

Characterization of a limonene- degrading methanogenic enrichment culture

Dissertation

zur Erlangung des Doktorgrades

der Naturwissenschaften

- Dr. rer. nat. -

Dem Fachbereich 2 Biologie/Chemie

der Universität Bremen

vorgelegt von

Almud Lonsing

Bremen, 7. November 2024

Die vorliegende Doktorarbeit wurde im Rahmen des
Programms „International Max Planck Research School of
Marine Microbiology (MarMic)“ in der Zeit von Oktober 2019
bis Oktober 2024 am Max Planck Institut für Marine Mikrobiologie
angefertigt.

This thesis was prepared under the framework of the
“International Max Planck Research School of Marine
Microbiology (MarMic)” at the Max Planck Institute for Marine
Microbiology from October 2019 until October 2024.

Gutachter: Prof. Dr. Jens Harder

Gutachter: Dr. Gunter Wegener

Prüfer: Prof. Dr. Michael Friedrich

Prüfer: Prof. Dr. Rudolf Amann

Datum des Promotionskolloquiums: 6.12.2024

"The closer you look, the more there is to see."

- unknown

Summary

In this study, a limonene-degrading methanogenic enrichment culture was analyzed, revealing new insights into the predatory microbial interactions and potential limonene degradation pathways. The enrichment culture, established in 1999 using activated sludge obtained from a wastewater treatment plant, sustains a complex microbial community despite being supplied only with limonene as a carbon source.

Using long-read PacBio sequencing, a metagenome-assembled genome assigned to the *Syntrophobacteraceae* family was identified as responsible for the initial step of limonene degradation under methanogenic conditions. Within the genome of the *Syntrophobacteraceae* MAG, two operons were identified, with metatranscriptomic and metaproteomic data verifying the transcription of the genes and the presence of the proteins in the metaproteome. The initial activation of limonene is catalyzed by an enzyme annotated as benzylsuccinate synthase, which forms a new phylogenetic branch within the known and characterized fumarate-adding enzymes and was therefore named limonenylsuccinate synthase. Limonenylsuccinate synthase is possibly the first enzyme detected for monoterpene degradation in a methanogenic environment.

With 32 MAGs assembled from the metagenome, the microbial community in the enrichment culture is a complex syntrophic community. Most members of the community can be assigned to syntrophic bacteria, methanogenic archaea, or fermenters, who work together to degrade limonene. However, there is a predator thriving in the culture. *Candidatus Velamenicoccus archaeovorus*, a member of the candidate phylum Omnitrophota, that was visualized using CARD-FISH in combination with high-resolution microscopy and was found attached to the methanogen *Methanotherix soehngenii* and other microorganisms within the enrichment culture. *Candidatus Velamenicoccus archaeovorus* transcribes genes that allow it to attach to its prey, lyse the cells, and consume their biomass, showing similarities to the mechanisms of other predatory bacteria. Furthermore, a group I intron was detected in the predator's 23S rRNA gene and could be visualized using CARD-FISH in its prey cells, showing the transfer of mobile genetic elements between prey and host.

The predation process introduces a new trophic level into the microbial community as it results in the release of necromass. The necromass, consisting of carbohydrates, lipids, and proteins

Summary

from dead cells, is expected to be fermented by other members of the community, such as *Anaerolineaceae* and *Lentimicrobium*, closing the microbial loop and facilitating further syntrophic interactions. The complex interplay between syntrophic bacteria, methanogens, and the predatory bacterium illustrates the intricate dynamics within this anaerobic system.

The discovery of the novel limonenylsuccinate synthase enzyme and the predatory role of *Candidatus Velamenicoccus archaeovorus* highlights the metabolic complexity of microbial biomass turnover in energy-limited ecosystems.

Zusammenfassung

In dieser Arbeit wurde eine Limonen-abbauende methanogene Anreicherung analysiert, wodurch neue Einblicke in den Lebensstil eines räuberischen Bakteriums sowie einen potenziellen Abbauweg von Limonen gewonnen wurden. Die Anreicherungskultur, die 1999 mit Belebtschlamm aus einer Kläranlage inokuliert wurde, enthält Limonen als einzige Kohlenstoffquelle, das von einer komplexen mikrobiellen Gemeinschaft abgebaut wird.

Durch PacBio-Sequenzierung konnte ein Metagenom-assembliertes Genom (MAG) der Familie *Syntrophobacteraceae* identifiziert werden, welches für den ersten Schritt des Limonenabbaus unter methanogenen Bedingungen verantwortlich ist. Innerhalb des *Syntrophobacteraceae*-MAGs wurden zwei Operons identifiziert, wobei metatranskriptomische und metaproteomische Daten die Transkription der Gene und das Vorhandensein der Proteine im Metaproteom bestätigten. Die initiale Aktivierung von Limonen wird von einem Enzym katalysiert, das als Benzylsuccinat-Synthase annotiert wurde und einen neuen phylogenetischen Zweig innerhalb der bekannten und charakterisierten Fumarat-addierenden Enzyme bildet. Daher wurde es als Limonenylsuccinat-Synthase benannt. Limonenylsuccinat-Synthase ist möglicherweise das erste nachgewiesene Enzym für den Monoterpenabbau in einer methanogenen Umgebung.

Mit 32 aus dem Metagenom assemblierten MAGs zeigt die mikrobielle Gemeinschaft in der Anreicherungskultur eine komplexe syntrophe Struktur. Die meisten Mitglieder der Gemeinschaft lassen sich syntrophen Bakterien, methanogenen Archaeen oder Fermentierern zuordnen, die gemeinsam Limonen abbauen. In der Kultur existiert jedoch ein räuberisches Bakterium. *Candidatus Velamenicoccus archaeovorus*, ein Mitglied des Kandidatenphylums Omnitrophota, wurde mittels CARD-FISH in Kombination mit hochauflösender Mikroskopie visualisiert und zeigte sich an den Methanogen *Methanotherix soehngenii* und andere Mikroorganismen der Anreicherungskultur angeheftet. *Candidatus Velamenicoccus archaeovorus* transkribiert Gene, die es ihm ermöglichen, sich an seine Beute anzuheften, deren Zellen zu lysieren und deren Biomasse zu konsumieren, wobei Mechanismen verwendet werden, die denen anderer räuberischer Bakterien ähneln. Darüber hinaus wurde ein Intron im 23S-rRNA-Gen des Räubers nachgewiesen und konnte mithilfe von CARD-FISH in den Beutezellen visualisiert werden, was auf den Transfer mobiler genetischer Elemente zwischen Beute und Wirt hinweist.

Zusammenfassung

Der räuberische Prozess führt zu einer neuen trophischen Ebene in der mikrobiellen Gemeinschaft, da er zur Freisetzung von Nekromasse führt. Diese Nekromasse, bestehend aus Kohlenhydraten, Lipiden und Proteinen abgestorbener Zellen, wird dann von anderen Mitgliedern der Gemeinschaft, wie *Anaerolineaceae* und *Lentimicrobium*, fermentiert. Dadurch wird der mikrobielle Kreislauf geschlossen und weitere syntrophe Interaktionen werden erleichtert. Das komplexe Zusammenspiel zwischen syntrophen Bakterien, Methanogenen und dem räuberischen Bakterium veranschaulicht die vielfältigen Dynamiken innerhalb dieses anaeroben Systems.

Die Entdeckung der neuen Limonenylsuccinat-Synthase und die räuberische Rolle von *Candidatus Velamenicoccus archaeovorus* zeigen, wie komplex mikrobielle Gemeinschaften und deren Interaktionen in methanogenen Systemen sein können.

Abbreviations

ACR	alkyl-coenzyme M reductases
ANI	average nucleotide identity
ANME	anaerobic methanotrophic archaea
Ass	alkylsuccinate synthase
ATP	adenosine triphosphate
BALOs	<i>Bdellovibrio</i> and Like Organisms
Bcr	benzyl-CoA reductase
BLAST	Basic Local Alignment Search Tool
Bp	base pair
Bss	benzylsuccinate synthase
<i>Cand.</i>	<i>Candidatus</i>
CARD	catalyzed reporter deposition
CDD	Conserved Domain Database
cds	coding sequences
CPR	candidate phyla radiation
CRISPR	clustered regularly interspaced short palindromic repeats
DAPI	4',6-diamidino-2-phenylindole
DIET	direct interspecies electron transfer
DIC	differential interference contrast
DNA	deoxyribonucleic acid
DOM	dissolved organic matter
FA	formamide
FDRs	false discovery rates
FDH	formate dehydrogenases
FISH	fluorescence <i>in situ</i> hybridization
GTDB	Genome Taxonomy Database
GPI	glycosylphosphatidylinositol
Hase	hydrogenases

Abbreviations

HMN	2,2,4,6,8,8-heptamethylnonane
HRP	horseradish peroxidase
kDa	kilodalton
kV	kilovolt
LPS	lipopolysaccharides
M	molar (mol/L)
MAG	metagenome-assembled genome
MAT	mosaic adhesive trimer
Mbp	megabase pair
Mcr	methyl-coenzyme M reductase
min	minute
mL	milliliter
NAD	nicotinamide adenine dinucleotide
NCBI	National Center for Biotechnology Information
ng	nanogram
NGS	next-generation sequencing
nm	nanometer
OD	optical density
OP	Obsidian Pool
ORF	open reading frame
ori	origin of replication
OTUs	operational taxonomic units
PBS	phosphate buffered saline
PCR	polymerase chain reaction
POM	particulate organic matter
ppm	parts per million
PVC	<i>Planctomycetota–Verrucomicrobiota–Chlamydiota</i>
rRNA	ribosomal ribonucleic acid
RNA	ribonucleic acid
rpm	revolutions per minute

SAM	s-adenosylmethionine
S	svedberg
SDS	sodium dodecyl sulfate
SEM	scanning electron microscopy
SR-SIM	super-resolution structured illumination microscopy
Tat	twin-arginine translocation
TE	tris-EDTA (buffer)
TEM	transmission electron microscopy
Tg	teragram
Tad	tight-adherence complex
tRNA	transfer ribonucleic acid
TPM	transcripts per million
μL	microliter
VOC	volatile organic compounds

Table of Contents

Summary	I
Zusammenfassung.....	III
Abbreviations	V
Chapter 1: Introduction	1
1.1 Limonene - an abundant monoterpene in the environment.....	1
1.2 Wastewater treatment and methanogenic communities	2
1.2.1 Fermenting bacteria and anaerobic hydrocarbon degradation in anoxic systems	2
1.2.2 Syntrophic bacteria utilize fermentation products	3
1.2.3 Methanogenic archaea produce methane on a limited number of substrates	4
1.3 The fate of biomass in methanic systems	5
1.3.1 Viruses	5
1.3.2 Predatory bacteria and the phylum Omnitrophota.....	6
1.4 A microbial loop in methanic systems - the recycling of cell biomass	11
1.5 A limonene-degrading methanogenic enrichment culture	12
1.6 Aim of this study	13
References	14
Chapter 2: <i>Methanosaeta</i> and “ <i>Candidatus Velamenicoccus archaeovor</i> us”	23
Abstract	24
2.1 Introduction	25
2.2 Results	26
2.3 Discussion	36
2.4 Material and Methods.....	38
References	44
Supplementary material.....	49
Chapter 3: Anaerobic limonene metabolism in a methanogenic enrichment involves a glycine radical enzyme	95
Abstract	96

3.1 Introduction.....	97
3.2 Material and Methods	100
3.3 Results.....	104
3.3.1 Cell fractionation and – omic datasets	104
3.3.2 Community members of the limonene-degrading culture.....	104
3.3.3 Evidence for hydrocarbon metabolism in the - omic datasets	108
3.3.4 Limonene degradation in <i>Syntrophobacteraceae</i> LiM6	109
3.3.5 Phylogenetic analysis of the benzylsuccinate synthase alpha subunit	115
3.4 Discussion.....	116
References.....	119
Supplementary material	128
Chapter 4: New insights into the biology of <i>Candidatus</i> Velamenicoccus archaeovor.....	132
Abstract.....	133
4.1 Introduction.....	134
4.2 Material and Methods	137
4.3 Results and Discussion	140
4.3.1 Structured illumination micrographs show <i>Candidatus</i> Velamenicoccus archaeovor..... attached to morphologically different cells.....	140
4.3.2 Structured illumination micrographs of Nile Red stained cells	143
4.3.3 SR-SIM images visualize the presence of a predator group I intron molecule in prey cells.....	146
4.3.4 Transcriptomic analysis of <i>Candidatus</i> Velamenicoccus archaeovor.....	151
References.....	158
Supplementary material	165
Chapter 5: Discussion.....	175
5.1 Trophic interactions in the limonene-degrading methanogenic enrichment culture	175
5.2 Necromass in methanogenic communities and its consumers.....	176
5.3 Comparison with other methanogenic hydrocarbon-degrading communities.....	179

Table of contents

5.4 Novel insights into anaerobic limonene degradation under methanogenic conditions	182
5.5 Omnitrophota: origins, classification, and characteristics.....	183
5.6 A group I Intron in the limonene-degrading enrichment culture	185
5.7 Outlook and future directions.....	186
References	189
Acknowledgments.....	193

Chapter 1: Introduction

1.1 Limonene - an abundant monoterpene in the environment

Volatile organic compounds (VOCs) are a group of chemicals characterized by their high vapor pressure at room temperature, which leads to their significant presence in the atmosphere. VOCs are emitted from a wide range of anthropogenic sources, e.g., paints, cleaning products, and cosmetics, and natural sources, e.g., vegetation and microbial activity (Atkinson and Arey, 2003; Krassi Rumchev *et al.*, 2007; Korpi *et al.*, 2009). In marine environments, the primary VOC producers are phytoplankton, e.g., the diatom *Thalassiosira pseudonana* or the green microalga *Dunaliella tertiolecta*, which produce VOCs as a byproduct of photosynthesis. These VOCs do not necessarily go into the atmosphere but have been shown to be consumed by bacteria, showing the importance of VOCs in the marine carbon cycle (Halsey *et al.*, 2017; Moore *et al.*, 2020). Not only phytoplankton but also higher plants produce a large variety of VOCs with thousands of compounds identified (Knudsen and Gershenzon, 2006). These VOCs let plants interact with their environment e.g., to communicate with other plants, attract pollinators, and repel herbivores (Dudareva *et al.*, 2006). Monoterpenes ($C_{10}H_{16}$) are a class of VOCs that consist of two isoprene (C_5H_8) units (Marmulla and Harder, 2014). They are secondary metabolites predominantly produced by plants and are significant contributors to the natural emissions of VOCs into the atmosphere (Loreto and Fares, 2013; Bouwmeester *et al.*, 2019). Plants are estimated to release 107.5 Tg of carbon per year as monoterpenes into the atmosphere (Messina *et al.*, 2016). Among these, limonene is an abundant monoterpene present in the essential oils of citrus fruits such as oranges, lemons, and mandarins (Espina *et al.*, 2011). In higher plants, limonene is produced and stored in the secretory cavities of the flavedo (the outermost layer) of citrus fruits, particularly in their peel (Lücker *et al.*, 2002). Its function can be complex, acting as a defense mechanism against herbivores and pathogens but also as an attractant for parasitoids, which help the plant control pests (Rodríguez *et al.*, 2011; Mohammed *et al.*, 2020).

