587 homologs represented in the phylogenetic tree (Figure 4.2).

Sequence	Phylum	Group	Source	Size	Clustered	Bifurcating	Protein ID	Located adjacent to	Organism
1033810.HLPCO_05850	unclassified Bacteria	ENOG4107QZ5	EggNOG	469	NA			NA	Haloplasma contractile str. SSD-17B
1033810.HLPCO_13749	unclassified Bacteria	ENOG4107QZ5	EggNOG	924	NA		5587-like protein	HypA, HypB	Haloplasma contractile str. SSD-17B
KWW26252.1	unclassified Bacteria	NA	NCBI	938	ClustRep		5587-like protein	HypA, HypB	bacterium F082
WP_081920190.1	unclassified Bacteria	NA	NCBI	968	ClustRep		5587-like protein	HypA, HypB	bacterium OL-1
WP_120471344.1	unclassified Bacteria	NA	NCBI	927	ClustRep		5587-like protein	HypA, HypB	bacterium 0.1xD8-82
243274.TM1217	Thermotogae	ENOG4107QZ5	EggNOG	618	NA			Nuo subunits NADH-quinone oxidoreductase	Thermotoga maritima
243274.TM1640	Thermotogae	ENOG4107QZ5	EggNOG	468	NA	Yes	NfnB	NfnA	**Thermotoga maritima**
309803.CTN_0818	Thermotogae	ENOG4107QZ5	EggNOG	467	NA		NfnB	NfnA	Thermotoga neapolitana DSM 4359
309803.CTN_1354	Thermotogae	ENOG4107QZ5	EggNOG	618	NA			Nuo subunits NADH-quinone oxidoreductase	Thermotoga neapolitana DSM 4359
381764.Fnod_0283	Thermotogae	ENOG4107QZ5	EggNOG	464	NA		NfnB	NfnA	Fervidobacterium nodosum
381764.Fnod_1607	Thermotogae	ENOG4107QZ5	EggNOG	467	NA		NfnB	NfnA	Fervidobacterium nodosum
381764.Fnod_1736	Thermotogae	ENOG4107QZ5	EggNOG	615	NA			Nuo subunits NADH-quinone oxidoreductase	Fervidobacterium nodosum
390874.Tpet_1151	Thermotogae	ENOG4107QZ5	EggNOG	468	NA		NfnB	NfnA	Thermotoga petrophila
391009.Tmel_1138	Thermotogae	ENOG4107QZ5	EggNOG	458	NA		NfnB	NfnA	Thermosipho melanesiensis (strain DSM 12029
391009.Tmel_1222	Thermotogae	ENOG4107QZ5	EggNOG	616	NA			Nuo subunits NADH-quinone oxidoreductase	Thermosipho melanesiensis (strain DSM 12029
403833.Pmob_0885	Thermotogae	ENOG4107QZ5	EggNOG	612	NA			Nuo subunits NADH-quinone oxidoreductase	Petrotoga mobilis (strain DSM 10674 / SJ95)
403833.Pmob_1795	Thermotogae	ENOG4107QZ5	EggNOG	506	NA			GltB	Petrotoga mobilis (strain DSM 10674 / SJ95)
403833.Pmob_1900	Thermotogae	ENOG4107QZ5	EggNOG	747	NA		NfnB	NfnA	Petrotoga mobilis (strain DSM 10674 / SJ95)
416591.Tlet_0421	Thermotogae	ENOG4107QZ5	EggNOG	471	NA		NfnB	NfnA	Pseudothermotoga lettingae
443254.Marpi_0625	Thermotogae	ENOG4107QZ5	EggNOG	617	NA			Nuo subunits NADH-quinone oxidoreductase	Marinitoga piezophila
484019.THA_1443	Thermotogae	ENOG4107QZ5	EggNOG	457	NA		NfnB	NfnA	Thermosipho africanus

484019.THA_1683	Thermotogae	ENOG4107QZ5	EggNOG	616	NA			Nuo subunits NADH-quinone oxidoreductase	Thermosipho africanus
521045.Kole_1828	Thermotogae	ENOG4107QZ5	EggNOG	466	NA		NfnB	NfnA	Kosmotoga olearia
521045.Kole_2051	Thermotogae	ENOG4107QZ5	EggNOG	604	NA			Nuo subunits NADH-quinone oxidoreductase	Kosmotoga olearia
660470.ThebaDRAFT_1816	Thermotogae	ENOG4107QZ5	EggNOG	472	NA			alone	Mesotoga prima MesG1.Ag.4.2
660470.ThebaDRAFT_2466	Thermotogae	ENOG4107QZ5	EggNOG	609	NA			Nuo subunits NADH-quinone oxidoreductase	Mesotoga prima MesG1.Ag.4.2
688269.Theth_0814	Thermotogae	ENOG4107QZ5	EggNOG	469	NA		NfnB	NfnA	Pseudothermotoga thermarum
PKL00671.1	Tenericutes	NA	NCBI	898	ClustRep		5587-like protein	HypA, HypB	Tenericutes bacterium HGW-Tenericutes-1
WP_038468797.1	Tenericutes	NA	NCBI	894	ClustRep		5587-like protein	HypA, HypB	Candidatus Izimaplasma sp. HR1
WP_095545124.1	Tenericutes	NA	NCBI	894	ClustRep		5587-like protein	HypA, HypB	Candidatus Izimaplasma sp. ZiA1
WP_106699703.1	Tenericutes	NA	NCBI	897	ClustRep		5587-like protein	HypA, HypB, formylmethanofuran dehydrogenase	Tenericutes bacterium MZ-XQ
WP_119016262.1	Tenericutes	NA	NCBI	980	ClustRep		5587-like protein	HypA, HypB	Anaeroplasma bactoclasticum
1104324.P186_1533	TACK group	arCOG01292	EggNOG	348	NA			cytidylate kinase (cmk)	Pyrobaculum ferrireducens
178306.PAE3227	TACK group	arCOG01292	EggNOG	347	NA			cytidylate kinase (cmk)	Pyrobaculum aerophilum (strain ATCC 51768)
272557.APE_0935	TACK group	arCOG01292	EggNOG	337	NA			hypothetical proteins	Aeropyrum pernix (strain ATCC 700893)
368408.Tpen_1766	TACK group	arCOG01292	EggNOG	331	NA			glyceraldehyde-3-phosphate dehydrogenase	Thermofilum pendens (strain DSM 2475 / Hrk 5)
374847.Kcr_0346	TACK group	arCOG01292	EggNOG	483	NA			oxidoreductase FAD/NAD(P)-binding domain prot	Korarchaeum cryptofilum (strain OPF8)
374847.Kcr_0781	TACK group	arCOG01292	EggNOG	361	NA			4Fe-4S ferredoxin iron-sulfur binding domain protein	Korarchaeum cryptofilum (strain OPF8)
384616.Pisl_0400	TACK group	arCOG01292	EggNOG	349	NA			cytidylate kinase (cmk)	Pyrobaculum islandicum (strain DSM 4184)
397948.Cmaq_0612	TACK group	arCOG01292	EggNOG	358	NA			hypothetical proteins	Caldivirga maquilingensis (strain ATCC 700844)
399550.Smar_1055	TACK group	arCOG01292	EggNOG	368	NA			4Fe-4S ferredoxin, iron-sulfur binding domain protein, ADH:ubiquinone oxidoreductase subunit	Staphylothermus marinus (strain ATCC 43588)
410359.Pcal_1108	TACK group	arCOG01292	EggNOG	352	NA			cytidylate kinase (cmk)	Pyrobaculum calidifontis (strain JCM 11548 / VA1)

415426.Hbut_0495	TACK group	arCOG01292	EggNOG	339	NA			Cytosol aminopeptidase	Hyperthermus butylicus (strain DSM 5456)
444157.Tneu_1728	TACK group	arCOG01292	EggNOG	349	NA			cytidylate kinase (cmk)	Pyrobaculum neutrophilum (strain DSM 2338)
490899.DKAM_0100	TACK group	arCOG01292	EggNOG	470	NA		NfnB	NfnA	Desulfurococcus amylolyticus (strain DSM 18924) (Desulfurococcus kamchatkensis)
490899.DKAM_1015	TACK group	arCOG01292	EggNOG	363	NA			4Fe-4S ferredoxin, iron-sulfur binding domain protein, ADH:ubiquinone oxidoreductase subunit 4 (chain M Formate hydrogenlyase subunit 3/Multisubunit Na+/H+ antiporter,	Desulfurococcus amylolyticus (strain DSM 18924 / JCM 16383 / VKM B-2413 / 1221n) (Desulfurococcus kamchatkensis)
572478.Vdis_1723	TACK group	arCOG01292	EggNOG	354	NA			5-carboxymethyl-2-hydroxymuconateDelta-isomerase	Vulcanisaeta distributa (strain DSM 14429)
591019.Shell_1408	TACK group	arCOG01292	EggNOG	368	NA			4Fe-4S ferredoxin, iron-sulfur binding domain protein	Staphylothermus hellenicus (strain DSM 12710)
633148.Tagg_0059	TACK group	arCOG01292	EggNOG	364	NA			4Fe-4S ferredoxin, iron-sulfur binding domain protein, ADH:ubiquinone oxidoreductase subunit	Thermosphaera aggregans (strain DSM 11486 / M11TL)
666510.ASAC_1426	TACK group	arCOG01292	EggNOG	361	NA			NUDIX hydrolase	Acidilobus saccharovorans (strain DSM 16705)
694429.Pyrfu_0077	TACK group	arCOG01292	EggNOG	352	NA			nickel-dependent hydrogenase large subunit	Pyrolobus fumarii (strain DSM 11204 / 1A)
698757.Pogu_0849	TACK group	arCOG01292	EggNOG	349	NA			cytidylate kinase, putative	Pyrobaculum oguniense (strain DSM 13380 / JCM 10595 / TE7)
765177.Desmu_0061	TACK group	arCOG01292	EggNOG	359	NA			4Fe-4S dicluster domain-containing protein Na+/H+ antiporter subunit D	Desulfurococcus mucosus (strain ATCC 35584)
765177.Desmu_0168	TACK group	arCOG01292	EggNOG	483	NA		NfnB	NfnA	Desulfurococcus mucosus (strain ATCC 35584)
768679.TTX_1987	TACK group	arCOG01292	EggNOG	358	NA			electron transfer flavoprotein alpha, beta-subunit (etfAB), Cytidylate kinase	Thermoproteus tenax Kra 1
985053.VMUT_0173	TACK group	arCOG01292	EggNOG	354	NA			5-carboxymethyl-2-hydroxymuconate Delta-isomerase	Vulcanisaeta moutnovskia (strain 768-28)
999630.TUZN_1489	TACK group	arCOG01292	EggNOG	358	NA			cytidylate kinase	Thermoproteus uzoniensis (strain 768-20)

352165.HMPREF7215_0151	Synergistetes	ENOG4107QZ5	EggNOG	480	NA		NfnB	NfnA	Pyramidobacter piscolens W5455
352165.HMPREF7215_1334	Synergistetes	ENOG4107QZ5	EggNOG	610	NA			hydrogenase, Fe-only	Pyramidobacter piscolens W5455
469381.Dpep_0069	Synergistetes	ENOG4107QZ5	EggNOG	1075	NA			NADH-quinone oxidoreductase	Dethiosulfovibrio peptidovorans DSM 11002
469381.Dpep_1443	Synergistetes	ENOG4107QZ5	EggNOG	393	NA			FAD dependent oxidoreductase (ENOG4106BIN)	Dethiosulfovibrio peptidovorans DSM 11002
525903.Taci_0423	Synergistetes	ENOG4107QZ5	EggNOG	1075	NA			Selenate Reductase	Thermanaerovibrio acidaminovorans
525903.Taci_1111	Synergistetes	ENOG4107QZ5	EggNOG	469	NA		NfnB	NfnA	Thermanaerovibrio acidaminovorans
572547.Amico_1369	Synergistetes	ENOG4107QZ5	EggNOG	471	NA		NfnB	NfnA	Aminobacterium colombiense
580340.Tlie_1249	Synergistetes	ENOG4107QZ5	EggNOG	1241	NA			Respiratory-chain NADH dehydrogenase domain 51 kDa subunit	Thermovirga lienii
584708.Apau_1568	Synergistetes	ENOG4107QZ5	EggNOG	467	NA		NfnB	NfnA	Aminomonas paucivorans DSM 12260
584708.Apau_2001	Synergistetes	ENOG4107QZ5	EggNOG	1080	NA			Selenate Reductase	Aminomonas paucivorans DSM 12260
592015.HMPREF1705_00257	Synergistetes	ENOG4107QZ5	EggNOG	1077	NA			protein HymB putative NADH dehydrogenase subunit E	Acetomicrobium hydrogeniformans
645512.GCWU000246_00972	Synergistetes	ENOG4107QZ5	EggNOG	481	NA			oxidoreductase NAD-binding domain protein	Jonquetella anthropi E3_33 E1 2
645512.GCWU000246_01098	Synergistetes	ENOG4107QZ5	EggNOG	635	NA			hydrogenase, Fe-only	Jonquetella anthropi E3_33 E1 2
573413.Spirs_0380	Spirochaetes	ENOG4107QZ5	EggNOG	834	NA			amidohydrolase	Sediminispirochaeta smaragdinae
OQY38447.1	Spirochaetes	NA	NCBI	924	ClustRep		5587-like protein	HypA, HypB	Spirochaetaceae bacterium 4572 7
PKL16097.1	Spirochaetes	NA	NCBI	897	ClustRep		5587-like protein	HypA, HypB	Spirochaetae bacterium HGW-Spirochaetae-5
PKL23510.1	Spirochaetes	NA	NCBI	890	ClustRep		5587-like protein	alone	Spirochaetae bacterium HGW-Spirochaetae-3
PKL23522.1	Spirochaetes	NA	NCBI	900	ClustRep		5587-like protein	alone	Spirochaetae bacterium HGW-Spirochaetae-3
PKL26155.1	Spirochaetes	NA	NCBI	900	ClustRep		5587-like protein	HypA, HypB, GNAT family acetyltransferase (ENOG4108SH0)	Spirochaetae bacterium HGW-Spirochaetae-3
PKL40640.1	Spirochaetes	NA	NCBI	903	ClustRep		5587-like protein	HypA, HypB	Spirochaetae bacterium HGW-Spirochaetae-1
WP_002680008.1	Spirochaetes	NA	NCBI	914	ClustRep		5587-like protein	HypA, HypB	Treponema denticola

WP_002705548.1	Spirochaetes	NA	NCBI	908	ClustRep		5587-like protein	HypA, HypB	Treponema saccharophilum
WP_014269999.1	Spirochaetes	NA	NCBI	894	ClustRep		5587-like protein	HypA, HypB, methyltransferase domain-containing protein (ENOG4108168)	Sphaerochaeta pleomorpha
WP_016525606.1	Spirochaetes	NA	NCBI	911	ClustRep		5587-like protein	HypB	Treponema maltophilum
WP_016526483.1	Spirochaetes	NA	NCBI	891	ClustRep		5587-like protein	HypA, Flavocytochrome c (ENOG4107EFR), NAD(P)H-dependent oxidoreductase	Treponema maltophilum
WP_080658424.1	Spirochaetes	NA	NCBI	568	ClustRep		5587-like protein	HypA	Treponema pedis
WP_114629844.1	Spirochaetes	NA	NCBI	896	ClustRep		5587-like protein	HypA, HypB, FOXRED(ENOG41061TH), helix-turn-helix domain-containing protein (ENOG4107MUM)	Oceanispirochaeta sp. M1
WP_114632307.1	Spirochaetes	NA	NCBI	894	ClustRep			alone	Oceanispirochaeta sp. M1
243090.RB5653	PVC group	ENOG4107QZ5	EggNOG	500	NA		GltD	GltB [EC 1.4.1.13]	Rhodopirellula baltica (strain DSM 10527 / NCIMB 13988 / SH1)
313628.LNTAR_18760	PVC group	ENOG4107QZ5	EggNOG	490	NA		GltD	GltB	Lentisphaera araneosa HTCC2155
314230.DSM3645_15700	PVC group	ENOG4107QZ5	EggNOG	480	NA		GltD	GltB	Blastopirellula marina DSM 3645
320771.Cflav_PD4336	PVC group	ENOG4107QZ5	EggNOG	493	NA		GltD	GltB [EC 1.4.7.1]	Pedosphaera parvula (strain Ellin514)
340101.Vvad_PD1103	PVC group	ENOG4107QZ5	EggNOG	675	NA			NA	Victivallis vadensis ATCC BAA-548
340101.Vvad_PD2971	PVC group	ENOG4107QZ5	EggNOG	462	NA			NA	Victivallis vadensis ATCC BAA-548
344747.PM8797T_04705	PVC group	ENOG4107QZ5	EggNOG	502	NA		GltD	GltB	Gimesia maris DSM 8797
382464.VDG1235_535	PVC group	ENOG4107QZ5	EggNOG	450	NA		GltD	GltB	Verrucomicrobiae bacterium DG1235
452637.Oter_0895	PVC group	ENOG4107QZ5	EggNOG	461	NA		GltD	GltB [EC 1.4.7.1]	Opitutus terrae (strain DSM 11246)
481448.Minf_0627	PVC group	ENOG4107QZ5	EggNOG	448	NA			Dihydroorotate dehydrogenase, Dihydroorotase, Inosine-uridine nucleoside N-ribohydrolase	Methylacidiphilum infernorum (isolate V4) (Methylokorus infernorum (strain V4))
521674.Plim_2693	PVC group	ENOG4107QZ5	EggNOG	489	NA		GltD	GltB	Planctopirus limnophila (strain ATCC 43296)

530564.Psta_0200	PVC group	ENOG4107QZ5	EggNOG	504	NA		GltD	GltB [EC 1.4.1.14]	Pirellula staleyi (strain ATCC 27377)
575540.Isop_2215	PVC group	ENOG4107QZ5	EggNOG	498	NA		GltD	GltB [EC1.4.1.14]	Isosphaera pallida (strain ATCC 43644)
583355.Caka_1870	PVC group	ENOG4107QZ5	EggNOG	495	NA		GltD	GltB [EC 1.4.7.1]	Coraliomargarita akajimensis (strain DSM 45221)
716544.wcw_1970	PVC group	ENOG4107QZ5	EggNOG	611	NA			Cytochrome b subunit of the bc complex, putative acylphosphatase-2	Waddlia chondrophila (strain ATCC VR-1470 / WSU 86-1044)
756272.Plabr_0567	PVC group	ENOG4107QZ5	EggNOG	506	NA		GltD	GltB [EC 1.4.1.14]	Rubinisphaera brasiliensis (strain ATCC 49424)
765952.PUV_26710	PVC group	ENOG4107QZ5	EggNOG	611	NA			alone	Parachlamydia acanthamoebae
OIP95693.1	Proteobacteria	NA	NCBI	1388	ClustRep		long version of 5587	HdrA (partial) homolog	Syntrophobacteraceae bacterium CG2_30_61_12
PKN51602.1	Proteobacteria	NA	NCBI	896	ClustRep			HypA, HypB, enoate reductase	Deltaproteobacteria bacterium HGW-Deltaproteobacteria-13
PNV85846.1	Proteobacteria	NA	NCBI	1397	ClustRep		long version of 5587	HdrA (partial) homolog, hydrogenase iron-sulfur subunit, 4Fe-4S Ferredoxin (ENOG4106TYF),	Desulfobacteraceae bacterium
RLB22586.1	Proteobacteria	NA	NCBI	520	ClustRep			HdrA (partial) homolog	Deltaproteobacteria bacterium
RLB37190.1	Proteobacteria	NA	NCBI	786	ClustRep			HdrA (partial) homolog	Deltaproteobacteria bacterium isolate B17_G16
RLB40401.1	Proteobacteria	NA	NCBI	1393	ClustRep		long version of 5587	HdrA (partial) homolog, HdrC homolog, hydA, nickel-dependent hydrogenase large subunit	Deltaproteobacteria bacterium
RLC27400.1	Proteobacteria	NA	NCBI	625	ClustRep			HdrA (partial) homolog	Deltaproteobacteria bacterium
ROQ90209.1	Proteobacteria	NA	NCBI	1404	ClustRep		long version of 5587	HdrA (partial), F420-non-reducing hydrogenase iron-sulfur subunit	Desulfosoma caldarium
WP_012142278.1	Proteobacteria	NA	NCBI	895	ClustRep		5587-like protein	HypA, HypB	Shewanella sediminis
WP_042497623.1	Proteobacteria	NA	NCBI	900	ClustRep		5587-like protein	HypA, HypB	Vibrio maritimus
WP_084882605.1	Proteobacteria	NA	NCBI	901	ClustRep		5587-like protein	HypA, HypB	Vibrio sp. qd031
WP_102524122.1	Proteobacteria	NA	NCBI	940	ClustRep		5587-like protein	HypA, HypB	Vibrio tapetis
WP_106818637.1	Proteobacteria	NA	NCBI	1525	ClustRep		long version	hydrogenase iron-sulfur subunit,	Syntrophobacter sp. SbD1

							of 5587	5,10-methylenetetrahydrofolate reductase	
WP_112712860.1	Proteobacteria	NA	NCBI	902	ClustRep		5587-like protein	HypA, HypB, another 5587-like protein	Moritella yayanosii
WP_112712864.1	Proteobacteria	NA	NCBI	895	ClustRep				Moritella yayanosii
330214.NIDE3383	Nitrospirae	ENOG4107QZ5	EggNOG	477	NA		GltD	GltB	Candidatus Nitrospira defluvii
KPK29963.1	Nitrospirae	NA	NCBI	1130	ClustRep			F420-non-reducing hydrogenase iron-sulfur subunit	Nitrospira bacterium SG8_3]
526218.Sterm_0257	Fusobacteria	ENOG4107QZ5	EggNOG	413	NA			alone	Sebaldella termitidis ATCC 33386
526218.Sterm_2339	Fusobacteria	ENOG4107QZ5	EggNOG	482	NA		GltD	GltB [1.4.7.1]	Sebaldella termitidis ATCC 33386
572544.Ilyop_0645	Fusobacteria	ENOG4107QZ5	EggNOG	643	NA		5587-like protein	nitroreductase, uncharacterized Fe-S protein	Ilyobacter polytropus (strain DSM 2926 / CuHBu1)
572544.Ilyop_1598	Fusobacteria	ENOG4107QZ5	EggNOG	414	NA		5587-like protein	alone	Ilyobacter polytropus (strain DSM 2926 / CuHBu1)
555088.DealDRAFT_2782	Firmicutes	ENOG4107QZ5	EggNOG	541	NA		5587-like protein	2-ketoisovalerate ferredoxin oxidoreductase subunit alpha and gamma	Dethiobacter alkaliphilus AHT 1
Acetobacterium_dehalogenans_434 WP_084504845.1	Firmicutes	NA	Genomes	949	NA		5587-like protein	HypA, HypB, FOXRED (ENOG41061TH), helix-turn-helix domain-containing protein (ENOG4107MUM)	**Acetobacterium dehalogenans**
Acetobacterium_HGW-4_698 PKM60046	Firmicutes	NA	Genomes	949	NA		5587-like protein	HypA, HypB	**Acetobacterium HGW-4**
Acetobacterium_KB1_466 WP_111886817	Firmicutes	NA	Genomes	924	NA		5587-like protein	HypA, HypB, FOXRED (ENOG41061TH), helix-turn-helix domain-containing protein (ENOG4107MUM)	**Acetobacterium sp. KB1**
Acetobacterium_UBA6819_364	Firmicutes	NA	Genomes	924	NA			NA	**Acetobacterium UBA6819**
ADL33778.1	Firmicutes	NA	NCBI	976	ClustRep		5587-like protein	alone	Butyrivibrio proteoclasticus B316
CBK73919.1	Firmicutes	NA	NCBI	754	ClustRep		5587-like protein	HypA, CO dehydrogenase/acetyl-CoA synthase beta subunit, Benzoyl-CoA reductase/2-hydroxyglutaryl-CoA dehydratase subunit, BcrC/BadD/HgdB	Butyrivibrio fibrisolvens 16 4
CCX45228.1	Firmicutes	NA	NCBI	846	ClustRep		5587-like protein	HypA, NADH flavin oxidoreductase	Firmicutes bacterium CAG 103

								(ENOG4105CCY)	
CCX45944.1	Firmicutes	NA	NCBI	671	ClustRep		5587-like protein	HypA, benzoyl-CoA reductase/2-hydroxyglutaryl-CoA dehydratase subunit BcrC/BadD/HgdB	Firmicutes bacterium CAG 103
CCX47149.1	Firmicutes	NA	NCBI	894	ClustRep		5587-like protein	HypA	Firmicutes bacterium CAG 103
CCX64814.1	Firmicutes	NA	NCBI	946	ClustRep		5587-like protein	MATE efflux family protein (ENOG4107QKB)	Firmicutes bacterium CAG 791
CCX71343.1	Firmicutes	NA	NCBI	905	ClustRep		5587-like protein	alone	Firmicutes bacterium CAG 555
CCX71362.1	Firmicutes	NA	NCBI	922	ClustRep		5587-like protein	HypA, benzoyl-CoA reductase/2-hydroxyglutaryl-CoA dehydratase subunit BcrC/BadD/HgdB	Firmicutes bacterium CAG 555
CCX71375.1	Firmicutes	NA	NCBI	913	ClustRep		5587-like protein	HypA	Firmicutes bacterium CAG 555
CCX71397.1	Firmicutes	NA	NCBI	867	ClustRep		5587-like protein	alone	Firmicutes bacterium CAG 555
CCX72141.1	Firmicutes	NA	NCBI	887	ClustRep		5587-like protein	HypA	Firmicutes bacterium CAG 555
CCX92855.1	Firmicutes	NA	NCBI	923	ClustRep		5587-like protein	HypA, HypB	Firmicutes bacterium CAG 110
CCY29071.1	Firmicutes	NA	NCBI	933	ClustRep		5587-like protein	alone	Roseburia inulinivorans CAG 15
CCY92058.1	Firmicutes	NA	NCBI	572	ClustRep		5587-like protein	HypA	Eubacterium sp. CAG 180
CDA82208.1	Firmicutes	NA	NCBI	935	ClustRep		5587-like protein	alone	Firmicutes bacterium CAG 176
CDB42314.1	Firmicutes	NA	NCBI	952	ClustRep		5587-like protein	HypA	Ruminococcus sp. CAG 177
CDB43749.1	Firmicutes	NA	NCBI	899	ClustRep		5587-like protein	alone	Firmicutes bacterium CAG 240
CDB68050.1	Firmicutes	NA	NCBI	925	ClustRep		5587-like protein	HypA, HypB	Eubacterium sp. CAG 252
CDB87265.1	Firmicutes	NA	NCBI	942	ClustRep		5587-like protein	HypB	Firmicutes bacterium CAG 170
CDC32581.1	Firmicutes	NA	NCBI	947	ClustRep		5587-like protein	HypA	Eubacterium sp. CAG 251
CDD08668.1	Firmicutes	NA	NCBI	952	ClustRep		5587-like protein	HypA	Clostridium sp. CAG 349
CDD22392.1	Firmicutes	NA	NCBI	950	ClustRep		5587-like protein	HypA, HypB	Firmicutes bacterium CAG 313

CDD23849.1	Firmicutes	NA	NCBI	958	ClustRep		5587-like protein	HypA, HypB	Firmicutes bacterium CAG 345
CDD68751.1	Firmicutes	NA	NCBI	888	ClustRep		5587-like protein	HypA	Firmicutes bacterium CAG 475
CDE15899.1	Firmicutes	NA	NCBI	954	ClustRep		5587-like protein	HypA, HypB	Clostridium sp. CAG 288
CKL_0460	Firmicutes	NA	Literature	463	NA	Yes	NfnB	NfnA	**Clostridium kluyveri (strain ATCC 8527)**
CUO55065.1	Firmicutes	NA	NCBI	903	ClustRep		5587-like protein	HypA, HypB	Roseburia hominis
CUP02535.1	Firmicutes	NA	NCBI	893	ClustRep		5587-like protein	HypA, HypB	Catenibacterium mitsuokai
EEC56269.1	Firmicutes	NA	NCBI	906	ClustRep		5587-like protein	HypA, HypB	Bacteroides pectinophilus ATCC 43243
EFW24436.1	Firmicutes	NA	NCBI	936	ClustRep		5587-like protein	HypA, HypB, HypB, MATE efflux family protein (ENOG4107QJQ)	Solobacterium moorei F0204
EGB91832.1	Firmicutes	NA	NCBI	908	ClustRep		5587-like protein	HypA, HypB, Fe-S oxidoreductase (ENOG4107QJX)	Clostridium sp. D5
EOS24024.1	Firmicutes	NA	NCBI	1093	ClustRep		5587-like protein	HypA, MATE efflux family protein (ENOG4107R47)	Lachnospiraceae bacterium 3-1
EOS77523.1	Firmicutes	NA	NCBI	908	ClustRep		5587-like protein	HypA, HypB	Lachnospiraceae bacterium 10-1
EOT23527.1	Firmicutes	NA	NCBI	912	ClustRep		5587-like protein	HypA, HypB	Eubacterium sp. 14-2
ISORED-2_4696	Firmicutes	NA	This work	491	NA		GltD	GltB	Acetobacterium wieringae ISORED-2
ISORED-2_5587	Firmicutes	NA	This work	901	NA		5587-like protein	HypA, HypA, HypB, 4Fe-4S Ferredoxin (ENOG4105DQ9)	Acetobacterium wieringae ISORED-2
OGS54416.1	Firmicutes	NA	NCBI	897	ClustRep		5587-like protein	HypA, HypB	Firmicutes bacterium GWF2 51 9
OJV65277.1	Firmicutes	NA	NCBI	899	ClustRep		5587-like protein	HypA, HypB	Clostridiales bacterium 38-18
OKZ56617.1	Firmicutes	NA	NCBI	901	ClustRep		5587-like protein	HypA, HypB, MATE efflux family protein (ENOG4105C33)	Clostridiales bacterium 44 9
OKZ81429.1	Firmicutes	NA	NCBI	641	ClustRep		5587-like protein	HypB	Clostridium sp.CAG 217 53 7
OLA41937.1	Firmicutes	NA	NCBI	910	ClustRep		5587-like protein	HypA	Firmicutes bacterium CAG 24053 14
OLA42393.1	Firmicutes	NA	NCBI	909	ClustRep		5587-like protein	HypA	Firmicutes bacterium CAG 24053 14
OLA50262.1	Firmicutes	NA	NCBI	894	ClustRep		5587-like protein	HypA	Oscillibacter sp. CAG 241 62 21

PHS34206.1	Firmicutes	NA	NCBI	918	ClustRep		5587-like protein	HypA, HypB	Alkaliphilus sp.
PKM39727.1	Firmicutes	NA	NCBI	910	ClustRep		5587-like protein	HypA, HypB	Firmicutes bacterium HGW-Firmicutes-9
PKM64713.1	Firmicutes	NA	NCBI	590	ClustRep		5587-like protein	HypA, HypB	Firmicutes bacterium HGW-Firmicutes-20
PKM71197.1	Firmicutes	NA	NCBI	595	ClustRep		5587-like protein	HypA, arylamine N-acetyltransferase (ENOG4108MUX)	Firmicutes bacterium HGW-Firmicutes-17
PKM71624.1	Firmicutes	NA	NCBI	874	ClustRep		5587-like protein	alone	Firmicutes bacterium HGW-Firmicutes-16
PKM73428.1	Firmicutes	NA	NCBI	851	ClustRep		5587-like protein	alone	Firmicutes bacterium HGW-Firmicutes-16
PKM73573.1	Firmicutes	NA	NCBI	895	ClustRep		5587-like protein	HypA	Firmicutes bacterium HGW-Firmicutes-16
PKM74107.1	Firmicutes	NA	NCBI	898	ClustRep		5587-like protein	HypA	Firmicutes bacterium HGW-Firmicutes-16
PKM74410.1	Firmicutes	NA	NCBI	895	ClustRep		5587-like protein	HypA	Firmicutes bacterium HGW-Firmicutes-16
PKM75487.1	Firmicutes	NA	NCBI	893	ClustRep		5587-like protein	MBL fold metallo-hydrolase, oleate hydratase	Firmicutes bacterium HGW-Firmicutes-17
PWJ21583.1	Firmicutes	NA	NCBI	931	ClustRep		5587-like protein	HypA, HypB	Faecalicatena orotica
PWJ48606.1	Firmicutes	NA	NCBI	909	ClustRep		5587-like protein	HypA, HypB	Faecalicatena contorta
PWL74023.1	Firmicutes	NA	NCBI	905	ClustRep		5587-like protein	HypA	Clostridiales bacterium
PWM07340.1	Firmicutes	NA	NCBI	926	ClustRep		5587-like protein	HypA, HypB	Clostridiales bacterium
PWM12463.1	Firmicutes	NA	NCBI	899	ClustRep		5587-like protein	HypA, HypB	Clostridiales bacterium
RBP66046.1	Firmicutes	NA	NCBI	909	ClustRep		5587-like protein	HypA, HypB	Alkalibaculum bacchi
RRF97071.1	Firmicutes	NA	NCBI	942	ClustRep		5587-like protein	HypA, another 5587-like protein	Lachnospiraceae bacterium
RRF97113.1	Firmicutes	NA	NCBI	940	ClustRep		5587-like protein		Lachnospiraceae bacterium
SBW07013.1	Firmicutes	NA	NCBI	906	ClustRep		5587-like protein	HypA, HypB	Clostridiales bacterium
SCH32240.1	Firmicutes	NA	NCBI	900	ClustRep		5587-like protein	HypA, HypB	Eubacterium sp.
SCH68246.1	Firmicutes	NA	NCBI	930	ClustRep		5587-like protein	HypA, HypB	Eubacterium sp.