In 2019, the global production of citrus fruit was 144 Gg, including 76 Gg of oranges, approximately 53% of citrus fruit production (FAO, 2021). Limonene is the most abundant volatile component in orange juice (Moufida and Marzouk, 2003), and the peel of citrus fruit

contains approximately 4% (w/w dry basis) limonene (Negro *et al.*, 2016), making it a substantial part of the waste from the orange juice industry. Because of its pleasant lemon-like scent and flavor, it is often added as a food additive to sweets, soft drinks, and ice cream, as well as to cosmetic products such as soaps and perfumes (Ozturk *et al.*, 2019; Li *et al.*, 2023). Due to its many applications in the human environment, limonene is an abundant VOC and is also present in wastewater, where concentrations of up to 420 µg/l were measured in Swedish wastewater (Potter *et al.*, 2014).

1.2 Wastewater treatment and methanogenic communities

Microbial processes such as activated sludge treatment play an important role in wastewater treatment as they remove a large variety of compounds. The organic matter in wastewater is degraded by microorganisms in aeration tanks under aerobic conditions. Dissolved organic matter is also transformed into bacterial biomass. Undegraded particles and biomass remain in the activated sludge of these aeration tanks. The sludge is then transferred into anaerobic digesters, where the organic matter is broken down in stages by different microorganisms and ultimately converted into methane and carbon dioxide. The methane can then be used as biogas, an energy source for the wastewater treatment plant (Yildiz, 2012; Venkiteshwaran *et al.*, 2015; Robles *et al.*, 2018).

1.2.1 Fermenting bacteria and anaerobic hydrocarbon degradation in anoxic systems

In anaerobic digesters, organic matter is initially degraded by fermenting microorganisms. Fermentation is a metabolic process to generate energy under anoxic conditions, where organic matter is converted to reduced compounds, i.e., alcohols or hydrogen, and oxidized compounds, i.e., organic acids or carbon dioxide, without an external electron acceptor. This process is crucial for bacteria in anoxic systems, allowing them to generate energy and regenerate NAD⁺ by reducing organic molecules (Schink, 1997; Ahring, 2003). Types of fermentation include lactic acid fermentation (e.g. *Lactobacillus*), which converts glucose to lactic acid; alcoholic fermentation (e.g. *Zymomonas mobilis*), which produces ethanol and CO₂ from glucose; mixed acid fermentation (e.g. *Escherichia coli*), yielding a mixture of acids and gases; and butyric acid fermentation (e.g. *Clostridium*), producing butyric acid and solvents like acetone and butanol (Bahl *et al.*, 1982; Barbosa *et al.*, 1994; Liu, 2003; Förster and Gescher, 2014). Fermentation

to lactate yields 2 ATP per glucose molecule, which is less than aerobic respiration with O₂ as the terminal electron acceptor, which yields up to 38 ATP per glucose molecule (Werner *et al.*, 2010; Rowe *et al.*, 2013). Fermentation is a metabolic process with a low free energy conservation per reaction (Schink, 1997; Gupta *et al.*, 2022).

Anaerobic hydrocarbon degradation is another process performed by microorganisms. One key mechanism is fumarate addition, where enzymes such as benzylsuccinate synthase catalyze the addition of fumarate to hydrocarbons like alkanes and aromatic compounds, initiating their breakdown without the need for oxygen (Heider *et al.*, 2016).

1.2.2 Syntrophic bacteria utilize fermentation products

Syntrophic bacteria degrade the fermentation products under anaerobic conditions and produce intermediate substrates like hydrogen, formate, and acetate. The anaerobic degradation of fermentation products, e.g., propionate, butyrate, and ethanol, is only thermodynamically favorable under specific conditions where the products - hydrogen, formate, and acetate - are constantly consumed in a terminal electron acceptor process. This cooperation is called syntrophic interaction (McInerney *et al.*, 2009; Schmidt *et al.*, 2016). The production of H₂ during the oxidation of fermentation products makes the reactions energetically unfavorable unless the hydrogen is rapidly removed (Schink and Friedrich, 1994). Here the methanogens play an important role by consuming the hydrogen to produce methane (CH₄), keeping the partial pressure of H₂ low and the reaction thermodynamically favorable (Schink, 1997). Studies have shown that not only substrates, e.g., hydrogen or formate, transfer electrons between species but that a direct interspecies electron transfer (DIET) occurs between syntrophic partners like *Geobacter metallireducens* and *Methanosarcina barkeri*. *Geobacter metallireducens* produces conductive pili to create a direct cell-to-cell contact to transfer electrons, allowing the conversion of ethanol to methane and carbon dioxide without a low molecular weight intermediate (Rotaru *et al.*, 2014; Lovley, 2017).

1.2.3 Methanogenic archaea produce methane on a limited number of substrates

Methanogenic archaea are responsible for the final step of anaerobic digestion, converting a small range of substrates like hydrogen, carbon dioxide, and acetate into methane and carbon dioxide. Living under anoxic conditions and using substrates that are already partially oxidized and have low energy yields, they live close to the thermodynamic limit, with one acetate producing less than one ATP under environmental conditions (Deppenmeier and Müller, 2008). Methanogens such as *Methanosarcina* and the acetoclastic *Methanothrix* (formerly published as *Methanosaeta*) are often found to be among the most common methanogenic archaea in biogas reactors, indicating their important role in anaerobic digestion (Guo *et al.*, 2015).

Methanogenic archaea use a limited range of substrates involving hydrogenotrophic, methylotrophic, and acetoclastic pathways (Conrad, 2020). Hydrogenotrophic methanogens use hydrogen (H₂) and carbon dioxide (CO₂) to form methane (CH₄), using H₂ as the electron donor and CO₂ as the terminal electron acceptor. Additionally, some hydrogenotrophic methanogens can utilize formate as an alternative electron donor (Thauer *et al.*, 2008). In this case, formate is first oxidized to produce H₂ and CO₂, which are then used in the standard hydrogenotrophic pathway (Long *et al.*, 2017; Conrad, 2020; Halim *et al.*, 2021). Hydrogenotrophic methanogenesis is performed by genera such as *Methanococcus*, *Methanobacterium*, *Methanothermobacter*, *Methanoculleus*, *Methanospirillum*, *Methanoregula*, and *Methanobrevibacter* (Bräuer *et al.*, 2006; Parshina *et al.*, 2014; Richards *et al.*, 2016; Joshi *et al.*, 2018). All these genera can also use formate but there are exceptions on the species level e.g. *Methanobacterium bryantii* which does not use formate (Peters *et al.*, 1999; Schink *et al.*, 2017).

Methylotrophic methanogens use substrates like methylamines, methanol, and various methylated sulfur compounds. This methylotrophic methanogenesis pathway is primarily found in members of the *Methanosarcinales* order. Methylotrophic methanogenesis is particularly important in saline environments, where methanogens are outcompeted by sulfate-reducing bacteria for substrates like hydrogen and acetate (Zhuang *et al.*, 2016; Kurth *et al.*, 2020; Li *et al.*, 2020).

Acetoclastic methanogens, such as *Methanosarcina* and *Methanothrix*, utilize acetate as their substrate for methane production (Welte and Deppenmeier, 2014). Acetate is an intermediate in the anaerobic breakdown of organic matter and its accumulation can inhibit other microbial processes, especially the activity of syntrophic bacteria. The conversion of acetate is, therefore, important for maintaining the flow of carbon through anaerobic ecosystems (Xu *et al.*, 2014).

The methane (CH₄) produced in these methanic environments is well known as a potent greenhouse gas with a global warming potential significantly higher than that of carbon dioxide (CO₂) (Badr *et al.*, 1991) and also a valuable energy source, particularly in the context of biogas production from wastewater treatment processes (Mao *et al.*, 2015). As methane is mostly of microbial origin, e.g., in animal digestive systems, it is also of interest in the search for extraterrestrial life and has been detected on Mars (Janssen and Kirs, 2008; Hook *et al.*, 2010; Webster *et al.*, 2015). In the environment, methane can be consumed as a carbon and energy source by methanotrophs (Hanson and Hanson, 1996).

1.3 The fate of biomass in methanic systems

Methanic systems are unique microbial ecosystems characterized by methanogenesis as the central process of their energy conservation and carbon cycling. While this process can be described with simplified categories, the reality remains complex. Questions regarding the fate of the formed biomass remain. To what extent do viruses and predatory bacteria degrade the biomass in methanic habitats? What role do they play in the microbial loop, a term that outlines how dissolved organic matter (DOM) and energy flow through microorganisms in aquatic systems?

1.3.1 Viruses

Viruses, particularly bacteriophages and archaeal viruses, consist of genetic material (DNA or RNA) encased in a protein coat and lack the cellular components typically found in living organisms. They cannot replicate on their own and infect a host cell to insert their genetic information and produce new viral particles. Viruses do not metabolize substrates, consume energy, or produce waste on their own, which has challenged the definition of a living organism (Weinbauer, 2004). In methanogenic communities, viruses play critical roles by infecting and lysing microbial cells, for example, methanogenic archaea, which contributes to the turnover of biomass and the release of organic matter and influences the population dynamics within these communities (Weinbauer, 2004; Thiroux *et al.*, 2021). A new lytic virus named Blf4 which infects archaea, specifically the hydrogenotrophic methanogen *Methanoculleus bourgensis*, was isolated from a biogas plant in Germany (Weidenbach *et al.*, 2021).

Stable isotope probing coupled with metagenomics has revealed multiple new viruses that infect methanogenic archaea. These viruses were described as spindle-shaped, of which several near-complete genomes were discovered, and head-tailed virus, of which a complete genome was discovered. The head-tailed virus infects *Methanobacterium* and has been proposed to belong to a new family named *Speroviridae*. This research highlights the extensive diversity of viral genomes in methanogenic environments and underscores their ecological significance (Ngo *et al.*, 2022).

A study analyzed the virome of mesophilic anaerobic digesters using metagenomic analysis found *Siphoviridae* as the dominant viral family, followed by *Myoviridae* and *Podoviridae* (Bhattarai *et al.*, 2024). These families are double-stranded DNA viruses with icosahedral head structure and variations in their tail structure (Ackermann, 2009). Most analyzed MAGs from the anaerobic digester contained at least one viral genome, indicating that lysogeny is prevalent. Lysogeny is a process where viral DNA is integrated into the host genome. This viral-mediated gene transfer can potentially increase the fitness of methanogens by introducing new functional genes into the host. In a *Methanotherix soehngenii*-affiliated MAG, the gene of a subunit of formate dehydrogenase (*fwdE*) was detected that originated based on sequence similarity from a hydrogenotrophic methanogen. This indicates the importance of viruses and virus-mediated gene transfer in anaerobic digesters, which can shape the metabolic pathways and ecological dynamics within anaerobic digesters (Bhattarai *et al.*, 2024). Recent studies on viruses attacking archaea complement phage studies in wastewater treatment plants, highlighting that all living biomass in methanic systems is subject to viral predation.

1.3.2 Predatory bacteria and the phylum Omnitrophota

Predatory bacteria gain carbon sources and energy by attacking and feeding on other bacteria and archaea. The first predatory bacterium discovered in 1962 was *Bdellovibrio bacteriovorus*, which preys on gram-negative bacteria by entering their periplasmic space and consuming their cellular contents (Socket, 2009). Other predatory bacteria with different predation strategies are, for example, *Micavibrio aeruginosavorus* which uses an epibiotic strategy, attaching to its prey's cell wall and consuming it from the outside (Pasternak *et al.*, 2014). *Myxococcus xanthus* has yet a different hunting strategy, forming complex multicellular structures and using tentacle-like packs to locate and attack their prey (Berleman and Kirby, 2009). Another predatory bacteria named *Vampirococcus*, which belongs to the Candidate Phyla Radiation

(CPR), is a gram-negative anaerobic bacterium that attaches to the surface of its prey - the phototrophic bacteria of the genus *Chromatium* (Guerrero *et al.*, 1986). *Vampirococcus lugosii*, a species within the genus *Vampirococcus*, is named after the actor Bela Lugosi, known for his portrayal of Dracula in the 1931 film (Moreira *et al.*, 2021). This bacterium uses a predatory mechanism reminiscent of vampires, attaching to its prey and consuming its cellular contents. It is a small (500–600 nm in diameter), non-flagellated round bacterium that has been described to have a reduced genome (1.3 Mbp), and genes encoding for proteins for biosynthetic pathways were missing (Moreira *et al.*, 2021). It has to acquire resources, e.g., lipids, vitamins, nucleotides, and amino acids from its prey. For example, as it cannot synthesize phospholipids, it is predicted to secrete a phospholipase that degrades the prey's membrane and generates a source of phospholipids, which it uses for cell membrane construction. Two murein DD-endopeptidases and two peptidoglycan hydrolases are also predicted to be secreted and, probably involved in degrading the host's cell wall. Of the two DD-endopeptidases, one is predicted to remodel the prey's cell wall, as has been described for *Bdellovibrio*. *Bdellovibrio* modifies the peptidoglycan layer of the prey bacterium, which turns it into a rounded, osmotically stable structure called a bdelloplast (Lerner *et al.*, 2012). The DD endopeptidases from *Bdellovibrio* are missing the regulatory domain III, which is associated with *Bdellovibrio*'s ability to predate and modify the cell walls of a variety of different bacteria. The presence of the regulatory domain III in the DD-endopeptidases from *Vampirococcus* possibly reflects its narrow host range which is limited to *Chromatiaceae* (Moreira *et al.*, 2021). The genome of *Vampirococcus lugosii* also contains genes for transporters and Type IV pili, which are likely involved in attachment to the host and the uptake of host DNA, which is then degraded by restriction endonucleases and recycled for DNA synthesis. Furthermore, defensive proteins which contain alpha-2-macroglobulin protease-inhibiting domains are predicted to protect *Vampirococcus lugosii* itself from host proteases.

The phylogeny of predatory bacteria has undergone significant revisions in recent years. Extensive phylogenetic analyses have led to the reclassification and establishment of new phyla, including *Bdellovibrionota* and *Myxococcota*. The new phyla were distinguished based on single-copy marker genes, rRNA genes, differences in their ecological roles, and metabolic capabilities, particularly focusing on their predatory strategies. For instance, *Bdellovibrionota* includes obligate predators that invade the periplasm of gram-negative bacteria, whereas *Myxococcota* is known for its group predation strategies. The phylogenetic analysis also indicated that the predatory behaviors observed in these groups evolved independently, and the

two new phyla reflect their different evolutionary pathways (Waite *et al.*, 2020; Kamada *et al.*, 2023).

In wastewater treatment plants, predatory bacteria like *Bdellovibrio bacteriovorus* influence microbial community structure in floccular and granular sludge (Feng *et al.*, 2017). Methanogens can also become the target and prey of predatory bacteria. *Candidatus Velamenicoccus archaeovorus* is an epibiotic anaerobic ultramicrobacterium that preys on *Methanotherix* (formerly also published as *Methanosaeta*) and bacteria by attaching to their surface and metabolizing their outer cell wall (Kizina, 2017). It was discovered in a limonene-degrading methanogenic enrichment culture inoculated from activated sludge. *Candidatus Velamenicoccus archaeovorus* was shown to have two lifestyles, one as free-living cells and one attached to its prey (Figure 1) (Kizina *et al.*, 2022). When it is attached to its prey, *Candidatus Velamenicoccus archaeovorus* shows higher metabolic activity visualized by stronger fluorescence *in situ* hybridization (FISH) signals, which indicates a higher ribosome content compared to the free-living cells (Kizina, 2017; Kizina *et al.*, 2022). Genomic and proteomic data suggest that the predatory mechanism of *Candidatus Velamenicoccus archaeovorus* involves a gene for a very large multienzyme surface protein, which encodes for hundreds of protein domains, of which most are located outside the cell and anchored by transmembrane helices. These protein domains were considered to degrade the prey cell wall and facilitate the uptake of the substrates. More specifically, enzymes such as glycosyl transferases, glycosyl hydrolases, and peptidases are involved in degrading the polysaccharides and proteins of the prey's cell wall (Kizina *et al.*, 2022). A type II secretion system and type IV pilus assembly system are involved in the attachment of *Candidatus Velamenicoccus archaeovorus* to its prey and ATP-binding domains and kinases regulate the process of predation. *Candidatus Velamenicoccus archaeovorus* has the metabolic capacity to ferment sugars to acetate and conserve energy via phosphodiester bonds in nucleic acids. Altogether, a complex interplay of cellular components is involved in the process of finding, attaching, and consuming its methanogenic prey (Rotaru *et al.*, 2012; Kizina *et al.*, 2022).

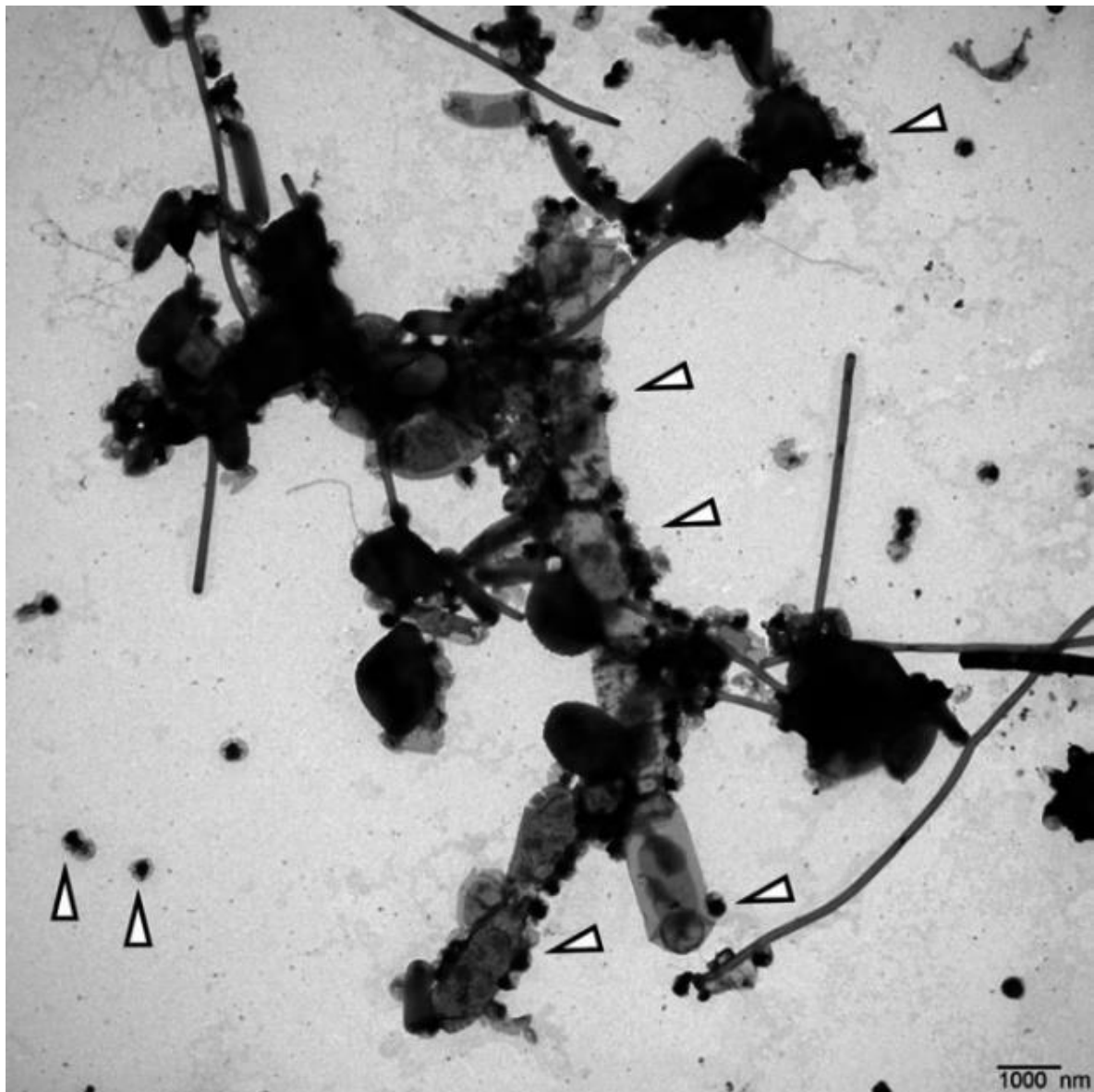


Figure 1: Transmission electron micrograph of the limonene-degrading methanogenic enrichment culture. The large filaments in the center, consisting of several cylindrical cells with spacer discs at their ends, were assigned to *Methanotherx* based on its characteristic morphology. The small coccoid-shaped cells, both attached to the filament and free-living, were assigned as *Candidatus Velamenicoccus archaeovorus* (marked by triangles) (Kizina *et al.*, 2022).

Phylogenetically, *Candidatus Velamenicoccus archaeovorus* is assigned to the candidate phylum OP3, recently termed Omnitrophota (Momper *et al.*, 2017; Parks *et al.*, 2018; Whitman *et al.*, 2018; Perez-Molphe-Montoya *et al.*, 2022). The phylum was initially identified from gene sequences recovered from Obsidian Pool in Yellowstone National Park, USA (Hugenholtz *et al.*, 1998). First insights based on 16S rRNA gene sequences identified at least five subdivision lineages and their placement in the PVC (*Planctomycetota–Verrucomicrobiota–*

Chlamydiota) superphylum was confirmed by its 23S rRNA sequence (Glöckner *et al.*, 2010). Reconstruction of 55 Omnitrophota MAGs from a German groundwater system revealed four new clades and genetic potentials for type IV pili, type II secretion systems, glycogen storage, and carbohydrate degradation, and one genome revealed a possible potential for magnetosome construction (Perez-Molphe-Montoya *et al.*, 2022).

Another study analyzed 349 existing and 72 newly sequenced genomes derived from metagenome-assembled genomes and single-cell amplified genomes (Seymour *et al.*, 2023). It was found that most Omnitrophota are ultra-small (~0.2 µm) bacteria, with gene clusters indicating a host-dependent lifestyle found in 64% of the genomes. A potential predatory lifestyle is supported by the presence of the tight-adherence (Tad) complex, which *Bdellovibrio* and like organisms (BALOs) use to attach to their prey or enter the prey cell (Seymour *et al.*, 2023). In an extension of this study, the very large multienzyme surface proteins were analyzed in environmental metagenomes (West-Roberts *et al.*, 2023). Proteins as large as 85,804 amino acids were identified in genomes and most prevalent in Omnitrophota, which have the highest average copy number of these very large genes (West-Roberts *et al.*, 2023). They are also prevalent in phyla such as *Firmicutes*, *Actinobacteriota*, *Candidatus* Elusimicrobiota, *Candidatus* Hinthialibacteriota, and *Planctomycetota* (West-Roberts *et al.*, 2023). Two categories of giant proteins were described, the archaeovorins and the dockerin-type giant proteins. Archaeovorins are characterized by domains involved in glycosylation and peptidase activity associated with predation on archaeal cell walls. Dockerin-type giant proteins are less common and characterized by domains related to extracellular matrix interactions, such as dockerin and thrombospondin type 3 repeats, which are suggested to be involved in binding to and degrading the cell walls of prey. These genes likely represent a key evolutionary adaptation that enables these bacteria to thrive as predators in various environments (West-Roberts *et al.*, 2023). Despite the advances made, studying Omnitrophota remains challenging due to their small size and the absence of laboratory cultures, with the exception of *Candidatus* Velamenicoccus archaeovorinus.

1.4 A microbial loop in methanic systems - the recycling of cell biomass

Current knowledge suggests that viruses and phages attack cells in methanic systems, contributing to a process known as the viral shunt (Suttle, 1994). Both, the viral shunt and predation by predatory bacteria lead to the release of dissolved organic matter (DOM) from cytosols, which again serve as substrate for microorganisms. This process is similar to the microbial loop observed in aqueous systems, where DOM is transformed into particulate organic matter (POM) (Fenchel, 2008). Although this process, along with the degradation of dead cells (necromass) in methanic systems, is crucial for understanding these environments, it has not yet been intensively studied.

A study has investigated the methanogenic microbial communities involved in degrading peptidoglycan, a component of bacterial cell walls (Quaiyum *et al.*, 2020). The microbial communities were enriched from rice paddy fields and anaerobic digesters and grown on purified peptidoglycan from *Escherichia coli* (gram-negative) and *Micrococcus luteus* (gram-positive) as the sole carbon and energy source. Methane production suggested that the *Escherichia coli* peptidoglycan was less degradable than that of *Micrococcus luteus* (Quaiyum *et al.*, 2020). Based on 16S rRNA gene analyses, the rice paddy soil and anaerobic digester enrichment cultures contained bacteria from the phyla *Armatimonadota*, *Bacteroidota*, *Cloacimonadota* (previously named WWE1), and *Verrucomicrobiota*, with many of these bacteria being uncultured and phylogenetically novel. Particularly within the *Bacteroidetes* phylum, several OTUs had less than 90% identity to known species. Although the enrichment cultures were overall similar regarding to their community structure, distinct microbial taxa were enriched depending on the peptidoglycan source. The *Escherichia coli* peptidoglycan culture was enriched with the phylum *Verrucomicrobiota* (OTU related to *Luteolibacter gellanilyticus*), while the *Micrococcus luteus* peptidoglycan culture was enriched with the phylum *Armatimonadota* (OTU related to *Oxalophagus oxalicus*). These differences highlight the substrate-specific selection of microbial communities involved in the degradation of peptidoglycan from gram-negative and gram-positive bacteria (Quaiyum *et al.*, 2020).

In other anoxic environments, such as deep sediments or subsurface groundwater, heterotrophic bacteria play a crucial role in recycling organic carbon from necromass. In Dead Sea sediments, members of the *Pseudomonadaceae* and *Flavobacteriaceae* families were found to recycle archaeal cell wall components into wax esters, forming a key survival strategy in nutrient-poor, hypersaline conditions (Thomas *et al.*, 2019). In groundwater systems, key players in

necromass degradation include bacteria from the genera *Flavobacterium*, *Massilia*, *Rheinheimera*, and *Rhodoferrax*, which rapidly metabolize necromass and support microbial communities across multiple trophic levels (Geesink *et al.*, 2022). This ability to utilize necromass provides a vital mechanism for sustaining microbial life in energy-limited environments, similar to methanogenic ecosystems.

1.5 A limonene-degrading methanogenic enrichment culture

In 1999, a limonene-degrading methanogenic enrichment culture, which is the focus of this study, was established by inoculating 1 μ L in a dilution-to-extinction series of an enrichment culture derived from activated sludge obtained at the local wastewater treatment plant in Lintel, Osterholz-Scharmbeck, Germany (Harder and Foss, 1999). After successfully establishing a stable culture on limonene, it was divided into twelve separate lineages in 2005 (Kizina, 2017). The cultures are grown in freshwater methanogenic media with 2,2,4,6,8,8-heptamethylnonane (HMN), which is the carrier phase for the R-(+)-limonene and transferred annually and produce methane for up to two years (Rotaru, 2009; Kizina, 2017). The microbial community's composition and structure were assessed using the full cycle rRNA approach in combination with catalyzed reporter deposition-fluorescence *in situ* hybridization (CARD-FISH) experiments (Amann *et al.*, 1995). Microscopic analysis revealed diverse and morphologically different cells within the enrichment, including small coccoid-shaped cells and long filaments. The community was composed of 40% bacteria and 33% archaea, as determined by CARD-FISH with domain-specific probes EUB338 I–III and ARCH-915. However, the EUB338 I–III probe mix does not target OP3 cells due to mismatches in the 16S rRNA sequence. OP3 cells identified by the OP3-565 probe accounted for 18% of all cells in the limonene-degrading methanogenic enrichment culture (Rotaru *et al.*, 2012; Kizina, 2017). Archaeal 16S rRNA gene sequences were mostly associated with members of the orders *Methanomicrobiales* and *Methanosarcinales*. Bacterial 16S rRNA gene sequences showed similarity to representatives from the *Bacteroidetes*, *Deltaproteobacteria*, Candidate Division OP3, and *Firmicutes* lineages (Rotaru *et al.*, 2012; Kizina, 2017). Although important insights have been made, many aspects of how this limonene-degrading methanogenic enrichment culture operates are still not fully understood.