SCI49916.1	Firmicutes	NA	NCBI	891	ClustRep		5587-like protein	HypA, HypB	Flavonifractor sp.
SCX06002.1	Firmicutes	NA	NCBI	931	ClustRep		5587-like protein	HypA, HypB	Lachnospiraceae bacterium YSD2013
SDB41628.1	Firmicutes	NA	NCBI	960	ClustRep		5587-like protein	HypA, CO dehydrogenase/acetyl-CoA synthase complex beta subunit, Benzoyl-CoA reductase/2-hydroxyglutaryl-CoA dehydratase subunit, BcrC/BadD/HgdB	Pseudobutyrivibrio sp. YE44
SDX27870.1	Firmicutes	NA	NCBI	535	ClustRep		5587-like protein	HypA, HypB, Pyridine nucleotide-disulphide oxidoreductase	Eubacterium barkeri
SEJ14715.1	Firmicutes	NA	NCBI	935	ClustRep		5587-like protein	alone	Lachnospiraceae bacterium A10
SEJ29012.1	Firmicutes	NA	NCBI	926	ClustRep		5587-like protein	alone	Sharpea azabuensis
SEP75351.1	Firmicutes	NA	NCBI	941	ClustRep		5587-like protein	HypB, HypB, HypA	Butyrivibrio sp. TB
SFM01331.1	Firmicutes	NA	NCBI	427	ClustRep		5587-like protein	HypA	Pelosinus propionicus DSM 13327
SFT47450.1	Firmicutes	NA	NCBI	957	ClustRep		5587-like protein	HypA	Lachnospiraceae bacterium XBD2001
SHI30339.1	Firmicutes	NA	NCBI	929	ClustRep		5587-like protein	HypB	Pseudobutyrivibrio xylanivorans DSM 14809
SHI66640.1	Firmicutes	NA	NCBI	929	ClustRep		5587-like protein	HypA, HypB, methyl-accepting chemotaxis protein	Lutispora thermophila DSM 19022
WP_004628155.1	Firmicutes	NA	NCBI	900	ClustRep		5587-like protein	HypA, HypB	Ruminiclostridium cellobioparum
WP_005358529.1	Firmicutes	NA	NCBI	922	ClustRep		5587-like protein	HypA, HypB	Eubacterium ventriosum
WP_005952153.1	Firmicutes	NA	NCBI	930	ClustRep		5587-like protein	HypA, HypB	Blautia hydrogenotrophica
WP_009644242.1	Firmicutes	NA	NCBI	930	ClustRep		5587-like protein	HypA, HypB	Mogibacterium sp. CM50
WP_010073488.1	Firmicutes	NA	NCBI	893	ClustRep		5587-like protein	HypA, HypB	Clostridium cellulovorans
WP_010294407.1	Firmicutes	NA	NCBI	902	ClustRep		5587-like protein	HypA, HypB, FOXRED(ENOG4107R52)	Clostridium senegalense
WP_011393026.1	Firmicutes	NA	Literature	463	NA	Yes	NfnB	NfnA	**Moorella thermoacetica DSM 521**
WP_012199592.1	Firmicutes	NA	NCBI	891	ClustRep		5587-like protein	HypA, HypB	Lachnoclostridium phytofermentans

WP_016297229.1	Firmicutes	NA	NCBI	903	ClustRep		5587-like protein	HypA, HypB	Lachnospiraceae bacterium M18-1
WP_017416881.1	Firmicutes	NA	NCBI	896	ClustRep		5587-like protein	HypA, HypB	Clostridium tunisiense
WP_021655858.1	Firmicutes	NA	NCBI	889	ClustRep		5587-like protein	HypA, HypB	Clostridiales bacterium oral taxon 876
WP_024621883.1	Firmicutes	NA	NCBI	895	ClustRep		5587-like protein	HypA, HypB	Clostridioides mangenotii
WP_024837914.1	Firmicutes	NA	NCBI	902	ClustRep		5587-like protein	HypA, HypB	Clostridium sp. 12 A
WP_025645307.1	Firmicutes	NA	NCBI	902	ClustRep		5587-like protein	HypA	Clostridium scindens
WP_026477472.1	Firmicutes	NA	NCBI	898	ClustRep		5587-like protein	HypA, HypB, FOXRED(ENOG41061TH), helix-turn-helix domain-containing protein (ENOG4107MUM)	Alkaliphilus transvaalensis
WP_026651298.1	Firmicutes	NA	NCBI	956	ClustRep		5587-like protein	HypA, HypB	Butyrivibrio proteoclasticus
WP_031369605.1	Firmicutes	NA	NCBI	703	ClustRep		5587-like protein	HypB	Clostridium botulinum
WP_031557378.1	Firmicutes	NA	NCBI	922	ClustRep		5587-like protein	HypA, MATE efflux family protein (ENOG4107QJQ),N-acetyltransferase (ENOG4107Y3V)	Lachnospira multipara
WP_034434763.1	Firmicutes	NA	NCBI	929	ClustRep		5587-like protein	HypA, HypB	Clostridiales bacterium S5-A14a
WP_035146502.1	Firmicutes	NA	NCBI	899	ClustRep		5587-like protein	HypA, HypB	Clostridium tetanomorphum
WP_035381103.1	Firmicutes	NA	NCBI	899	ClustRep		5587-like protein	HypA, HypB	Fervidicella metallireducens
WP_039679487.1	Firmicutes	NA	NCBI	894	ClustRep		5587-like protein	HypA, HypB	Terrisporobacter othiniensis
WP_044947775.1	Firmicutes	NA	NCBI	895	ClustRep		5587-like protein	HypA, HypB	Blautia schinkii
WP_044959082.1	Firmicutes	NA	NCBI	927	ClustRep		5587-like protein	HypA	Shuttleworthia sp. MSX8B
WP_051205918.1	Firmicutes	NA	NCBI	930	ClustRep		5587-like protein	HypA	Butyrivibrio sp. FC2001
WP_051637902.1	Firmicutes	NA	NCBI	921	ClustRep		5587-like protein	HypA	Lachnospiraceae bacterium AC2014
WP_051665539.1	Firmicutes	NA	NCBI	929	ClustRep		5587-like protein	alone	Erysipelotrichaceae bacterium NK3D112
WP_051667144.1	Firmicutes	NA	NCBI	920	ClustRep		5587-like	alone	Lachnospiraceae bacterium

							protein		FE2018
WP_053984118.1	Firmicutes	NA	NCBI	895	ClustRep		5587-like protein	HypA, HypB	Niameybacter massiliensis
WP_059069370.1	Firmicutes	NA	NCBI	902	ClustRep		5587-like protein	HypA, HypB	Clostridiales
WP_066826621.1	Firmicutes	NA	NCBI	901	ClustRep		5587-like protein	HypA, HypB	Clostridium tepidiprofundi
WP_069874311.1	Firmicutes	NA	NCBI	891	ClustRep		5587-like protein	HypA, HypB, HAMP domain-containing histidine kinase	Fusibacter sp. 3D3
WP_072514724.1	Firmicutes	NA	NCBI	962	ClustRep		5587-like protein	HypA, HypB	Ndongobacter massiliensis
WP_072744605.1	Firmicutes	NA	NCBI	887	ClustRep		5587-like protein	HypA, HypB	Sporanaerobacter acetigenes
WP_072985950.1	Firmicutes	NA	NCBI	898	ClustRep		5587-like protein	alone	Clostridium cavendishii
WP_072991077.1	Firmicutes	NA	NCBI	896	ClustRep		5587-like protein	HypA, HypB, iron ABC transporter substrate-binding protein	Clostridium cavendishii]
WP_073006780.1	Firmicutes	NA	NCBI	895	ClustRep		5587-like protein	alone	Clostridium amylolyticum
WP_073008353.1	Firmicutes	NA	NCBI	892	ClustRep		5587-like protein	HypA, HypB	Clostridium amylolyticum
WP_073279569.1	Firmicutes	NA	NCBI	884	ClustRep		5587-like protein	HypA, HypB	Anaerocolumna jejuensis
WP_073285520.1	Firmicutes	NA	NCBI	893	ClustRep		5587-like protein	HypA, HypB	Anaerosporobacter mobilis
WP_073337465.1	Firmicutes	NA	NCBI	894	ClustRep		5587-like protein	HypB	Clostridium grantii
WP_073587367.1	Firmicutes	NA	NCBI	888	ClustRep		5587-like protein	HypA, HypB	Anaerocolumna xylanovorans
WP_074349430.1	Firmicutes	NA	NCBI	900	ClustRep		5587-like protein	HypA, HypB, FOXRED(ENOG41061TH), helix-turn-helix domain-containing protein (ENOG4107MUM)	Proteiniborus sp. DW1
WP_074913038.1	Firmicutes	NA	NCBI	894	ClustRep		5587-like protein	alone	Proteiniclasticum ruminis
WP_075718552.1	Firmicutes	NA	NCBI	922	ClustRep		5587-like protein	HypA, HypB	Roseburia sp. 499
WP_077391239.1	Firmicutes	NA	NCBI	920	ClustRep		5587-like protein	HypA, HypB	Mobilibacterium timonense
WP_077610348.1	Firmicutes	NA	NCBI	913	ClustRep		5587-like protein	HypA, HypB	Clostridium sp. Marseille-P2415

WP_078696869.1	Firmicutes	NA	NCBI	897	ClustRep		5587-like protein	HypA, HypB	Caloramator quimbayensis
WP_078768050.1	Firmicutes	NA	NCBI	925	ClustRep		5587-like protein	alone	Eubacterium coprostanoligenes
WP_078784693.1	Firmicutes	NA	NCBI	923	ClustRep		5587-like protein	HypA, HypB	Gemmiger formicilis
WP_081827563.1	Firmicutes	NA	NCBI	924	ClustRep		5587-like protein	HypA, HypB, FOXRED(ENOG41061TH), helix-turn-helix domain-containing protein (ENOG4108UV8)	Proteiniclasticum ruminis
WP_081921717.1	Firmicutes	NA	NCBI	924	ClustRep		5587-like protein	HypA, HypB, FOXRED(ENOG41061TH), helix-turn-helix domain-containing protein (ENOG4107MUM)	Clostridium sp. CL-2
WP_082669683.1	Firmicutes	NA	NCBI	922	ClustRep		5587-like protein	HypA, HypB	Fournierella massiliensis
WP_084150164.1	Firmicutes	NA	NCBI	912	ClustRep		5587-like protein	HypA, HAD-IIB family hydrolase	Erysipelotrichaceae bacterium NK3D112
WP_084150218.1	Firmicutes	NA	NCBI	920	ClustRep		5587-like protein	alone	Erysipelotrichaceae bacterium NK3D112
WP_087150363.1	Firmicutes	NA	NCBI	914	ClustRep		5587-like protein	HypA, HypB	Lachnoclostridium sp. An298
WP_087160847.1	Firmicutes	NA	NCBI	923	ClustRep		5587-like protein	HypA, HypB	Lachnoclostridium sp. An169
WP_087171445.1	Firmicutes	NA	NCBI	889	ClustRep		5587-like protein	HypA, HypB	Gemmiger sp. An120
WP_089868507.1	Firmicutes	NA	NCBI	925	ClustRep		5587-like protein	HypA, HypB	Lachnospiraceae bacterium G11
WP_089901345.1	Firmicutes	NA	NCBI	927	ClustRep		5587-like protein	HypA	Lachnospiraceae bacterium KHCPX20
WP_089967047.1	Firmicutes	NA	NCBI	903	ClustRep		5587-like protein	HypA	Clostridium gasigenes
WP_090307713.1	Firmicutes	NA	NCBI	908	ClustRep		5587-like protein	HypA, HypB	Eubacterium aggregans
WP_090939676.1	Firmicutes	NA	NCBI	586	ClustRep		5587-like protein	HypA, glutamate racemase	Pelosinus propionicus
WP_091541169.1	Firmicutes	NA	NCBI	894	ClustRep		5587-like protein	GNAT family N-acetyltransferase (ENOG4106FPB)	Alkaliphilus peptidifermentans
WP_093373425.1	Firmicutes	NA	NCBI	896	ClustRep		5587-like protein	HypA, HypB, FOXRED(ENOG41061TH), helix-turn-helix domain-	Tindallia magadiensis

								containing protein (ENOG4107MUM), class I SAM-dependent methyltransferase	
WP_094607538.1	Firmicutes	NA	NCBI	884	ClustRep		5587-like protein	zinc-binding dehydrogenase	Sporomusa silvacetica
WP_097015370.1	Firmicutes	NA	NCBI	892	ClustRep		5587-like protein	HypA, HypB, enoate reductase	Anaerocolumna aminovalerica
WP_100305336.1	Firmicutes	NA	NCBI	906	ClustRep		5587-like protein	HypA, HypB	Clostridium celerecrescens
WP_101698357.1	Firmicutes	NA	NCBI	899	ClustRep		5587-like protein	HypA, HypB	Clostridium minihomine
WP_102048917.1	Firmicutes	NA	NCBI	906	ClustRep		5587-like protein	HypA, HypB, SAM-dependent methyltransferase	Pygmaiobacter massiliensis
WP_102343445.1	Firmicutes	NA	NCBI	962	ClustRep		5587-like protein	HypA, HypB	Lactomassilus timonensis
WP_102400348.1	Firmicutes	NA	NCBI	896	ClustRep		5587-like protein	HypA, HypB	Khelaifiella massiliensis
WP_105304952.1	Firmicutes	NA	NCBI	930	ClustRep		5587-like protein	alone	Anaerolactibacter massiliensis
WP_105304957.1	Firmicutes	NA	NCBI	920	ClustRep		5587-like protein	HypA,HypB, HypB	Anaerolactibacter massiliensis
WP_105614785.1	Firmicutes	NA	NCBI	924	ClustRep		5587-like protein	HypA, HypB, FOXRED(ENOG41061TH), helix-turn-helix domain-containing protein (ENOG4107MUM)	Vallitalea sp. S15
WP_106086518.1	Firmicutes	NA	NCBI	908	ClustRep		5587-like protein	HypA, HypB	Lachnospiraceae bacterium oral taxon 500
WP_108775678.1	Firmicutes	NA	NCBI	925	ClustRep		5587-like protein	HypA, HypB	Lactimicrobium massiliense
WP_110290665.1	Firmicutes	NA	NCBI	889	ClustRep		5587-like protein	HypA, HypB	Lachnotalea glycerini
WP_111930779.1	Firmicutes	NA	NCBI	892	ClustRep		5587-like protein	HypA, HypB	Clostridium tertium
WP_113673380.1	Firmicutes	NA	NCBI	892	ClustRep		5587-like protein	HypA, HypB, FOXRED(ENOG41061TH), helix-turn-helix domain-containing protein (ENOG4107MUM)	Vallitalea guaymasensis
WP_115640882.1	Firmicutes	NA	NCBI	900	ClustRep		5587-like protein	HypA, HypB, HypB	Clostridium putrefaciens
WP_117447008.1	Firmicutes	NA	NCBI	901	ClustRep		5587-like protein	HypA, HypB	Faecalicoccus pleomorphus

WP_117870784.1	Firmicutes	NA	NCBI	580	ClustRep		5587-like protein	HypA	Coprobacillus sp. AF24-1LB
WP_117913544.1	Firmicutes	NA	NCBI	922	ClustRep		5587-like protein	HypA, MATE family efflux transporter	Eubacterium
WP_117922285.1	Firmicutes	NA	NCBI	909	ClustRep		5587-like protein	HypA, HypB, [FeFe]-hydrogenase	Roseburia sp. OF03-24
WP_117995087.1	Firmicutes	NA	NCBI	893	ClustRep		5587-like protein	HypA, HypB	Eubacterium sp. AF22-8LB
WP_118282340.1	Firmicutes	NA	NCBI	957	ClustRep		5587-like protein	HypA, carbon monoxide dehydrogenase	Roseburia
WP_118336469.1	Firmicutes	NA	NCBI	901	ClustRep		5587-like protein	HypA, HypB	Emergencia timonensis
WP_118444154.1	Firmicutes	NA	NCBI	952	ClustRep		5587-like protein	HypA	Ruminococcus gnavus
WP_118573169.1	Firmicutes	NA	NCBI	937	ClustRep		5587-like protein	HypA, HypB, HypB	Coprobacillus
WP_118661508.1	Firmicutes	NA	NCBI	909	ClustRep		5587-like protein	HypA	Coprobacillus
WP_118679210.1	Firmicutes	NA	NCBI	907	ClustRep		5587-like protein	HypA	Eisenbergiella sp. OF01-20
WP_119866010.1	Firmicutes	NA	NCBI	893	ClustRep		5587-like protein	HypA, HypB, HAMP domain-containing histidine kinase	Clostridium isatidis
WP_120406691.1	Firmicutes	NA	NCBI	908	ClustRep		5587-like protein	HypA, HypB	Roseburia sp. 1XD42-69
WP_124755280.1	Firmicutes	NA	NCBI	906	ClustRep		5587-like protein	HypA, HypB	Clostridiales bacterium COT073 COT-073
194439.CT0473	FCB group	ENOG4107QZ5	EggNOG	472	NA		NfnB	NfnA	Chlorobaculum tepidum
194439.CT0854	FCB group	ENOG4107QZ5	EggNOG	577	NA			sulfite reductase, dissimilatory-type, alpha beta gamma subunit	Chlorobaculum tepidum
194439.CT2247	FCB group	ENOG4107QZ5	EggNOG	577	NA			sulfite reductase, dissimilatory-type, alpha beta gamma subunit	Chlorobaculum tepidum
290315.Clim_2002	FCB group	ENOG4107QZ5	EggNOG	480	NA		NfnB	NfnA	Chlorobium limicola
290317.Cpha266_0489	FCB group	ENOG4107QZ5	EggNOG	482	NA		NfnB	NfnA	Chlorobium phaeobacteroides
290512.Paes_0031	FCB group	ENOG4107QZ5	EggNOG	578	NA			sulfite reductase, dissimilatory-type, alpha beta subunit and nitrate reductase gamma subunit	Prosthecochloris aestuarii
290512.Paes_0436	FCB group	ENOG4107QZ5	EggNOG	482	NA		NfnB	NfnA	Prosthecochloris aestuarii
319225.Plut_0038	FCB group	ENOG4107QZ5	EggNOG	578	NA			sulfite reductase, dissimilatory-type, alpha beta subunit and nitrate reductase gamma subunit, 4fe 4s ferredoxin,sulfur	Chlorobium luteolum

								relay protein TusB/DsrH transfer protein DsrM	
319225.Plut_1693	FCB group	ENOG4107QZ5	EggNOG	483	NA		NfnB	NfnA	Chlorobium luteolum
324925.Ppha_0493	FCB group	ENOG4107QZ5	EggNOG	477	NA		NfnB	NfnA	Pelodictyon phaeoclathratiforme
324925.Ppha_2319	FCB group	ENOG4107QZ5	EggNOG	578	NA			sulfite reductase, dissimilatory-type, alpha beta subunit hydrogenobyrinic acid a,c-diamide synthase putative sulfite reductase-associated electron	Pelodictyon phaeoclathratiforme
340177.Cag_0537	FCB group	ENOG4107QZ5	EggNOG	480	NA		NfnB	NfnA	Chlorobium chlorochromatii
340177.Cag_0580	FCB group	ENOG4107QZ5	EggNOG	653	NA			penicillin tolerance protein LytB	Chlorobium chlorochromatii
340177.Cag_1954	FCB group	ENOG4107QZ5	EggNOG	579	NA			sulfite reductase, dissimilatory-type, alpha beta subunit hydrogenobyrinic acid a,c-diamide synthase putative sulfite reductase-associated electron	Chlorobium chlorochromatii
379066.GAU_2380	FCB group	ENOG4107QZ5	EggNOG	593	NA		NADP-specific KGOR. A	Next to 2-oxoacid:acceptor oxidoreductase subunit alpha and beta	Gemmatimonas aurantiaca
517417.Cpar_0030	FCB group	ENOG4107QZ5	EggNOG	578	NA			sulfite reductase, dissimilatory-type, alpha beta subunit and nitrate reductase gamma subunit, 4fe 4s ferredoxin,	Chlorobaculum parvum
517417.Cpar_0468	FCB group	ENOG4107QZ5	EggNOG	482	NA		NfnB	NfnA	Chlorobaculum parvum
517418.Ctha_1864	FCB group	ENOG4107QZ5	EggNOG	482	NA			alone	Chloroherpeton thalassium
59374.Fisuc_0898	FCB group	ENOG4107QZ5	EggNOG	499	NA			pseudo oxidoreductase, NAD-binding	Fibrobacter succinogenes (strain ATCC 19169 / S85)
59374.FSU_1347	FCB group	ENOG4107QZ5	EggNOG	499	NA			pseudo oxidoreductase, NAD-binding	Fibrobacter succinogenes (strain ATCC 19169 / S85)
59374.FSU_2321	FCB group	ENOG4107QZ5	EggNOG	454	NA		GltD	GltB [EC 1.4.1.13]	Fibrobacter succinogenes (strain ATCC 19169 / S85)
626522.GCWU000325_01712	FCB group	ENOG4107QZ5	EggNOG	303	NA			oxidoreductase NAD-binding domain protein	Alloprevotella tannerae ATCC 51259
700598.Niako_1260	FCB group	ENOG4107QZ5	EggNOG	540	NA			pyruvate ferredoxin/flavodoxin oxidoreductase, hydrogenase maturation protease Ni/Fe-hydrogenase, b-type cytochrome subunit Nickel- dependent hydrogenase (NiFe) small subunit HydA hypA hypB hypC	Niastella koreensis

								hypD	
714943.Mucpa_7139	FCB group	ENOG4107QZ5	EggNOG	542	NA			pyruvate ferredoxin/flavodoxin oxidoreductase	Mucilaginibacter paludis
742766.HMPREF9455_01179	FCB group	ENOG4107QZ5	EggNOG	434	NA			hypothetical proteins	Dysgonomonas gadei ATCC BAA-286
752555.PBR_1352	FCB group	ENOG4107QZ5	EggNOG	782	NA			serine--tRNA ligase	Prevotella bryantii B14
760192.Halhy_3670	FCB group	ENOG4107QZ5	EggNOG	598	NA		NADP-specific KGOR. A	2-oxoacid:acceptor oxidoreductase subunit alpha and beta	Haliscomenobacter hydrossis
760192.Halhy_6512	FCB group	ENOG4107QZ5	EggNOG	553	NA			pyruvate ferredoxin/flavodoxin oxidoreductase	Haliscomenobacter hydrossis
761193.Runsl_4914	FCB group	ENOG4107QZ5	EggNOG	598	NA		NADP-specific KGOR. A	2-oxoacid:acceptor oxidoreductase subunit alpha	Runella slithyformis
880070.Cycma_0108	FCB group	ENOG4107QZ5	EggNOG	593	NA		NADP-specific KGOR. A	2-oxoacid:acceptor oxidoreductase subunit alpha and beta	Cyclobacterium marinum
880070.Cycma_4553	FCB group	ENOG4107QZ5	EggNOG	549	NA			pyruvate ferredoxin/flavodoxin oxidoreductase	Cyclobacterium marinum
908937.HMPREF9136_1824	FCB group	ENOG4107QZ5	EggNOG	799	NA		NfnB	NfnA	Prevotella dentalis
997353.HMPREF9144_1191	FCB group	ENOG4107QZ5	EggNOG	784	NA			serine--tRNA ligase	Prevotella dentalis
1042877.GQS_02370	Euryarchaeota	arCOG01292	EggNOG	352	NA			4Fe-4S cluster-binding protein, Coenzyme F420 hydrogenase alpha beta gamma subunit, iron-containing alcohol dehydrogenase, formate hydrogenlyase subunit 3/Multisubunit Na+/H+ antiporter, MnhD subunit	Thermococcus sp. (strain CGMCC 1.5172 / 4557)
1042877.GQS_04960	Euryarchaeota	arCOG01292	EggNOG	950	NA		NsoC	NADH-quinone oxidoreductase subunit NuoF	Thermococcus sp. (strain CGMCC 1.5172 / 4557)
1042877.GQS_08715	Euryarchaeota	arCOG01292	EggNOG	478	NA		NfnB	NfnA	Thermococcus sp. (strain CGMCC 1.5172 / 4557)
1042877.GQS_09215	Euryarchaeota	arCOG01292	EggNOG	479	NA		NfnB	NfnA	Thermococcus sp. (strain CGMCC 1.5172 / 4557)
1110509.Mhar_0690	Euryarchaeota	arCOG01292	EggNOG	496	NA			alone	Methanosaeta harundinacea
186497.PF1327	Euryarchaeota	arCOG01292	EggNOG	474	NA	Yes	NfnI	NfnA	**Pyrococcus furiosus**
186497.PF1852	Euryarchaeota	arCOG01292	EggNOG	357	NA			hypothetical proteins	Pyrococcus furiosus

186497.PF1910	Euryarchaeota	arCOG01292	EggNOG	476	NA		NfnII	sulfide/dihydroorotate dehydrogenase-like FAD/NAD-binding protein, ferredoxin	Pyrococcus furiosus
188937.MA3787	Euryarchaeota	arCOG01292	EggNOG	469	NA		NfnB	NfnA	Methanosarcina acetivorans
192952.MM_0664	Euryarchaeota	arCOG01292	EggNOG	469	NA		NfnB	NfnA	Methanosarcina mazei strain Goe1
246969.TAM4_108	Euryarchaeota	arCOG01292	EggNOG	949	NA		NsoC	NsoAB	Thermococcus sp. AM4
246969.TAM4_1090	Euryarchaeota	arCOG01292	EggNOG	471	NA			Dihydroorotate dehydrogenase electron transfer subunit	Thermococcus sp. AM4
246969.TAM4_594	Euryarchaeota	arCOG01292	EggNOG	350	NA			4Fe-4S cluster-binding protein on each side, Carbon monoxide dehydrogenase CooS, Fe-4S cluster-binding protein CooF, Carbon monoxide dehydrogenase accessory nickel-insertion protein CooC	Thermococcus sp. AM4
246969.TAM4_608	Euryarchaeota	arCOG01292	EggNOG	348	NA			Deoxyhypusine synthase	Thermococcus sp. AM4
246969.TAM4_859	Euryarchaeota	arCOG01292	EggNOG	479	NA				Thermococcus sp. AM4
259564.Mbur_1328	Euryarchaeota	arCOG01292	EggNOG	455	NA		NfnB	NfnA	Methanococcoides burtonii
269797.Mbar_A0227	Euryarchaeota	arCOG01292	EggNOG	469	NA		NfnB	NfnA	Methanosarcina barkeri strain Fusaro
272844.PAB1214	Euryarchaeota	arCOG01292	EggNOG	475	NA			hydG-like hydrogenase, gamma chain fdxA ferredoxin	Pyrococcus abyssi (strain GE5 / Orsay)
272844.PAB1738	Euryarchaeota	arCOG01292	EggNOG	474	NA			ferredoxin--NADP(+) reductase subunit alpha	Pyrococcus abyssi (strain GE5 / Orsay)
273075.Ta0414	Euryarchaeota	arCOG01292	EggNOG	484	NA			2,3-bisphosphoglycerate-independent phosphoglycerate mutase	Thermoplasma acidophilum
273116.TVN1157	Euryarchaeota	arCOG01292	EggNOG	484	NA		NA		
304371.MCP_0989	Euryarchaeota	arCOG01292	EggNOG	468	NA		NfnB	NfnA	Methanocella paludicola
323259.Mhun_1117	Euryarchaeota	arCOG01292	EggNOG	447	NA		NfnB	NfnA	Methanospirillum hungatei JF-1
342949.PNA2_0462	Euryarchaeota	arCOG01292	EggNOG	476	NA		NfnB	NfnA	Pyrococcus sp. (strain NA2)
342949.PNA2_1557	Euryarchaeota	arCOG01292	EggNOG	475	NA			hydrogenase (cytochrome-c3 hydrogenase) delta gamma beta chain	Pyrococcus sp. (strain NA2)
349307.Mthe_1163	Euryarchaeota	arCOG01292	EggNOG	457	NA			alone	Methanothrix thermoacetophila
351160.RRC282	Euryarchaeota	arCOG01292	EggNOG	466	NA			predicted oxidoreductase (cytochrome-c3 hydrogenase,	Methanocella arvoryzae

								gamma subunit family)	
368407.Memar_1282	Euryarchaeota	arCOG01292	EggNOG	447	NA		NfnB	NfnA	Methanoculleus marisnigri
391623.TERMP_00065	Euryarchaeota	arCOG01292	EggNOG	470	NA		NfnB	NfnA	Thermococcus barophilus (strain DSM 11836 / MP)
391623.TERMP_00541	Euryarchaeota	arCOG01292	EggNOG	351	NA			4Fe-4S cluster-binding protein on each side	Thermococcus barophilus (strain DSM 11836 / MP)
391623.TERMP_01523	Euryarchaeota	arCOG01292	EggNOG	480	NA		NfnB	NfnA	Thermococcus barophilus (strain DSM 11836 / MP)
410358.Mlab_1112	Euryarchaeota	arCOG01292	EggNOG	446	NA		NfnB	NfnA	Methanocorpusculum labreanum
439481.Aboo_0084	Euryarchaeota	arCOG01292	EggNOG	476	NA		NfnB	NfnA	Aciduliprofundum boonei (strain DSM 19572 / T469)
439481.Aboo_0977	Euryarchaeota	arCOG01292	EggNOG	1105	NA		annotated as heterodisulfide A , but is much longer and has the 5587 domain in middle	Heterodisulfide reductase, subunit B and C	Aciduliprofundum boonei (strain DSM 19572 / T469)
519442.Huta_0609	Euryarchaeota	arCOG01292	EggNOG	492	NA		GltD	GltB [EC 1.4.7.1]	Halorhabdus utahensis (strain DSM 12940)
521011.Mpal_1828	Euryarchaeota	arCOG01292	EggNOG	448	NA		NfnB	NfnA	Methanosphaerula palustris
523850.TON_0057	Euryarchaeota	arCOG01292	EggNOG	474	NA		NfnB	NfnA	Thermococcus onnurineus (strain NA1)
523850.TON_0542	Euryarchaeota	arCOG01292	EggNOG	352	NA			3 x 4Fe-4S cluster-binding protein, formate dehydrogenase alpha subunit	Thermococcus onnurineus (strain NA1)
523850.TON_0702	Euryarchaeota	arCOG01292	EggNOG	348	NA			deoxyhypusine synthase	Thermococcus onnurineus (strain NA1)
523850.TON_1336	Euryarchaeota	arCOG01292	EggNOG	483	NA		NfnB	NfnA	Thermococcus onnurineus (strain NA1)
523850.TON_1376	Euryarchaeota	arCOG01292	EggNOG	482	NA		NfnB	NfnA	Thermococcus onnurineus (strain NA1)
529709.PYCH_08330	Euryarchaeota	arCOG01292	EggNOG	472	NA		NfnB	NfnA	Pyrococcus yayanosii
529709.PYCH_11210	Euryarchaeota	arCOG01292	EggNOG	351	NA			Pyridine nucleotide-disulfide oxidoreductase, putative	Pyrococcus yayanosii
547558.Mmah_1633	Euryarchaeota	arCOG01292	EggNOG	458	NA		NfnB	NfnA	Methanohalophilus mahii (strain ATCC 35705
593117.TGAM_0070	Euryarchaeota	arCOG01292	EggNOG	356	NA			2x 4Fe-4S cluster-binding protein, Coenzyme F420 hydrogenase subunit alpha (frh	Thermococcus gammatolerans (strain DSM 15229 / JCM 11827 / EJ3)

								alpha gamma beta), formate hydrogenlyase II subunit A (Mhy2-I)	
593117.TGAM_0401	Euryarchaeota	arCOG01292	EggNOG	471	NA		NfnB	NfnA	Thermococcus gammatolerans (strain DSM 15229)
593117.TGAM_0826	Euryarchaeota	arCOG01292	EggNOG	350	NA			4Fe-4S ferredoxin, iron-sulfur binding, 7Fe ferredoxin, Carbon monoxide dehydrogenase (cooS), Hydrogenase maturation protease (HycI)	Thermococcus gammatolerans (strain DSM 15229 / JCM 11827 / EJ3)
593117.TGAM_0896	Euryarchaeota	arCOG01292	EggNOG	482	NA		NfnB	NfnA	Thermococcus gammatolerans (strain DSM 15229)
593117.TGAM_1019	Euryarchaeota	arCOG01292	EggNOG	366	NA			Deoxyhypusine synthase (dhys), FAD/FMN-containing dehydrogenase, Iron-sulfur binding reductase, anaerobic glycerol-3-phosphate dehydrogenase subunit A, putative (glpA) Glycerol kinase (glpK)	Thermococcus gammatolerans (strain DSM 15229 / JCM 11827 / EJ3)
604354.TSIB_1465	Euryarchaeota	arCOG01292	EggNOG	480	NA		NfnB	NfnA	Thermococcus sibiricus
604354.TSIB_1517	Euryarchaeota	arCOG01292	EggNOG	963	NA		NsoC	NADH:ubiquinone oxidoreductase, NADH-binding subunit F E NiFe II hydrogenase alpha beta gamma	Thermococcus sibiricus
679901.Mzhil_0362	Euryarchaeota	arCOG01292	EggNOG	722	NA			Aldehyde Dehydrogenase	Methanosalsum zhilinae (strain DSM 4017)
679926.Mpet_2094	Euryarchaeota	arCOG01292	EggNOG	448	NA		NfnB	NfnA	Methanolacinia petrolearia
69014.TK0672	Euryarchaeota	arCOG01292	EggNOG	348	NA			deoxyhypusine synthase	Thermococcus kodakarensis (strain ATCC BAA-918 / JCM 12380 / KOD1) (Pyrococcus kodakaraensis (strain KOD1))
69014.TK1325	Euryarchaeota	arCOG01292	EggNOG	481	NA		NfnB	NfnA	Thermococcus kodakarensis (strain ATCC BAA-918)
69014.TK1612	Euryarchaeota	arCOG01292	EggNOG	952	NA			NsoAB	Thermococcus kodakarensis (strain ATCC BAA-918)
69014.TK1684	Euryarchaeota	arCOG01292	EggNOG	481	NA		NfnB	NfnA	Thermococcus kodakarensis (strain ATCC BAA-918)
69014.TK2074	Euryarchaeota	arCOG01292	EggNOG	351	NA			4x 4Fe-4S cluster-binding protein probable formate	Thermococcus kodakarensis (strain ATCC BAA-918)

								dehydrogenase, alpha subunit	
693661.Arcve_0800	Euryarchaeota	arCOG01292	EggNOG	387	NA			pyruvate/ketoisovalerate oxidoreductase, gamma subunit, oxidoreductase FAD/NAD(P)-binding domain protein, pyruvate ferredoxin/flavodoxin oxidoreductase, delta subunit	Archaeoglobus veneficus
70601.PH0876	Euryarchaeota	arCOG01292	EggNOG	472	NA			hypothetical proteins	Pyrococcus horikoshii
70601.PH1873	Euryarchaeota	arCOG01292	EggNOG	476	NA		NfnB	NfnA	Pyrococcus horikoshii
990316.MCON_2737	Euryarchaeota	arCOG01292	EggNOG	473	NA		NfnB	NfnA	Methanothrix soehngenii
WP_004069400.1	Euryarchaeota	NA	Literature	955	NA		NsoC	NsoABD	Thermococcus litoralis
445932.Emin_0167	Elusimicrobia	ENOG4107QZ5	EggNOG	465	NA		NfnB	NfnA	Elusimicrobium minutum (strain Pei191)
515635.Dtur_1068	Dictyoglomi	ENOG4107QZ5	EggNOG	461	NA		NfnB	NfnA	Dictyoglomus turgidum
522772.Dacet_0223	Deferribacteres	ENOG4107QZ5	EggNOG	495	NA		NfnB	NfnA	Denitrovibrio acetiphilus
522772.Dacet_2494	Deferribacteres	ENOG4107QZ5	EggNOG	657	NA			NADH ubiquinone oxidoreductase subunit G iron-sulfur binding protein	Denitrovibrio acetiphilus
639282.DEFDS_1613	Deferribacteres	ENOG4107QZ5	EggNOG	651	NA			NADH ubiquinone oxidoreductase subunit G iron-sulfur binding protein	Deferribacter desulfuricans
717231.Flexsi_0433	Deferribacteres	ENOG4107QZ5	EggNOG	648	NA			NADH ubiquinone oxidoreductase subunit G iron-sulfur binding protein	Flexistipes sinusarabici
768670.Calni_1870	Deferribacteres	ENOG4107QZ5	EggNOG	646	NA			NADH ubiquinone oxidoreductase subunit G iron-sulfur binding protein	Calditerrivibrio nitroreducens
1148.SYNGTS_0201	Cyanobacteria	ENOG4107QZ5	EggNOG	494	NA		GltD	alone gltB different location: ALJ66774.1	Synechocystis sp. PCC 6803
118168.MC7420_7676	Cyanobacteria	ENOG4107QZ5	EggNOG	495	NA		GltD	alone gltB different location: EDX77962	Coleofasciculus chthonoplastes PCC 7420
329726.AM1_6355	Cyanobacteria	ENOG4107QZ5	EggNOG	494	NA		GltD	alone gltB different location: ABW31296.1	Acaryochloris marina (strain MBIC 11017)
449447.MAE_14900	Cyanobacteria	ENOG4107QZ5	EggNOG	494	NA		GltD	alone gltB different location: BAG00578	Microcystis aeruginosa (strain NIES-843)
497965.Cyan7822_2990	Cyanobacteria	ENOG4107QZ5	EggNOG	494	NA		GltD	GltB [EC 1.4.7.1]	Cyanothece sp. (strain PCC 7822)
65393.PCC7424_3656	Cyanobacteria	ENOG4107QZ5	EggNOG	494	NA		GltD	GltB [EC 1.4.7.1]	Cyanothece sp. (strain PCC 7424)

756067.MicvaDRAFT_4566	Cyanobacteria	ENOG4107QZ5	EggNOG	494	NA		GltD	alone gltB different location EGK89946.1	Microcoleus vaginatus FGP-2
91464.S7335_2259	Cyanobacteria	ENOG4107QZ5	EggNOG	494	NA		GltD	GltB	Synechococcus sp. (strain ATCC 29403 / PCC 7335)
653733.Selin_0533	Chrysiogenetes	ENOG4107QZ5	EggNOG	478	NA		NfnB	NfnA	Desulfurispirillum indicum
653733.Selin_1387	Chrysiogenetes	ENOG4107QZ5	EggNOG	460	NA		GltD	GltB [EC 1.4.7.1]	Desulfurispirillum indicum (strain ATCC BAA-1389 / S5)
243164.DET0038	Chloroflexi	ENOG4107QZ5	EggNOG	465	NA		NfnB	NfnA	Dehalococcoides mccartyi
255470.cbdb_A46	Chloroflexi	ENOG4107QZ5	EggNOG	465	NA		NA		
311424.DhcVS_36	Chloroflexi	ENOG4107QZ5	EggNOG	465	NA		NfnB	NfnA	Dehalococcoides mccartyi
485913.Krac_4219	Chloroflexi	ENOG4107QZ5	EggNOG	801	NA			transcriptional regulator	Ktedonobacter racemifer DSM 44963
485913.Krac_7555	Chloroflexi	ENOG4107QZ5	EggNOG	801	NA			NADH dehydrogenase (quinone), 4Fe-4S ferredoxin iron-sulfur binding domain protein	Ktedonobacter racemifer DSM 44963
485913.Krac_8409	Chloroflexi	ENOG4107QZ5	EggNOG	601	NA			urocanate hydratase	Ktedonobacter racemifer DSM 44963
525904.Tter_0473	Chloroflexi	ENOG4107QZ5	EggNOG	489	NA			GltB [EC 1.4.7.1]	Thermobaculum terrenum (strain ATCC BAA-798 / YNP1)
552811.Dehly_0707	Chloroflexi	ENOG4107QZ5	EggNOG	1119	NA			alone	Dehalogenimonas lykanthroporepellens
552811.Dehly_1208	Chloroflexi	ENOG4107QZ5	EggNOG	466	NA		NfnB	NfnA	Dehalogenimonas lykanthroporepellens
765420.OSCT_2802	Chloroflexi	ENOG4107QZ5	EggNOG	475	NA			alone	Oscillochloris trichoides DG-6
WP_075062802.1	Chloroflexi	NA	NCBI	896	ClustRep			HypA, HypB	Ornatilinea apprima
123214.PERMA_0627	Aquificae	ENOG4107QZ5	EggNOG	484	NA			ornithine carbamoyltransferase	Persephonella marina (strain DSM 14350 / EX-H1)
204536.SULAZ_0810	Aquificae	ENOG4107QZ5	EggNOG	483	NA			ornithine carbamoyltransferase	Sulfurihydrogenibium azorense (strain Az-Fu1 / DSM 15241 / OCM 825)
224324.aq_2064	Aquificae	ENOG4107QZ5	EggNOG	476	NA			alone	Aquifex aeolicus VF5
380749.HY04AAS1_1069	Aquificae	ENOG4107QZ5	EggNOG	465	NA			2-polyprenylphenol hydroxylase and related flavodoxin oxidoreductase-like	Hydrogenobaculum sp. (strain Y04AAS1)
436114.SYO3AOP1_0539	Aquificae	ENOG4107QZ5	EggNOG	483	NA			ornithine carbamoyltransferase	Sulfurihydrogenibium sp. YO3AOP1

608538.Hydth_0948	Aquificae	ENOG4107QZ5	EggNOG	475	NA			cysteine synthase	Hydrogenobacter thermophilus (strain DSM 6534)
638303.Thal_1076	Aquificae	ENOG4107QZ5	EggNOG	474	NA			cytochrome c class I	Thermocrinis albus DSM 14484]
648996.Theam_0967	Aquificae	ENOG4107QZ5	EggNOG	482	NA			radical SAM domain-containing protein	Thermovibrio ammonificans HB-1,
648996.Theam_1343	Aquificae	ENOG4107QZ5	EggNOG	475	NA			Phosphoglycerate mutase	Thermovibrio ammonificans HB-1,
868864.Dester_0691	Aquificae	ENOG4107QZ5	EggNOG	466	NA		NfnB	NfnA	Desulfurobacterium thermolithotrophum
868864.Dester_1157	Aquificae	ENOG4107QZ5	EggNOG	488	NA			Methionine synthase	Desulfurobacterium thermolithotrophum
BBH50463.1	Actinobacteria	NA	NCBI	929	ClustRep		5587-like protein	HypA, HypB, GltB, GltD	Parolsenella catena
CDD59612.1	Actinobacteria	NA	NCBI	993	ClustRep		5587-like protein	HypA	Eggerthella sp. CAG 298
CUN50020.1	Actinobacteria	NA	NCBI	925	ClustRep		5587-like protein	HypA	Collinsella aerofaciens
EEA90901.1	Actinobacteria	NA	NCBI	931	ClustRep		5587-like protein	HypA	Collinsella stercoris DSM 13279
PWM30825.1	Actinobacteria	NA	NCBI	969	ClustRep		5587-like protein	HypA, HypB	Coriobacteriaceae bacterium
RRF95867.1	Actinobacteria	NA	NCBI	966	ClustRep		5587-like protein	HypA, HypB	Coriobacteriaceae bacterium
SDC41049.1	Actinobacteria	NA	NCBI	1021	ClustRep		5587-like protein	alone	Olsenella umbonata
WP_003148339.1	Actinobacteria	NA	NCBI	1001	ClustRep		5587-like protein	HypA, HypB, GNAT family N-acetyltransferase (ENOG4105MVQ)	Atopobium rimae
WP_013708018.1	Actinobacteria	NA	NCBI	924	ClustRep		5587-like protein	HypA, HypB	Coriobacterium glomerans
WP_077598784.1	Actinobacteria	NA	NCBI	970	ClustRep		5587-like protein	HypA, class II fumarate hydratase, nadh-dependent malic enzyme	Olsenella urininfantis
WP_087430615.1	Actinobacteria	NA	NCBI	932	ClustRep		5587-like protein	HypA, HypB	Collinsella sp. An307
WP_123185963.1	Actinobacteria	NA	NCBI	959	ClustRep		5587-like protein	HypA	Parvibacter caecicola
204669.Acid345_2796	Acidobacteria	ENOG4107QZ5	EggNOG	477	NA		NfnB	NfnA	Koribacter versatilis (strain Ellin345)
204669.Acid345_3679	Acidobacteria	ENOG4107QZ5	EggNOG	471	NA			GltB [EC 1.4.1.14]	Koribacter versatilis (strain Ellin345)

234267.Acid_3510	Acidobacteria	ENOG4107QZ5	EggNOG	671	NA			alone	Solibacter usitatus (strain Ellin6076)
234267.Acid_3815	Acidobacteria	ENOG4107QZ5	EggNOG	466	NA			GltB [EC1.4.1.14]	Solibacter usitatus (strain Ellin6076)
234267.Acid_7661	Acidobacteria	ENOG4107QZ5	EggNOG	437	NA			dihydropyrimidine dehydrogenase, dihydropyrimidinase, aminotransferase, Nitrilase	Solibacter usitatus (strain Ellin6076)
240015.ACP_1387	Acidobacteria	ENOG4107QZ5	EggNOG	478	NA			alone	Acidobacterium capsulatum (strain ATCC 51196)
682795.AciX8_2687	Acidobacteria	ENOG4107QZ5	EggNOG	452	NA			dihydroorotate dehydrogenase family protein, dihydropyrimidinase	Granulicella mallensis (strain ATCC BAA-1857)
940615.AciX9_0264	Acidobacteria	ENOG4107QZ5	EggNOG	448	NA			dihydroorotate dehydrogenase family protein	Granulicella tundricola (strain ATCC BAA-1859)

Table S4 2 InterProScan results of 5587 and characterized NfnB subunits (Figure 4.4).

Start	Stop	Analysis	Signature Accession	Signature description	Interpro Annotation-Accession	Interpro Annotation-description	COG	Score (e-value)	Coverage
5587									
270	320			Glutamate synthase GltD			ENOG4107QZ5	8.75e-196	0.055
345	895			Glutamate synthase GltD			ENOG4107QZ5	8.75e-196	0.61
269	310	Gene3D	G3DSA:3.30.70.20					1.8E-6	
280	309	ProSiteProfiles	PS51379	4Fe-4S ferredoxin-type iron-sulfur binding domain profile.	IPR017896	4Fe-4S ferredoxin-type, iron-sulphur binding domain		11.482	
284	304	Pfam	PF00037	4Fe-4S binding domain	IPR017896	4Fe-4S ferredoxin-type, iron-sulphur binding domain		6.0E-8	
284	314	SUPERFAMILY	SSF54862	4Fe-4S ferredoxins superfamily				2.18E-7	
289	300	ProSitePatterns	PS00198	4Fe-4S ferredoxin-type iron-sulfur binding region signature.	IPR017900	4Fe-4S ferredoxin, iron-sulphur binding, conserved site		-	
337	781	PANTHER	PTHR11938	FAD NADPH dehydrogenase, oxidoreductase				8.7E-96	
337	781	PANTHER	PTHR11938:SF112	FAD NADPH dehydrogenase, oxidoreductase				8.7E-96	
339	440	Gene3D	G3DSA:1.10.1060.10	Alpha-helical ferredoxin	IPR009051	Alpha-helical ferredoxin		2.5E-25	
350	454	SUPERFAMILY	SSF46548	alpha-helical ferredoxin superfamily	IPR009051	Alpha-helical ferredoxin		1.05E-18	
351	428	Pfam	PF14691	Dihydroprymidine dehydrogenase domain II, 4Fe-4S cluster	IPR028261	Dihydroprymidine dehydrogenase domain II		1.4E-18	
441	524	Gene3D	G3DSA:3.50.50.60	FAD/NAD(P)-binding domain	IPR036188	FAD/NAD(P)-binding domain superfamily		5.4E-13	
443	687	SUPERFAMILY	SSF51971	Nucleotide-binding domain superfamily				6.47E-58	
444	702	Pfam	PF07992	Pyridine nucleotide-disulphide oxidoreductase	IPR023753	FAD/NAD(P)-binding domain		5.2E-18	
445	467	PRINTS	PR00419	Adrenodoxin reductase family signature				3.9E-20	
468	481	PRINTS	PR00419	Adrenodoxin reductase family signature				3.9E-20	
510	520	PRINTS	PR00419	Adrenodoxin reductase family signature				3.9E-20	
543	686	Gene3D	G3DSA:3.50.50.60	FAD/NAD(P)-binding domain	IPR036188	FAD/NAD(P)-binding domain superfamily		4.9E-38	

577	591	PRINTS	PR00419	Adrenodoxin reductase family signature				3.9E-20	
611	624	PRINTS	PR00419	Adrenodoxin reductase family signature				3.9E-20	
667	783	SUPERFAMILY	SSF51971	Nucleotide-binding domain superfamily				7.3E-5	
822	875	SUPERFAMILY	SSF46548	alpha-helical ferredoxin superfamily	IPR009051	Alpha-helical ferredoxin		5.75E-10	
828	898	Gene3D	G3DSA:3.30.70.3270					1.7E-5	
848	877	ProSiteProfiles	PS51379	4Fe-4S ferredoxin-type iron-sulfur binding domain profile.	IPR017896	4Fe-4S ferredoxin-type, iron-sulphur binding domain		10.289	
TM_1640									
1	140	CATH-Gene3D	G3DSA:1.10.1060.10	Alpha-helical ferredoxin	IPR009051	Alpha-helical ferredoxin		2.5E-43	
2	460	PANTHER	PTHR11938:SF130	PTHR11938:SF130				0	
2	460	PANTHER	PTHR11938	FAD NADPH dehydrogenase, oxidoreductase				0	
7	151	SUPERFAMILY	SSF46548	alpha-helical ferredoxin superfamily	IPR009051	Alpha-helical ferredoxin		2.82E-34	
13	459	TIGRFAMs	TIGR01316	gltA glutamate synthase (NADPH), homotetrameric	IPR006004	Glutamate synthase (NADPH), homotetrameric		0	
16	127	Pfam	PF14691	Dihydroprymidine dehydrogenase domain II, 4Fe-4S cluster	IPR028261	Dihydroprymidine dehydrogenase domain II		2.3E-44	
141	266	CATH-Gene3D	G3DSA:3.50.50.60	FAD/NAD(P)-binding domain	IPR036188	FAD/NAD(P)-binding domain superfamily		4.7E-37	
141	393	SUPERFAMILY	SSF51971	Nucleotide-binding domain superfamily				2.24E-72	
142	448	Pfam	PF07992	Pyridine nucleotide-disulphide oxidoreductase	IPR023753	FAD/NAD(P)-binding domain		2.6E-39	
142	164	PRINTS	PR00419	Adrenodoxin reductase family signature				1.8E-21	
165	178	PRINTS	PR00419	Adrenodoxin reductase family signature				1.8E-21	
207	217	PRINTS	PR00419	Adrenodoxin reductase family signature				1.8E-21	
277	466	CATH-Gene3D	G3DSA:3.50.50.60	FAD/NAD(P)-binding domain	IPR036188	FAD/NAD(P)-binding domain superfamily		2.7E-34	
282	296	PRINTS	PR00419	Adrenodoxin reductase family signature				1.8E-21	
316	329	PRINTS	PR00419	Adrenodoxin reductase family signature				1.8E-21	
CKL_0460									
4	140	CATH-Gene3D	G3DSA:1.10.1060.10	Alpha-helical ferredoxin	IPR009051	Alpha-helical ferredoxin		2.2E-43	

4	463	PANTHER	PTHR11938	FAD NADPH dehydrogenase, oxidoreductase				0	
4	463	PANTHER	PTHR11938:SF130	FAD NADPH dehydrogenase/oxidoreductase				0	
7	150	SUPERFAMILY	SSF46548	alpha-helical ferredoxin superfamily	IPR009051	Alpha-helical ferredoxin		8.37E-36	
16	460	TIGRFAMs	TIGR01316	gltA glutamate synthase (NADPH), homotetrameric	IPR006004	Glutamate synthase (NADPH), homotetrameric		0	
19	129	Pfam	PF14691	Dihydroprymidine dehydrogenase domain II, 4Fe-4S cluster	IPR028261	Dihydroprymidine dehydrogenase domain II		1.5E-44	
140	393	SUPERFAMILY	SSF51971	Nucleotide-binding domain superfamily				2.3E-68	
141	448	Pfam	PF07992	Pyridine nucleotide-disulphide oxidoreductase	IPR023753	FAD/NAD(P)-binding domain		1.4E-32	
141	279	CATH-Gene3D	G3DSA:3.50.50.60	FAD/NAD(P)-binding domain	IPR036188	FAD/NAD(P)-binding domain superfamily		3.7E-38	
142	164	PRINTS	PR00419	Adrenodoxin reductase family signature				4.5E-19	
165	178	PRINTS	PR00419	Adrenodoxin reductase family signature				4.5E-19	
208	218	PRINTS	PR00419	Adrenodoxin reductase family signature				4.5E-19	
280	463	CATH-Gene3D	G3DSA:3.50.50.60	FAD/NAD(P)-binding domain	IPR036188	FAD/NAD(P)-binding domain superfamily		2.7E-30	
284	298	PRINTS	PR00419	Adrenodoxin reductase family signature				4.5E-19	
317	330	PRINTS	PR00419	Adrenodoxin reductase family signature				4.5E-19	
PF1327 SUDHA_PYRFU									
5	165	SUPERFAMILY	SSF46548	alpha-helical ferredoxin superfamily	IPR009051	Alpha-helical ferredoxin		8.37E-35	
5	472	PANTHER	PTHR11938	FAD NADPH dehydrogenase, oxidoreductase				0	
5	472	PANTHER	PTHR11938:SF130	FAD NADPH dehydrogenase, oxidoreductase				0	
6	147	CATH-Gene3D	G3DSA:1.10.1060.10	Alpha-helical ferredoxin	IPR009051	Alpha-helical ferredoxin		3.2E-41	
17	470	TIGRFAMs	TIGR01316	gltA glutamate synthase (NADPH), homotetrameric	IPR006004	Glutamate synthase (NADPH), homotetrameric		0	
20	137	Pfam	PF14691	Dihydroprymidine dehydrogenase domain II, 4Fe-4S cluster	IPR028261	Dihydroprymidine dehydrogenase domain II		1.7E-39	
151	263	CATH-Gene3D	G3DSA:3.40.50.720	G3DSA:3.40.50.720				1.8E-25	
153	405	SUPERFAMILY	SSF51971	Nucleotide-binding domain superfamily				1E-74	

155	455	Pfam	PF07992	Pyridine nucleotide-disulphide oxidoreductase	IPR023753	FAD/NAD(P)-binding domain		7.1E-44	
156	178	PRINTS	PR00411	Pyridine nucleotide disulphide reductase class-I signature				4.7E-16	
157	176	PRINTS	PR00368	FADPNR				1E-23	
264	404	CATH-Gene3D	G3DSA:3.50.50.60	FAD/NAD(P)-binding domain	IPR036188	FAD/NAD(P)-binding domain superfamily		6.2E-51	
295	320	PRINTS	PR00411	Pyridine nucleotide disulphide reductase class-I signature	IPR009051	Alpha-helical ferredoxin		4.7E-16	
295	313	PRINTS	PR00368	FAD-dependent pyridine nucleotide reductase signature				1E-23	
396	412	PRINTS	PR00368	FAD-dependent pyridine nucleotide reductase signature				1E-23	
397	411	PRINTS	PR00411	Pyridine nucleotide disulphide reductase class-I signature	IPR009051	Alpha-helical ferredoxin		4.7E-16	
424	446	PRINTS	PR00368	FAD-dependent pyridine nucleotide reductase signature	IPR006004	Glutamate synthase (NADPH), homotetrameric		1E-23	
439	446	PRINTS	PR00411	Pyridine nucleotide disulphide reductase class-I signature	IPR028261	Dihydroprymidine dehydrogenase domain II		4.7E-16	
WP_011393026									
5	463	PANTHER	PTHR11938	FAD NADPH dehydrogenase, oxidoreductase				0	
5	463	PANTHER	PTHR11938:SF130	FAD NADPH dehydrogenase, oxidoreductase				0	
6	138	CATH-Gene3D	G3DSA:1.10.1060.10	Alpha-helical ferredoxin	IPR009051	Alpha-helical ferredoxin		4.9E-43	
9	151	SUPERFAMILY	SSF46548	alpha-helical ferredoxin superfamily	IPR009051	Alpha-helical ferredoxin		1.2E-36	
15	460	TIGRFAMs	TIGR01316	gltA glutamate synthase (NADPH), homotetrameric	IPR006004	Glutamate synthase (NADPH), homotetrameric		0	
18	127	Pfam	PF14691	Dihydroprymidine dehydrogenase domain II, 4Fe-4S cluster	IPR028261	Dihydroprymidine dehydrogenase domain II		2E-44	
139	248	CATH-Gene3D	G3DSA:3.40.50.720	G3DSA:3.40.50.720				6.7E-25	
140	394	SUPERFAMILY	SSF51971	Nucleotide-binding domain superfamily				2.06E-71	
141	448	Pfam	PF07992	Pyridine nucleotide-disulphide oxidoreductase	IPR023753	FAD/NAD(P)-binding domain		8.7E-38	
142	164	PRINTS	PR00419	Adrenodoxin reductase family signature				3.7E-19	
165	178	PRINTS	PR00419	Adrenodoxin reductase family signature				3.7E-19	
249	393	CATH-Gene3D	G3DSA:3.50.50.60	FAD/NAD(P)-binding domain	IPR036188	FAD/NAD(P)-binding domain superfamily		2.9E-49	

283	297	PRINTS	PR00419	Adrenodoxin reductase family signature				3.7E-19	
317	330	PRINTS	PR00419	Adrenodoxin reductase family signature				3.7E-19	
***Acetobacterium dehalogenans* WP_084504845**									
280	401	SUPERFAMILY	SSF46548	alpha-helical ferredoxin superfamily	IPR009051	Alpha-helical ferredoxin		2.35E-9	
304	343	Gene3D	G3DSA:3.30.70.20					6.3E-6	
311	340	ProSiteProfiles	PS51379	4Fe-4S ferredoxin-type iron-sulfur binding domain profile.	IPR017896	4Fe-4S ferredoxin-type, iron-sulphur binding domain		12.036	
314	336	Pfam	PF00037	4Fe-4S binding domain	IPR017896	4Fe-4S ferredoxin-type, iron-sulphur binding domain		1.2E-9	
320	331	ProSitePatterns	PS00198	4Fe-4S ferredoxin-type iron-sulfur binding region signature.	IPR017900	4Fe-4S ferredoxin, iron-sulphur binding, conserved site		-	
344	474	Gene3D	G3DSA:1.10.1060.10		IPR009051	Alpha-helical ferredoxin		7.8E-28	
367	788	PANTHER	PTHR11938:SF112	FAD NADPH dehydrogenase, oxidoreductase				1.0E-89	
367	788	PANTHER	PTHR11938	FAD NADPH dehydrogenase, oxidoreductase				1.0E-89	
380	485	SUPERFAMILY	SSF46548	alpha-helical ferredoxin superfamily	IPR009051	Alpha-helical ferredoxin		3.24E-19	
381	460	Pfam	PF14691	Dihydroprymidine dehydrogenase domain II, 4Fe-4S cluster	IPR028261	Dihydroprymidine dehydrogenase domain II		6.1E-20	
475	605	Gene3D	G3DSA:3.50.50.60	FAD/NAD(P)-binding domain	IPR036188	FAD/NAD(P)-binding domain superfamily		8.9E-37	
475	719	SUPERFAMILY	SSF51971	Nucleotide-binding domain superfamily				8.24E-53	
476	774	Pfam	PF07992	Pyridine nucleotide-disulphide oxidoreductase	IPR023753	FAD/NAD(P)-binding domain		4.3E-26	
477	499	PRINTS	PR00419	Adrenodoxin reductase family signature				6.2E-20	
500	513	PRINTS	PR00419	Adrenodoxin reductase family signature				6.2E-20	
542	552	PRINTS	PR00419	Adrenodoxin reductase family signature				6.2E-20	
606	807	Gene3D	G3DSA:3.50.50.60	FAD/NAD(P)-binding domain	IPR036188	FAD/NAD(P)-binding domain superfamily		3.8E-21	
609	623	PRINTS	PR00419	Adrenodoxin reductase family signature				6.2E-20	
703	787	SUPERFAMILY	SSF51971	Nucleotide-binding domain superfamily				7.65E-6	
850	904	SUPERFAMILY	SSF54862	4Fe-4S ferredoxins superfamily				1.28E-7	
857	886	ProSitePr	PS51379	4Fe-4S ferredoxin-type iron-sulfur binding	IPR017896	4Fe-4S ferredoxin-type, iron-sulphur binding		9.735	

		ofiles		domain profile.		domain			
860	882	Pfam	PF00037	4Fe-4S binding domain	IPR017896	4Fe-4S ferredoxin-type, iron-sulphur binding domain		3.6E-5	
5588									
1	112						ENOG41082RJ	3.2E-32	0.94
1	111	Pfam	PF01155	Hydrogenase/urease nickel incorporation, metallochaperone, hypA	IPR000688	Hydrogenase maturation factor HypA/HybF		2.2E-24	
1	115	PANTHER	PTHR34535		IPR000688	Hydrogenase maturation factor HypA/HybF		8.1E-24	
1	115	PANTHER	PTHR34535:SF3					8.1E-24	
1	117	PIRSF	PIRSF004761		IPR000688	Hydrogenase maturation factor HypA/HybF		4.2E-24	
1	115	Gene3D	G3DSA:3.30.2320.80					2.9E-25	
1	114	Hamap	MF_00213	Hydrogenase maturation factor HypA [hypA].	IPR000688	Hydrogenase maturation factor HypA/HybF		17.66	
5589									
6	218						ENOG4107T0V	2.6E-101	0.94
8	217	TIGRFAM	TIGR00073	hypB: hydrogenase accessory protein HypB	IPR004392	Hydrogenase maturation factor HypB		1.2E-73	
16	220	Gene3D	G3DSA:3.40.50.300					3.2E-59	
37	197	Pfam	PF02492	CobW/HypB/UreG, nucleotide-binding domain	IPR003495	CobW/HypB/UreG, nucleotide-binding domain		2.7E-25	
6	217	SUPERFAMILY	SSF52540		IPR027417	P-loop containing nucleoside triphosphate hydrolase		6.83E-38	
4	219	PANTHER	PTHR30134:SF2		IPR004392	Hydrogenase maturation factor HypB		1.1E-85	
3	223	PIRSF	PIRSF005624		IPR012202	[NiFe]-hydrogenase/urease maturation factor, Ni2-binding GTPase		7.9E-78	
4	219	PANTHER	PTHR30134					1.1E-85	
5590									
2	348						ENOG4105DQ9	7.73E-85	0.98
272	317	Pfam	PF13237	4Fe-4S dicluster domain	IPR017896	4Fe-4S ferredoxin-type, iron-sulphur binding domain		1.4E-7	
213	233	Coils	Coil					-	

267	327	SUPERFAMILY	SSF54862					4.05E-17	
278	289	ProSitePatterns	PS00198	4Fe-4S ferredoxin-type iron-sulfur binding region signature.	IPR017900	4Fe-4S ferredoxin, iron-sulphur binding, conserved site		-	
269	298	ProSiteProfiles	PS51379	4Fe-4S ferredoxin-type iron-sulfur binding domain profile.	IPR017896	4Fe-4S ferredoxin-type, iron-sulphur binding domain		10.314	
270	350	Gene3D	G3DSA:3.30.70.20					1.6E-13	
182	323	PANTHER	PTHR24960					2.9E-12	
300	327	ProSiteProfiles	PS51379	4Fe-4S ferredoxin-type iron-sulfur binding domain profile.	IPR017896	4Fe-4S ferredoxin-type, iron-sulphur binding domain		8.81	
5591									
1	115						ENOG4105WMM	1.9E-24	0.97
1	115	Pfam	PF01155	Hydrogenase/urease nickel incorporation, metallochaperone, hypA	IPR000688	Hydrogenase maturation factor HypA/HybF		8.1E-25	
1	117	Gene3D	G3DSA:3.30.2320.80					1.6E-27	
1	115	PANTHER	PTHR34535		IPR000688	Hydrogenase maturation factor HypA/HybF		5.4E-24	
1	117	PIRSF	PIRSF004761		IPR000688	Hydrogenase maturation factor HypA/HybF		1.2E-25	
1	116	Hamap	MF_00213	Hydrogenase maturation factor HypA [hypA].	IPR000688	Hydrogenase maturation factor HypA/HybF		18.439	
1	115	PANTHER	PTHR34535:SF3					5.4E-24	
TM_1640									
1	140	CATH-Gene3D	G3DSA:1.10.1060.10	G3DSA:1.10.1060.10				2.5E-43	
2	460	PANTHER	PTHR11938:SF130	PTHR11938:SF130				0	
2	460	PANTHER	PTHR11938	PTHR11938				0	
7	151	SUPERFAMILY	SSF46548	SSF46548				2.82E-34	
13	459	TIGRFAMs	TIGR01316	gltA				0	
16	127	Pfam	PF14691	Fer4_20				2.3E-44	
141	266	CATH-Gene3D	G3DSA:3.50.50.60	G3DSA:3.50.50.60				4.7E-37	
141	393	SUPERFAMILY	SSF51971	SSF51971				2.24E-72	

142	448	Pfam	PF07992	Pyr_redox_2				2.6E-39	
142	164	PRINTS	PR00419	ADXRDTASE				1.8E-21	
165	178	PRINTS	PR00419	ADXRDTASE				1.8E-21	
207	217	PRINTS	PR00419	ADXRDTASE				1.8E-21	
277	466	CATH-Gene3D	G3DSA:3.50.50.60	G3DSA:3.50.50.60				2.7E-34	
282	296	PRINTS	PR00419	ADXRDTASE				1.8E-21	
316	329	PRINTS	PR00419	ADXRDTASE				1.8E-21	

Table S4 3 List of the first 300 most abundant proteins detected in *Acetobacterium wieringae* ISORED-2 proteom analysis that were not differently expressed. Enzymes of the Wood-Ljungdhal pathway are marked in orange, subunits of the [FeFe]-hydrogenase are marked in blue, the RNF complex is marked in green and subunits of the ATP synthase are marked in purple.