1.6 Aim of this study

The limonene-degrading methanogenic enrichment culture was established in 1999 from a dilution to extinction series, initially inoculated from activated sludge obtained from a wastewater treatment plant in Lintel, Osterholz- Scharmbeck, Germany (Harder and Foss, 1999; Rotaru *et al.*, 2012). Although the culture is provided with a single substrate, limonene, it has a substantial microbial diversity (Rotaru *et al.*, 2012).

This study aims to improve our understanding of the biology and lifestyle of *Candidatus Velamenicoccus archaeovorus* and to identify the enzyme and organism responsible for limonene degradation in a limonene-degrading methanogenic enrichment culture. To achieve these aims, two primary approaches were used: High-resolution microscopy combined with various staining methods and a bioinformatic approach using metagenomic, metatranscriptomic, and metaproteomic data.

Confocal laser scanning microscopy was combined with fluorescence *in situ* hybridization (FISH) using probes targeting *Methanothrix* and *Candidatus Velamenicoccus archaeovorus*, as well as non-coding rRNA and an intron, together with live/dead staining and lipid staining with Nile Red. This approach visualized the limonene-degrading methanogenic enrichment culture as a whole and gave detailed insights into the predatory lifestyle of *Candidatus Velamenicoccus archaeovorus*.

Long-read PacBio sequencing was performed to generate high-quality metagenome-assembled genomes (MAGs), which formed the basis for metatranscriptomic and metaproteomic analyses. These analyses led to the identification of a new limonenylsuccinate synthase in a MAG of the *Syntrophobacteraceae* family and the proposal of a new limonene degradation pathway under methanogenic conditions. The analysis yielded more MAGs than expected for a hydrocarbon-degrading methanogenic enrichment culture, indicating the presence of necromass utilization.

References

- Ackermann, H.-W. (2009) Phage classification and characterization. In *Bacteriophages: Methods and Protocols, Volume 1: Isolation, Characterization, and Interactions*. Clokie, M.R.J. and Kropinski, A.M. (eds). Totowa, NJ: Humana Press, pp. 127–140.
- Ahring, B.K. (2003) Perspectives for anaerobic digestion. In *Biomethanation I*. Ahring, B.K., Angelidaki, I., de Macario, E.C., Gavala, H.N., Hofman-Bang, J., Macario, A.J.L., et al. (eds). Berlin, Heidelberg: Springer Berlin Heidelberg, pp. 1–30.
- Amann, R.L., Ludwig, W., and Schleifer, K.H. (1995) Phylogenetic identification and *in situ* detection of individual microbial cells without cultivation. *Microbiological reviews* **59**: (1) 143–69.
- Atkinson, R. and Arey, J. (2003) Atmospheric degradation of volatile organic compounds. *Chemical Reviews* **103**: (12) 4605–4638.
- Badr, O., Probert, S.D., and O’Callaghan, P.W. (1991) Atmospheric methane: Its contribution to global warming. *Applied Energy* **40**: (4) 273–313.
- Bahl, H., Andersch, W., Braun, K., and Gottschalk, G. (1982) Effect of pH and butyrate concentration on the production of acetone and butanol by *Clostridium acetobutylicum* grown in continuous culture. *European Journal of Applied Microbiology and Biotechnology* **14**: (1) 17–20.
- Barbosa, M.F., Yomano, L.P., and Ingram, L.O. (1994) Cloning, sequencing and expression of stress genes from the ethanol-producing bacterium *Zymomonas mobilis*: the groESL operon. *Gene* **148**: (1) 51–57.
- Berleman, J.E. and Kirby, J.R. (2009) Deciphering the hunting strategy of a bacterial wolfpack. *FEMS Microbiology Reviews* **33**: (5) 942–957.
- Bhattacharai, B., Bhattacharjee, A.S., Coutinho, F.H., and Goel, R. (2024) Investigating the viral ecology and contribution to the microbial ecology in full-scale mesophilic anaerobic digesters. *Chemosphere* **349**: 140743.
- Bouwmeester, H., Schuurink, R.C., Bleeker, P.M., and Schiestl, F. (2019) The role of volatiles in plant communication. *The Plant Journal* **100**: (5) 892–907.
- Bräuer, S.L., Cadillo-Quiroz, H., Yashiro, E., Yavitt, J.B., and Zinder, S.H. (2006) Isolation of a novel acidiphilic methanogen from an acidic peat bog. *Nature* **442**: (7099) 192–194.

- Conrad, R. (2020) Importance of hydrogenotrophic, acetoclastic and methylotrophic methanogenesis for methane production in terrestrial, aquatic and other anoxic environments: a mini review. *Pedosphere* **30**: (1) 25–39.
- Deppenmeier, U. and Müller, V. (2008) Life close to the thermodynamic limit: how methanogenic archaea conserve energy. In *Bioenergetics: Energy Conservation and Conversion*. Schäfer, G. and Penefsky, H.S. (eds). Berlin, Heidelberg: Springer Berlin Heidelberg, pp. 123–152.
- Dudareva, N., Negre, F., Nagegowda, D.A., and Orlova, I. (2006) Plant volatiles: recent advances and future perspectives. *Critical Reviews in Plant Sciences* **25**: (5) 417–440.
- Espina, L., Somolinos, M., Lorán, S., Conchello, P., García, D., and Pagán, R. (2011) Chemical composition of commercial citrus fruit essential oils and evaluation of their antimicrobial activity acting alone or in combined processes. *Food Control* **22**: (6) 896–902.
- FAO (2021) Citrus Fruit Statistical Compendium 2020. Available at: <https://www.fao.org/economic/est/est-commodities/citrus-fruit/en/>.
- Fenchel, T. (2008) The microbial loop - 25 years later. *Journal of Experimental Marine Biology and Ecology* **366**: (1) 99–103.
- Feng, S., Tan, C.H., Constancias, F., Kohli, G.S., Cohen, Y., and Rice, S.A. (2017) Predation by *Bdellovibrio bacteriovorus* significantly reduces viability and alters the microbial community composition of activated sludge flocs and granules. *FEMS Microbiology Ecology* **93**: (4).
- Förster, A.H. and Gescher, J. (2014) Metabolic engineering of *Escherichia coli* for production of mixed-acid fermentation end products. *Frontiers in Bioengineering and Biotechnology* **2**: 16.
- Geesink, P., Taubert, M., Jehmlich, N., Bergen, M. von, and Küsel, K. (2022) Bacterial necromass is rapidly metabolized by heterotrophic bacteria and supports multiple trophic levels of the groundwater microbiome. *Microbiology Spectrum* **10**: (4) e00437-22.
- Glöckner, J., Kube, M., Shrestha, P.M., Weber, M., Glöckner, F.O., Reinhardt, R., and Liesack, W. (2010) Phylogenetic diversity and metagenomics of candidate division OP3. *Environmental Microbiology* **12**: (5) 1218–1229.
- Guerrero, R., Pedros-Alio, C., Esteve, I., Mas, J., Chase, D., and Margulis, L. (1986) Predatory prokaryotes: predation and primary consumption evolved in bacteria. *Proceedings of the National Academy of Sciences* **83**: (7) 2138–42.

- Guo, J., Peng, Y., Ni, B.-J., Han, X., Fan, L., and Yuan, Z. (2015) Dissecting microbial community structure and methane-producing pathways of a full-scale anaerobic reactor digesting activated sludge from wastewater treatment by metagenomic sequencing. *Microbial Cell Factories* **14**: (1) 33.
- Gupta, K., Bardhan, P., Saikia, D., Rather, M.A., Loying, S., Mandal, M., and Kataki, R. (2022) Microbial fermentation: basic fundamentals and its dynamic prospect in various industrial applications. In *Industrial Microbiology and Biotechnology*. Verma, P. (ed). Singapore: Springer Singapore, pp. 107–128.
- Halim, M.F.A., Day, L.A., and Costa, K.C. (2021) Formate-dependent heterodisulfide reduction in a Methanomicrobiales archaeon. *Applied and Environmental Microbiology* **87**: (6) e02698-20.
- Halsey, K.H., Giovannoni, S.J., Graus, M., Zhao, Y., Landry, Z., Thrash, J.C., et al. (2017) Biological cycling of volatile organic carbon by phytoplankton and bacterioplankton. *Limnology and Oceanography* **62**: (6) 2650–2661.
- Hanson, R.S. and Hanson, T.E. (1996) Methanotrophic bacteria. *Microbiological Reviews* **60**: (2) 439–71.
- Harder, J. and Foss, S. (1999) Anaerobic formation of the aromatic hydrocarbon p-cymene from monoterpenes by methanogenic enrichment cultures. *Geomicrobiology Journal* **16**: (4) 295–305.
- Heider, J., Szaleniec, M., Martins, B.M., Seyhan, D., Buckel, W., and Golding, B.T. (2016) Structure and function of benzylsuccinate synthase and related fumarate-adding glyceryl radical enzymes. *Microbial Physiology* **26**: (1–3) 29–44.
- Hook, S.E., Wright, A.-D.G., and McBride, B.W. (2010) Methanogens: methane producers of the rumen and mitigation strategies. *Archaea* **2010**: 1–11.
- Hugenholtz, P., Pitulle, C., Hershberger Karen, L., and Pace Norman, R. (1998) Novel division level bacterial diversity in a Yellowstone hot spring. *Journal of Bacteriology* **180**: (2) 366–376.
- Janssen, P.H. and Kirs, M. (2008) Structure of the archaeal community of the rumen. *Applied and Environmental Microbiology* **74**: (12) 3619–3625.
- Joshi, A., Lanjekar, V., Dhakephalkar, P.K., and Dagar, S.S. (2018) Cultivation of multiple genera of hydrogenotrophic methanogens from different environmental niches. *Anaerobe* **50**: 64–68.
- Kamada, S., Wakabayashi, R., and Naganuma, T. (2023) Phylogenetic revisit to a review on predatory bacteria. *Microorganisms* **11**: (7) 1673.

- Kizina, J. (2017) Insights into the biology of Candidate Division OP3 LiM populations. *PhD Thesis*.
- Kizina, J., Jordan Sebastian, F.A., Martens Gerrit, A., Lonsing, A., Probian, C., Kolovou, A., et al. (2022) *Methanosaeta* and “*Candidatus Velamenicoccus archaeovor*us”. *Applied and Environmental Microbiology* **88**: (7) e02407-21.
- Knudsen, J. and Gershenzon, J. (2006) The chemical diversity of floral scent. In *Biology of floral scent*. Boca Raton, Florida: CRC Press, pp. 27–52.
- Korpi, A., Järnberg, J., and Pasanen, A.-L. (2009) Microbial volatile organic compounds. *Critical Reviews in Toxicology* **39**: (2) 139–193.
- Krassi Rumchev, Helen Brown, and Jeffery Spickett (2007) Volatile organic compounds: do they present a risk to our health? *Reviews on Environmental Health* **22**: (1) 39–56.
- Kurth, J.M., Op den Camp, H.J.M., and Welte, C.U. (2020) Several ways one goal - methanogenesis from unconventional substrates. *Applied Microbiology and Biotechnology* **104**: (16) 6839–6854.
- Lerner, T.R., Lovering, A.L., Bui, N.K., Uchida, K., Aizawa, S., Vollmer, W., and Sockett, R.E. (2012) Specialized peptidoglycan hydrolases sculpt the intra-bacterial niche of predatory *Bdellovibrio* and increase population fitness. *PLoS pathogens* **8**: (2) e1002524.
- Li, H., Yang, Q., and Zhou, H. (2020) Niche differentiation of sulfate- and iron-dependent anaerobic methane oxidation and methylotrophic methanogenesis in deep sea methane seeps. *Frontiers in Microbiology* **11**: 1409.
- Li, Q., Putra, N.R., Rizkiyah, D.N., Abdul Aziz, A.H., Irianto, I., and Qomariyah, L. (2023) Orange pomace and peel extraction processes towards sustainable utilization: A short review. *Molecules* **28**: (8) 3550.
- Liu, S.Q. (2003) Practical implications of lactate and pyruvate metabolism by lactic acid bacteria in food and beverage fermentations. *International Journal of Food Microbiology* **83**: (2) 115–131.
- Long, F., Wang, L., Lupa, B., and Whitman, W.B. (2017) A flexible system for cultivation of *Methanococcus* and other formate-utilizing methanogens. *Archaea* **2017**: 7046026.
- Loreto, F. and Fares, S. (2013) Biogenic volatile organic compounds and their impacts on biosphere–atmosphere interactions. In *Developments in Environmental Science*. Matyssek, R., Clarke, N., Cudlin, P., Mikkelsen, T.N., Tuovinen, J.P., Wieser, G., and Paoletti, E. (eds). Elsevier, pp. 57–75.
- Lovley, D.R. (2017) Syntrophy goes electric: direct interspecies electron transfer. *Annual Review of Microbiology* **71**: (Volume 71, 2017) 643–664.

- Lücker, J., El Tamer, M.K., Schwab, W., Verstappen, F.W.A., van der Plas, L.H.W., Bouwmeester, H.J., and Verhoeven, H.A. (2002) Monoterpene biosynthesis in lemon (*Citrus limon*). *European Journal of Biochemistry* **269**: (13) 3160–3171.
- Mao, C., Feng, Y., Wang, X., and Ren, G. (2015) Review on research achievements of biogas from anaerobic digestion. *Renewable and Sustainable Energy Reviews* **45**: 540–555.
- Marmulla, R. and Harder, J. (2014) Microbial monoterpene transformations - a review. *Frontiers in Microbiology* **5**: 346.
- McInerney, M.J., Sieber, J.R., and Gunsalus, R.P. (2009) Syntrophy in anaerobic global carbon cycles. *Current Opinion in Biotechnology* **20**: (6) 623–632.
- Messina, P., Lathièrre, J., Sindelarova, K., Vuichard, N., Granier, C., Ghattas, J., et al. (2016) Global biogenic volatile organic compound emissions in the ORCHIDEE and MEGAN models and sensitivity to key parameters. *Atmospheric Chemistry Physics* **16**: (22) 14169–14202.
- Mohammed, K., Agarwal, M., Li, B., Newman, J., Liu, T., and Ren, Y. (2020) Evaluation of d-limonene and β -ocimene as attractants of *Aphytis melinus* (Hymenoptera: Aphelinidae), a parasitoid of *Aonidiella aurantii* (Hemiptera: Diaspididae) on *Citrus* spp. *Insects* **11**: (1) 44.
- Momper, L., Jungbluth, S.P., Lee, M.D., and Amend, J.P. (2017) Energy and carbon metabolisms in a deep terrestrial subsurface fluid microbial community. *The ISME Journal* **11**: (10) 2319–2333.
- Moore, E.R., Davie-Martin, C.L., Giovannoni, S.J., and Halsey, K.H. (2020) *Pelagibacter* metabolism of diatom-derived volatile organic compounds imposes an energetic tax on photosynthetic carbon fixation. *Environmental Microbiology* **22**: (5) 1720–1733.
- Moreira, D., Zivanovic, Y., López-Archilla, A.I., Iniesto, M., and López-García, P. (2021) Reductive evolution and unique predatory mode in the CPR bacterium *Vampirococcus lugosii*. *Nature Communications* **12**: (1) 2454.
- Moufida, S. and Marzouk, B. (2003) Biochemical characterization of blood orange, sweet orange, lemon, bergamot and bitter orange. *Phytochemistry* **62**: (8) 1283–1289.
- Negro, V., Mancini, G., Ruggeri, B., and Fino, D. (2016) Citrus waste as feedstock for bio-based products recovery: review on limonene case study and energy valorization. *Bioresource Technology* **214**: 806–815.
- Ngo, V.Q.H., Enault, F., Midoux, C., Mariadassou, M., Chapleur, O., Mazéas, L., et al. (2022) Diversity of novel archaeal viruses infecting methanogens discovered through coupling of stable isotope probing and metagenomics. *Environmental Microbiology* **24**: (10) 4853–4868.