Protein ID	UniProtKB	EggNOG	Protein/enzyme function	Classification
149	349161.Dred_2604	ENOG4108M4Y	Phage capsid family	Function unknown
154	903814.ELI_1304	ENOG4107FYT	phage major tail protein, phi13 family	Function unknown
155	499229.TepRe1_2415	ENOG41061FZ	NA	Function unknown
156	931626.Awo_c35010	ENOG4108FTH	tail tape measure protein	Cell wall/membrane/envelope biogenesis
176	428125.CLOLEP_02977	ENOG4106TRD	domain protein	Function unknown
234	445971.ANASTE_00093	ENOG410653W	NA	Function unknown
237	428125.CLOLEP_02977	ENOG4106TRD	domain protein	Function unknown
257	931626.Awo_c07570	ENOG4108UGT	Methyl-accepting chemotaxis	Cell motility, Signal transduction mechanisms
263	931626.Awo_c09390	ENOG4105CRK	lysyL-tRNA synthetase	Translation, ribosomal structure and biogenesis
265	931626.Awo_c09370	ENOG4105CGR	Catalyzes the attachment of serine to tRNA(Ser). Is also able to aminoacylate tRNA(Sec) with serine, to form the misacylated tRNA L-seryl-tRNA(Sec), which will be further converted into selenocysteinyl-tRNA(Sec) (By similarity)	Translation, ribosomal structure and biogenesis
273	931626.Awo_c09310	ENOG4107UPV	Methylenetetrahydrofolate reductase	Amino acid transport and metabolism
274	931626.Awo_c09300	ENOG4107ZBE	zinc-finger protein	Function unknown
275	476272.RUMHYD_01187	ENOG4107QTR	Required for nitrogen fixation. May be part of a membrane complex functioning as an intermediate in the electron transport to nitrogenase (By similarity)	Energy production and conversion
276	931626.Awo_c09280	ENOG4105CN0	Catalyzes the oxidation of 5,10- methylenetetrahydrofolate to 5,10-methenyltetrahydrofolate and then the hydrolysis of 5,10-methenyltetrahydrofolate to 10- formyltetrahydrofolate (By similarity)	Coenzyme transport and metabolism
277	931626.Awo_c09270	ENOG410908R	Methenyltetrahydrofolate cyclohydrolase	Amino acid transport and metabolism
278	931626.Awo_c08040	ENOG4105CKU	formyltetrahydrofolate synthetase	Nucleotide transport and metabolism
356	536227.CLCAR_3604	ENOG4108FSY	Domain of unknown function (DUF336)	Function unknown
657	931626.Awo_c06270	ENOG41079AY	NA	Function unknown

662	931626.Awo_c06340	ENOG4105C70	Catalyzes the reversible conversion of 2- phosphoglycerate into phosphoenolpyruvate. It is essential for the degradation of carbohydrates via glycolysis (By similarity)	Carbohydrate transport and metabolism
785	931626.Awo_c08040	ENOG4105CKU	formyltetrahydrofolate synthetase	Nucleotide transport and metabolism
796	457396.CSBG_01074	ENOG4105CTQ	phosphohexokinase	Carbohydrate transport and metabolism
814	1029718.SFBM_0247	ENOG4105C07	amino acids such as valine, to avoid such errors it has two additional distinct tRNA(Ile)-dependent editing activities. One activity is designated as 'pretransfer' editing and involves the hydrolysis of activated Val-AMP. The other activity is designated 'posttransfer' editing and involves deacylation of mischarged Val-tRNA(Ile) (By similarity)	Translation, ribosomal structure and biogenesis
1169	931626.Awo_c34930	ENOG4107H3G	N-acetylmuramoyl-L-alanine amidase	Function unknown
1170	931626.Awo_c30880	ENOG4105XSJ	toxin secretion phage lysis holin	Function unknown
1175	931626.Awo_c18200	ENOG4108UGT	Methyl-accepting chemotaxis	Cell motility, Signal transduction mechanisms
1184	903814.ELI_2637	ENOG4105C5H	ABC transporter	Function unknown
1258	903814.ELI_3127	ENOG4108X0E	Endonuclease that resolves Holliday junction intermediates in genetic recombination. Cleaves mobile four-strand junctions by introducing symmetrical nicks in paired strands. Promotes annealing of linear ssDNA with homologous dsDNA. Required for DNA repair, homologous recombination and chromosome segregation (By similarity)	Function unknown
1266	349161.Dred_2608	ENOG4106ZB8	NA	Function unknown
1268	349161.Dred_2606	ENOG419058	Phage portal protein	Function unknown
1270	349161.Dred_2604	ENOG4108M4Y	Phage capsid family	Function unknown
1313	931626.Awo_c20650	ENOG4105ET5	solute-binding protein	Function unknown
1320	1148.SYNGTS_1829	ENOG4107QM6	fructose-bisphosphate aldolase	Carbohydrate transport and metabolism
1736	931626.Awo_c09880	ENOG4105CN5	Metallo-Beta-Lactamase	Post-translational modification, protein turnover, and chaperones
1739	931626.Awo_c09860	ENOG4107NZZ	Protein of unknown function (DUF1292)	Function unknown
1743	195103.CPF_2033	ENOG4105CIM	Catalyzes the attachment of alanine to tRNA(Ala) in a two-step reaction alanine is first activated by ATP to form Ala- AMP and then transferred to the acceptor end of tRNA(Ala). Also edits incorrectly charged Ser-tRNA(Ala) and Gly-tRNA(Ala) via its editing domain (By similarity)	Translation, ribosomal structure and biogenesis
1748	931626.Awo_c09780	ENOG4105C3J	Cysteine desulfurase	Amino acid transport and metabolism
1750	903814.ELI_2506	ENOG4105CAV	thus facilitating recognition of the initiation point. It is needed to translate mRNA with a short Shine-Dalgarno (SD) purine-rich sequence (By similarity)	Translation, ribosomal structure and biogenesis
1763	931626.Awo_c09640	ENOG4105CDK	Essential cell division protein that forms a contractile ring structure (Z ring) at the future cell division site. The regulation of the ring assembly controls the timing and the location of cell division. One of the functions of the FtsZ ring is to recruit	Cell cycle control, cell division, chromosome partitioning

			other cell division proteins to the septum to produce a new cell wall between the dividing cells. Binds GTP and shows GTPase activity (By similarity)	
1782	448385.sce0244	ENOG4108KD8	secreted protein	Inorganic ion transport and metabolism
1846	931626.Awo_c14060	ENOG4108UGT	Methyl-accepting chemotaxis	Cell motility, Signal transduction mechanisms
1849	931626.Awo_c02850	ENOG4105CQI	Catalyzes the isomerization between 2-isopropylmalate and 3-isopropylmalate, via the formation of 2-isopropylmaleate (By similarity)	Amino acid transport and metabolism
1870	272563.CD0728	ENOG4105E8X	CO dehydrogenase acetyl-CoA synthase complex beta subunit	Energy production and conversion
1976	931626.Awo_c30830	ENOG4105PGF	domain containing) protein	Energy production and conversion
1977	931626.Awo_c30840	ENOG4105DQB	NA	Function unknown
2074	203119.Cthe_3158	ENOG4108I0Z	Aconitate hydratase	Energy production and conversion
2233	931626.Awo_c03270	ENOG4105DZS	Phenylacetate-CoA ligase	Secondary metabolites biosynthesis, transport, and catabolism
2236	931626.Awo_c03240	ENOG4105CUP	Catalyzes the synthesis of alpha-ribazole-5'-phosphate from nicotinate mononucleotide (NAMN) and 5,6-dimethylbenzimidazole (DMB) (By similarity)	Coenzyme transport and metabolism
2241	931626.Awo_c03170	ENOG4105D01	(LipO)protein	Inorganic ion transport and metabolism
2252	931626.Awo_c03000	ENOG4105CHT	Allows the formation of correctly charged Asn-tRNA(Asn) or Gln-tRNA(Gln) through the transamidation of misacylated Asp- tRNA(Asn) or Glu-tRNA(Gln) in organisms which lack either or both of asparaginyl-tRNA or glutaminyl-tRNA synthetases. The reaction takes place in the presence of glutamine and ATP through an activated phospho-Asp-tRNA(Asn) or phospho-Glu-tRNA(Gln) (By similarity)	Translation, ribosomal structure and biogenesis
2253	931626.Awo_c02990	ENOG4105C3P	Allows the formation of correctly charged Gln-tRNA(Gln) through the transamidation of misacylated Glu-tRNA(Gln) in organisms which lack glutaminyl-tRNA synthetase. The reaction takes place in the presence of glutamine and ATP through an activated gamma-phospho-Glu-tRNA(Gln) (By similarity)	Translation, ribosomal structure and biogenesis
2262	931626.Awo_c02900	ENOG4105CNW	domain protein	Energy production and conversion
2268	931626.Awo_c02820	ENOG4105CFH	Aspartokinase	Amino acid transport and metabolism
2315	931626.Awo_c09010	ENOG4105CMT	Dihydroxyacetone kinase	Carbohydrate transport and metabolism
2321	931626.Awo_c09070	ENOG4106ADB	ABC transporter substrate-binding protein	Amino acid transport and metabolism
2324	1000569.HMPREF1040_1416	ENOG4108IIJ	DNA-dependent RNA polymerase catalyzes the transcription of DNA into RNA using the four ribonucleoside triphosphates as substrates (By similarity)	Transcription
2325	1000569.HMPREF1040_1417	ENOG4105D27	DNA-dependent RNA polymerase catalyzes the transcription of DNA into RNA using the four ribonucleoside triphosphates as substrates (By similarity)	Transcription
2327	931626.Awo_c09130	ENOG4108UKE	Interacts with and stabilizes bases of the 16S rRNA that are involved in tRNA selection in the A site and with the mRNA backbone. Located at the interface of the 30S and 50S subunits, it traverses the body of the 30S subunit contacting proteins	Translation, ribosomal structure and biogenesis

			on the other side and probably holding the rRNA structure together. The combined cluster of proteins S8, S12 and S17 appears to hold together the shoulder and platform of the 30S subunit (By similarity)	
2328	931626.Awo_c09140	ENOG4108UHY	One of the primary rRNA binding proteins, it binds directly to 16S rRNA where it nucleates assembly of the head domain of the 30S subunit. Is located at the subunit interface close to the decoding center, probably blocks exit of the E-site tRNA (By similarity)	Translation, ribosomal structure and biogenesis
2329	1000569.HMPREF1040_1421	ENOG4105CEJ	Catalyzes the GTP-dependent ribosomal translocation step during translation elongation. During this step, the ribosome changes from the pre-translocational (PRE) to the post- translocational (POST) state as the newly formed A-site-bound peptidyl-tRNA and P-site-bound deacylated tRNA move to the P and E sites, respectively. Catalyzes the coordinated movement of the two tRNA molecules, the mRNA and conformational changes in the ribosome (By similarity)	Translation, ribosomal structure and biogenesis
2509	338966.Ppro_3520	ENOG4108HIZ	formate dehydrogenase alpha subunit	Energy production and conversion
2526	931626.Awo_c09160	ENOG4105CGV	This protein promotes the GTP-dependent binding of aminoacyl-tRNA to the A-site of ribosomes during protein biosynthesis (By similarity)	Translation, ribosomal structure and biogenesis
2528	931626.Awo_c28640	ENOG4108Z10	Involved in the binding of tRNA to the ribosomes (By similarity)	Translation, ribosomal structure and biogenesis
2529	931626.Awo_c28630	ENOG4105EEE	One of the primary rRNA binding proteins, it binds directly near the 3'-end of the 23S rRNA, where it nucleates assembly of the 50S subunit (By similarity)	Translation, ribosomal structure and biogenesis
2530	931626.Awo_c28620	ENOG4106U5A	One of the primary rRNA binding proteins, this protein initially binds near the 5'-end of the 23S rRNA. It is important during the early stages of 50S assembly. It makes multiple contacts with different domains of the 23S rRNA in the assembled 50S subunit and ribosome (By similarity)	Translation, ribosomal structure and biogenesis
2531	931626.Awo_c28610	ENOG41080UE	One of the early assembly proteins it binds 23S rRNA. One of the proteins that surrounds the polypeptide exit tunnel on the outside of the ribosome. Forms the main docking site for trigger factor binding to the ribosome (By similarity)	Translation, ribosomal structure and biogenesis
2532	903814.ELI_4032	ENOG4105CFD	One of the primary rRNA binding proteins. Required for association of the 30S and 50S subunits to form the 70S ribosome, for tRNA binding and peptide bond formation. It has been suggested to have peptidyltransferase activity	Translation, ribosomal structure and biogenesis
2533	903814.ELI_4031	ENOG4105K7S	Protein S19 forms a complex with S13 that binds strongly to the 16S ribosomal RNA (By similarity)	Translation, ribosomal structure and biogenesis
2534	903814.ELI_4030	ENOG4105KAP	The globular domain of the protein is located near the polypeptide exit tunnel on the outside of the subunit, while an extended beta-hairpin is found that lines the wall of the exit tunnel in the center of the 70S ribosome (By similarity)	Translation, ribosomal structure and biogenesis
2535	931626.Awo_c28580	ENOG4105CKE	Binds the lower part of the 30S subunit head. Binds mRNA in the 70S ribosome, positioning it for translation (By similarity)	Translation, ribosomal structure and biogenesis
2536	931626.Awo_c28570	ENOG4108R70	Binds 23S rRNA and is also seen to make contacts with the A and possibly P site tRNAs (By similarity)	Translation, ribosomal structure and biogenesis
2537	931626.Awo_c28560	ENOG41082SA	50s ribosomal protein l29	Translation, ribosomal structure and biogenesis
2538	931626.Awo_c28550	ENOG4105K87	One of the primary rRNA binding proteins, it binds specifically to the 5'-end of 16S ribosomal	Translation, ribosomal structure and biogenesis
2539	931626.Awo_c28540	ENOG4108UNN	Binds to 23S rRNA. Forms part of two intersubunit bridges in the 70S ribosome (By similarity)	Translation, ribosomal structure and biogenesis
2540	931626.Awo_c28530	ENOG4105KAR	One of the proteins that surrounds the polypeptide exit tunnel on the outside of the subunit (By similarity)	Translation, ribosomal structure and biogenesis
2541	931626.Awo_c28520	ENOG4105CW6	This is 1 of the proteins that binds and probably mediates the attachment of the 5S RNA into the large ribosomal subunit, where it forms part of the central protuberance. In the 70S ribosome it contacts protein S13 of the 30S subunit (bridge B1b), connecting the 2 subunits	Translation, ribosomal structure and biogenesis
2543	931626.Awo_c2	ENOG4108UJY	One of the primary rRNA binding proteins, it binds directly to 16S rRNA central domain where it helps coordinate	Translation, ribosomal structure

	8500		assembly of the platform of the 30S subunit (By similarity)	and biogenesis
2544	931626.Awo_c28490	ENOG4108R5J	This protein binds to the 23S rRNA, and is important in its secondary structure. It is located near the subunit interface in the base of the L7 L12 stalk, and near the tRNA binding site of the peptidyltransferase center (By similarity)	Translation, ribosomal structure and biogenesis
2545	931626.Awo_c28480	ENOG4105K4C	This is one of the proteins that binds and probably mediates the attachment of the 5S RNA into the large ribosomal subunit, where it forms part of the central protuberance (By similarity)	Translation, ribosomal structure and biogenesis
2546	931626.Awo_c28470	ENOG4108RA9	Located at the back of the 30S subunit body where it stabilizes the conformation of the head with respect to the body (By similarity)	Translation, ribosomal structure and biogenesis
2547	931626.Awo_c28460	ENOG41084RK	50S ribosomal protein L30	Translation, ribosomal structure and biogenesis
2548	931626.Awo_c28450	ENOG4108UZ0	Binds to the 23S rRNA (By similarity)	Translation, ribosomal structure and biogenesis
2550	931626.Awo_c28430	ENOG4105CC8	Catalyzes the reversible transfer of the terminal phosphate group between ATP and AMP. Plays an important role in cellular energy homeostasis and in adenine nucleotide metabolism (By similarity)	Nucleotide transport and metabolism
2555	931626.Awo_c28390	ENOG4108Z04	Located at the top of the head of the 30S subunit, it contacts several helices of the 16S rRNA. In the 70S ribosome it contacts the 23S rRNA (bridge B1a) and protein L5 of the 50S subunit (bridge B1b), connecting the 2 subunits	Translation, ribosomal structure and biogenesis
2556	931626.Awo_c28380	ENOG4108UHH	Located on the platform of the 30S subunit, it bridges several disparate RNA helices of the 16S rRNA. Forms part of the Shine-Dalgarno cleft in the 70S ribosome (By similarity)	Translation, ribosomal structure and biogenesis
2557	931626.Awo_c28370	ENOG4105G6W	One of the primary rRNA binding proteins, it binds directly to 16S rRNA where it nucleates assembly of the body of the 30S subunit (By similarity)	Translation, ribosomal structure and biogenesis
2558	931626.Awo_c28360	ENOG4105CTF	DNA-dependent RNA polymerase catalyzes the transcription of DNA into RNA using the four ribonucleoside triphosphates as substrates (By similarity)	Transcription
2559	931626.Awo_c28350	ENOG4108ZT0	50S ribosomal protein l17	Translation, ribosomal structure and biogenesis
2564	931626.Awo_c28300	ENOG4108UM5	This protein is one of the early assembly proteins of the 50S ribosomal subunit, although it is not seen to bind rRNA by itself. It is important during the early stages of 50S assembly (By similarity)	Translation, ribosomal structure and biogenesis
2565	931626.Awo_c28290	ENOG4108UJD	30S ribosomal protein S9	Translation, ribosomal structure and biogenesis
2576	931626.Awo_c28180	ENOG4107QM1	Catalyzes the reversible conversion of 3- phosphohydroxypyruvate to phosphoserine and of 3-hydroxy-2-oxo-4-phosphonooxybutanoate to phosphohydroxythreonine (By similarity)	Amino acid transport and metabolism
2577	665571.STHERM_c01490	ENOG4107RDV	Dehydrogenase	Amino acid transport and metabolism
2578	931626.Awo_c28160	ENOG4105CGQ	conserved protein UCP033563	Function unknown
2583	931626.Awo_c28080	ENOG4105MS2	ribosomal subunit Interface protein	Translation, ribosomal structure and biogenesis
2584	931626.Awo_c28070	ENOG4108ZP4	NA	Function unknown
2585	1009370.ALO_19082	ENOG4105CI6	Part of the Sec protein translocase complex. Interacts with the SecYEG preprotein conducting channel. Has a central role in coupling the hydrolysis of ATP to the transfer of proteins into and across the cell membrane, serving	Intracellular trafficking, secretion, and vesicular transport
2592	931626.Awo_c27990	ENOG4108TCH	flavin reductase domain protein	Function unknown
2630	592015.HMPRE	ENOG4107R0V	Pyruvate dehydrogenase	Energy production and

	F1705_00473			conversion
2631	931626.Awo_c01710	ENOG4105CPP	Dehydrogenase, E1 component	Energy production and conversion
2633	445335.CBN_1771	ENOG4105C7S	Dehydrogenase	Energy production and conversion
2655	903814.ELI_0843	ENOG410902H	NADH dehydrogenase (Ubiquinone), 24 kDa subunit	Energy production and conversion
2657	931626.Awo_c26990	ENOG4105M1A	NADP-reducing hydrogenase, subunit B	Energy production and conversion
2658	203119.Cthe_0341	ENOG4107QIZ	NADH dehydrogenase	Energy production and conversion
2659	903814.ELI_0847	ENOG4107QHI	-hydrogenase	Energy production and conversion
2673	931626.Awo_c26820	ENOG4108SIQ	Methyl-accepting chemotaxis	Transcription, Signal transduction mechanisms
2676	886872.LargK3_010100007289	ENOG4106TRD	domain protein	Function unknown
2689	931626.Awo_c10670	ENOG4107FEI	Cobyrinic acid ac-diamide synthase	Cell cycle control, cell division, chromosome partitioning
2690	903814.ELI_3608	ENOG4107QT9	Ferredoxin	Energy production and conversion
2691	931626.Awo_c10690	ENOG4108Z7N	NA	Function unknown
2693	931626.Awo_c10710	ENOG4106HYN	CO dehydrogenase acetyl-CoA synthase delta subunit	Energy production and conversion
2694	931626.Awo_c10720	ENOG4105EYX	Acetyl-CoA decarbonylase synthase complex subunit gamma	Energy production and conversion
2695	931626.Awo_c10730	ENOG41073UA	Dihydropteroate synthase, DHPS	Amino acid transport and metabolism
2696	903814.ELI_3603	ENOG4105CJG	carbon-monoxide dehydrogenase catalytic subunit	Energy production and conversion
2697	931626.Awo_c10750	ENOG41080XS	Cobyrinic acid ac-diamide synthase	Cell cycle control, cell division, chromosome partitioning
2698	272563.CD0728	ENOG4105E8X	CO dehydrogenase acetyl-CoA synthase complex beta subunit	Energy production and conversion
2706	931626.Awo_c10820	ENOG4105WDJ	50S ribosomal protein L33	Translation, ribosomal structure and biogenesis
2709	931626.Awo_c10850	ENOG4108UIK	This protein binds directly to 23S ribosomal RNA (By similarity)	Translation, ribosomal structure and biogenesis
2710	931626.Awo_c10860	ENOG4105C64	Binds directly to 23S rRNA. The L1 stalk is quite mobile in the ribosome, and is involved in E site tRNA release (By similarity)	Translation, ribosomal structure and biogenesis
2711	931626.Awo_c10870	ENOG4108VZM	50s ribosomal protein L10	Translation, ribosomal structure and biogenesis

2712	931626.Awo_c10880	ENOG4105KBC	Seems to be the binding site for several of the factors involved in protein synthesis and appears to be essential for accurate translation (By similarity)	Translation, ribosomal structure and biogenesis
2713	699184.BCSJ1_20403	ENOG41053TK	Catalyzes the attachment of proline to tRNA(Pro) in a two-step reaction proline is first activated by ATP to form Pro- AMP and then transferred to the acceptor end of tRNA(Pro) (By similarity)	Translation, ribosomal structure and biogenesis
2734	931626.Awo_c11070	ENOG4105PBX	(rubredoxin-type Fe(Cys)4 protein)	Function unknown
3187	213810.RUM_13040	ENOG4105C8T	Leucyl-tRNA synthetase	Translation, ribosomal structure and biogenesis
3189	931626.Awo_c16770	ENOG4105DG4	malate L-lactate dehydrogenase	Energy production and conversion
3193	931626.Awo_c16730	ENOG4108V0P	LemA family	Function unknown
3210	931626.Awo_c16610	ENOG4105KQV	Binds directly to 16S ribosomal RNA (By similarity)	Translation, ribosomal structure and biogenesis
3216	931626.Awo_c16540	ENOG4107K2P	HTH_XRE	Transcription
3220	903814.ELI_1042	ENOG4105CJ9	Prevents misfolding and promotes the refolding and proper assembly of unfolded polypeptides generated under stress conditions (By similarity)	Post-translational modification, protein turnover, and chaperones
3221	931626.Awo_c16470	ENOG4105K5Y	Binds to Cpn60 in the presence of Mg-ATP and suppresses the ATPase activity of the latter (By similarity)	Post-translational modification, protein turnover, and chaperones
3233	696747.NIES39_A02060	ENOG4107QPR	SAICAR synthetase	Nucleotide transport and metabolism
3242	931626.Awo_c14060	ENOG4108UGT	Methyl-accepting chemotaxis	Cell motility, Signal transduction mechanisms
3243	931626.Awo_c23520	ENOG4108UGT	Methyl-accepting chemotaxis	Cell motility, Signal transduction mechanisms
3472	931626.Awo_c14060	ENOG4108UGT	Methyl-accepting chemotaxis	Cell motility, Signal transduction mechanisms
3593	931626.Awo_c17270	ENOG4105KW6	This is one of the proteins that binds to the 5S RNA in the ribosome where it forms part of the central protuberance (By similarity)	Translation, ribosomal structure and biogenesis
3626	887929.HMP0721_1837	ENOG4105CKH	Is required not only for elongation of protein synthesis but also for the initiation of all mRNA translation through initiator tRNA(fMet) aminoacylation (By similarity)	Translation, ribosomal structure and biogenesis
3628	431943.CKL_0086	ENOG4107QYN	fad dependent oxidoreductase	Function unknown
3630	931626.Awo_c35470	ENOG4105K9Q	Binds to the 23S rRNA (By similarity)	Translation, ribosomal structure and biogenesis
3633	903814.ELI_4585	ENOG4105VH8	Binds as a heterodimer with protein S6 to the central domain of the 16S rRNA, where it helps stabilize the platform of the 30S subunit (By similarity)	Translation, ribosomal structure and biogenesis
3634	931626.Awo_c35500	ENOG4108UUM	single-stranded DNA-binding protein	Replication, recombination and repair
3635	931626.Awo_c35510	ENOG4108ZDX	Binds together with S18 to 16S ribosomal RNA (By similarity)	Translation, ribosomal structure and biogenesis

3646	931626.Awo_c35610	ENOG4108Z7S	Single-stranded nucleic acid binding R3H domain-containing protein	Function unknown
3647	931626.Awo_c35620	ENOG4105DHW	Required for the insertion and or proper folding and or complex formation of integral membrane proteins into the membrane. Involved in integration of membrane proteins that insert both dependently and independently of the Sec translocase complex, as well as at least some lipoproteins	Intracellular trafficking, secretion, and vesicular transport
3654	887929.HMP0721_1866	ENOG4105C7D	DNA gyrase negatively supercoils closed circular double- stranded DNA in an ATP-dependent manner and also catalyzes the interconversion of other topological isomers of double-stranded DNA rings, including catenanes and knotted rings (By similarity)	Replication, recombination and repair
3655	1000569.HMPREF1040_1159	ENOG4105C24	DNA gyrase negatively supercoils closed circular double- stranded DNA in an ATP-dependent manner and also catalyzes the interconversion of other topological isomers of double-stranded DNA rings, including catenanes and knotted rings (By similarity)	Replication, recombination and repair
3700	931626.Awo_c00610	ENOG4105KGY	integral membrane protein	Function unknown
3970	903814.ELI_3133	ENOG4108U1D	NA	Function unknown
4213	931626.Awo_c06010	ENOG4105EVN	Prephenate dehydrogenase	Amino acid transport and metabolism
4214	931626.Awo_c06000	ENOG4108JPM	phospho-2-dehydro-3-deoxyheptonate aldolase	Amino acid transport and metabolism
4332	742767.HMPREF9456_01091	ENOG410569P	Involved in the synthesis of meso-diaminopimelate (m-DAP or DL-DAP), required for both lysine and peptidoglycan biosynthesis. Catalyzes the direct conversion of tetrahydrodipicolinate to LL-diaminopimelate, a reaction that requires three enzymes in E.coli (By similarity)	Amino acid transport and metabolism
4333	931626.Awo_c17640	ENOG4105CJX	Molecular chaperone. Has ATPase activity (By similarity)	Post-translational modification, protein turnover, and chaperones
4336	279714.FuraDRAFT_2302	ENOG4108ESE	uses NADH to detoxify nitric oxide (NO), protecting several 4Fe-4S NO-sensitive enzymes. Has at least 2 reductase partners, only one of which (NorW, flavorubredoxin reductase) has been identified. NO probably binds to the di-iron center	Energy production and conversion
4337	545694.TREPR_1703	ENOG4108DFG	Nitrite and sulphite reductase 4Fe-4S domain	Energy production and conversion
4339	931626.Awo_c17680	ENOG4105D5B	Sulfite reductase, subunit A	Energy production and conversion
4340	931626.Awo_c17690	ENOG4105EJ7	Catalyzes the reduction of hydroxylamine to form NH(3) and H(2)O (By similarity)	Energy production and conversion
4383	641491.DND132_2395	ENOG4105WE2	aldehyde ferredoxin oxidoreductase	Energy production and conversion
4397	588581.Cpap_1886	ENOG4105CK7	Sodium proline symporter	Amino acid transport and metabolism
4405	931626.Awo_c20530	ENOG4105J51	Nitroreductase	Function unknown
4431	931626.Awo_c20330	ENOG4105CMT	Dihydroxyacetone kinase	Carbohydrate transport and metabolism
4514	931626.Awo_c04630	ENOG4105CPA	NA	Function unknown
4618	428125.CLOLE	ENOG4106TRD	domain protein	Function unknown

	P_02977			
4697	411903.COLAER_01397	ENOG410141K	Glutamate synthase central domain	Amino acid transport and metabolism
4714	931626.Awo_c07580	ENOG4105EYY	glutamate synthase	Amino acid transport and metabolism
4721	931626.Awo_c07630	ENOG4108JEB	The beta subunit is responsible for the synthesis of L- tryptophan from indole and L-serine (By similarity)	Amino acid transport and metabolism
4758	931626.Awo_c09260	ENOG4105CKU	formyltetrahydrofolate synthetase	Nucleotide transport and metabolism
4759	931626.Awo_c26710	ENOG4108AV3	Inherit from NOG: hydrolase family 43	Carbohydrate transport and metabolism
4760	398767.Glov_1486	ENOG4105F21	two component regulator three y	Function unknown
4787	931626.Awo_c29580	ENOG4105CM0	Catalyzes the synthesis of GMP from XMP (By similarity)	Nucleotide transport and metabolism
4788	931626.Awo_c29590	ENOG4105CP4	Catalyzes the conversion of inosine 5'-phosphate (IMP) to xanthosine 5'-phosphate (XMP), the first committed and rate-limiting step in the de novo synthesis of guanine nucleotides, and therefore plays an important role in the regulation of cell growth (By similarity)	Nucleotide transport and metabolism
4800	931626.Awo_c29710	ENOG4105C1J	atp-dependent rna helicase	Replication, recombination and repair
4801	931626.Awo_c29720	ENOG4107K2P	HTH_XRE	Transcription
4803	903814.ELI_4101	ENOG4105C8D	Catalyzes the ATP-dependent amination of UTP to CTP with either L-glutamine or ammonia as the source of nitrogen (By similarity)	Nucleotide transport and metabolism
4811	887929.HMP0721_1149	ENOG4105C3H	Acts as a processive, ATP-dependent zinc metallopeptidase for both cytoplasmic and membrane proteins. Plays a role in the quality control of integral membrane proteins (By similarity)	Post-translational modification, protein turnover, and chaperones
4817	931626.Awo_c29940	ENOG4105K70	DNA-binding protein	Replication, recombination and repair
4820	931626.Awo_c29970	ENOG4107Y2S	PPIC-type PPIASE domain	Post-translational modification, protein turnover, and chaperones
4825	931626.Awo_c30020	ENOG4105K4W	Could be involved in septation (By similarity)	Cell wall/membrane/envelope biogenesis
4902	981383.OTW25_010100005190	ENOG4104YEB	s-layer domain-containing protein	Function unknown
5052	931626.Awo_c18790	ENOG4105CDP	Catalyzes the condensation of (S)-aspartate-beta- semialdehyde (S)-ASA and pyruvate to 4-hydroxy- tetrahydrodipicolinate (HTPA) (By similarity)	Amino acid transport and metabolism
5053	931626.Awo_c18800	ENOG4105CM3	Catalyzes the NADPH-dependent formation of L-aspartate- semialdehyde (L-ASA) by the reductive dephosphorylation of L- aspartyl-4-phosphate (By similarity)	Amino acid transport and metabolism
5091	1033810.HLPCO_07489	ENOG4107QMW	pyridine nucleotide-disulfide oxidoreductase	Inorganic ion transport and metabolism
5139	903814.ELI_216	ENOG4108HW3	(ABC) transporter	Amino acid transport and

	0			metabolism, Signal transduction mechanisms
5143	931626.Awo_c19380	ENOG4105C4N	this subunit has chaperone activity. The binding of ATP and its subsequent hydrolysis by HslU are essential for unfolding of protein substrates subsequently hydrolyzed by HslV. HslU recognizes the N-terminal part of its protein substrates and unfolds these before they are guided to HslV for hydrolysis (By similarity)	Post-translational modification, protein turnover, and chaperones
5144	931626.Awo_c19390	ENOG4105CE9	30S ribosomal protein S2	Translation, ribosomal structure and biogenesis
5145	931626.Awo_c19400	ENOG4105CU7	Associates with the EF-Tu.GDP complex and induces the exchange of GDP to GTP. It remains bound to the aminoacyl-tRNA.EF- Tu.GTP complex up to the GTP hydrolysis stage on the ribosome (By similarity)	Translation, ribosomal structure and biogenesis
5147	931626.Awo_c19420	ENOG4108VCV	Responsible for the release of ribosomes from messenger RNA at the termination of protein biosynthesis. May increase the efficiency of translation by recycling ribosomes from one round of translation to another (By similarity)	Translation, ribosomal structure and biogenesis
5149	931626.Awo_c19440	ENOG4105CHV	Transcription elongation factor NusA	Transcription
5151	887929.HMP0721_0633	ENOG4107EEM	One of the essential components for the initiation of protein synthesis. Protects formylmethionyl-tRNA from spontaneous hydrolysis and promotes its binding to the 30S ribosomal subunits. Also involved in the hydrolysis of GTP during the formation of the 70S ribosomal complex (By similarity)	Translation, ribosomal structure and biogenesis
5156	931626.Awo_c19500	ENOG4105K77	Forms an intersubunit bridge (bridge B4) with the 23S rRNA of the 50S subunit in the ribosome (By similarity)	Translation, ribosomal structure and biogenesis
5157	293826.Amet_2667	ENOG4105C62	Involved in mRNA degradation. Hydrolyzes single-stranded polyribonucleotides processively in the 3'- to 5'-direction (By similarity)	Translation, ribosomal structure and biogenesis
5163	931626.Awo_c19560	ENOG4105C68	Can catalyze the hydrolysis of ATP in the presence of single-stranded DNA, the ATP-dependent uptake of single-stranded DNA by duplex DNA, and the ATP-dependent hybridization of homologous single-stranded DNAs. It interacts with LexA causing its activation and leading to its autocatalytic cleavage (By similarity)	Replication, recombination and repair
5164	931626.Awo_c19570	ENOG4105E4Y	Endoribonuclease that initiates mRNA decay (By similarity)	Function unknown
5167	931626.Awo_c19610	ENOG4105BZN	citrate synthase	Energy production and conversion
5169	903814.ELI_2052	ENOG4105BZI	indolepyruvate ferredoxin oxidoreductase	Energy production and conversion
5173	1029718.SFBM_1205	ENOG4107RTZ	reductase	Nucleotide transport and metabolism
5177	1000569.HMPREF1040_0355	ENOG4105CA4	amino acids such as threonine, to avoid such errors, it has a posttransfer editing activity that hydrolyzes mischarged Thr-tRNA(Val) in a tRNA-dependent manner (By similarity)	Translation, ribosomal structure and biogenesis
5200	545694.TREPR_0963	ENOG4107YX4	Amino acid ABC transporter substrate-binding protein	Amino acid transport and metabolism
5206	931626.Awo_c19730	ENOG4105DKZ	GTPase that plays an essential role in the late steps of ribosome biogenesis (By similarity)	Nucleotide transport and metabolism
5226	931626.Awo_c19830	ENOG41060KF	Activator of cell division through the inhibition of FtsZ GTPase activity, therefore promoting FtsZ assembly into bundles of protofilaments necessary for the formation of the division Z ring. It is recruited early at mid-cell but it is not essential for cell division (By similarity)	Function unknown
5228	903814.ELI_2213	ENOG4105C6A	phenylalanyl-tRNA synthetase (beta subunit)	Translation, ribosomal structure and biogenesis
5229	931626.Awo_c19860	ENOG4105CSS	phenylalanyl-tRNA synthetase (alpha subunit)	Translation, ribosomal structure and biogenesis

5234	903814.ELI_2218	ENOG4108YZX	Binds directly to 23S ribosomal RNA and is necessary for the in vitro assembly process of the 50S ribosomal subunit. It is not involved in the protein synthesizing functions of that subunit (By similarity)	Translation, ribosomal structure and biogenesis
5236	931626.Awo_c19900	ENOG4108UUX	IF-3 binds to the 30S ribosomal subunit and shifts the equilibrum between 70S ribosomes and their 50S and 30S subunits in favor of the free subunits, thus enhancing the availability of 30S subunits on which protein synthesis initiation begins (By similarity)	Translation, ribosomal structure and biogenesis
5239	931626.Awo_c19930	ENOG4105DAZ	ABC transporter	Function unknown
5240	203119.Cthe_1228	ENOG4105C22	threonyL-tRNA synthetase	Translation, ribosomal structure and biogenesis
5248	931626.Awo_c20020	ENOG4105K46	50S ribosomal protein l27	Translation, ribosomal structure and biogenesis
5250	931626.Awo_c20040	ENOG4105KK9	This protein binds to 23S rRNA in the presence of protein L20 (By similarity)	Translation, ribosomal structure and biogenesis
5252	931626.Awo_c20060	ENOG4105BZG	amino acid carrier protein	Amino acid transport and metabolism
5309	931626.Awo_c23510	ENOG41087YS	Phage shock protein A	Transcription, Signal transduction mechanisms
5310	931626.Awo_c23500	ENOG4105J0G	E3 Ubiquitin ligase	Function unknown
5348	931626.Awo_c15490	ENOG4105Z0S	NA	Function unknown
5357	931626.Awo_c02520	ENOG4108HW3	(ABC) transporter	Amino acid transport and metabolism, Signal transduction mechanisms
5374	697282.Mettu_1785	ENOG4108HHM	Glycosyl transferase, family 2	Cell wall/membrane/envelope biogenesis
5377	190304.FN1667	ENOG4107QJP	DTDP-glucose 4-6-dehydratase	Cell wall/membrane/envelope biogenesis
5379	744872.Spica_0528	ENOG4107QPT	Catalyzes the formation of dTDP-glucose, from dTTP and glucose 1-phosphate, as well as its pyrophosphorolysis (By similarity)	Cell wall/membrane/envelope biogenesis
5385	931626.Awo_c02240	ENOG4105KNM	Produces ATP from ADP in the presence of a proton gradient across the membrane (By similarity)	Energy production and conversion
5386	931626.Awo_c02230	ENOG4105C4J	Produces ATP from ADP in the presence of a proton gradient across the membrane. The catalytic sites are hosted primarily by the beta subunits (By similarity)	Energy production and conversion
5387	931626.Awo_c02220	ENOG4105J80	Produces ATP from ADP in the presence of a proton gradient across the membrane. The gamma chain is believed to be important in regulating ATPase activity and the flow of protons through the CF(0) complex (By similarity)	Energy production and conversion
5388	931626.Awo_c02210	ENOG4105CDG	Produces ATP from ADP in the presence of a proton gradient across the membrane. The alpha chain is a regulatory subunit (By similarity)	Energy production and conversion
5389	931626.Awo_c02200	ENOG4107YZM	F(1)F(0) ATP synthase produces ATP from ADP in the presence of a proton or sodium gradient. F-type ATPases consist of two structural domains, F(1) containing the extramembraneous catalytic core and F(0) containing the membrane proton channel, linked together by a central stalk and a peripheral stalk. During catalysis, ATP synthesis in the catalytic domain of F(1) is coupled via a rotary mechanism of the central stalk subunits to proton translocation (By similarity)	Energy production and conversion
5390	931626.Awo_c02190	ENOG4107ZFM	Component of the F(0) channel, it forms part of the peripheral stalk, linking F(1) to F(0) (By similarity)	Energy production and conversion

5392	457570.Nther_2849	ENOG41085E0	F(1)F(0) ATP synthase produces ATP from ADP in the presence of a proton or sodium gradient. F-type ATPases consist of two structural domains, F(1) containing the extramembraneous catalytic core and F(0) containing the membrane proton channel, linked together by a central stalk and a peripheral stalk. During catalysis, ATP synthesis in the catalytic domain of F(1) is coupled via a rotary mechanism of the central stalk subunits to proton translocation (By similarity)	Energy production and conversion
5411	931626.Awo_c02000	ENOG41087BM	Prokaryotic N-terminal methylation motif	Function unknown
5412	931626.Awo_c02000	ENOG41087BM	Prokaryotic N-terminal methylation motif	Function unknown
5414	931626.Awo_c02000	ENOG41087BM	Prokaryotic N-terminal methylation motif	Function unknown
5418	1000565.METUNv1_00376	ENOG4105CU6	carbamoyl-phosphate synthetase ammonia chain	Nucleotide transport and metabolism
5420	931626.Awo_c01940	ENOG4105C91	Plays an important role in the de novo pathway of purine nucleotide biosynthesis	Nucleotide transport and metabolism
5424	213810.RUM_16590	ENOG4108HMU	glutamine synthetase	Amino acid transport and metabolism
5512	768670.Calni_0256	ENOG4107VRK	Methyl-accepting chemotaxis	Function unknown
5679	931626.Awo_c24310	ENOG4105DG1	Sigma factors are initiation factors that promote the attachment of RNA polymerase to specific initiation sites and are then released (By similarity)	Transcription
5682	1000569.HMPREF1040_0917	ENOG4105D95	Oxidoreductase required for the transfer of electrons from pyruvate to flavodoxin (By similarity)	Energy production and conversion
5687	931626.Awo_c24380	ENOG4105C0C	Catalyzes the oxidation of 3-carboxy-2-hydroxy-4- methylpentanoate (3-isopropylmalate) to 3-carboxy-4-methyl-2-oxopentanoate. The product decarboxylates to 4-methyl-2 oxopentanoate (By similarity)	Amino acid transport and metabolism
5694	1027396.LMOSA_1950	ENOG4104S60	ATP-dependent CLP protease ATP-binding subunit	Post-translational modification, protein turnover, and chaperones
5696	293826.Amet_3574	ENOG4105EAY	pump that utilizes the energy of pyrophosphate hydrolysis as the driving force for	Energy production and conversion
5699	931626.Awo_c24510	ENOG4105CP7	Triose-phosphate isomerase	Carbohydrate transport and metabolism
5700	931626.Awo_c24520	ENOG4105BZA	Phosphoglycerate kinase	Carbohydrate transport and metabolism
5701	931626.Awo_c24530	ENOG4105C17	glyceraldehyde-3-phosphate dehydrogenase	Carbohydrate transport and metabolism
5708	931626.Awo_c24600	ENOG4105CS3	FliG is one of three proteins (FliG, FliN, FliM) that forms the rotor-mounted switch complex (C ring), located at the base of the basal body. This complex interacts with the CheY and CheZ chemotaxis proteins, in addition to contacting components of the motor that determine the direction of flagellar rotation (By similarity)	Cell motility
5717	931626.Awo_c24690	ENOG4105CBS	Chea signal transduction histidine kinase	Signal transduction mechanisms
5731	643648.Slip_1903	ENOG4105I03	Flagellin	Cell motility
5738	589865.DaAHT2_0659	ENOG4108V3W	Chemotaxis protein, CheW	Cell motility, Signal transduction mechanisms

5762	931626.Awo_c25220	ENOG4105D9H	Catalyzes the attachment of glycine to tRNA(Gly) (By similarity)	Translation, ribosomal structure and biogenesis
5763	931626.Awo_c25230	ENOG4105GZW	4fe-4S ferredoxin, iron-sulfur binding domain protein	Energy production and conversion
5764	931626.Awo_c25240	ENOG4105C8K	Peptide chain release factor 1 directs the termination of translation in response to the peptide chain termination codons UAG and UAA (By similarity)	Translation, ribosomal structure and biogenesis
5795	931626.Awo_c25550	ENOG4105NAX	ABC-type nitrate sulfonate bicarbonate transport	Inorganic ion transport and metabolism
5804	887929.HMP0721_2005	ENOG4105CEJ	Catalyzes the GTP-dependent ribosomal translocation step during translation elongation. During this step, the ribosome changes from the pre-translocational (PRE) to the post- translocational (POST) state as the newly formed A-site-bound peptidyl-tRNA and P-site-bound deacylated tRNA move to the P and E sites, respectively. Catalyzes the coordinated movement of the two tRNA molecules, the mRNA and conformational changes in the ribosome (By similarity)	Translation, ribosomal structure and biogenesis
5808	903814.ELI_0800	ENOG41072G5	Bacterial protein of unknown function (DUF948)	Function unknown
5811	931626.Awo_c25710	ENOG4105BZ3	General (non sugar-specific) component of the phosphoenolpyruvate-dependent sugar phosphotransferase system (sugar PTS). This major carbohydrate active-transport system catalyzes the phosphorylation of incoming sugar substrates concomitantly with their translocation across the cell membrane. Enzyme I transfers the phosphoryl group from phosphoenolpyruvate (PEP) to the phosphoryl carrier protein (HPr) (By similarity)	Carbohydrate transport and metabolism
5817	931626.Awo_c25750	ENOG4105GM9	Membrane	Function unknown
5829	931626.Awo_c25870	ENOG4108ZFT	Dehydratase small subunit	Secondary metabolites biosynthesis, transport, and catabolism
5830	931626.Awo_c25880	ENOG4105YCI	dehydratase medium subunit	Secondary metabolites biosynthesis, transport, and catabolism
5831	887929.HMP0721_1963	ENOG4105DMB	glycerol dehydratase	Secondary metabolites biosynthesis, transport, and catabolism
5832	931626.Awo_c25900	ENOG4105EHR	propanediol utilization protein pduB	Amino acid transport and metabolism
5866	931626.Awo_c26090	ENOG41080S6	Alpha-L-glutamate ligase, RimK family	H, Translation, ribosomal structure and biogenesis
5873	1000569.HMPREF1040_0358	ENOG4108JIJ	Catalyzes a 2-step reaction, involving the ATP-dependent carboxylation of the covalently attached biotin in the first step and the transfer of the carboxyl group to pyruvate in the second (By similarity)	Energy production and conversion
5882	931626.Awo_c26250	ENOG4107RMD	Basic membrane lipoprotein	Function unknown
5901	903814.ELI_0913	ENOG4105VCC	30S ribosomal protein S21	Translation, ribosomal structure and biogenesis
5911	931626.Awo_c26620	ENOG4105BZ5	ATP binding to DnaK triggers the release of the substrate protein, thus completing the reaction cycle. Several rounds of ATP-dependent interactions between DnaJ, DnaK and GrpE are required for fully efficient folding. Also involved, together with DnaK and GrpE, in the DNA replication of plasmids through activation of initiation proteins (By similarity)	Post-translational modification, protein turnover, and chaperones
5912	903814.ELI_3543	ENOG4105CFG	Acts as a chaperone (By similarity)	Post-translational modification, protein turnover, and

				chaperones
5919	347256.MHO_2260	ENOG4105DRM	Cell surface protein	Cell motility
5920	931626.Awo_c26710	ENOG4108AV3	Inherit from NOG: hydrolase family 43	Carbohydrate transport and metabolism
5921	931626.Awo_c26710	ENOG4108AV3	Inherit from NOG: hydrolase family 43	Carbohydrate transport and metabolism
5929	195103.CPF_0552	ENOG41060VT	NA	Function unknown
5958	931626.Awo_c12770	ENOG4108HMR	Key enzyme in the regulation of glycerol uptake and metabolism (By similarity)	Energy production and conversion
5961	931626.Awo_c12800	ENOG4105CA9	Pyruvate kinase	Carbohydrate transport and metabolism
5977	931626.Awo_c12950	ENOG4105D8I	Catalyzes the ATP-dependent transfer of a sulfur to tRNA to produce 4-thiouridine in position 8 of tRNAs, which functions as a near-UV photosensor. Also catalyzes the transfer of sulfur to the sulfur carrier protein ThiS, forming ThiS-thiocarboxylate. This is a step in the synthesis of thiazole, in the thiamine biosynthesis pathway. The sulfur is donated as persulfide by IscS (By similarity)	Coenzyme transport and metabolism
5998	545693.BMQ_5171	ENOG4105VFB	Binds the 23S rRNA (By similarity)	Translation, ribosomal structure and biogenesis
6000	931626.Awo_c13080	ENOG4105DRH	Involved in peptide bond synthesis. Stimulates efficient translation and peptide-bond synthesis on native or reconstituted 70S ribosomes in vitro. Probably functions indirectly by altering the affinity of the ribosome for aminoacyl-tRNA, thus increasing their reactivity as acceptors for peptidyl transferase (By similarity)	Translation, ribosomal structure and biogenesis
6012	931626.Awo_c13200	ENOG4108VPX	Involved in 1,2-propanediol (1,2-PD) degradation by catalyzing the conversion of propanoyl-CoA to propanoyl-phosphate (By similarity)	Secondary metabolites biosynthesis, transport, and catabolism
6019	931626.Awo_c13540	ENOG4105CDH	Citrulline--aspartate ligase	Amino acid transport and metabolism
6036	903814.ELI_2160	ENOG4108HW3	(ABC) transporter	Amino acid transport and metabolism, Signal transduction mechanisms
6043	1037409.BJ6T_72810	ENOG4101SU2	pyruvate phosphate dikinase	Carbohydrate transport and metabolism
6062	931626.Awo_c13930	ENOG4105C6M	Alpha-keto-beta-hydroxylacyl reductoisomerase	Amino acid transport and metabolism
6063	931626.Awo_c13940	ENOG4105CYQ	Catalyzes the condensation of the acetyl group of acetyl-CoA with 3-methyl-2-oxobutanoate (2-oxoisovalerate) to form 3-carboxy-3-hydroxy-4-methylpentanoate (2-isopropylmalate) (By similarity)	Amino acid transport and metabolism
6065	931626.Awo_c13980	ENOG4105C0C	Catalyzes the oxidation of 3-carboxy-2-hydroxy-4- methylpentanoate (3-isopropylmalate) to 3-carboxy-4-methyl-2-oxopentanoate. The product decarboxylates to 4-methyl-2 oxopentanoate (By similarity)	Amino acid transport and metabolism
6083	903814.ELI_1211	ENOG4105C8W	ABC transporter	Inorganic ion transport and metabolism
6085	931626.Awo_c14160	ENOG4107QXD	PAS PAC sensor protein	Energy production and conversion
6099	931626.Awo_c14260	ENOG4105C65	Catalyzes the reversible interconversion of serine and glycine with tetrahydrofolate (THF) serving as the one-carbon carrier. This reaction serves as the major source of one-carbon groups required for the biosynthesis of purines, thymidylate,	Amino acid transport and metabolism

			methionine, and other important biomolecules. Also exhibits THF- independent aldolase activity toward beta-hydroxyamino acids, producing glycine and aldehydes, via a retro-aldol mechanism (By similarity)	
6115	931626.Awo_c14490	ENOG4105DZ1	carboxypeptidase	Cell wall/membrane/envelope biogenesis
6128	931626.Awo_c14620	ENOG4105CUP	Catalyzes the synthesis of alpha-ribazole-5'-phosphate from nicotinate mononucleotide (NAMN) and 5,6-dimethylbenzimidazole (DMB) (By similarity)	Coenzyme transport and metabolism
6130	931626.Awo_c14640	ENOG4105CBF	Catalyzes the synthesis of the hydroxymethylpyrimidine phosphate (HMP-P) moiety of thiamine from aminoimidazole ribotide (AIR) in a radical S-adenosyl-L-methionine (SAM)-dependent reaction (By similarity)	Coenzyme transport and metabolism
6132	931626.Awo_c14710	ENOG4105C36	SPFH domain, Band 7 family protein	Post-translational modification, protein turnover, and chaperones
6136	931626.Awo_c14750	ENOG4105MAX	Superoxide reductase	Energy production and conversion
6166	931626.Awo_c15270	ENOG4107T76	Transporter	Carbohydrate transport and metabolism
6246	931626.Awo_c22960	ENOG4105F72	Domain of unknown function (DUF362)	Function unknown
6269	760011.Spico_0879	ENOG4107RWU	extracellular ligand-binding receptor	Amino acid transport and metabolism
6368	498761.HM1_0998	ENOG4105WE2	aldehyde ferredoxin oxidoreductase	Energy production and conversion
6372	931626.Awo_c22270	ENOG4105K8F	Phosphoribosyl-amp cyclohydrolase	Amino acid transport and metabolism
6387	931626.Awo_c22140	ENOG4107XMQ	Peptidyl-prolyl cis-trans isomerase	Post-translational modification, protein turnover, and chaperones
6389	887929.HMP0721_1036	ENOG4105C6P	ATP-dependent serine protease that mediates the selective degradation of mutant and abnormal proteins as well as certain short-lived regulatory proteins. Required for cellular homeostasis and for survival from DNA damage and developmental changes induced by stress. Degrades polypeptides processively to yield small peptide fragments that are 5 to 10 amino acids long. Binds to DNA in a double-stranded, site-specific manner (By similarity)	Post-translational modification, protein turnover, and chaperones
6394	931626.Awo_c22070	ENOG4105DFM	M18 family aminopeptidase	Amino acid transport and metabolism
6395	931626.Awo_c22060	ENOG4107QTR	Required for nitrogen fixation. May be part of a membrane complex functioning as an intermediate in the electron transport to nitrogenase (By similarity)	Energy production and conversion
6397	931626.Awo_c22040	ENOG41080WF	Electron transport complex, RnfABCDGE type, G subunit	Energy production and conversion
6399	931626.Awo_c22020	ENOG4105DGN	Electron transport complex	Energy production and conversion
6400	931626.Awo_c22010	ENOG4107SVV	electron transport complex, RnfABCDGE type, B subunit	Energy production and conversion
6402	931626.Awo_c21980	ENOG4105BZ4	penicillin-binding protein	Cell wall/membrane/envelope biogenesis
6405	931626.Awo_c21950	ENOG4105CP9	extracellular ligand-binding receptor	Amino acid transport and metabolism

6410	931626.Awo_c21900	ENOG4105DEA	Involved in protein export. Acts as a chaperone by maintaining the newly synthesized protein in an open conformation	Post-translational modification, protein turnover, and chaperones
6412	931626.Awo_c21880	ENOG4105CHN	ATP-dependent specificity component of the Clp protease. It directs the protease to specific substrates. Can perform chaperone functions in the absence of ClpP (By similarity)	Post-translational modification, protein turnover, and chaperones
6413	887929.HMP0721_0405	ENOG4105C6P	ATP-dependent serine protease that mediates the selective degradation of mutant and abnormal proteins as well as certain short-lived regulatory proteins. Required for cellular homeostasis and for survival from DNA damage and developmental changes induced by stress. Degrades polypeptides processively to yield small peptide fragments that are 5 to 10 amino acids long. Binds to DNA in a double-stranded, site-specific manner (By similarity)	Post-translational modification, protein turnover, and chaperones
6417	469381.Dpep_1903	ENOG4107QN8	Part of the Sec protein translocase complex. Interacts with the SecYEG preprotein conducting channel. SecDF uses the proton motive force (PMF) to complete protein translocation after the ATP-dependent function of SecA (By similarity)	Intracellular trafficking, secretion, and vesicular transport
6418	931626.Awo_c21820	ENOG4107SBW	Part of the Sec protein translocase complex. Interacts with the SecYEG preprotein conducting channel. SecDF uses the proton motive force (PMF) to complete protein translocation after the ATP-dependent function of SecA (By similarity)	Intracellular trafficking, secretion, and vesicular transport
6425	903814.ELI_3503	ENOG4105C9M	aspartyl-trna synthetase	Translation, ribosomal structure and biogenesis
6444	944560.HMPREF9058_2220	ENOG4101415	serine threonine protein kinase	Signal transduction mechanisms
6447	887929.HMP0721_2082	ENOG4105VB9	50S ribosomal protein l28	Translation, ribosomal structure and biogenesis
6448	931626.Awo_c21530	ENOG4108ZAA	Alkaline-shock protein	Function unknown
6449	931626.Awo_c21520	ENOG4105DP4	dak2 domain fusion protein ylov	Function unknown
6456	931626.Awo_c21450	ENOG4105C7K	acetolactate synthase	Amino acid transport and metabolism
6460	316275.VSAL_II0398	ENOG4108EFR	Formate dehydrogenase Alpha subunit	Energy production and conversion
6471	931626.Awo_c21290	ENOG4108EDY	Chemotaxis sensory transducer	Cell motility, Signal transduction mechanisms
6474	931626.Awo_c21260	ENOG4105C6H	Catalyzes the formation of acetyl phosphate from acetate and ATP. Can also catalyze the reverse reaction (By similarity)	Energy production and conversion
6476	903814.ELI_2891	ENOG41067VY	50S ribosomal protein L32	Translation, ribosomal structure and biogenesis
6484	931626.Awo_c21170	ENOG4105K5M	30s ribosomal protein S16	Translation, ribosomal structure and biogenesis
6488	931626.Awo_c21130	ENOG4108YY1	This protein is located at the 30S-50S ribosomal subunit interface and may play a role in the structure and function of the aminoacyl-tRNA binding site (By similarity)	Translation, ribosomal structure and biogenesis

Chapter 5 Isoprene and methylbutenes in natural environments: Rethinking the role of isoprene in biogeochemical cycles

5.1 Introduction

Previously it was shown that isoprene is used as an electron acceptor by the homoacetogen *Acetobacterium wieringae* ISORED-2 found in anaerobic sludge from a wastewater treatment plant (Kronen *et al.*, 2019). While studying isoprene for its anaerobic biodegradability it appeared that isoprene also inhibited biogenic methane formation. In sludge samples supplied with isoprene/lactate no methane formation occurred and no methanogens were detected in the community analysis, whereas in samples without isoprene methane increased to 400 μmols/microcosms [Section **2.3.1, 2.3.2**] (Kronen *et al.*, 2019). The inhibitory effect of isoprene on methanogenesis has been previously observed, where isoprene partially inhibited methanogenesis in sediment slurries (Schink, 1985). This Chapter describes preliminary experimentation to address whether isoprene reduction and methanogen inhibition have ecological relevance in anaerobic ecosystems containing isoprene.

Terrestrial plants are considered the main source of (atmospheric) isoprene in the isoprene budget (Sharkey *et al.*, 2008; McGenity *et al.*, 2018). But isoprene emission has also been measured in wetlands, mostly subarctic peatlands (Janson and De Serves, 1998; Harley *et al.*, 1999a; Janson *et al.*, 1999; Tiiva *et al.*, 2007; Ekberg *et al.*, 2009; Lindwall *et al.*, 2016) where isoprene was described to be emitted from different mosses (Janson and De Serves, 1998; Janson *et al.*, 1999; Tiiva *et al.*, 2007) or sedge species (Tiiva *et al.*, 2007; Ekberg *et al.*, 2009). However, in these studies it was difficult to determine whether isoprene was coming from the plant components above or below ground or from the soil itself. Many soil bacteria and fungi are known to produce isoprene (Kuzma *et al.*, 1995; Schöller *et al.*, 1997, 2002; Wagner *et al.*, 1999; Fall and Copley, 2000; Bäck *et al.*, 2010; Effmert *et al.*, 2012) but soils are also known sources of other volatile hydrocarbons e.g. ethane, ethene, propane and propene, which are derived from bacteria and fungi or plant roots (Smith and Cook, 1974; Van Cleemput and El-Sebaay, 1986; Peñuelas *et al.*, 2014).

Wetlands are complex ecosystems with exceptional ecological value where key anaerobic processes take place such as denitrification (Jordan *et al.*, 2011), iron and sulfate reduction (Pester *et al.*, 2012), and methanogenesis (Cao *et al.*, 1998; Wuebbles and Hayhoe, 2002). They are also environments with high microbial activity and plant-microbe interactions and they facilitate the simultaneous activity of aerobic as well as anaerobic microbial communities (Bodelier and Dedysh, 2013). These factors point to wetlands as a natural environment where isoprene reduction is likely to occur. In the present Chapter, information is provided on isoprene reduction in anaerobic wetland sediments and on the inhibitory effects of isoprene on methanogenesis in the same wetland sediment samples, and on pure and enriched methanogen cultures.

5.2 Material and Methods

5.2.1 Sample sites

Three wetlands were selected for sample collection in Sydney, NSW, Australia. Two sampling sites were located within in the hanging swamp wetlands on North Head; site one (longitude 151°18'3.06", latitude -33°49'6.668") and site two (longitude 151°18'5.278", latitude -33°48'52.438"). The third sample site (longitude 151°14'10.2"E, latitude 33°53'59.2"S) was in the Lachlan Swamps in Centennial Park, Sydney which contain natural springs that feed the wetland and are inhabited by large numbers of grey-headed flying-foxes (*Pteropus poliocephalus*). Samples were taken in June 2018.

5.2.2 Sample collection

At each site, four individual mud/sediment samples were taken 50-100 cm apart. For each sample 40 ml mud/sediment was collected with a 60 ml plastic syringe with a cut injection tip and injected into a 120 ml glass flask with an O-ring sealing the connection between the bottle and the syringe. Wetland water (80 ml) was added on top and the flask was immediately crimp-sealed with Teflon faced rubber septa. For each site a water-only control was taken. Samples were then transferred to the laboratory where 10 ml of the water phase was removed with a syringe to allow headspace analysis of the samples. Sediment samples (300 µl) were removed from the microcosms with a syringe for community analysis and VFA analysis. The microcosms were stored in the dark at ambient temperature during the course of the experiment.

5.2.3 Triple Quadrupole GC/MS measurements of isoprene and methylbutene

The headspace of the wetland samples was monitored over a period of three months with GC-MS analysis performed on an Agilent 7890A GC interfaced to an Agilent 7000 triple quad mass spectrometer (Agilent Technologies, USA). Isoprene and methylbutenes were separated on a GasPro PLOT column (60 m x 0.32 mm, Agilent Technologies) the oven temperature was 150°C for 30 s and was increased by 20°C min^{-1} to a final temperature of 250°C. The Helium flow was 2.25 ml min^{-1} and the temperature of the transfer line was 290°C. The solvent delay time was 3 min. Isoprene and methylbutene had two transitions per target; isoprene precursor ion 70 m/z

(product ions 55.1 m/z and 29.1 m/z) and methylbutene precursor ion 67 m/z (product ions 41.1 m/z and 39.1 mz) with collision energy of 20 eV. Gas samples (500 µl) were withdrawn from the flask via a pressure-lockable gas tight syringe and directly injected into the GC with a split ratio of 1:10. The syringe was flushed with air to clean it after every injection and a blank air sample was injected between every sample to ensure there was no analyte carryover between samples or standards.

5.2.4 Isoprene, methylbutene and methane quantification with GC-FID

Isoprene and methylbutenes were measured as described in Chapter 2 or (Kronen *et al.*, 2019).

Methane amounts were measured by GC-FID using a GS-Q column (30 m x 0.32 mm × 0.25 mm) at 100 °C for 2 min with helium as the carrier gas (3 ml min^{-1}). Gas samples (100 µl) were withdrawn from the flask via a gas lock syringe and directly injected into a Shimadzu GC-2010.

5.2.5 Preparation of stock solutions

Stock solutions were made in 120 ml glass flasks completely filled with anaerobic minimal media to avoid partitioning of volatile substances into the head space. Isoprene, 2-methyl-1-butene, 3-methyl-1-butene and 2-methyl-2-butene were added as neat liquid with a glass syringe (10 µl) to achieve nominal concentrations of 9 mM and 2 mM. More dilute aqueous solutions (i.e 20 µM nominal) were made in the same way except diluting the appropriate 2 mM stock solution rather than neat liquid. Dimensionless Henry constants for isoprene, methylbutenes, and methane were calculated from Sander (2015).

5.2.6 DNA extraction and Illumina sequencing

DNA was extracted from 100 µl of swamp sediment samples diluted with 200 µl sterile water or 300 µl culture medium in the case of enrichment cultures. DNA extraction was performed as described in Chapter 2 (Kronen *et al.*, 2019). Regions of 16S rRNA genes were amplified by PCR from extracted DNA with the Q5 high-fidelity DNA polymerase (New England BioLabs) using the universal primers 515F (5′-TCGTCGGCAGCGTCAGATGTGTATAAGAGACAG-[GTGYCAGCMGCCGCGGTAA]-3′) and 806R (5′-

GTCTCGTGGGCTCGGAGATGTGTATAAGAGACAG - [GGACTACNVGGGTWTCTAAT-3′) targeting bacteria and archaea (Caporaso *et al.*, 2011). The samples were sequenced on an Illumina MiSeq Sequencer (Illumina, USA) using V4 chemistry at the Next Generation Sequencing Facility at the Hawkesbury Institute for the Environment (Western Sydney University, Sydney, Australia). The 16S rRNA gene amplicon sequences were analysed as described in Chapter 2 (Kronen *et al.*, 2019).