- Ozturk, B., Winterburn, J., and Gonzalez-Miquel, M. (2019) Orange peel waste valorisation through limonene extraction using bio-based solvents. *Biochemical Engineering Journal* **151**: 107298.
- 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. *Nature Biotechnology* **36**: (10) 996–1004.
- Parshina, S.N., Ermakova, A.V., Bomberg, M., and Detkova, E.N. (2014) *Methanospirillum stamsii* sp. nov., a psychrotolerant, hydrogenotrophic, methanogenic archaeon isolated from an anaerobic expanded granular sludge bed bioreactor operated at low temperature. *International Journal of Systematic and Evolutionary Microbiology* **64**: (Pt_1) 180–186.
- Pasternak, Z., Njagi, M., Shani, Y., Chanyi, R., Rotem, O., Lurie-Weinberger, M.N., et al. (2014) In and out: an analysis of epibiotic vs periplasmic bacterial predators. *The ISME Journal* **8**: (3) 625–635.
- Perez-Molphe-Montoya, E., Küsel, K., and Overholt, W.A. (2022) Redefining the phylogenetic and metabolic diversity of phylum Omnitrochota. *Environmental Microbiology* **24**: (11) 5437–5449.
- Peters, V., Janssen, P.H., and Conrad, R. (1999) Transient production of formate during chemolithotrophic growth of anaerobic microorganisms on hydrogen. *Current Microbiology* **38**: (5) 285–289.
- Potter, A., Andersson, J., Sjöblom, A., Junedahl, E., Cousins, A., Brorström-Lundén, E., and Cato, I. (2014) Results from the swedish national screening programme 2004 - subreport 3: limonene.
- Quaiyum, S., Igarashi, K., Narihiro, T., and Kato, S. (2020) Microbial community analysis of anaerobic enrichment cultures supplemented with bacterial peptidoglycan as the sole substrate. *Microbes and Environments* **35**: (3).
- Richards, M.A., Lie, T.J., Zhang, J., Ragsdale, S.W., Leigh, J.A., and Price, N.D. (2016) Exploring hydrogenotrophic methanogenesis: a genome scale metabolic reconstruction of *Methanococcus maripaludis*. *Journal of Bacteriology* **198**: (24) 3379–3390.
- Robles, G., Nair, R.B., Kleinstuber, S., Nikolausz, M., and Sárvári Horváth, I. (2018) Biogas production: microbiological aspects. In *Biogas: Fundamentals, Process, and Operation*. Tabatabaei, M. and Ghanavati, H. (eds). Cham: Springer International Publishing, pp. 163–198.
- Rodríguez, A., San Andrés, V., Cervera, M., Redondo, A., Alquézar, B., Shimada, T., et al. (2011) Terpene down-regulation in orange reveals the role of fruit aromas in

- mediating interactions with insect herbivores and pathogens. *Plant Physiology* **156**: (2) 793–802.
- Rotaru, A.-E. (2009) Anaerobic degradation of limonene and p-xylene in freshwater enrichment cultures. *PhD Thesis*.
- Rotaru, A.E., Schauer, R., Probian, C., Musmann, M., and Harder, J. (2012) Visualization of candidate division OP3 cocci in limonene-degrading methanogenic cultures. *Journal of Microbiology and Biotechnology* **22**: (4) 457–61.
- Rotaru, A.-E., Shrestha, P.M., Liu, F., Markovaite, B., Chen, S., Nevin, K.P., and Lovley, D.R. (2014) Direct interspecies electron transfer between *Geobacter metallireducens* and *Methanosarcina barkeri*. *Applied and Environmental Microbiology* **80**: (15) 4599–4605.
- Rowe, I., Chiaravalli, M., and Boletta, A. (2013) ATP and lactate quantification. *Bio-protocol* **3**: (23) e986.
- Schink, B. (1997) Energetics of syntrophic cooperation in methanogenic degradation. *Microbiology and molecular biology reviews* **61**: (2) 262–80.
- Schink, B. and Friedrich, M. (1994) Energetics of syntrophic fatty acid oxidation. *FEMS Microbiology Reviews* **15**: (2–3) 85–94.
- Schink, B., Montag, D., Keller, A., and Müller, N. (2017) Hydrogen or formate: alternative key players in methanogenic degradation. *Environmental Microbiology Reports* **9**: (3) 189–202.
- Schmidt, O., Hink, L., Horn, M.A., and Drake, H.L. (2016) Peat: home to novel syntrophic species that feed acetate- and hydrogen-scavenging methanogens. *The ISME Journal* **10**: (8) 1954–1966.
- Seymour, C.O., Palmer, M., Becraft, E.D., Stepanauskas, R., Friel, A.D., Schulz, F., et al. (2023) Hyperactive nanobacteria with host-dependent traits pervade Omnitrophota. *Nature Microbiology* **8**: (4) 727–744.
- Sockett, R.E. (2009) Predatory lifestyle of *Bdellovibrio bacteriovorus*. *Annual Review of Microbiology* **63**: (Volume 63, 2009) 523–539.
- Suttle, C.A. (1994) The significance of viruses to mortality in aquatic microbial communities. *Microbial Ecology* **28**: (2) 237–243.
- Thauer, R.K., Kaster, A.-K., Seedorf, H., Buckel, W., and Hedderich, R. (2008) Methanogenic archaea: ecologically relevant differences in energy conservation. *Nature Reviews Microbiology* **6**: (8) 579–591.

- Thioux, S., Dupont, S., Nesbø, C.L., Bienvenu, N., Krupovic, M., L'Haridon, S., et al. (2021) The first head-tailed virus, MFTV1, infecting hyperthermophilic methanogenic deep-sea archaea. *Environmental Microbiology* **23**: (7) 3614–3626.
- Thomas, C., Grossi, V., Antheaume, I., and Ariztegui, D. (2019) Recycling of archaeal biomass as a new strategy for extreme life in Dead Sea deep sediments. *Geology* **47**: (5) 479–482.
- Venkiteshwaran, K., Bocher, B., Maki, J., and Zitomer, D. (2015) Relating anaerobic digestion microbial community and process function: supplementary issue: water microbiology. *Microbiology Insights* **8s2**: MBI.S33593.
- Waite, D.W., Chuvochina, M., Pelikan, C., Parks, D.H., Yilmaz, P., Wagner, M., et al. (2020) Proposal to reclassify the proteobacterial classes *Deltaproteobacteria* and *Oligoflexia*, and the phylum *Thermodesulfobacteria* into four phyla reflecting major functional capabilities. *International Journal of Systematic and Evolutionary Microbiology* **70**: (11) 5972–6016.
- Webster, C.R., Mahaffy, P.R., Atreya, S.K., Flesch, G.J., Mischna, M.A., Meslin, P.-Y., et al. (2015) Mars methane detection and variability at Gale crater. *Science* **347**: (6220) 415–417.
- Weidenbach, K., Wolf, S., Kupczok, A., Kern, T., Fischer, M.A., Reetz, J., et al. (2021) Characterization of Blf4, an archaeal lytic virus targeting a member of the methanomicrobiales. *Viruses* **13**: (10) 1934.
- Weinbauer, M.G. (2004) Ecology of prokaryotic viruses. *FEMS Microbiology Reviews* **28**: (2) 127–181.
- Welte, C. and Deppenmeier, U. (2014) Bioenergetics and anaerobic respiratory chains of acetoclastic methanogens. *Biochimica et Biophysica Acta (BBA) - Bioenergetics* **1837**: (7) 1130–1147.
- Werner, S., Diekert, G., and Schuster, S. (2010) Revisiting the thermodynamic theory of optimal ATP stoichiometries by analysis of various ATP-producing metabolic pathways. *Journal of Molecular Evolution* **71**: (5) 346–355.
- West-Roberts, J., Valentin-Alvarado, L., Mullen, S., Sachdeva, R., Smith, J., Hug, L.A., et al. (2023) Giant genes are rare but implicated in cell wall degradation by predatory bacteria [Preprint]. *bioRxiv* 2023.11.21.568195.
- Whitman, W.B., Oren, A., Chuvochina, M., Da Costa, M.S., Garrity, G.M., Rainey, F.A., et al. (2018) Proposal of the suffix -ota to denote phyla. Addendum to 'Proposal to include the rank of phylum in the International Code of Nomenclature of Prokaryotes.' *International Journal of Systematic and Evolutionary Microbiology* **68**: (3) 967–969.

- Xu, Z., Zhao, M., Miao, H., Huang, Z., Gao, S., and Ruan, W. (2014) *In situ* volatile fatty acids influence biogas generation from kitchen wastes by anaerobic digestion. *Bioresource Technology* **163**: 186–192.
- Yildiz, B.S. (2012) 18 - Water and wastewater treatment: biological processes. In *Metropolitan Sustainability*. Zeman, F. (ed). Woodhead Publishing, pp. 406–428.
- Zhuang, G.-C., Elling, F.J., Nigro, L.M., Samarkin, V., Joye, S.B., Teske, A., and Hinrichs, K.-U. (2016) Multiple evidence for methylotrophic methanogenesis as the dominant methanogenic pathway in hypersaline sediments from the Orca Basin, Gulf of Mexico. *Geochimica et Cosmochimica Acta* **187**: 1–20.

Chapter 2:

***Methanosaeta* and “*Candidatus Velamenicoccus archaeovor*us”**

Jana Kizina¹, Sebastian F. A. Jordan¹, Gerrit Alexander Martens¹, **Almud Lonsing**¹, Christina Probian¹, Androniki Kolovou², Rachel Santarella-Mellwig², Erhard Rhiel³, Sten Littmann¹, Stephanie Markert⁴, Kurt Stüber⁵, Michael Richter¹, Thomas Schweder⁴, Jens Harder¹

Published in:

Applied and Environmental Microbiology, 2022

Volume 88, Issue 7, e02407-21

¹Max Planck Institute for Marine Microbiology, Bremen, Germany.

²Electron Microscopy Core Facility, EMBL Heidelberg, Heidelberg, Germany.

³Institute for Chemistry and Biology of the Marine Environment, Carl von Ossietzky University of Oldenburg, Oldenburg, Germany.

⁴Department of Pharmaceutical Biotechnology, Institute for Pharmacy, University of Greifswald, Greifswald, Germany.

⁵Max Planck-Genome-Centre Cologne, Cologne, Germany.

Contribution of the Ph.D. candidate in % of the total work:

Experimental concept and design – 10%

Experimental work/acquisition of experimental data – 10%

Data analysis and interpretation – 10%

Preparation of figures and tables – 10%

Drafting of the manuscript – 10%

***Methanosaeta* and “*Candidatus Velamenicoccus archaeovor*us”**

Jana Kizina^a, Sebastian F. A. Jordan^a, Gerrit Alexander Martens^a, Almud Lonsing^a, Christina Probian^a, Androniki Kolovou^b, Rachel Santarella-Mellwig^b, Erhard Rhiel^c, Sten Littmann^a, Stephanie Markert^d, Kurt Stüber^e, Michael Richter^a, Thomas Schweder^d, Jens Harder^a

^aMax Planck Institute for Marine Microbiology, Bremen, Germany

^bElectron Microscopy Core Facility, EMBL Heidelberg, Heidelberg, Germany

^cInstitute for Chemistry and Biology of the Marine Environment, Carl von Ossietzky University of Oldenburg, Oldenburg, Germany

^dDepartment of Pharmaceutical Biotechnology, Institute for Pharmacy, University of Greifswald, Greifswald, Germany

^eMax Planck-Genome-Centre Cologne, Cologne, Germany

ABSTRACT The phylum “*Candidatus Omnitrophica*” (candidate division OP3) is ubiquitous in anaerobic habitats but is currently characterized only by draft genomes from metagenomes and single cells. We had visualized cells of the phylotype OP3 LiM in methanogenic cultures on limonene as small epibiotic cells. In this study, we enriched OP3 cells by double density gradient centrifugation and obtained the first closed genome of an apparently clonal OP3 cell population by applying metagenomics and PCR for gap closure. Filaments of acetoclastic *Methanosaeta*, the largest morphotype in the culture community, contained empty cells, cells devoid of rRNA or of both rRNA and DNA, and dead cells according to transmission electron microscopy (TEM), thin-section TEM, scanning electron microscopy (SEM), catalyzed reporter deposition-fluorescence *in situ* hybridization (CARD-FISH), and LIVE/DEAD imaging. OP3 LiM cells were ultramicrobacteria (200 to 300 nm in diameter) and showed two physiological stages in CARD-FISH fluorescence signals: strong signals of OP3 LiM cells attached to *Bacteria* and to *Archaea* indicated many rRNA molecules and an active metabolism, whereas free-living OP3 cells had weak signals. Metaproteomics revealed that OP3 LiM lives with highly expressed secreted proteins involved in depolymerization and uptake of macromolecules and an active glycolysis and energy conservation by the utilization of pyruvate via a pyruvate:ferredoxin oxidoreductase and an Rnf complex (ferredoxin:NAD oxidoreductase). Besides sugar fermentation, a nucleotidyl transferase may contribute to energy conservation by phosphorylation, the phosphate-dependent depolymerization of nucleic acids. Thin-section TEM showed distinctive structures of predation. Our study demonstrated a predatory metabolism for OP3 LiM cells, and therefore, we propose the name “*Candidatus Velamenicoccus archaeovor*us” gen. nov., sp. nov., for OP3 LiM.