5.2.7 Isoprene and methylbutenes in inhibitory experiments

Culture flasks (120 ml) containing 80 ml of anaerobic minimal media were amended with different isoprene concentrations ranging between 130 nM, 0.013 mM, 0.13 mM and 1.3 mM concentrations in the aquatic phase. Lower concentrations were achieved by adding isoprene from stock solutions made in minimal media of 20 µM, 2 mM or 9 mM. The isomer 2-methyl-1-butene was added from the neat solvent.

5.2.8 Acetate analysis

Wetland sediment samples (300 µl) were centrifuged for 2 min at 10 000 g and 40 µl of supernatant was used for acetate analysis. Acetate was analysed as its ethyl ester derivative by GC-FID using a DB-FFAP column (30 m x 0.32 mm × 0.25 mm, Agilent Technologies) at 40 °C for 6 min with helium as the carrier gas (3 ml min−1). Samples (40 µl) were supplied with 100% ethanol (20 µl) and undiluted sulfuric acid (20 µl) for esterification, sealed immediately and incubated at 60 °C for 45 min. Before injection into the GC, samples were incubated at 80 °C for 5 min at 500 rpm, 250 µl of headspace sample was withdrawn from the flask via an automatic sampler (Shimadzu AOC-5000 plus) and directly injected into a Shimadzu Plus GC-2010 at 500 µl s^{-1}.

5.3 Results

5.3.1 Natural evolution of methylbutenes in wetland samples

Wetland sediment samples were collected from three sites, transferred to microcosms, incubated under anaerobic conditions and the headspace of the samples was monitored by triple quadrupole GC/MS for isoprene and methylbutene formation. No methylbutene or isoprene was detected straight after collection of the samples (time point zero). However, over a period of three months, evolution of methylbutenes was detected in the headspace of samples from each site [**Figure 5.1**]. In wetland sediment samples from site one and two, 2-methyl-1-butene, 2-methyl-2-butene and 3-methyl-1-butene were detected. In samples from site three, 2-methyl-1-butene and 2-methyl-2-butene were detected [**Figure 5.1 ABC**]. The highest total methylbutene amounts measured in headspace samples ranged between 2-4 nmol per microcosm. No methylbutenes could be detected in water only controls from each samples site. Isoprene could not be detected in any sample. .

5.3.2 Addition of isoprene to wetland samples

To see whether isoprene reduction could be the source of the naturally evolved methylbutenes, isoprene (150 µmols/microcosm) was added after 101 days of incubation to all wetland sediment containing microcosms [**Figure 5.1 DEF**]. All microcosms showed isoprene reducing activity. In microcosms from Site 1 and Site 3 isoprene was depleted after 60 days of incubation and methylbutene was produced stoichiometrically in samples from Site 1 [**Figure 5.1 D**]. In samples from Site 3 [**Figure 5.1 F**] only half of the expected methylbutene amounts were measured which could be due to differences in sediment properties e.g. absorbance to organic matter, possible metabolism of methylbutene. Samples from Site 2 [**Figure 5.1 E**] showed little isoprene reduction activity, after 60 days most of the isoprene (75 µmols/microcosm) was still present in the incubation with minor amounts of methylbutene (5 µmols/microcosm) detected.

After exogenous addition of isoprene to all samples from all Sites, 2-methyl-1-butene was the main product of the reduction process.

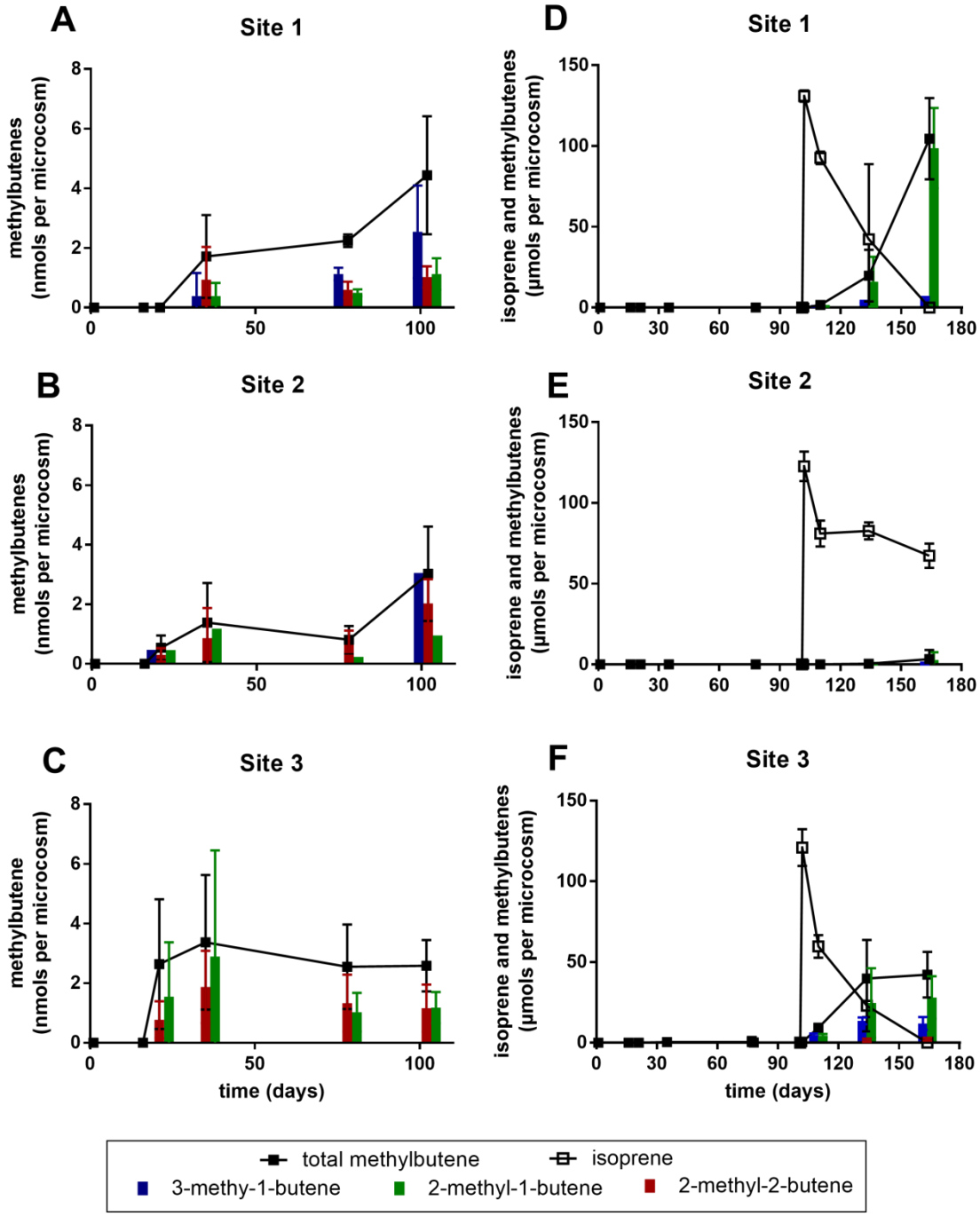

Figure 5.1 Evolution of methylbutenes in microcosms containing wetland samples from three different sites. Naturally evolving methylbutene amounts were measured with highly sensitive TQ GS/MS (**ABC**). After 100 days of incubation 1.3 mM isoprene nominal concentration was added to the wetland sediment samples (**DEF**), isoprene and methylbutene concentrations were measured with GC-FID. Error bars represent one standard deviation (n = 4).

5.3.3 Effect of isoprene addition on methanogenesis, actogenesis and microbial community composition in wetland samples

Wetland samples from the three different sites were monitored for the effect of isoprene on methane and acetate production and on the microbial assemblage. Bacterial and archaeal community analysis was performed using 16S rRNA gene amplicons derived from DNA extracted from wetland samples at different time points before and after isoprene addition and depletion [**Figure 5.2**, **Figure 5.3**, **Figure 5.4**].

In samples from Site 1 [**Figure 5.2**], methane concentration increased from 0 to 500 μmols/microcosm within 100 days and plateaued at day 130. Acetate amounts increased upon addition of isoprene from 13 to 270 μmols/microcosm. Acidobacteria (~10-20%), Firmicutes (~2-10%), Verrucomicrobia (~2-12%) and diverse Proteobacteria (~7-30%) constituted the major phylogenetic groups, with minor representation of Archaea (~0.5-2%), mainly methanogens. Shifts in the structure of the bacterial community were not significant before and after addition of isoprene and only one organism (*Candidatus Koribacter*) significantly decreased in relative abundance between the four time points [Supplement **Table S5 1**].

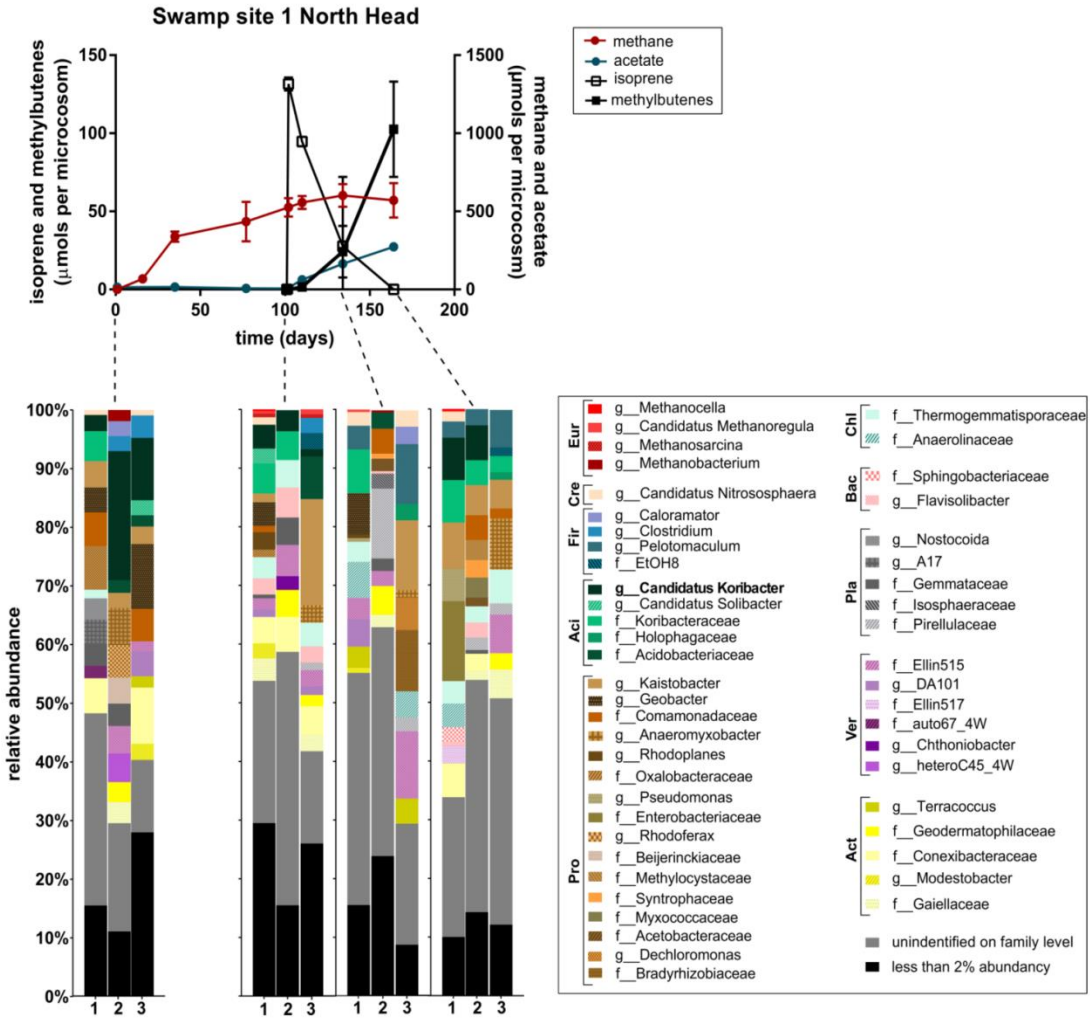

Figure 5.2 Bacterial and archaeal community analysis together with methane and acetate formation in wetland sediment samples collected from sample site one. Samples were incubated for 164 days in anaerobic microcosm and populations classified at Family level and if possible on Genus level were analysed at time points 0, 101, 134 and 164 days. Classifications in the legend are clustered according to their phylum (from top to bottom, left to right); Euryarchaeota, Crenarchaeota, Firmicutes, Acidobacteria, Proteobacteria, Chloroflexi, Bacteroidetes, Planctomycetes, Verrucomicrobia, Actinobacteria. Organism groups whose relative abundance changed significantly between the time points are marked in bold [see Supplement **Table S5 1** for details]. Error bars represent one standard deviation (n = 4). Only three out of four replicates were analysed by Illumina sequencing**.**

In samples from Site 2 [**Figure 5.3**] methane emissions increased from 0 to 400 µmols/microcosm until the addition of isoprene where it plateaued. Acetate was present at t0 in Site 2 microcosms and ranged around 500 µmols/microcosm throughout the incubation period. Similar to Site 1, Acidobacteria (~4-33%), Firmicutes (~9-38%), Verrucomicrobia (~2-15%) and diverse Proteobacteria (~3-28%) were the dominating phylogenetic groups, with moderate representation of Archaea, mainly *Methanobacterium* sp. (~2-12%). The relative abundance within the bacterial community composition changed over time with some bacterial groups *Acidobacteriaceae* (from ~10% to ~22%), *Ethanoligenens* (from 0% to 7-14%), *Ruminococcaceae* (from 0% to 7-18%), *Oxobacter* (from 2% to 9%) becoming more abundant [Supplement **Table S5 1**]. The relative abundance of *Methanobacterium* sp. increased in samples until day 134 (from 0% to 12%) [Supplement **Table S5 1**] but decreased down to 2% again after 164 days of incubation.

Site 3 microcosms [**Figure 5.4**] differed in their microbial community composition compared to Site 1 and 2 in that Firmicutes and Acidobacteria group members were not detected. These samples primarily contain diverse Proteobacteria strains and the relative abundance of two bacterial groups *Syntrophaceae* (from ~3% to ~29%) and *Dechloromonas* (from ~0% to ~25%) increased significantly during the incubation time. Methane emission increased from 0 to 750 µmols/microcosms and after addition of isoprene it increased significantly further from 750 to 1200 µmols/microcosm. Acetate dropped from 50 µmols/microcosm initially to 0 µmols/microcosm by 101 days but increased upon addition isoprene to around 15 µmols/microcosm.

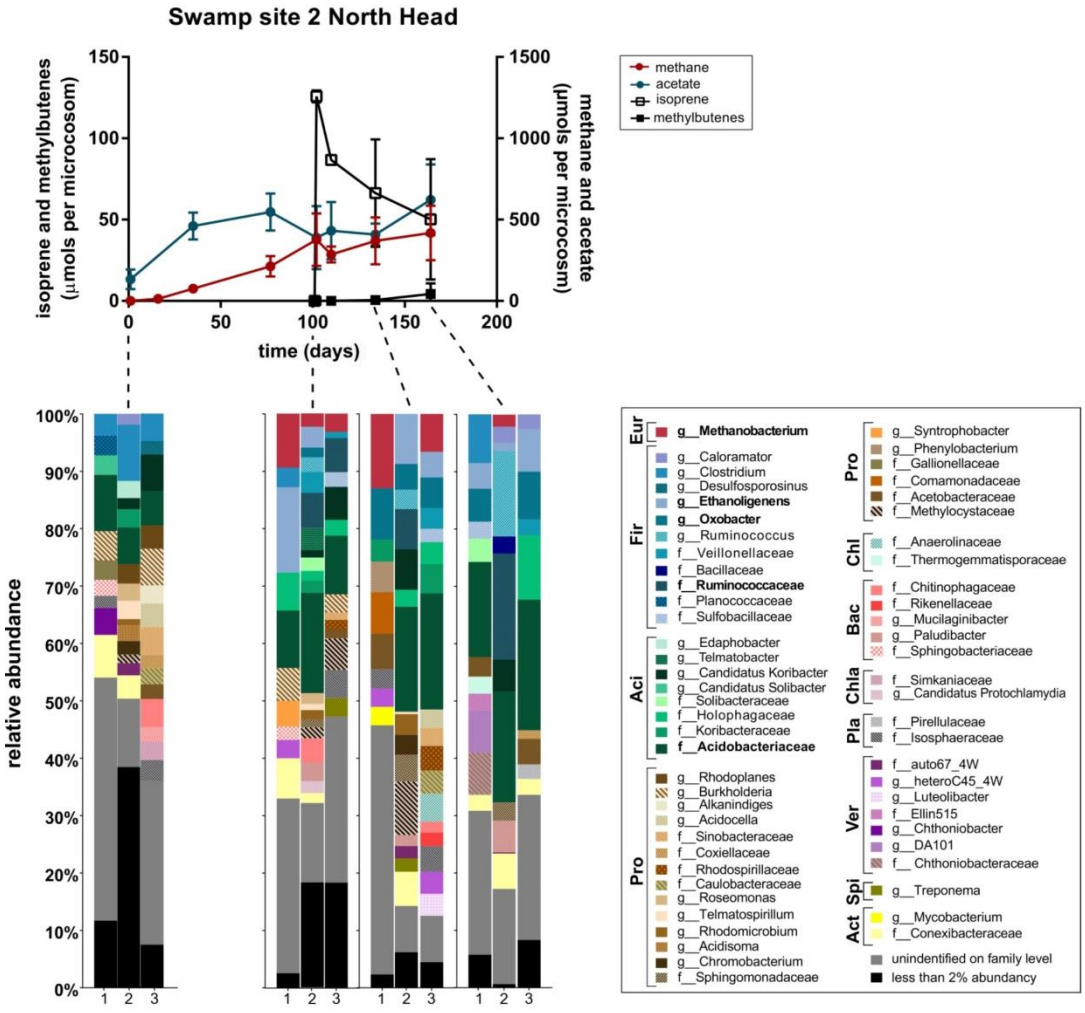

Figure 5.3 Bacterial and archaeal community analysis together with methane and acetate formation in wetland sediment samples collected from sample site two. Samples were incubated for 164 days in anaerobic microcosm and populations classified at Family level and if possible on Genus level were analysed at time points 0, 101, 134 and 164 days. Classifications in the legend are clustered according to their phylum (from top to bottom, left to right); Euryarchaeota, Firmicutes, Acidobacteria, Proteobacteria, Chloroflexi, Bacteroidetes, Chlamydiae, Planctomycetes, Verrucomicrobia, Spirochaetes, Actinobacteria. Organism groups whose relative abundance changed significantly between the time points are marked in bold [see Supplement **Table S5 1** for details]. Error bars represent one standard deviation (n = 4). Only three out of four replicates were analysed by Illumina sequencing.

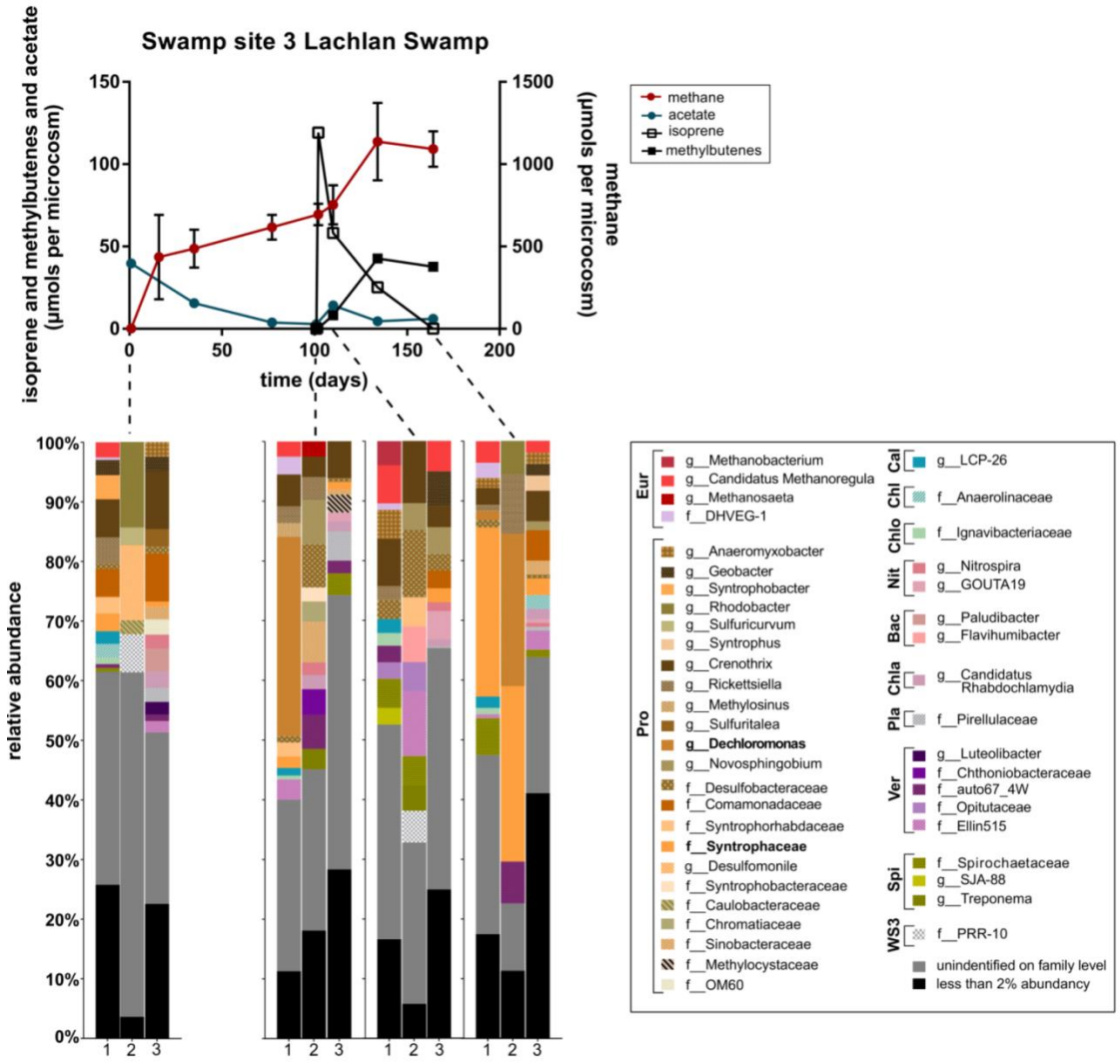

Figure 5.4 Bacterial and archaeal community analysis together with methane and acetate formation in wetland sediment samples collected from sample site three. Samples were incubated for 164 days in anaerobic microcosm and populations classified at Family level and if possible on Genus level were analysed at time points 0, 101, 110 and 164 days. Classifications in the legend are clustered according to their phylum (from top to bottom, left to right); Euryarchaeota, Proteobacteria, Caldithrix, Chloroflexi, Chlorobi, Nitrospirae, Bacteroidetes, Chlamydiae, Planctomycetes, Verrucomicrobia, Spirochaetes, WS3. Organism groups whose relative abundance changed significantly between the time points are marked in bold [see Supplement **Table S5** 1 for details]. Error bars represent one standard deviation (n = 4). Only three out of four replicates were analysed by Illumina sequencing.

5.3.4 Inhibitory effect of different isoprene concentrations on methanogenesis in axenic *Methanosarcina barkeri* and *Methanosarcina mazei* cultures

Isoprene concentrations ranging from 130 nM to 1.3 mM in aqueous phase were tested on axenic *M. barkeri* and *M. mazei* cultures grown in DSMZ media 120 for its inhibitory effect on methanogenesis over 16 days [**Figure 5.5**]. In isoprene free controls methane emission started after 8 days of incubation where as in the presence of 1.3 mM isoprene methane formation was completely inhibited in *M. barkeri* and *M. mazei*. In the presence of 0.13 mM isoprene, 80% (*M. mazei*) and 60% (*M. barkeri*) less methane was formed compared to isoprene free controls.

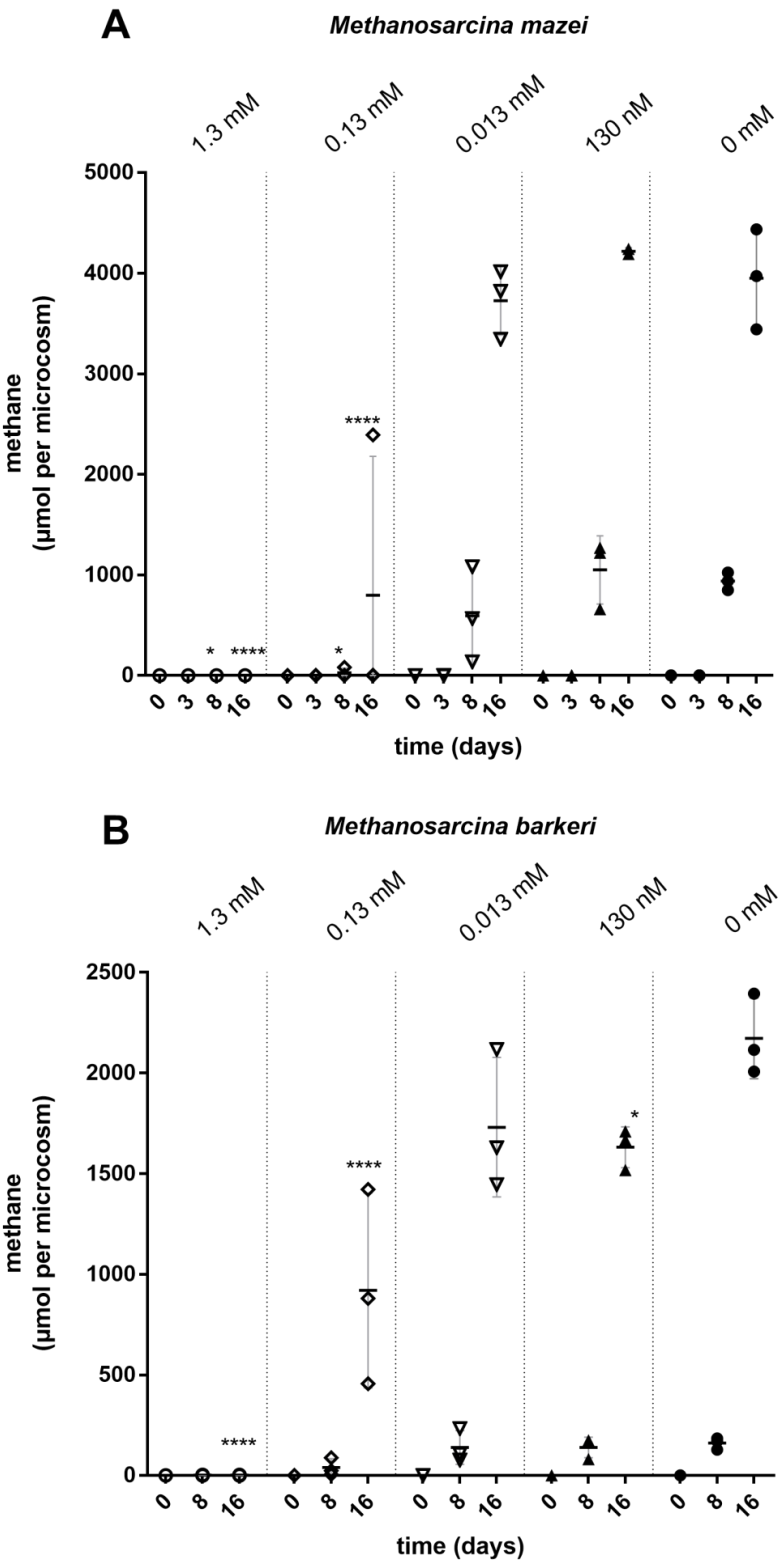

Figure 5.5 Effect of different isoprene concentration on methane formation in pure *M. mazei* (A) and *M. barkeri* (B) cultures grown in DSMZ medium 120. Isoprene concentrations refer to concentration in the aqueous phase. Shown are triplicate measurements for each isoprene concentration and the significance was calculated with a 2way ANOVA multiple comparison between each concentration and the 0 mM isoprene control. (****)=p value <0.0001, (*)=p value<0.05 analysed with 2way ANOVA.

5.3.5 Inhibitory effect of isoprene on methanogenesis in wetland hydrogenotrophic methanogen enrichment culture

Samples obtained from Site 2 were incubated for three months in minimal media supplied with H_2/HCO_3^- to enrich a mixed methanogen culture. This methanogen enrichment culture was used as an inoculum to test the effect of isoprene on methane formation in mixed methanogen communities grown in H_2/HCO_3^-. The methanogen group present in the sample was *Methanobacterium* sp. representing 30-40% of the community at the time of inoculation.

After 40 days of incubation methane in isoprene free control microcosms increased to around 250 μmols/microcosm whereas methane formation was completely inhibited in 1.3 mM and 0.13 mM isoprene supplied samples [**Figure 5.6**]. In 0.013 mM isoprene supplied culture one out of three replicates produced methane whereas the two other replicates showed no methane production and community analysis supported this data. In the non-methane producing replicates, *Methanobacterium* sp. relative abundance decreased to 2-3% whilst in the methane producing replicate relative abundance increased to 70%. In samples without isoprene the community structure was dominated by *Methanobacterium* sp. in all three replicates representing 70% relative abundance of the community. However, in samples supplied with 1.3 mM and 0.13 mM isoprene the relative abundance of *Methanobacterium* sp. decreased to 2-3% and *Pseudomonadaceae* became the dominant organism.

After 146 days of incubation there was still no methanogenesis in 1.3 mM supplied cultures however in those with 0.13 mM two replicates showed methane formation and also an increase in *Methanobacterium* sp. to 70% relative abundance in their community structure. The same applied for samples with 0.013 mM isoprene. Isoprene reduction i.e. methylbutene formation did not occur in any of these samples. However, isoprene concentrations decreased over time (after 146 days), a phenomenon which also occurs in uninoculated controls after long incubation times.

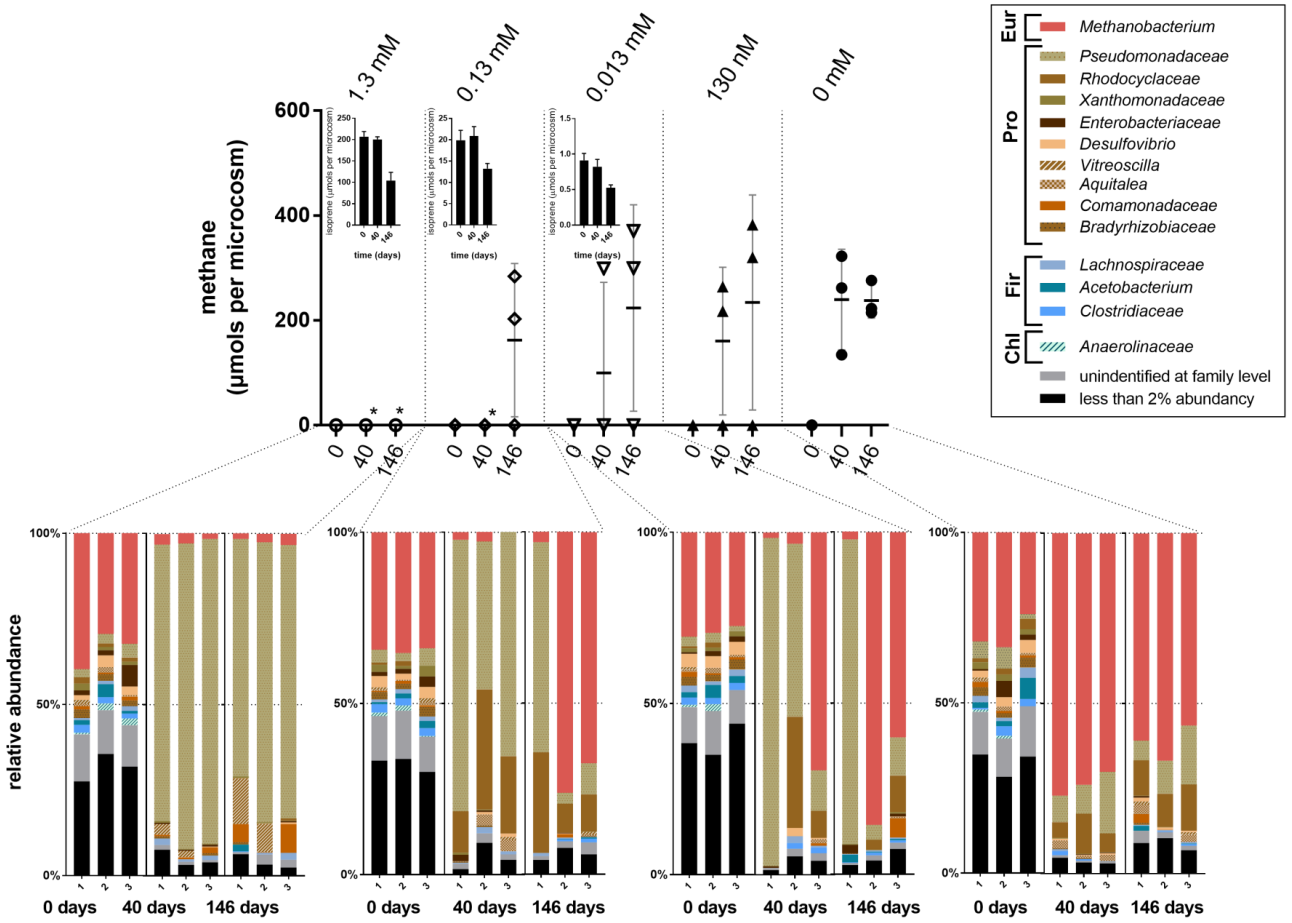

Figure 5.6 Effect of different isoprene concentrations on a methanogen mixed community enriched from wetland samples collected at Site 2. Cultures were grown in minimal media on H_2 and HCO_3^- over 146 days. Isoprene (black bar) amounts are shown in small graphs for concentrations 1.3 mM, 0.13 mM and 0.013 mM. Classifications in the legend are clustered according to their phylum (from top to bottom); Euryarchaeota, Proteobacteria, Firmicutes, Chloroflexi. (*)=p value<0.05 analysed with 2way ANOVA.

5.3.6 Inhibitory effect of different isoprene concentrations on acetoclastic and hydrogenotrophic methanogenesis in Site 3 methanogenic enrichment cultures

The inhibitory effect of different isoprene concentrations was also tested on enrichment cultures from Site 3 since these samples produced the highest amounts of methane of the three wetland sites [**Figure 5.4**]. Minimal media was inoculated with 2% of freshly obtained wetland samples. The cultures were grown on acetate [**Figure 5.7 A**] or H_2/HCO_3^- [**Figure 5.7 B**] and were amended with different isoprene concentrations and compared to isoprene free controls at different time points.

With acetate supplied cultures, methane formation started after 43 days of incubation in isoprene free controls and was completely inhibited in the presence of 1.3 mM and 0.13 mM isoprene even after 79 days of incubation. Since isoprene in 0.013 mM cultures had disappeared after 43 days of incubation nothing can be concluded about its inhibitory effect.

With H_2/HCO_3^- supplied cultures methane formation started after 43 days of incubation in isoprene free controls and was significantly inhibited in the presence of 1.3 mM and 0.13 mM. After 79 days of incubation, 94% less methane was formed in 1.3 mM supplied cultures and 80% less methane in 0.13 mM supplies cultures [**Figure 5.7 B**]. Isoprene reduction occurred in 0.13 mM and 0.013 mM in H_2/HCO_3^- supplied samples so isoprene amounts decreased over time which would influence the inhibitory effect.

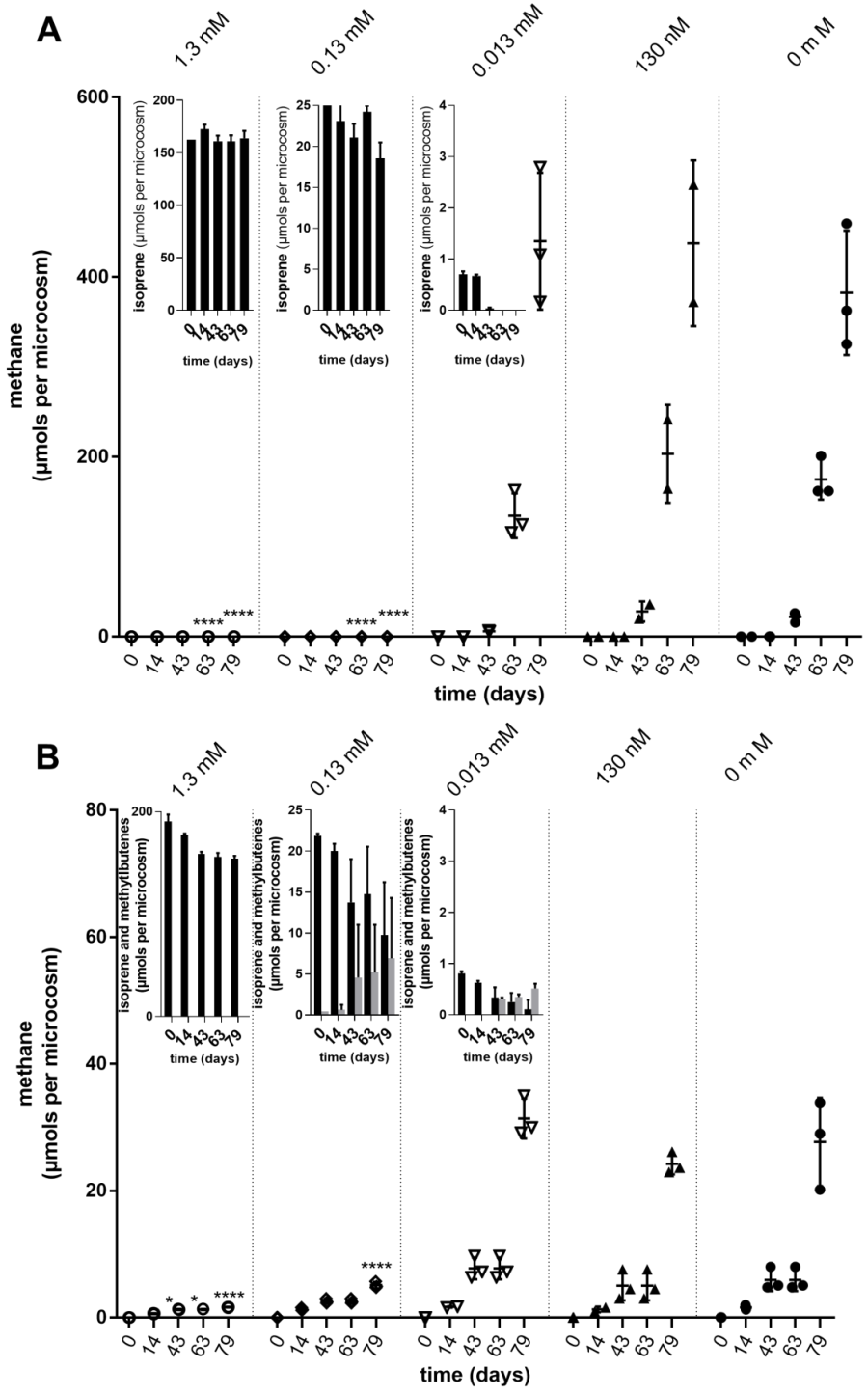

Figure 5.7 Effect of different isoprene concentration on methane formation in minimal media inoculated with 2% wetland sample from sample site three grown on 10 mM acetate (A) and 0.5 bar H_2 plus 30 mM HCO_3^- (B). Isoprene (black bar) and methylbutene (grey bar) amounts are shown in small graphs for concentrations 1.3 mM, 0.13 mM and 0.013 mM. Isoprene concentrations refer to concentration in the aqueous phase. Shown are triplicate measurements for each isoprene concentration and the significance was calculated with a 2way ANOVA multiple comparison between each concentration and the 0 mM isoprene control. (****)=p value <0.0001, (*)=p value<0.05 analysed with 2way ANOVA.

5.3.7 Isoprene dependent inhibition of methanogen activity in a mixed community including methanogens and isoprene reducing organisms

Previously a mixed community including methanogens and isoprene reducing organisms was enriched from waste water treatment plant sludge. The aim of this experiment was to test if methane formation can be regulated by isoprene addition and removal in natural environments in the presence of H_2 and HCO_3^-. In one scenario isoprene was periodically added [**Figure 5.8 A**] to counter removal via reductive transformation. In another scenario isoprene was added at the beginning of the experiment and after methane formation was detected [**Figure 5.8 B**]. Cultures only grown with H_2/HCO_3^- were used as a control. Continual resupply of isoprene led to complete inhibition of methane formation [**Figure 5.8 A**]. When isoprene was added from the beginning and removed via reductive transformation, methanogenesis occurred normally as in control samples [**Figure 5.8 B**]. When methane formation started and isoprene was added constantly over four days, methane formation did not significantly decrease compared to isoprene free controls.

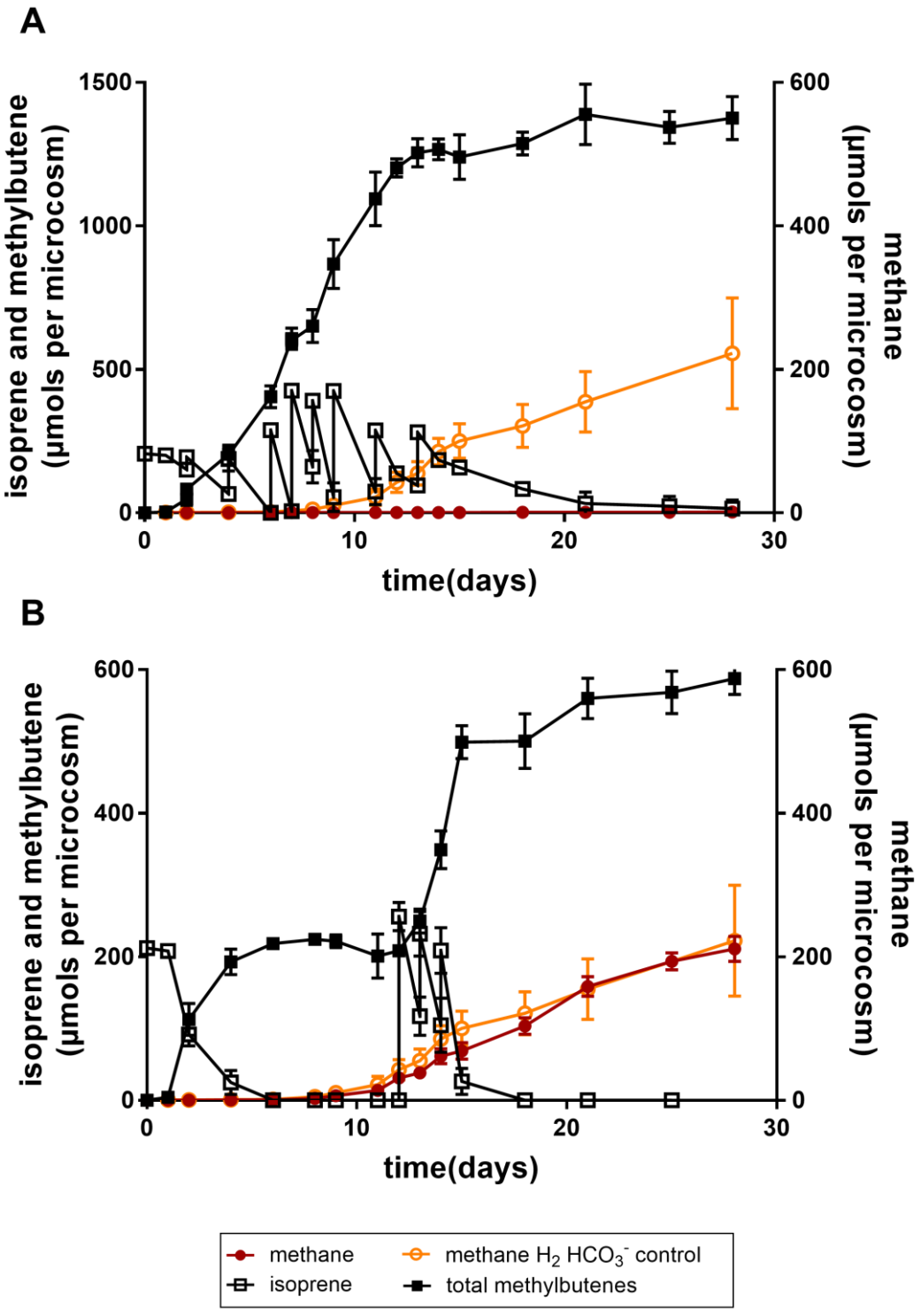

Figure 5.8 Mixed community cultures including methanogens and isoprene reducing organisms grown on H_2 and HCO_3^-. In one scenario (**A**) isoprene was constantly added throughout the experiment and in another scenario isoprene was added at the beginning of the experiment and after methane formation had already started (**B**). Shown in orange are control samples only grown on H_2/HCO_3^-. Note the scale difference between left y-axis from A and B. Error bars represent one standard deviation (n = 4).

5.3.8 Inhibitory effect of 2-methly-1-butene

M. barkeri grown on DMSZ media 120 grown on acetate were tested for inhibition by methylbutenes. Inhibitory concentrations tested were 1 and 5 mM 2-methyl-1-butene. Methane formation in *M. barkeri* was not effected by addition of 2-methyl-1-butene [**Figure 5.9**].

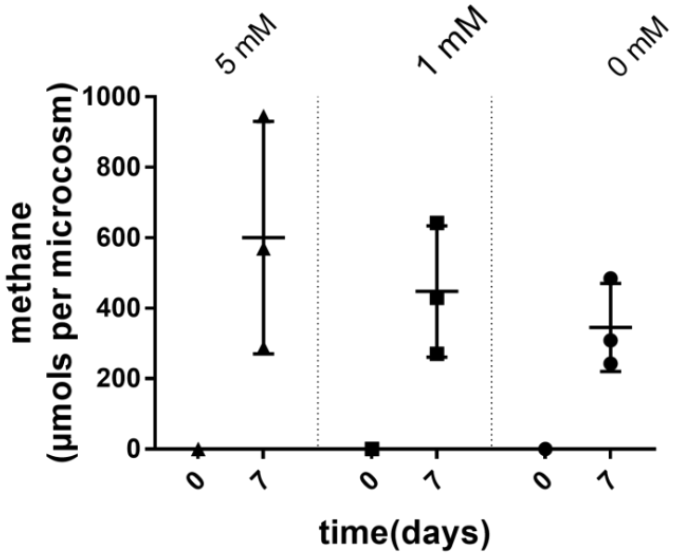

Figure 5.9 Effect of different 2-methyl-1-butene concentration on methane formation in pure *M. barkeri* cultures grown in DSMZ medium 120. Methylbutene concentrations refer to concentration in the aqueous phase. Shown are triplicate measurements for each 2-methyl-1-butene concentration.

5.4 Discussion

5.4.1 Isoprene reduction and methylbutene formation in natural environments

The natural evolution of the three methylbutene isomers (i.e. 2-methyl-1-butene, 2-methyl-2-butene and 3-methyl-1-butene) could be measured in the headspace of wetlands samples obtained from the three different locations in Sydney Australia [**Figure 5.1 ABC**]. Upon addition of isoprene, methylbutene generation increased in wetland samples within 60 days of incubation [**Figure 5.1 DEF**] indicating that isoprene reduction can at least in part be responsible for methylbutene formation. Natural sources of methylbutenes have not been reported before these experiments. Methylbutenes have previously been obtained from thermocatalytic conversion of petroleum hydrocarbons and used as intermediate products for industrial isoprene production (Pavlov *et al.*, 2010; Lybarger, 2014).

Isoprene could not be measured in the wetland samples. If there were soil bacteria and fungi producing isoprene in these samples (Kuzma *et al.*, 1995; Schöller *et al.*, 1997, 2002; Wagner *et al.*, 1999; Fall and Copley, 2000; Bäck *et al.*, 2010; Effmert *et al.*, 2012) it may be consumed/reduced before it can be measured in the head-space (McGenity *et al.*, 2018). Methylbutene amounts ranging between 2-4 nmols/microcosm were measured in the headspace but it is unknown how much of the hydrophobic substances methylbutene and isoprene (water solubility 642 mg/l (25° C) (Hernandez, 2005)) is adsorbed to incumbent organic matter. To obtain accurate concentrations of isoprene and methylbutene in wetland and soil sediments, quantifying adsorption to soil is required since organic matter can act as a cosolvent for hydrophobic substances (Gawlik *et al.*, 1997; Site, 2001). For the evaluation of the effects of isoprene in natural environments it will be necessary to generate data on its behaviour in soil and soil/water mixtures which requires determination of comparable soil sorption data.

Even though isoprene reduction was clearly detected in all samples, no *Acetobacterium* spp. were identified in the community analysis suggesting other organisms apart from *Acetobacterium* are capable of isoprene reduction. Homologs of the FAD-dependent oxidoreductase protein, proposed to be the isoprene reductase, are found in many phyla (e.g. Firmicutes, Actinobacteria and Proteobacteria [Chapter 4, Section **4.3.2**]).

Across all three sample sites *Syntrophaceae* was the only lineage that changed (increased) in relative abundance after the addition of isoprene [**Figure 5.4**]. Whether the increase is related to isoprene, through energy harvesting from isoprene reduction or indirectly through inhibition of methangenesis or other mechanisms, is difficult to say especially since isoprene was already consumed at the time point community analysis was conducted (164 days). *Syntrophaceae* contain four genera *Syntrophus*, *Smithella*, *Desulfobacca*, and *Desulfomonile* which encode the long version of putative isoprene reductase like proteins in their genome, located next to heterodisulfide reductase subunits [**Figure 4.4**]. Further enrichment experiments are needed to clarify which organism(s) are responsible for isoprene reduction in the wetland environments.

5.4.2 Inhibitory effect of isoprene on microbial methanogenesis

When isoprene was first studied for anaerobic biodegradability it appeared that isoprene also inhibited methane formation from methanogens present in the inoculum since methane formation only occurred in samples without isoprene (Kronen *et al.*, 2019). By testing an isoprene concentration range on axenic cultures of *Methanosarcina mazei* and *Methanosarcina barkeri*, on a natural enrichment culture of *Methanobacterium* sp. and on wetland samples from Site 3, it was revealed that 1.3 mM isoprene in the aquatic phase led to complete inhibition of methane formation in all tested methanogenic cultures [**Figure 5.5, Figure 5.6, Figure 5.7**]. Lower concentrations like 0.13 mM isoprene in pure cultures of *M. mazei* and *M. barkeri* seemed to be inhibitory at first (after 8 days) but lost the inhibitory effect over time (after 16 days) [**Figure 5.5**]. This was also the case in samples enriched with hydrogenotrophic *Methanobacterium* sp. where over time the inhibitory effect of isoprene decreased at concentrations of 0.13 mM or below (from 40 days to 146 days).

In samples inoculated from wetland Site 3, methane formation was completely inhibited at 1.3 mM and 0.13 mM isoprene when grown on acetate but when grown in H_2 HCO_3^- 0.13 mM only led to a decrease in methane formation and not a complete inhibition. There are many other known chemical substances inhibiting methane formation with different specificities and inhibition concentrations that are categorized into specific and non-specific inhibitors (Liu *et al.*, 2011). Specific inhibition only affects methanogen enzymes, while nonspecific inhibitors influence the activity of both methanogens and non-methanogens (Liu *et al.*, 2011). Examples of inhibitors that are considered as

specific are structural analogues of coenzyme M (CoM) (2-bromoethanesulfonate (BES), 2-chloroethanesulfonate (CES), 2-mercaptoethanesulfonate (MES), 3-nitrooxypropanol (3-NOP) (Duin *et al*., 2016) and lumazine (Bouwer and McCarty, 1983; Nagar-Anthal *et al.*, 1996; Nollet *et al.*, 1997), hydroxymethylglutaryl-SCoA (HMG-CoA) reductase inhibitors (statins) (Wolin and Miller, 2006) and ethene and acetylene (Schink, 1985). Chloroform ($CHCl_3$) is a classic example of nonspecific inhibitor for methanogens since it is known to block the function of corrinoid enzymes (Chidthaisong and Conrad, 2000).

Whether isoprene is a specific or non-specific inhibitor for methanogens and what mechanism the inhibition follows requires further investigation. When isoprene was added to cultures that were already actively producing methane, methane formation was not significantly inhibited [**Figure 5.8 B**]. Hence, one could speculate that isoprene inhibits cell growth rather than enzymes involved in methane formation. Because of isoprene's structural analogy to isoprenoids it could interfere with the enzymes in the mevalonate production pathway which produces precursors for isoprenoid synthesis in methanogens. Inhibition of isoprenoid synthesis would inhibit membrane formation and therefore growth, since archaea have unique membrane lipids that contain glycerol joined by ether linkages to long chain isoprenoid alcohols (Mancuso *et al.*, 1985; De Rosa *et al.*, 1986; Kates, 1993). Otherwise isoprene could intercalate into the methanogen membrane lipids made of isoprene units and influence membrane functions and properties. Other fermentative bacteria would not be inhibited because bacterial lipids are mostly glycerol esters of long chain fatty acids (Sohlenkamp and Geiger, 2015).

When 1.3 mM isoprene was added to wetland samples from Site 1 and 2, methanogenesis plateaued [**Figure 5.2, Figure 5.3**]. However, in samples from Site 3 methane increased after addition of isoprene [**Figure 5.4**]. Since these scenarios are contradictory to one another it is difficult to draw a conclusion. However, in all wetland samples addition of isoprene led to increased acetate accumulation. This has been observed before in multiple reports that studied the effect of the methanogen inhibitors BES and chloroform on complex microbial communities (Wüst *et al.*, 2009; Xu, Liu, and Chen, 2010; Xu, Liu, Li, *et al.*, 2010). Methanogen inhibition in a complex trophic system indirectly affected the syntrophic interaction between hydrogen-producing fatty

acid degraders and reductive homoacetogens (Xu, Liu, and Chen, 2010; Xu, Liu, Li, *et al.*, 2010). Acetogenesis from H_2/CO_2 was stimulated by inhibition of methanogenesis (Xu, Liu, and Chen, 2010) since in the absence of inorganic electron acceptors other than carbonate, methanogenic archaea and homoacetogenic bacteria are the main H_2 consumers.

A hypothetical model of trophic interactions in wetland samples (Drake *et al.*, 2009; Wüst *et al.*, 2009; Liu *et al.*, 2011) was adapted to organisms identified in wetland samples from this study [**Figure 5.10**]. The microbial communities found in wetland samples from this study resemble typical wetland communities mostly dominated by Proteobacteria, Acidobacteria Chloroflexi, Bacteroidetes, and Euryarchaeota (Wüst *et al.*, 2009; Kanokratana *et al.*, 2011; He *et al.*, 2015; Dalcin Martins *et al.*, 2019). Site 1 and 2 samples contained many primary fermenters from Proteobacteria and Acidobacteria groups that together with Firmicutes play key roles in plant biomass degradation (Kanokratana *et al.*, 2011). Site 3 samples did not contain any members of Acidobacteria or Firmicutes but a lot of syntrophic Proteobacteria involved in secondary fermentation processes which correlates with highest methane amounts measured in Site 3 [**Figure 5.4**]. Overall the structure of the active microbial community was marginally affected by addition of isoprene in all wetland samples.

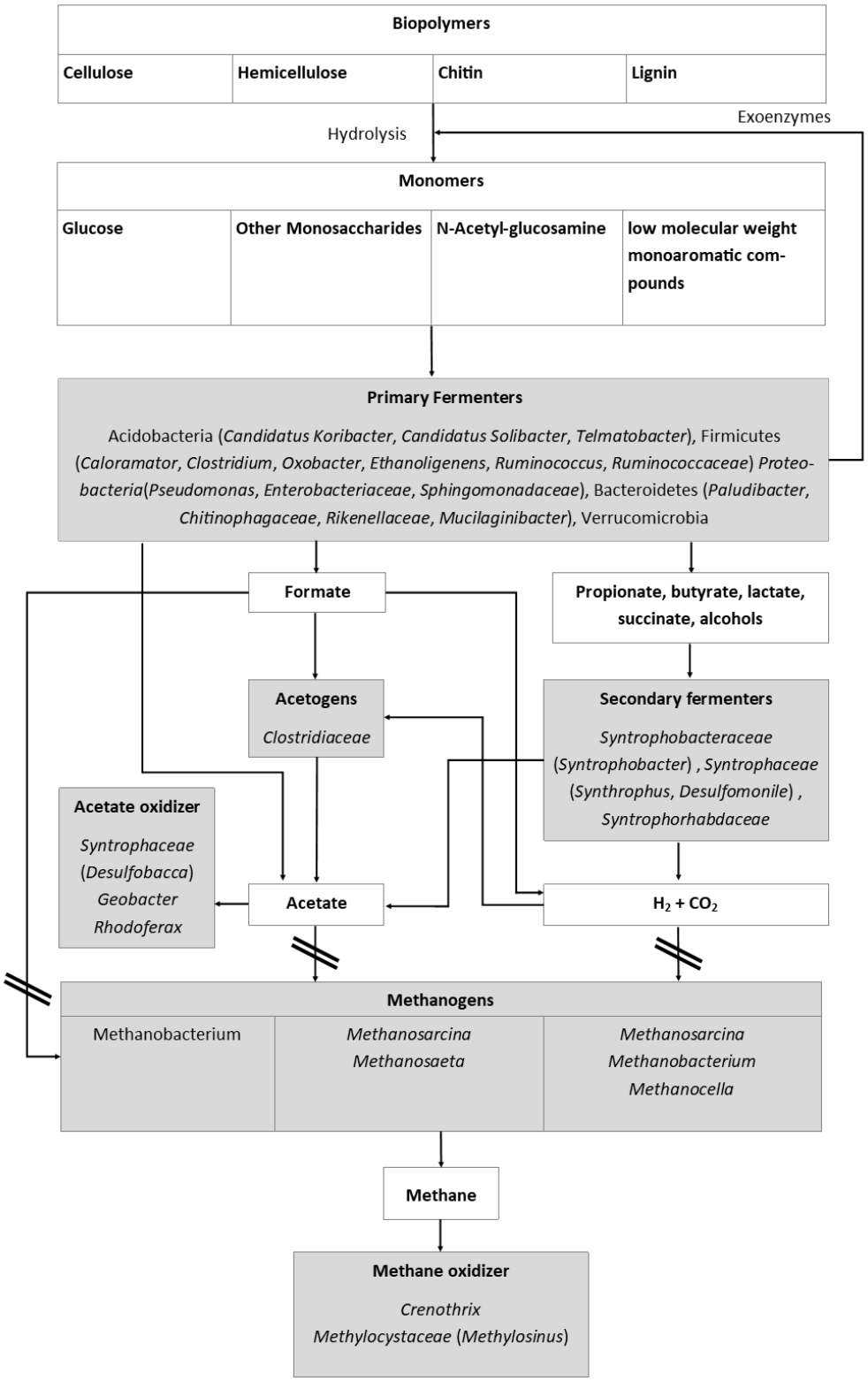

Figure 5.10 Hypothetical model of trophic interactions in wetland samples based on the metabolism of known identified organisms in the samples and based on a previously described model (Drake *et al.*, 2009; Wüst *et al.*, 2009)**.** Note that the microbial community structure of samples from Site 3 is missing Acidobacteria and Firmicutes in their community. Refer to **Figure 5.2**, **Figure 5.3**, **Figure 5.4** to see which organism can be found in each wetland site. (≠ isoprene mediated methanogen inhibition).

5.4.3 Environmental and ecological relevance of isoprene mediated methanogenesis inhibition and isoprene reduction

It is now widely recognised that biological processes in terrestrial ecosystems broadly affect the atmosphere and climate systems of Earth (Arneth *et al.*, 2010; Penuelas *et al.*, 2010). Apart from influencing the atmosphere directly, BVOC compounds also have important effects within plants, between plants and other organisms and they are mediator molecules of inter- and intra-organismic relationships in microorganisms (Laothawornkitkul *et al.*, 2009; Effmert *et al.*, 2012; Zhao *et al*., 2017).

BVOCs can act in two major ways on microbial interactions, as: "(I) infochemical molecules affecting the behaviour, population dynamic and gene expression in the responding microorganism, and (II) competitive tools directly exerting antimicrobial activity, providing an advantage by suppressing or eliminating potential enemies" (Kai *et al.*, 2009; Effmert *et al.*, 2012; Garbeva *et al.*, 2014; Schmidt *et al.*, 2015). An enormous range of BVOCs is produced by cultivated soil microbes (Lemfack *et al*. 2017), and many of these VOCs were found to have biocidal or biostatic effects.

BVOCs like isoprene may act on a broader scale as an inhibitor or negative feedback regulator for methanogensis (Insam and Seewald, 2010; Wenke *et al.*, 2010; Peñuelas *et al.*, 2014). Warmer temperatures lead to higher emission of BVOCs e.g. isoprene from plants and wetlands (Harley *et al.*, 1999b; Pacifico *et al.*, 2009; Peñuelas and Staudt, 2010; Lindwall *et al.*, 2016). In atmospheric chemistry methane and isoprene compete for OH radicals which are the major methane sink in the biogeochemical cycle but in the presence of isoprene OH radicals react preferentially with isoprene, leading indirectly to an elevation of methane in the atmosphere (Collins *et al.*, 2002; Pacifico *et al.*, 2009; Pike and Young, 2009). One could hypothesize that at the same time, BVOCs (e.g. isoprene, ethene and acetylene) emitted by fungi, bacteria and plants in the rhizosphere and soil (Insam and Seewald, 2010; Wenke *et al.*, 2010; Peñuelas *et al.*, 2014) might also inhibit methane formation in the first place. The inhibition could be both direct, by interfering with the methanogenic metabolism or growth, and indirect by introducing alternative electron acceptors (unsaturated hydrocarbons) into the system that enable other organism to compete for common methanogen substrates H_2 and acetate (Achtnich *et al.*, 1995; Bodegom and Stams, 1999). However, whether isoprene mediated methanogenesis inhibition takes place *in situ* remains uncertain since

inhibitory concentrations ranged around 0.013-1.3 mM and it is not yet clear whether these concentrations are reached in natural environments. Further, acquired data in this Chapter showed that isoprene may be rapidly transformed to methylbutene before it can directly inhibit methanogens. But with this in mind one has to consider that *in situ* environmental conditions are not static but vary constantly and the biotic and abiotic complexity of soil environments detains the extrapolation of the production and suppressing activity of BVOCs by microcosmic processes (de Boer *et al*. 2019). Soil microbial communities (interactions, diversity and composition) and abiotic factors (e.g. organic matter quality and moisture, temperature, pH) determine the quantity and composition of BVOC emission (de Boer *et al*. 2019). *In situ* there may be constant isoprene flux from plants roots, from bacteria and fungi in deeper layers in soils or wetlands. Also unlike non-volatiles, BVOCs can act over a wider range of spatial scale since they can diffuse through both the liquid and gaseous phases of soil matrices which can lead to local niches with higher isoprene concentrations (Garbeva *et al.*, 2014; Schmidt *et al.*, 2015). It is also unclear how much of the atmospheric isoprene diffuses into the soil. Moreover, not all wetland samples exhibited high isoprene reduction activity (Site 2) so isoprene might accumulate to cause methanogenesis inhibition. Isoprene is also not the only BVOC that inhibitites methane formation and in case isoprene is reduced ethene or acetylene might be inhibiting methanogens (Schink, 1985). Generally, complex BVOCs blends are what to expect in soil and wetland environments and it is highly unlikely that microorganisms will be exposed to only one BVOC in nature (de Boer *et al*. 2019).

There are other examples were BVOCs regulate microbial processes in soils. Ethene has previously been suggested to be an integral part of a self-regulatory cycle in soil, controlling microbial activity to maintain the biological balance (Smith and Cook, 1974). Ethene is produced by anaerobic spore-forming bacteria and arrests the growth of aerobic bacteria (Smith and Cook, 1974). The concept of an ethene-oxygen cycle prevents microbiologically active soils from becoming predominantly anaerobic since when aerobic bacteria are inhibited by ethene, oxygen can diffuse back into the soil. When oxygen is present however it inhibits ethene production until it is used up again by aerobic bacteria (Smith and Cook, 1974). Furthermore, there is increasing evidence that BVOCs play a role in the suppression of plant-pathogenic soil fungi (de Boer *et al*. 2019).

Terrestrial wetlands are among the largest biogenic sources of methane (190 Tg y-1) (Kirschke *et al.*, 2013; Nisbet, Euan G; Dlugokencky, Edward J; Bousquet, 2014; Dean *et al.*, 2018) contributing to growing atmospheric CH_4 concentrations and are, in turn, highly sensitive to climate change (Arneth *et al.*, 2010; Zhang *et al.*, 2017). Since microbes are the engines that drive Earth's biogeochemical cycles (Falkowski *et al.*, 2008) they also control the balance of microbial methane production and consumption, which are critical to methane climate feedbacks in all environments (Dean *et al.*, 2018). The role of BVOCs like isoprene in regulating biogeochemical cycles needs major validation and requires extensive field research especially to find out if environmental concentrations of isoprene, ethene and acetylene would be inhibitory to methane formation. Understanding biogeochemical controls on methane and isoprene production is important to predict spatial and temporal patterns of the emission to the atmosphere of these important greenhouse gases.