IMPORTANCE Epibiotic bacteria are known to live on and off bacterial cells. Here, we describe the ultramicrobacterial anaerobic epibiont OP3 LiM living on *Archaea* and *Bacteria*. We detected sick and dead cells of the filamentous archaeon *Methanosaeta* in slowly growing methanogenic cultures. OP3 LiM lives

as a sugar fermenter, likely on polysaccharides from outer membranes, and has the genomic potential to live as a syntroph. The predatory lifestyle of OP3 LiM was supported by its genome, the first closed genome for the phylum “*Candidatus Omnitrophica*,” and by images of cell-to-cell contact with prey cells. We propose naming OP3 LiM “*Candidatus Velamenicoccus archaeovorus*.” Its metabolic versatility explains the ubiquitous presence of “*Candidatus Omnitrophica*” 3 in anoxic habitats and gives ultramicrobacterial epibionts an important role in the recycling and remineralization of microbial biomass. The removal of polysaccharides from outer membranes by ultramicrobacteria may also influence biological interactions between pro- and eukaryotes.

KEYWORDS “*Candidatus Omnitrophica*”, predatory bacteria, LPS, methanogenic enrichment culture, candidate division OP3, limonene, lipopolysaccharide, methanogens

Received 6 December 2021

Accepted 15 February 2022

Published 21 March 2022

Address correspondence to Jens Harder, jharder@mpi-bremen.de.

The authors declare no conflict of interest.

Editor Nicole R. Buan, University of Nebraska-Lincoln

Copyright © 2022 American Society for Microbiology. All Rights Reserved.

Life may have begun in a primordial soup (1) with the evolution of an RNA world (2). Once the dissolved organic carbon in this soup was depleted, chemical energy was available only in the form of reduced inorganic compounds or as particulate organic matter in the form of cells. Geochemically produced hydrogen has been identified as a potential energy source for ancient acetogenic and methanogenic organisms (3). The origin and evolution of predation have been underestimated as important processes in early times of the earth (4) but are now established as a widespread mode of interaction among living organisms in many ecosystems (5, 6). Predators have been identified in groundwater, rivers, estuaries, the open ocean, sewage, soils, plant roots, and animal feces (7, 8). Predators have been classified as obligate (unable to grow in the absence of prey) or facultative (able to grow as a pure culture without the presence of prey) (8).

Bacterial predators attack their prey in groups (*Myxobacteria*) or individually. Epibiotic species attach to the prey, and some species penetrate the periplasm or the cytoplasm (8, 9). Members of the genus *Bdellovibrio* and related organisms, summarized as “*Bdellovibrio* and like organisms” (BALOs), are the most-studied group of predatory bacteria (10). They prey exclusively on Gram-negative cells and have a dimorphic life cycle. Motile cells with a single polar flagellum find prey cells and attach to the outer membrane (7). After an irreversible attachment (11), invading BALOs such as *Bdellovibrio bacteriovorus* enter the prey’s periplasmic space and proliferate at the expense of the prey’s cytoplasmic content. Motile progeny cells release themselves from the remnants of the prey cell to start a new cycle (11).

Epibiotic predators remain attached to the outer membrane while deriving nourishment from the prey (11). This lifestyle has been described for *Bdellovibrio exovorus*, formerly *Bdellovibrio* sp. strain JSS and a novel predator of *Caulobacter crescentus*, and *Micavibrio aeruginosavorus* (12, 13). The eukaryotic microalga *Chlorella* has an epibiotic predator, *Vampirovibrio chlorellavorus* (14). Recently, the ultramicrobacterium “*Nanosynbacter lyticus*” of the phylum “*Candidatus Saccharibacteria*”

(formerly candidate division TM7) was isolated in coculture with its host/prey, an *Actinomyces* strain (15).

Absence of molecular oxygen characterized early evolution. The anaerobe *Vampirococcus*, which was defined by microscopic cell counts and electron micrographs of cells on anoxygenic phototrophs in environmental samples (16), has stimulated discussion on the early evolution of predators. Recently, a phototrophic enrichment from an athalassic salt lake was found to contain anoxygenic phototrophs related to *Halochromatium* and epibiotic predators which were described as “*Candidatus Vampirococcus lugosii*” (17). Its genome revealed that it is a member of the phylum “*Candidatus Absconditabacteria*.”

Here, we describe a novel anaerobic predatory bacterium with a coccoid morphology. We had observed cells of candidate division OP3, also named phylum “*Candidatus Omnitrphica*” (18), in high abundances in a methanogenic enrichment culture on limonene. The phylotype OP3 LiM originated from a 16S rRNA gene clone library. Catalyzed reporter deposition-fluorescence *in situ* hybridization (CARD-FISH) with the specific OP3 LiM probe OP3-565 revealed that 18% of all cells in the enrichment culture were OP3 LiM cells (19). The micrographs showed small round cells, either free living or attached to larger cells. Physical cell separations and a range of visualizations as well as metagenomes and metaproteomes provided insight into the biology of OP3 LiM cells. Based on our observations, the name “*Candidatus Velamenicoccus archaeovor*us” gen. nov. sp. nov. is proposed for OP3 LiM.

RESULTS

Maintenance of a methanogenic enrichment culture on limonene. Methanogenic enrichment cultures (20) were used as inocula in 1999 in liquid dilution series. An inoculum of 1 μ L grew, and this culture was maintained in 12 parallel lineages by an annual transfer of 10% (vol/vol), which corresponds to three to four generations each year. These cultures showed active methanogenesis for more than 2 years. Template dilution OP3 LiM-specific PCR and CARD-FISH experiments with probe OP3-565 guided the selection of a lineage containing a high cell number of OP3 LiM for experiments.

Visualization of *Methanosaeta* in methanogenic limonene enrichment cultures. *Methanosaeta* was identified in 16S rRNA gene amplicon libraries as one of the abundant methanogenic *Archaea* in the limonene enrichment culture. FISH studies using the *Methanosaetaceae*-specific probe MX-825 (21) confirmed the filamentous cell morphology, which was also evident in electron micrographs. *Methanosaeta* cells are located within a filamentous sheath in chambers separated by spacer plugs (22). Sheath and cells were visible in filaments of enrichment cultures in transmission electron micrographs (TEM) (Fig. 1; also, see Fig. S1 in the supplemental material). The enrichment was morphologically very diverse, also containing smaller filaments (Fig. S2 to S4). TEM images also showed large sheaths without cells (Fig. 2). To investigate the absence of cells or of cellular components in *Methanosaeta* filaments, we combined DNA staining by DAPI (4',6-diamidino-2-phenylindole) with rRNA visualization by FISH, lipid staining by Nile red, and a relief view of the filament morphology by differential interference contrast microscopy applying confocal laser scanning microscopy. *Methanosaeta* filaments contained healthy cells with a full biovolume and the presence of lipids, rRNA, and DNA. However, the filaments also had cells with a reduced biovolume, and either rRNA or rRNA and DNA were not present according to the staining intensity. In the absence of a cell biovolume in the differential interference contrast (DIC) graph, a faint Nile red

stain identified the shape of cells in the filament. We annotated these shapes as empty cells (Fig. 3; Fig. S5 to S7). For control experiments, *Methanosaeta concilii* GP6^T was obtained as *Methanotherx soehngeni* DSM3671. Cells in stationary phase showed loss of rRNA in only a few Nile red-stained cells. Nile red also stained spacer plugs, albeit with less sensitivity than lipids. In some cases, overstaining revealed the presence of pairs of spacer plugs between normal-sized cells, suggesting the presence of small cells as precursors of filament breakage (Fig. S8 and S9). *Methanosaeta* performs, besides its equal cell division into two viable cells, an unequal cell division into a normal and a short cell, followed by lysis of the short daughter cell, a biological strategy to control the filament length (22).

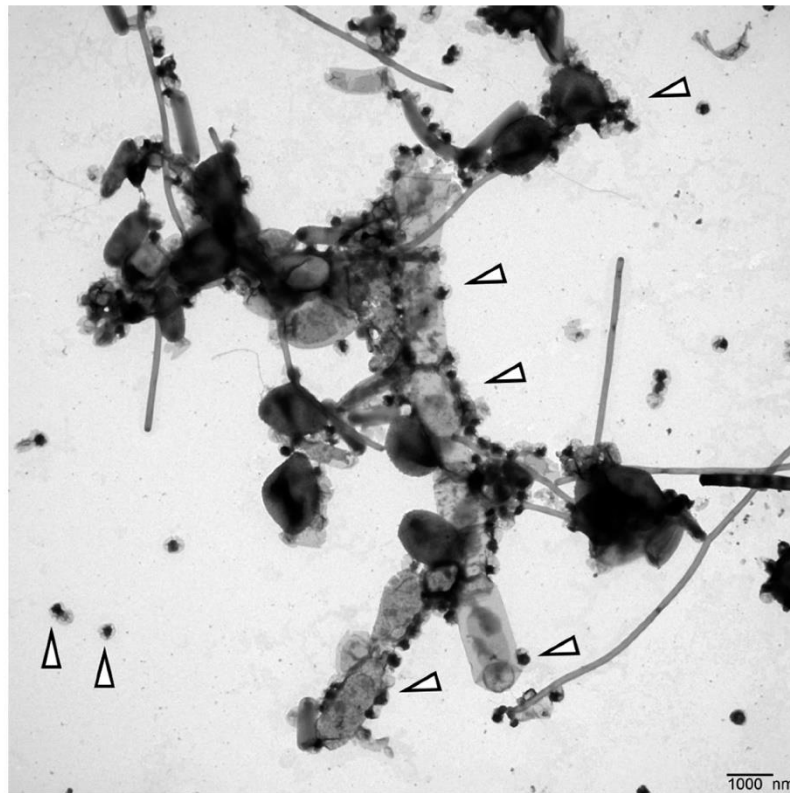


FIG 1 Transmission electron micrograph of a resuspended cell pellet of the methanogenic limonene enrichment culture. We assigned the large filament to *Methanosaeta*, based on the spacer disc at the lower end of the filament and the presence and structure of a sheath (see Fig. S1). The cells may have been shrunk by the TEM preparation. The small black cells with a gray cape were assigned to the ultramicrobacterium OP3 LiM (triangles indicate OP3 LiM cells). Bar, 1 μ m.

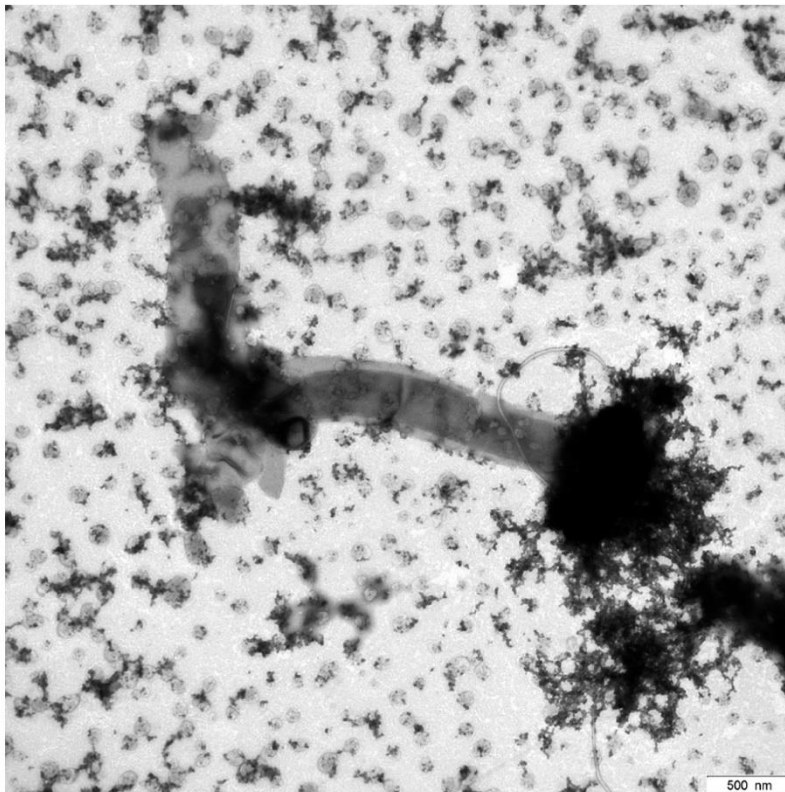


FIG 2 Transmission electron micrograph of a dried culture droplet of the methanogenic limonene enrichment culture showing sheaths with cells and without cells. We assigned the large filament to *Methanosaeta*. The small cells were assigned to OP3 LiM and other ultramicrobacteria in the culture. Bar, 500 nm.

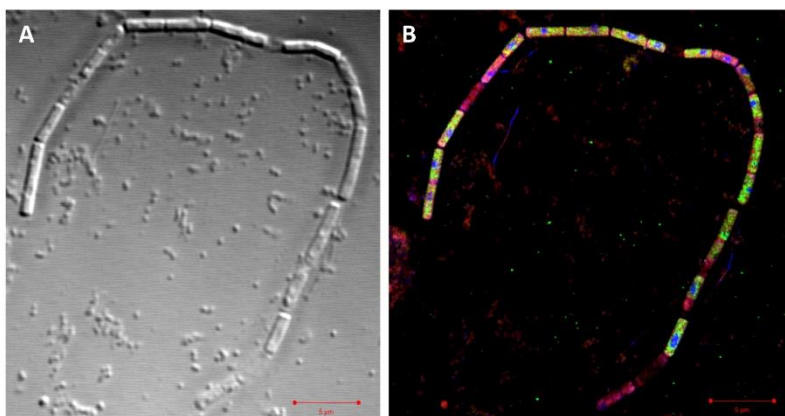


FIG 3 Differential interference contrast micrograph (A) and superresolution structured illumination microscopy (SR-SIM) overlay image (B) of a filament assigned to *Methanosaeta*. Staining was achieved using Nile red (red) for lipids, probe Arch915 labeled with four 6-carboxyfluorescein molecules (green) for rRNA, and DAPI (blue) for DNA. Bar, 5 µm. For individual and dual overlay SR-SIM images, see Fig. S5.