5.5 Conclusion

For the first time, methylbutenes were measured evolving from natural environments. Three methylbutene isomers were measured in the headspace of wetland samples collected from three different sample sites and results suggest that isoprene reduction could be at least one of the responsible metabolisms resulting in methylbutene formation. Isoprene has an inhibitory effect on methanogens but it remains to be determined if isoprene, together with other BVOCs, play a regulatory role in methane climate feedback. Regardless, these experiments show that isoprene reduction is a metabolism found in natural environments like wetlands and suggests its presence in other anaerobic environments.

5.6 References

Achtnich, C., Bak, F., and Conrad, R. (1995) Competition for electron donors among nitrate reducers, ferric iron reducers, sulfate reducers, and methanogens in anoxic paddy soil. *Biol. Fertil. Soils* **19**: 65–72.

Arneth, A., Harrison, S.P., Zaehle, S., Tsigaridis, K., Menon, S., Bartlein, P.J., et al. (2010) Terrestrial biogeochemical feedbacks in the climate system. *Nat. Geosci.* **3**: 525–532.

Bäck, J., Aaltonen, H., Hellén, H., Kajos, M.K., Patokoski, J., Taipale, R., et al. (2010) Variable emissions of microbial volatile organic compounds (MVOCs) from root-associated fungi isolated from Scots pine. *Atmos. Environ.* **44**: 3651–3659.

Bodegom, P.M. Van and Stams, A.J.M. (1999) Effects of alternative electron acceptors and temperature on methanogenesis in rice paddy soils. *Chemosphere* **39**: 167–182.

Bodelier, P.L.E. and Dedysh, S.N. (2013) Microbiology of wetlands. *Front. Microbiol.* **4**: 2–5.

de Boer, W., Li, X., Meisner, A., and Garbeva, P. (2019) Pathogen suppression by microbial volatile organic compounds in soils. FEMS Microbiol. Ecol. 95: 1–10.

Bouwer, E.J. and McCarty, P.L. (1983) Transformation of trace halogenated organic aliphatic organic compounds under methanogenic conditions. *Appl. Environ. Microbiol.* **42**: 1286–1294.

Cao, M., Gregson, K., and Marshall, S. (1998) Global methane emission from wetlands and its sensitivity to climate change. **32**: 1–7.

Caporaso, J.G., Lauber, C.L., Walters, W.A., Berg-Lyons, D., Lozupone, C.A., Turnbaugh, P.J., et al. (2011) Global patterns of 16S rRNA diversity at a depth of millions of sequences per sample. *Proc. Natl. Acad. Sci.* **108**: 4516–4522.

Chidthaisong, A. and Conrad, R. (2000) Specificity of chloroform, 2-bromoethanesulfonate and fluoroacetate to inhibit methanogenesis and other anaerobic processes in anoxic rice field soil. *Soil Biol. Biochem.* **32**: 977–988.

Van Cleemput, O. and El-Sebaay, A.S. (1986) Gaseous hydrocarbons in soil. *Adv. Agron.* **38**: 159–181.

Collins, W.J., Derwent, R.G., Johnson, C.E., and Stevenson, D.S. (2002) The oxidation of organic compounds in the troposphere and their global warming potentials. *Clim. Change* **52**: 453–479.

Dalcin Martins, P., Frank, J., Mitchell, H., Markillie, L.M., and Wilkins, M.J. (2019) Wetland sediments host diverse microbial taxa capable of cycling alcohols. *Appl. Environ. Microbiol.*

Dean, J.F., Middelburg, J.J., Röckmann, T., Aerts, R., Blauw, L.G., Egger, M., et al. (2018) Methane feedbacks to the global climate system in a warmer world. *Rev. Geophys.* **56**: 207–250.

Drake, H.L., Horn, M.A., and Wüst, P.K. (2009) Intermediary ecosystem metabolism as a main driver of methanogenesis in acidic wetland soil. *Environ. Microbiol. Rep.* **1**: 307–318.

Duin, E.C., Wagner, T., Shima, S., Prakash, D., Cronin, B., Yáñez-Ruiz, D.R., et al. (2016) Mode of action uncovered for the specific reduction of methane emissions from ruminants by the small molecule 3-nitrooxypropanol. *Proc. Natl. Acad. Sci. U. S. A.* **113**: 6172–6177.

Effmert, U., Kalderás, J., Warnke, R., and Piechulla, B. (2012) Volatile mediated interactions between bacteria and fungi in the soil. *J. Chem. Ecol.* **38**: 665–703.

Ekberg, A., Arneth, A., Hakola, H., Hayward, S., and Holst, T. (2009) Isoprene emission from wetland sedges. *Biogeosciences* **6**: 601–613.

Falkowski, P.G., Fenchel, T., and Delong, E.F. (2008) The microbial engines that drive Earth's biogeochemical cycles. *Science* **320**: 1034–1039.

Fall, R. and Copley, S.D. (2000) Bacterial sources and sinks of isoprene, a reactive atmospheric hydrocarbon. *Environ. Microbiol.* **2**: 123–130.

Garbeva, P., Hordijk, C., Gerards, S., and De Boer, W. (2014) Volatile-mediated interactions between phylogenetically different soil bacteria. *Front. Microbiol.* **5**: 1–9.

Gawlik, B.M., Sotiriou, N., Feicht, E., Schulte-Hostede, S., and Kettrup, A. (1997) Alternatives for the determination of the soil adsorption coefficient, Koc, of non-ionicorganic compounds-a review. *Chemosphere* **34**: 2525–2551.

Harley, P.C., Monson, R.K., and Lerdau, M.T. (1999a) Ecological and evolutionary aspects of isoprene emission from plants. *Oecologia* **118**: 109–123.

Harley, P.C., Monson, R.K., and Lerdau, M.T. (1999b) Ecological and evolutionary aspects of isoprene emission from plants. *Oecologia* **118**: 109–123.

He, S., Malfatti, S.A., McFarland, J.W., Anderson, F.E., Pati, A., Huntemann, M., et al. (2015) Patterns in wetland microbial community composition and functional gene repertoire associated with methane emissions. *MBio* **6**: 1–15.

Hernandez, O. (2005) SIDS Initial Assessment Report.

Insam, H. and Seewald, M.S.A. (2010) Volatile organic compounds (VOCs) in soils. *Biol. Fertil. Soils* **46**: 199–213.

Janson, R. and De Serves, C. (1998) Isoprene emissions from boreal wetlands in Scandinavia. *J. Geophys. Res. Atmos.* **103**: 25513–25517.

Janson, R., De Serves, C., and Romero, R. (1999) Emission of isoprene and carbonyl compounds from a boreal forest and wetland in Sweden. *Agric. For. Meteorol.* **98**–**99**: 671–681.

Jordan, S.J., Stoffer, J., and Nestlerode, J.A. (2011) Wetlands as sinks for reactive nitrogen at continental and global scales: A meta-analysis. *Ecosystems* **14**: 144–155.

Kai, M., Haustein, M., Molina, F., Petri, A., Scholz, B., and Piechulla, B. (2009) Bacterial volatiles and their action potential. *Appl. Microbiol. Biotechnol.* **81**: 1001–1012.

Kanokratana, P., Uengwetwanit, T., Rattanachomsri, U., Bunterngsook, B., Nimchua, T., Tangphatsornruang, S., et al. (2011) Insights into the phylogeny and metabolic potential of a primary tropical peat swamp forest microbial community by metagenomic analysis. *Microb. Ecol.* **61**: 518–528.

Kates, M. (1993) Lipids of Archaea. In, Kates,M., Kushner,D.J., and Matheson,A.T. (eds), *The Biochemistry of Archaea (Archaebacteria)*. Elsevier, New York, pp. 261–295.

Kirschke, S., Bousquet, P., Ciais, P., Saunois, M., Canadell, J.G., Dlugokencky, E.J., et al. (2013) Three decades of global methane sources and sinks. *Nat. Geosci.* **6**: 813–823.

Kronen, M., Lee, M., Jones, Z.L., and Manefield, M.J. (2019) Reductive metabolism of the important atmospheric gas isoprene by homoacetogens. *ISME J.* **13**: 1168–1182.

Kuzma, J., Nemecek-Marshall, M., Pollock, W.H., and Fall, R. (1995) Bacteria produce the volatile hydrocarbon isoprene. *Curr. Microbiol.* **30**: 97–103.

Laothawornkitkul, J., Taylor, J.E., Paul, N.D., and Hewitt, C.N. (2009) Biogenic volatile organic compounds in the Earth system. *New Phytol.* **183**: 27–51.

Lemfack, M.C., Gohlke, B.O., Toguem, S.M.T., Preissner, S., Piechulla, B., and Preissner, R. (2017) mVOC 2.0: A database of microbial volatiles. *Nucleic Acids Res.* **46**: D1261–D1265.

Lindwall, F., Svendsen, S.S., Nielsen, C.S., Michelsen, A., and Rinnan, R. (2016) Warming increases isoprene emissions from an arctic fen. *Sci. Total Environ.* **553**: 297–304.

Liu, H., Wang, J., Wang, A., and Chen, J. (2011) Chemical inhibitors of methanogenesis and putative applications. *Appl. Microbiol. Biotechnol.* **89**: 1333–1340.

Lybarger, H.M. (2014) Isoprene. *Kirk-Othmer Encycl. Chem. Technol.*

Mancuso, C.A., Odham, G., Westerdahl, G., Reeve, J.N., and White, D.C. (1985) C_{15}, C_{20}, and C_{25} isoprenoid homologues in glycerol diether phospholipids of methanogenic archaebacteria. *Lipids* **26**: 1120–1125.

McGenity, T.J., Crombie, A.T., and Murrell, J.C. (2018) Microbial cycling of isoprene, the most abundantly produced biological volatile organic compound on Earth. *ISME J.* **12**: 931–941.

Nagar-Anthal, K.R., Worrell, V.E., Teal, R., and Nagle, D.P. (1996) The pterin lumazine inhibits growth of methanogens and methane formation. *Arch. Microbiol.* **166**: 136–140.

Nisbet, E.G., Dlugokencky, E.J., and Bousquet, P. (2014) Methane on the rise-again. *Science* **343**: 493–496.

Nollet, L., Demeyer, D., and Verstraete, W. (1997) Effect of 2-bromoethanesulfonic acid and *Peptostreptococcus productus* ATCC 35244 addition on stimulation of reductive acetogenesis in ruminal ecosystem by selective inhibition of methanogenesis. *Appl. Environ. Microbiol.* **63**: 194–200.

Pacifico, F., Harrison, S.P., Jones, C.D., and Sitch, S. (2009) Isoprene emissions and climate. *Atmos. Environ.* **43**: 6121–6135.

Pavlov, O.S., Karsakov, S.A., and Pavlov, S.Y. (2010) 2-Methyl-2-butene production using positional isomerization catalyzed bv sulfonic acid cation exchangers. *Theor. Found. Chem. Eng.* **44**: 536–539.

Peñuelas, J., Asensio, D., Tholl, D., Wenke, K., Rosenkranz, M., Piechulla, B., and Schnitzler, J.P. (2014) Biogenic volatile emissions from the soil. *Plant, Cell Environ.* **37**: 1866–1891.

Penuelas, J., Rutishauser, T., and Filella, I. (2010) Phenology feedbacks on climate change. *Science* **324**: 887–888.

Peñuelas, J. and Staudt, M. (2010) BVOCs and global change. *Trends Plant Sci.* **15**: 133–144.

Pester, M., Knorr, K.H., Friedrich, M.W., Wagner, M., and Loy, A. (2012) Sulfate-reducing microorganisms in wetlands - fameless actors in carbon cycling and climate change. *Front. Microbiol.* **3**: 1–19.

Pike, R.C. and Young, P.J. (2009) How plants can influence tropospheric chemistry: The role of isoprene emissions from the biosphere. *Weather* **64**: 332–336.

De Rosa, M., Gambacorta, A., and Gliozzi, A. (1986) Structure, biosynthesis, and physicochemical properties of archaebacterial lipids. *Microbiol. Rev.* **50**: 70–80.

Sander, R. (2015) Compilation of Henry's law constants (version 4.0) for water as solvent. *Atmos. Chem. Phys.* **15**: 4399–4981.

Schink, B. (1985) Inhibition of methanogenesis by ethylene and other unsaturated hydrocarbons. *FEMS Microbiol. Lett.* **31**: 63–68.

Schmidt, R., Cordovez, V., De Boer, W., Raaijmakers, J., and Garbeva, P. (2015) Volatile affairs in microbial interactions. *ISME J.* **9**: 2329–2335.

Schöller, C., Molin, S., and Wilkins, K. (1997) Volatile metabolites from some Gram-negative bacteria. *Chemosphere* **35**: 1487–1495.

Schöller, C.E.G., Gürtler, H., Pedersen, R., Molin, S., and Wilkins, K. (2002) Volatile metabolites from actinomycetes. *J. Agric. Food Chem.* **50**: 2615–2621.

Sharkey, T.D., Wiberley, A.E., and Donohue, A.R. (2008) Isoprene emission from plants: Why and how. *Ann. Bot.* **101**: 5–18.

Site, A.D. (2001) Factors affecting sorption of organic compounds in natural sorbent/water systems and sorption coefficients for selected pollutants . A review. *J. Phys. Chem. Ref. Data* **30**: 187–439.

Smith, A.M. and Cook, R.J. (1974) Implications of ethylene production by bacteria for biological balance of soil. *Nature* **252**: 703–705.

Sohlenkamp, C. and Geiger, O. (2015) Bacterial membrane lipids: Diversity in structures and pathways. *FEMS Microbiol. Rev.* **40**: 133–159.

Tiiva, P., Rinnan, R., Faubert, P., Räsänen, J., Holopainen, T., Kyrö, E., and Holopainen, J.K. (2007) Isoprene emission from a subarctic peatland under enhanced UV-B radiation. *New Phytol.* **176**: 346–355.

Wagner, W.P., Nemecek-Marshall, M., and Fall, R. (1999) Three distinct phases of isoprene formation during growth and sporulation of Bacillus subtilis. *J. Bacteriol.* **181**: 4700–4703.

Wenke, K., Kai, M., and Piechulla, B. (2010) Belowground volatiles facilitate interactions between plant roots and soil organisms. *Planta* **231**: 499–506.

Wolin, M.J. and Miller, T.L. (2006) Control of rumen methanogenesis by inhibiting the

growth and activity of methanogens with hydroxymethylglutaryl-SCoA inhibitors. *Int. Congr. Ser.* **1293**: 131–137.

Wuebbles, D.J. and Hayhoe, K. (2002) Atmospheric methane and global change. *Earth-Science Rev.* **57**: 177–210.

Wüst, P.K., Horn, M.A., and Drake, H.L. (2009) Trophic links between fermenters and methanogens in a moderately acidic fen soil. *Environ. Microbiol.* **11**: 1395–1409.

Xu, K., Liu, H., and Chen, J. (2010) Effect of classic methanogenic inhibitors on the quantity and diversity of archaeal community and the reductive homoacetogenic activity during the process of anaerobic sludge digestion. *Bioresour. Technol.* **101**: 2600–2607.

Xu, K., Liu, H., Li, X., Chen, J., and Wang, A. (2010) Typical methanogenic inhibitors can considerably alter bacterial populations and affect the interaction between fatty acid degraders and homoacetogens. *Appl. Microbiol. Biotechnol.* **87**: 2267–2279.

Zhang, Z., Zimmermann, N.E., Stenke, A., Li, X., Hodson, E.L., Zhu, G., et al. (2017) Emerging role of wetland methane emissions in driving 21st century climate change. *Proc. Natl. Acad. Sci.* **114**: 9647–9652.

Zhao, D.F., Buchholz, A., Tillmann, R., Kleist, E., Wu, C., and Rubach, F. (2017) Environmental conditions regulate the impact of plants on cloud formation. **8**: 1–35.

5.7 Supplementary material

Table S5 1 Significant results of 2way ANOVA multicomparison between means of bacterial and archaeal structure analysis at each time point in each sample site. Please compare with [**Figure 5.2, Figure 5.3, Figure 5.4**]. P value = 0.05.

	Mean Diff.	95.00% CI of diff.	Significant?	Summary	Adj. P value
Sample site 1 (Figure 5.2)					
g__Candidatus Koribacter					
T0 0 days vs. T1 101 days	8.865	3.057 to 14.67	Yes	***	0.0006
T0 0 days vs. T2 134 days	11.84	6.036 to 17.65	Yes	****	<0.0001
T0 0 days vs. T3 164 days	7.447	1.639 to 13.25	Yes	**	0.0056
T1 101 days vs. T2 134 days	2.979	-2.829 to 8.787	No	ns	0.5487
T1 101 days vs. T3 164 days	-1.418	-7.227 to 4.39	No	ns	0.9223
T2 134 days vs. T3 164 days	-4.397	-10.21 to 1.411	No	ns	0.2077
Sample site 2 (Figure 5.3)					
g__Methanobacterium					
T0 0 days vs. T1 101 days	-4.965	-11.11 to 1.179	No	ns	0.1601
T0 0 days vs. T2 134 days	-6.525	-12.67 to -0.3808	Yes	*	0.0324
T0 0 days vs. T3 164 days	-0.7092	-6.853 to 5.435	No	ns	0.9908
T1 101 days vs. T2 134 days	-1.56	-7.704 to 4.584	No	ns	0.9139
T1 101 days vs. T3 164 days	4.255	-1.889 to 10.4	No	ns	0.2816
T2 134 days vs. T3 164 days	5.816	-0.3284 to 11.96	No	ns	0.0711
g__Oxobacter					
T0 0 days vs. T1 101 days	-0.5674	-6.711 to 5.577	No	ns	0.9952
T0 0 days vs. T2 134 days	-6.241	-12.39 to -0.09713	Yes	*	0.0449
T0 0 days vs. T3 164 days	-4.681	-10.82 to 1.463	No	ns	0.2032
T1 101 days vs. T2 134 days	-5.674	-11.82 to 0.4702	No	ns	0.0822
T1 101 days vs. T3 164 days	-4.113	-10.26 to 2.031	No	ns	0.3113
T2 134 days vs. T3 164 days	1.56	-4.584 to 7.704	No	ns	0.9139
g__Ethanoligenens					
T0 0 days vs. T1 101 days	-6.17	-12.31 to -0.02621	Yes	*	0.0486
T0 0 days vs. T2 134 days	-4.397	-10.54 to 1.747	No	ns	0.2536
T0 0 days vs. T3 164 days	-4.468	-10.61 to 1.676	No	ns	0.2403
T1 101 days vs. T2 134 days	1.773	-4.371 to 7.917	No	ns	0.8792
T1 101 days vs. T3 164 days	1.702	-4.442 to 7.846	No	ns	0.8915
T2 134 days vs. T3 164 days	-0.07092	-6.215 to 6.073	No	ns	>0.9999
f__Acidobacteriaceae					
T0 0 days vs. T1 101 days	-5.177	-11.32 to 0.9667	No	ns	0.1324
T0 0 days vs. T2 134 days	-5.674	-11.82 to 0.4702	No	ns	0.0822
T0 0 days vs. T3 164 days	-12.2	-18.34 to -6.055	Yes	****	<0.0001
T1 101 days vs. T2 134 days	-0.4965	-6.64 to 5.648	No	ns	0.9968
T1 101 days vs. T3 164 days	-7.021	-13.17 to -0.8773	Yes	*	0.0177
T2 134 days vs. T3 164 days	-6.525	-12.67 to -0.3808	Yes	*	0.0324
f__Ruminococcaceae					
T0 0 days vs. T1 101 days	-4.043	-10.19 to 2.101	No	ns	0.3268
T0 0 days vs. T2 134 days	-2.34	-8.484 to 3.804	No	ns	0.7598
T0 0 days vs. T3 164 days	-6.17	-12.31 to -0.02621	Yes	*	0.0486
T1 101 days vs. T2 134 days	1.702	-4.442 to 7.846	No	ns	0.8915
T1 101 days vs. T3 164 days	-2.128	-8.272 to 4.016	No	ns	0.8086
T2 134 days vs. T3 164 days	-3.83	-9.974 to 2.314	No	ns	0.3756
Sample site 3 (Figure 5.4)					
g__Dechloromonas					
T0 0 days vs. T1 101 days	-11.13	-18.18 to -4.085	Yes	***	0.0003
T0 0 days vs. T2 110 days	0	-7.05 to 7.05	No	ns	>0.9999
T0 0 days vs. T3 164 days	-9.078	-16.13 to -2.028	Yes	**	0.0054
T1 101 days vs. T2 110 days	11.13	4.085 to 18.18	Yes	***	0.0003
T1 101 days vs. T3 164 days	2.057	-4.993 to 9.107	No	ns	0.8755
T2 110 days vs. T3 164 days	-9.078	-16.13 to -2.028	Yes	**	0.0054

f__Syntrophaceae					
T0 0 days vs. T1 101 days	0.2128	-6.837 to 7.263	No	ns	0.9998
T0 0 days vs. T2 110 days	0.4965	-6.554 to 7.547	No	ns	0.9979
T0 0 days vs. T3 164 days	-19.08	-26.13 to -12.03	Yes	****	<0.0001
T1 101 days vs. T2 110 days	0.2837	-6.767 to 7.334	No	ns	0.9996
T1 101 days vs. T3 164 days	-19.29	-26.34 to -12.24	Yes	****	<0.0001
T2 110 days vs. T3 164 days	-19.57	-26.62 to -12.52	Yes	****	<0.0001

Chapter 6 Future Experiments and Perspectives

6.1 Summary of findings

Even though isoprene is an abundant, climate-active atmospheric gas, produced in large quantities from mostly terrestrial plants, little is known about its biogeochemical processing. Bacteria, able to grow on isoprene aerobically, have been isolated from soils and sediments but anaerobic environments have remained entirely unexplored. This project was set up to test for anaerobic metabolism of isoprene to gain further insights into its fate and role in the biogeochemical cycle.

Under anaerobic conditions isoprene was found to act as an electron acceptor which can be reduced to three different methylbutene isomers 2-methyl-1-butene, 3-methyl-1-butene and 2-methyl-2-butene. Isoprene reduction was carried out by the newly discovered homoacetogenic *Acetobacterium wieringae* strain ISORED-2 which reduced isoprene in addition to CO_2 upon growth on H_2 and CO_2. Comparative proteogenomics identified a five gene operon in *A. wieringae* ISORED-2 likely involved in anaerobic isoprene metabolism. The operon encodes a putative FAD-dependent oxidoreductase (5587), three pleiotropic nickel chaperones HypA (5588, 5591), HypB (5589) and one 4Fe-4S ferredoxin (5590). Whilst the hydrogenase maturation factors and ferredoxin are common amongst bacteria, the FAD-dependent oxidoreductase 5587 is proposed as a putative nickel-dependent isoprene reductase with a binding site for NADH, FAD and a 4Fe-4S ferredoxin.

The putative isoprene reductase shares sequence homology with the FAD-dependent oxidoreductase family to which the flavin-dependent electron bifurcating NADH-dependent ferredoxin-$NADP^+$-oxidoreductase subunit B also belongs. Even though other closely related *Acetobacterium* strains e.g. *A. woodii*, *A. wieringae*, *A. malicum* could not reduce isoprene, evidence is presented suggesting the isoprene reduction ability is not restricted to *Acetobacterium*. Isoprene reduction activity was found in wetlands samples from three different sites but no *Acetobacterium* sp. was identified in the community analysis. Isoprene also had an inhibitory effect in methane formation suggesting it might play a role in the methane climate feedback.

6.2 Cloning of putative isoprene operon into *Acetobacterium woodii*

Comparative proteogenomics identified a five gene operon in the *A. wieringae* ISORED-2 genome likely involved in isoprene metabolism but direct evidence remains to be gathered. For direct demonstration of the isoprene reducing activity of the operon a plasmid encoding the operon could be introduced into *A. woodii* and tested for isoprene reduction. Since *A. woodii* does not reduce isoprene and does not encode this operon, measuring isoprene reduction after successful transformation would directly link the five gene operon to consumption of the major biogenic volatile organic compound in the Earth's atmosphere.

6.3 Purification and energetics of putative isoprene reductase from *A. wieringae* ISORED-2.

The FAD-dependent oxidoreductase 5587 is predicted to be a nickel-dependent enzyme that contains a binding site for NADH, FAD and a 4Fe-4S ferredoxin. Further experiments to explore its properties and enzymatic function could involve purification from wildtype *A. wieringae* ISORED-2 cells by using ion exchange chromatography followed by gel filtration (Bertsch *et al.*, 2013). Alternatively, protein 5587 could be overexpressed from plasmid transformed *A. woodii* or *A. wieringae* DSM 1911 cells.

Activity assays for measuring isoprene reducing activity in fractions could involve methyl viologen as an artificial electon donor (methyl viologen oxidation assay) and isoprene as the electron acceptor. The physiological isoprene reductase activity could be measured with NADH or NADPH as electron donor and addition of at least one of the following electron acceptors: isoprene, ferredoxin, FAD or FMN. Subsequent experiments should then test whether ferredoxin is reduced alongside isoprene, supporting the bifurcation contention (Bertsch *et al.*, 2013). Once purified, analytical gel filtration, as well as separation on a native gel together with peptide mass finger-printing of the isoprene reductase can be applied to see whether the isoprene reductase is composed of different subunits.

Follow up experiments would involve revealing the protein structure of the isoprene reductase or isoprene reductase complex to gain insights into its reduction mechanism. Biocatalytic hydrogenation of isoprene to methylbutene is an enzymatic challenge since

non-activated C=C bonds are difficult targets for a nucleophilic attack. No enzyme capable of performing a direct reduction of a non-activated C=C bond has been isolated or characterized so far which makes the mechanism in the isoprene reduction process highly interesting.

6.4 Enrichment of other isoprene reducing organisms

Isoprene reduction was found in wetlands samples from three different sites but no *Acetobacterium* sp. were identified in the community analysis suggesting the isoprene hydrogenation reaction is catalysed by other organisms also. The fact that phylogenetic analysis of 5587 homologs that are also associated with the Hyp proteins were widely distributed among anaerobic organisms supports this hypothesis. The long version of 5587 homologs found in Proteobacteria could also be involved in isoprene reduction. Further enrichment and isolation steps from isoprene reducing wetland cultures are needed to identify additional organisms capable of isoprene reduction.

6.5 Measuring methylbutene and isoprene concentration in anaerobic sediments and soil

Methylbutene isomers at nM levels were detected in the head space of wetland samples from three different locations around Sydeny, however, isoprene could not be detected. Exogenous addition of isoprene to these samples resulted in reciprocal methylbutene production. This finding strengthens the argument that naturally occurring low concentrations methylbutenes originated from endogenous undetectable isoprene production. Future experiments will explore isoprene and methylbutene amounts in sediment samples by extracting isoprene and methylbutene from sediment/soil directly. Soil adsorption experiments of isoprene and other BVOC have to be carried out to determine environmental concentrations of these compounds.

Measurements also need to be conducted using the static chamber technique (Livingston and Hutchinson, 1995; Pihlatie *et al.*, 2013) in different wetlands field sites covering wetland sediments only, to measure isoprene and methane gas fluxes over years. Isoprene has an inhibitory effect on methane formation in *in vitro* samples but whether it actually affects methane emissions in wetlands *in situ* remains to be determined. To explore whether BVOCs shown to be inhibitory to methanogens e.g. ethene, acetylene, isoprene have regulatory effects on methane emissions, field studies need to be carried

out. BVOCs can be artificially added to chambers, into the headspace and into the sediment, and changes in methane emissions can be compared to untreated controls.

Isoprene is thought to be mostly produced by oxygenic phototrophs. However, it is possible that anaerobic isoprene sources in soils and wetlands are not properly explored and that isoprene quantities are higher than expected in these environments, given the standing concentration is a function of production and consumption. Unsaturated soil void space could contain locally higher concentrations of gases and BVOC (Garbeva *et al*., 2014).

6.6 Fate of methylbutenes

Data from this study showded that methylbutenes occurred naturally. In future work the fate of methylbutene isomers will be explored under aerobic and anaerobic culturing conditions using stable isotope probing.

6.7 Toxicity effect of isoprene and methylbutenes

Whether isoprene is a specific or non-specific inhibitor for methanogens and what mechanism the inhibition follows requires further investigation. Data from this study suggest a possible inhibition mechanism through interference with isoprenoid synthesis and therefore could inhibit archaeal growth. To exclude that it is not a general toxicity effect different isoprene concentration will be tested on multiple pure bacterial cultures to see if the inhibition is restricted to methanogens. Furthermore, different concentrations of all three methylbutenes (2-methyl-1-butene, 3-methyl-1-butene, 2-methyl-2-butene) as well as 2-methylbutane will be tested for their inhibitory effects on methanogens e.g. multiple pure cultures and mixed communities and on pure bacterial cultures.

6.8 Reduction of other unsaturated hydrocarbons

Ethene (C_2H_4), which contains a single double bond and is another volatile alkene emitted by plants, was found to act as an electron acceptor under anaerobic conditions which can be reduced to ethane (De Bruin *et al.*, 1992; Koene-Cottaar and Schraa, 1998; Mundle *et al.*, 2011; Elsgaard, 2013). A pure bacterial culture or responsible enzymes have not been identified yet and *A. wieringae* ISORED-2 did not reduce ethene (Kronen *et al.*, 2019). In future work ethene reducing enrichment cultures could be enriched and, like in the case of isoprene, proteogenomic analysis can be used to

identify putative ethene reducing enzymes for subsequent demonstration of direct involvement. Depending on which enzymes are involved it would be of great interest to compare them to the putative isoprene reductase and find out if these enzyme(s) belong to the same protein family.

In addition, enrichment cultures with other unsaturated hydrocarbons that commonly occur in anaerobic environments can be set up and tested for their reducibility. Examples include monoterpenes, which consist out of two isoprene subunits ($C_{10}H_{16}$) and are also BVOCs produced by plants e.g. alpha terpinene, alpha pinene, limonene and carveol as well as longer terpenes. A wide range of anaerobic bacteria contain 5587-like homologs but their amino acid sequence identity to 5587 is low, the highest being 60%. This raises the question of whether these proteins are able to reduce C=C double bonds from substrates similar to isoprene. Perhaps unsaturated hydrocarbons (C=C bonds) in general, are actually important, broadly used electron acceptors in anaerobic environments and isoprene, ethene and caffeate just happen to be the first ones to be identified.

6.9 Final conclusion

This study is the first demonstration of anaerobic consumption of the important atmospheric gas isoprene. Surprisingly isoprene is reduced rather than oxidized in anaerobic environments. This is also the first demonstration of natural methylbutene production which shows that isoprene reduction is a metabolism found in natural environments. The putative isoprene reductase may be the first putative enzyme capable of reducing non-activated double bonds. This study also shows that unsaturated hydrocarbons may be important, commonly used electron acceptors for anaerobic respiration processes.

6.10 References

Bertsch, J., Parthasarathy, A., Buckel, W., and Müller, V. (2013) An electron-bifurcating caffeyl-CoA reductase. *J. Biol. Chem.* **288**: 11304–11311.

De Bruin, W.P., Kotterman, M.J.J., Posthumus, M.A., Schraa, G., and Zehnder, A.J.B. (1992) Complete biological reductive transformation of tetrachloroethene to ethane. *Appl. Environ. Microbiol.* **58**: 1996–2000.

Elsgaard, L. (2013) Reductive transformation and inhibitory effect of ethylene under methanogenic conditions in peat-soil. *Soil Biol. Biochem.* **60**: 19–22.

Garbeva, P., Hordijk, C., Gerards, S., and De Boer, W. (2014) Volatile-mediated interactions between phylogenetically different soil bacteria. *Front. Microbiol.* **5**: 1–9.

Koene-Cottaar, F.H.M. and Schraa, G. (1998) Anaerobic reduction of ethene to ethane in an enrichment culture. *FEMS Microbiol. Ecol.* **25**: 251–256.

Kronen, M., Lee, M., Jones, Z.L., and Manefield, M.J. (2019) Reductive metabolism of the important atmospheric gas isoprene by homoacetogens. *ISME J.* **13**: 1168–1182.

Livingston, G.. and Hutchinson, G.L. (1995) Enclosure-based measurement of trace gas exchange: application and sources of error. In, Matson, P.A., Harriss,R.C. (ed), Biogenic Trace Gases: Measuring Emissions from Soil and Water. Backwell Science, Cambridge, pp. 14–50.

Mundle, S.O.C., Johnson, T., Lacrampe-Couloume, G., Perez-de-Mora, A., Duhamel, M., Edwards, E.A., et al. (2011) Monitoring biodegradation of ethene and bioremediation of chlorinated ethenes at a contaminated site using compound-specific isotope analysis (CSIA). *Environ. Sci. Technol.* **46**: 1731–1738.

Pihlatie, M.K., Christiansen, J.R., Aaltonen, H., Korhonen, J.F.J., Nordbo, A., Rasilo, T., et al. (2013) Comparison of static chambers to measure CH_4 emissions from soils. *Agric. For. Meteorol*. **171–172**: 124–136.