A decrease in cellular content of *Methanosaeta* cells was also observed in thin-section TEM. For visibility in TEM images, cells require the presence of electron-diffracting heavy metal ions that are bound to anionic groups of proteins, lipids, and nucleic acids. Thus, the darkness of the cellular content in the image is a proxy for cellular integrity. *Methanosaeta* filaments were identified by their unique morphology and contained cells that appeared as full, intermediate, or empty with respect to the cellular content (Fig. 4; Fig. S10). Some cells contained faintly stained intracellular local structures of unknown biological identity, likely neutral storage compounds or unknown infectious viruses. Very short cells within filaments were annotated as short daughter cells enabling filament segregation. These cells did not contain cytoplasm, and we expected them to be dead based on their biological function. In addition, we detected large cells in filaments without cellular contents. To confirm their physiological status, we investigated the integrity of *Methanosaeta* membranes with a LIVE/DEAD stain. In confocal laser scanning micrographs, *Methanosaeta* cells showed a condensed DNA using optimized laser intensities. The green stain colored the whole cell at higher laser intensities, which is equivalent to overexposed images (Fig. 5). Besides live cells, *Methanosaeta* filaments also contained dead cells of ordinary size, accounting for 11% of all *Methanosaeta* cells in a 3-year-old, still-methanogenic, active enrichment culture. We annotated shorter cells as recently divided cells. The LIVE/DEAD stain also visualized very short dead cells with a very small thickness between normal-sized cells (Fig. 5). The presence of these short cells was also visualized in thin-section TEM images, where they appeared to be empty (Fig. 4; Fig. S10), and gained additional support from pairs of spacer plugs that were stained by Nile red (Fig. S8 and S9). We suggest that these short dead cells result from unequal cell division preceding filament segregation.

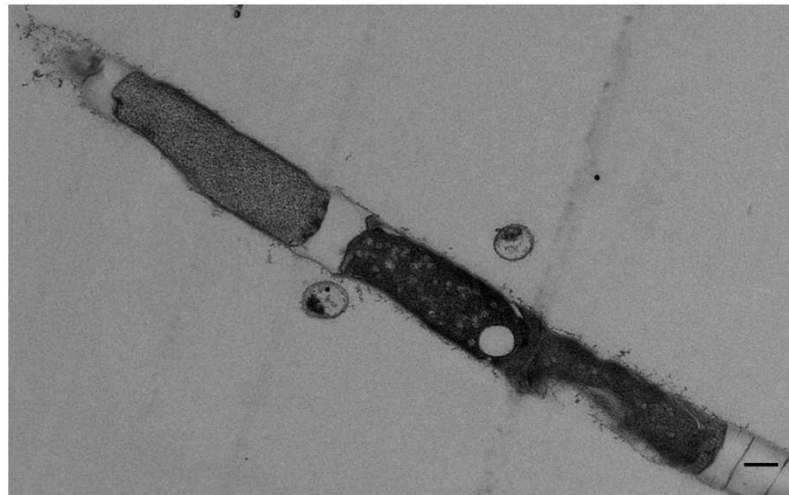


FIG 4 TEM image of a thin section preparation showing a *Methanosaeta* filament with microcells and damaged regular cells. Bar, 300 nm.

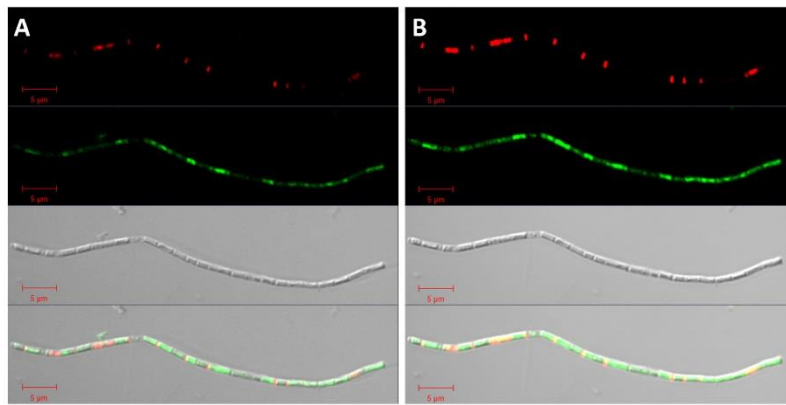


FIG 5 LIVE/DEAD staining of a filament assigned to *Methanosaeta* in images with optimal exposure (A) and overexposure (B) to visualize regions of weak staining. In the DIC micrograph, dead (red) cells had less biovolume than live (green) cells. Overexposure showed a weak green staining of cells and short red cells separating green cells similar to spacer plugs. We assigned the short dead cells to unequally produced daughter cells that precede filament separation of *Methanosaeta* based on the observation of two spacer plugs between cells visible by Nile red staining (Fig. S8 and S9) and *Methanosaeta* microcells in thin-section TEM images (Fig. S11, S15, and S16). Bar, 5 µm.

Methanosaeta cells have a sheath made of protein and carbohydrates (22). To exclude a problem with accessibility of FISH probes to cellular rRNA molecules because of the rigid sheath, we tested harsh cell lysis treatments with proteinase K to improve the detection of archaeal rRNA in all cells of filaments. Small cells were lysed, but the detection efficiency of rRNA-containing cells in filaments did not increase, similar to the results of Kubota et al. (23). We conclude that the lack of CARD-FISH signals of cells in *Methanosaeta* filaments was caused by the absence of rRNA molecules in the cells, but not by a sheath that is impermeable to probes. In nature, dead cells within filaments may be widespread, as *Methanosaeta* filaments in environmental samples had low *in situ* hybridization detection rates of cells independent of the applied detection technique (FISH, CARD-FISH, or hybridization chain reaction [HCR]-FISH) (24). In summary, our visualization experiments confirmed previous observations of the *Methanosaeta* cell biology and detected the presence of sickly and dead cells of ordinary cell size in *Methanosaeta* filaments. Here, we present initial evidence that a biological cause is responsible for the sickness and death of *Methanosaeta*.

OP3 LiM cell enrichment in density gradients yielded an ultramicrobacterium. The small size of OP3 LiM cells in CARD-FISH images (19) suggested a separation of cell populations by density gradients. The density centrifugation of concentrated cells from the limonene-degrading methanogenic enrichment culture yielded two visible bands close to the top and to the bottom in a Percoll gradient (Fig. S11). PCR and CARD-FISH analyses specific for OP3 LiM detected the highest relative abundance of OP3 cells in the gradient in a macroscopically clear layer above the visible bottom band of cells. This macroscopically clear fraction was collected from 10 gradients and further concentrated in a second Percoll gradient, yielding a fraction with over 80% OP3 LiM cells according to CARD-FISH. The twice-enriched cells were used for transmission electron microscopy and an OP3 LiM-enriched metagenome.

TEM images showed a dominant morphotype of small cells surrounded by a surface structure weakly stained with uranyl acetate (Fig. 6A). The cell size was 200 to 300 nm in diameter, which is in the range of ultramicrobacteria. The abundance of this morphotype correlated with the abundance of

OP3 LiM cells detected by CARD-FISH experiments in the cell population. This observation related the small cells with a cape in TEM images to the phylotype OP3 LiM.

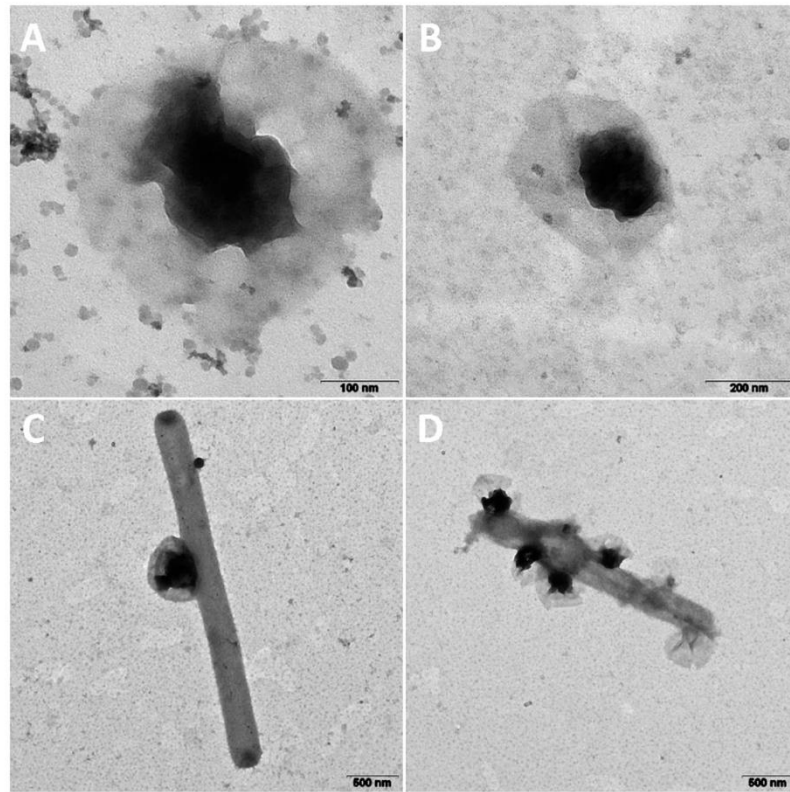


FIG 6 TEM images of OP3 LiM cells obtained from the second density gradient (A) (the small particles are colloidal silica particles 15 to 30 nm in diameter originating from Percoll) and in the limonene enrichment culture as free-living cells (B) or attached to other cells (C and D). Bars, 100 nm (A), 200 nm (B) and 500 nm (C and D).

TEM images of enrichment cultures were prepared together with those of the OP3 LiM-enriched gradient samples, and on the basis of the characteristic morphotype, they revealed that OP3 LiM cells were mainly attached to larger cells (Fig. 6B to D). The attached lifestyle detected earlier by FISH (19) was confirmed in these first TEM pictures recorded in 2012. TEM images in 2017 showed the same ultramicrobacterial morphotype, characterized by a veil around a strongly stained cell (Fig. 1; Fig. S1 to S4).

As ultramicrobacteria are too small to be visualized by phase-contrast microscopy, we initially used scanning electron microscopy to confirm the presence of ultramicrobacteria on *Methanosaeta* filaments (Fig. S12 and S13). Differential interference contrast microscopy can resolve 100-nm objects and has become the preferred technology to quickly demonstrate the presence of ultramicrobacteria in samples of the limonene enrichment cultures (Fig. 3 and 5; Fig. S6 and S7).

Thin-section TEM images visualized cell-to-cell interactions. TEM of 70-nm thin sections provided insights into the cell-to-cell interactions in the methanogenic enrichment culture (Fig. 7; Fig. S14 to S16). *Methanosaeta* was identified by its morphotype. A second morphotype of small, irregular,

pentagonal cells with strong cellular staining and small dark dots on the outer surface (Fig. 7B) was annotated as a methanogenic archaeon with an S-layer on the surface, identified as *Methanoculleus*-related cells based on an archaeal 16S rRNA gene library of the enrichment (19) and published descriptions of *Methanoculleus* strains (25). The dominant cell type in thin-section TEM images was an ultramicrobacterium with uneven staining of the cytoplasm and with strong staining of an intracellular complex. These cells were annotated as OP3 LiM. The cells interacted with other cells in an attack mode, with the intracellular complex close to the attack site (Fig. 7). Similar images of predation were reported in the first “*Candidatus Vampirococcus*” publication, which described it as “sucking the innards of its prey in a fashion reminiscent of vampires” (16). Prey cells include bacterial cells (Fig. 7A) as well as *Methanoculleus* (Fig. 7B). Stationary-phase cultures contained many remains of cell membranes, which suggested intensive predation in the enrichment culture. The cell morphotype assigned to OP3 LiM (Fig. S15 and S16) was still present in these stationary-phase cultures.

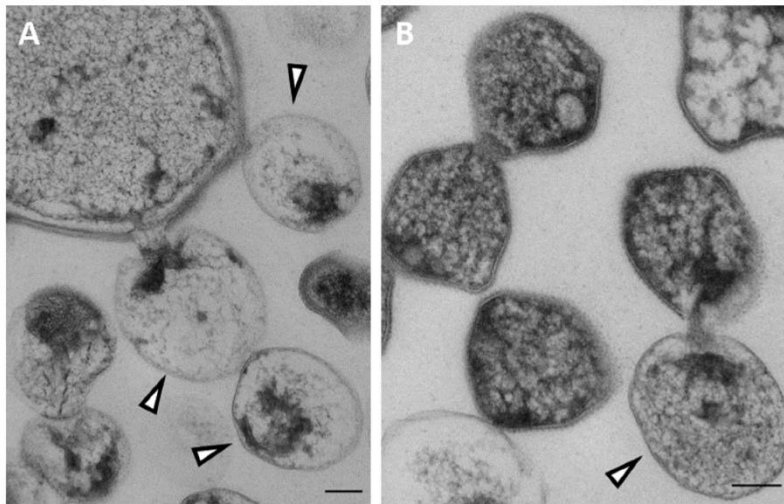


FIG 7 TEM images of thin sections of the methanogenic limonene enrichment culture showing the attack by OP3 LiM ultramicrobacteria of a large bacterial cell (A) and of a small archaeal cell (B). Ultramicrobial cells with a strongly stained part of the cytosol were assigned to OP3 LiM (triangles) based on the frequency of occurrence of this cell type in the thin-section TEM images. The small irregular cells with intense staining and a characteristic surface structure, likely S-layer protein, were assigned to archaeal cells, likely *Methanoculleus*-related species. Bars, 100 nm.

Visualization of the phylotype OP3 LiM. CARD-FISH detection of OP3 LiM cells with the probe OP3-565 was improved by introducing four helper oligonucleotides (26). In cultures, small coccoid OP3 LiM cells presented up to 30% of DAPI-stained cells in two different signal intensities (Fig. S16). Stronger detection signals were observed for attached OP3 LiM cells, suggesting a larger ribosome content and a higher metabolic activity of the attached cells than the free-living cells. The latter had weaker signals, suggesting a state of low metabolic activity and evidence of starvation. OP3 LiM cells attached to archaeal and bacterial cells, according to double-hybridization CARD-FISH experiments with probes for OP3 LiM cells and for *Archaea* or for *Bacteria*, respectively (Fig. 8; Fig. S17). This is possible, as the bacterial probe mix EUB338I-III does not detect OP3 LiM due to mismatches in the targeted 16S rRNA sequence (19). A range of morphologies was identified as *Bacteria*, including vibrios of different sizes, large coccoid cells, and short thin rod-shaped cells (Fig. S17).

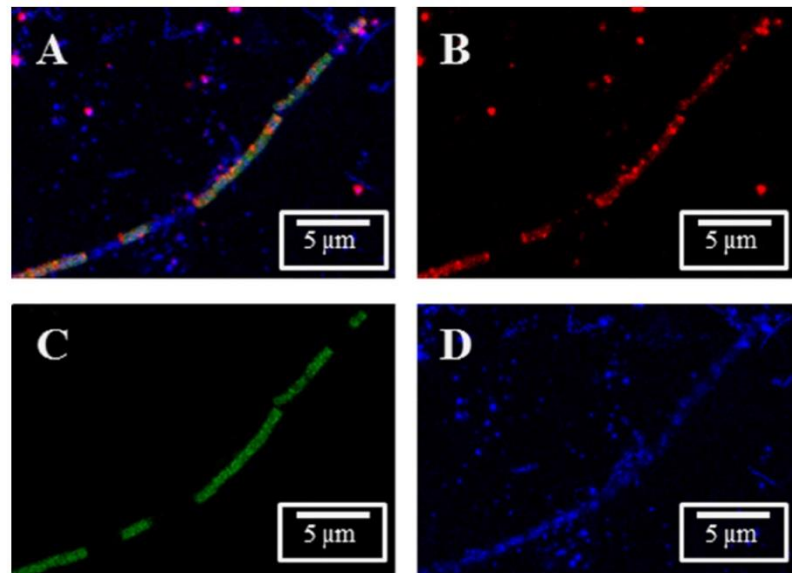


FIG 8 Detection of rRNA from OP3 LiM by fluorescence *in situ* hybridization. Archaea and OP3 LiM were detected in a methanogenic enrichment culture with probes ARCH-915 and OP3-565, respectively, in confocal laser scanning microscopy images: overlay (A) and individual signals of OP3 LiM rRNA (B, red), rRNA of *Archaea* (C, green), and DNA (D, blue, obtained by DAPI staining). Bars, 5 µm.

OP3 LiM cells were frequently attached to *Methanosaeta* filaments, either as single cells (Fig. 8) or as a group of cells (Fig. S17 and S18). Besides staining small cells of OP3 LiM, the CARD-FISH signal of probe OP3-565 faintly stained *Methanosaeta* cells (Fig. 8). To exclude an archaeal target sequence for the probe, BLASTn was performed with the probe sequence against sequences affiliated with *Methanosaeta* and the *Methanosaeta concilii* GP-6 genome. The best match was a 13-mer oligonucleotide with a G+C content of 54%. The predicted melting temperature was below the hybridization temperature of the CARD-FISH experiments, which excluded a false-positive signal and suggested that the weak signal of the probe OP3-565 originated from the hybridization to rRNA of OP3 LiM present on or in the archaeal cells.

Secondary electron micrographs from the scanning electron microscope showed healthy and sickly *Methanosaeta* filaments as well as ultramicrobacteria on the filaments (Fig. S12 and S13). Scanning electron microscopy (SEM) in combination with CARD-FISH detected some intact *Methanosaeta* cells in filaments together with OP3 LiM cells attached to a degraded part of the filament (Fig. S19).

Closed genome of OP3 LiM. In density gradients, OP3 LiM cells enriched in density gradients were the basis for 454 and Illumina metagenomes. A first closed draft genome had uncertainties in repetitive regions, indicated by 454 read mappings. Combinatorial PCRs integrated additional contigs of repetitive elements and confirmed unusual regions, including the presence of three repeats of about 2,000 bp with small sequence variations mixed with two completely identical sequences of 2,506 bp within the open reading frame (ORF) of the very large multienzyme surface protein.

The assembled OP3 LiM genome is a single chromosome of 1,974,201 bp with a GC content of 52.9%. We placed the genome start point at the origin of replication (*ori*) at the N terminus of *dnaA*.

The asymmetry of the nucleotide composition between leading and lagging strands indicated the first base of the start codon of *dnaA* coding for the chromosomal replication initiator protein DnaA (27) as base 1. The terminus is about 950,000 bp away from *ori*, based on the shift points of the GC skew graphs (28). Gene orientation is highly ordered. Eighty-five percent of the genes are oriented in the direction of DNA replication, and only 15% are oriented opposite to the direction of DNA replication. We circulated the genome in a GC-homopolymeric region close to the terminus; however, reads covering both regions adjacent to the GC-homopolymer were not obtained in both 454 and Illumina MiSeq read sets. The genome has 1,851 protein-coding open reading frames, 45 tRNA genes, and one rRNA operon, with a group I intron in the 23S rRNA gene. The intron encodes a homing endonuclease in the peripheral stem-loop regions of the group I ribozyme (29, 30). The identified insertion position 1,917 (referring to the *Escherichia coli* sequence) has been reported as a preferred insertion site (31), e.g., in *Coxiella burnetii* (32), *Thermotoga subterranea* (32), *Thiomargarita* sp. clone NAM092 (33), *Synechococcus* sp. strain C9 (29), and groundwater-associated bacteria (31). The 16S rRNA gene of the genome has a sequence identity of 99% to 100% to 16S rRNA gene clone sequences obtained in the initial characterization of the limonene enrichment culture (19). Only one other 16S rRNA gene sequence is known with an identity above 97% with the genomic gene sequence, a clone sequence from a methanogenic benzene degradation (GenBank number [KT028835](#)). The identity to 16S rRNA genes of validly described bacteria was below 80%, confirming a genetic distance on the level of phyla. Besides 45 tRNA genes, the genome codes for two noncoding RNA molecules, assigned to SsrS and Rnp, likely with regulating (inhibitory) functions. The genome encodes a type I restriction-modification cluster (*hsdMSR*), a type II restriction-modification system, and a subunit of a type III restriction-modification system. Three candidate loci for clustered regularly interspaced short palindromic repeats (CRISPR) providing acquired immunity against foreign genetic elements (34) were identified at positions 875578 to 875727, 1062302 to 1062425, and 1625552 to 1625657. Each CRISPR locus was separated by only one spacer.

Proteomic insights into the physiology of OP3 LiM. A second method to enrich OP3 LiM cells was differential centrifugation. A 10,000S pellet collected all large and aggregated cells (10kS cells). Small cells in the supernatant were collected in a 100S pellet that after resuspension was separated by a second 10,000S centrifugation into aggregated cells (second 10,000S pellet [100S aggregates]) and free-living cells (supernatant of second 10,000S centrifugation [100S cells]). SEM images of the different fractions confirmed the size fractionation. Small single cells accounted for more than 99% of all cells in the 100S aggregate and 100S cell fractions. All fractions (10kS cells, 100S aggregates, and 100S cells) were sources of metaproteomes. A total of 1,279 proteins encoded in the OP3 LiM genome were identified in the proteomes of two biological replicates and gave insight into the metabolism (Table S1). Most frequently detected was a 18-stranded beta-sheet outer membrane pore protein that facilitates the uptake and export of macromolecules. Three other large proteins with transmembrane helices are among the most frequently detected proteins, a glycosylphosphatidylinositol (GPI) membrane-anchored actin/protein binding protein with Kelch repeats (499 kDa), a protein with unknown function (326 kDa), and a very large multienzyme surface protein (4,384 kDa). Its gene codes for 39,678 amino acids, containing hundreds of protein domains. Most were predicted to be outside the cell, anchored by 42 predicted transmembrane helices. Among 44 conserved enzyme domains were 7 glycosyl transferases, 3 glycosyl hydrolases, and a sugar epimerase, likely catalyzing the degradation of external polysaccharides. Three peptidases and a

DnaJ chaperone may act on proteins. A phosphatase, a dehydrogenase, a methyltransferase, and an acetyltransferase completed the identified range of degradative enzyme domains. A signal receiver domain and a cellular signaling response via cyclic guanine nucleotides by a diguanylate cyclase/phosphodiesterase (GGDEF and EAL domains) may serve as an environmental signal sensor. The nucleotide metabolism may be influenced by three ATP-binding domains, a protein kinase and two nucleotide kinases, two ppGpp synthases/hydrolases, and two single-strand-DNA (ssDNA)-binding and two double-strand-DNA (dsDNA)-binding domains. The gene for the secreted very large multienzyme surface protein is located on the leading strand directly after the genes for proteins involved in DNA replication and in a secretion pathway. This is the expected position for highly expressed proteins minimizing collisions of replicating and transcribing polymerases (35, 36). Together with a pilus secretin, this large protein was induced 3-fold in attached OP3 LiM cells (10kS cells) in comparison to free-living cells (100S cells), suggesting a function as an “attack” protein. We propose that the very large surface protein stains as a coat that is visible in TEM images.

Among the detected transporter proteins were sugar ABC transporters. Monosaccharide degradation likely involves a class II fructose-bisphosphate aldolase. A pyruvate:ferredoxin oxidoreductase may provide reducing equivalents for energy conservation via a membrane-integrated Rnf complex (ferredoxin:NAD oxidoreductase). The proteome supported a polysaccharide fermentation to acetate, carbon dioxide, and eventually hydrogen or formate. OP3 LiM has the potential to conserve energy from the phosphoester bonds in nucleic acids. An abundant polyribonucleotide nucleotidyl transferase is expected to move a nucleoside monophosphate from RNA onto phosphate, thus synthesizing a nucleoside diphosphate. Other potentially energy-conserving enzymes were pyruvate phosphate dikinase and a proton-translocating pyrophosphatase.

Based on N-terminally encoded signal peptides in genes, the periplasm also contains a peptidyl prolyl *cis-trans* isomerase, a Clp protease and an ATPase that has, in addition to the CpaF/VirB11-like ATPase domain, an N-terminal CobQ/CobB/MinD/ParA nucleotide-binding domain, suggesting participation in nucleic acid transfer. Although OP3 has no flagellum, an abundant protein was MotD, a stator ring protein for a flagellum forming a pore in the inner membrane that is known to be filled by a type III secretion system and accessory proteins. The second stator protein, MotB, was also expressed. Other abundant proteins were an outer membrane protein assembly factor, the Sec translocase SecD/SecE, and the bacterial and archaeal forms of the chromosome segregation protein SMC, which binds nucleic acids. The bacterial actin MreB was also highly expressed.

We searched for other proteins potentially involved in the uptake of nucleic acids or other macromolecules. Besides the aforementioned extracellular nucleotide-binding CpaF-related ATPase, three PilT ATPases were present to export/import macromolecules. Of the *pil* genes, CpaB, PilM, PilO, PilQ, and two PilZ proteins were expressed. A PilA-related pseudopilin, PulG, was the most abundant pilin. OP3 LiM has an active Sec secretion system and a type II secretion/type IV pilus assembly system. Several loci contained *pil*-related genes, including one locus in front of the very large multienzyme surface protein with several export proteins of a type II secretion system.

A FIC protein was expressed. FIC proteins modify proteins by posttranslational AMPylation, which results in the activation of toxins of pathogenic bacteria and inhibition of host/prey small GTPases, thereby reducing metabolic activities (37). The FIC protein was shown to induce hibernation (38). It may be responsible for the low 16S rRNA content of OP3 LiM cells in the free-living state.

The OP3 LiM proteomes highlighted the presence of large outer membrane pores and of polymer-binding and -depolymerizing domains located extracellularly and in the periplasm. The metabolism of OP3 LiM was dominated by sugar fermentation. However, the genome also encodes a syntrophic life on hydrogen or formate, with carbon dioxide fixation on the Wood–Ljungdahl pathway using formate dehydrogenase, the reduction of a formyl group on tetrahydrofolate, and methyl transfer to an acetyl coenzyme A (acetyl-CoA) synthase, but these enzymes were not highly expressed. Among the four hydrogenases, an uptake hydrogenase (hydrogen:NAD oxidoreductase) was the most highly expressed. Interestingly, high expression of serine hydroxymethyltransferase suggested serine and glycine as precursors of C₁ compounds. Building blocks (amino acids, ribose, nucleotides, and fatty acids) were synthesized *de novo* according to expressed enzymes. OP3 LiM may eventually salvage these monomers from the environment, but it did not seem to depend on salvage pathways. D-Alanine and D-glutamate racemases as well as peptidoglycan biosynthetic enzymes were expressed. Together with the presence of peptidoglycan-binding domains in several secreted proteins, this suggested the presence of a murein sacculus in OP3 LiM. Genes for the synthesis of the LPS core glycolipid suggested the well-known lipid asymmetry in the outer membrane.

DISCUSSION

Methanosaeta is likely one of the most abundant microbes on Earth (39). Individual cells of this anaerobic acetoclastic archaeon are strongly separated from the environment and from the neighboring cells in a filament. They are enclosed in an amorphous matrix and shielded by a protein- and sugar-containing sheath and by spacer plugs that grow out from the sheath during cell division (22). Together with the demand for only one transporter for substrate (acetate), this morphology offers few entrance points for predators, which can be interpreted as a necessity to balance the slow growth of the archaeon in the environment. In this study, using a variety of microscopic investigations and specific stains for RNA, DNA and lipids as well as the integrity of the cellular lipid layer, we found that *Methanosaeta* cells were individually attacked. We detected cells with a decreased biovolume, a decreased concentration of cellular components in general, and decreased concentrations of RNA and DNA. Dead cells were present in filaments. This predation pressure on individual *Methanosaeta* cells plausibly explains the evolution of spacer plugs and the isolation of cells. The filament can continue to live although individual cells are dead.

These observations ask for the identification of the biological predator. Phages have been observed in methanogenic reactors (40) but not convincingly shown in *Methanosaeta* cells. Our thin-section electron micrographs show nonstaining structures within *Methanosaeta* cells. The absence of stain indicates a neutral or cationic surface that does not bind the heavy metal cations used for staining. Whether these structures are polysaccharide storage compounds or viral coats may be answered in future studies. In this study, we focused on the characterization of the phylotype candidate division OP3 LiM. It is a chemoheterotroph attaching to and living off *Bacteria* and *Archaea*. As an ultramicrobacterium (0.2 to 0.3 μm in diameter), it is barely or not visible by phase-contrast microscopy (optical resolution > 0.3 μm). In addition, it is not detectable by the standard FISH probes EUBI to -III for *Bacteria* (19). To detect it, we used a range of electron-microscopic techniques in combination with *in situ* hybridizations.

The presence of the candidate division OP3 in many anaerobic habitats and cultures has been established in hundreds of studies using gene-based, cultivation-independent detection by 16S rRNA amplicon sequences and metagenome-assembled genomes (MAGs). For example, over 100 MAGs

for the phylum “*Candidatus Omnitrophica*” (18) have been established, but a closed genome was lacking due to the absence of a pure strain. We developed a physical enrichment of OP3 LiM cells and obtained after PCR-based integration of repetitive regions in the genome the first closed genome for the phylum. Metaproteomes enriched in attached OP3 LiM cells and in free-living OP3 LiM cells gave insights into the biology of OP3 LiM. It may grow preferentially as a sugar fermenter.

Most bacteria are Gram-negative; they have two membranes, and the outer layer of the outer membrane consists of lipopolysaccharides (LPS) with a lipid anchor, a polysaccharide core, and a long polysaccharide chain directed to the environment. The biosynthesis of LPS is well understood, but it is a one-way synthesis: lipopolysaccharides are not recycled by their producers. They provide the substrate for ultramicrobial and small epibionts, for example, “*Nanosynbacter lyticus*” from the phylum “*Candidatus Saccharibacteria*,” the TM7 phylum (41, 42). The expressed proteins identified in this study suggested that OP3 LiM also grazes the polysaccharide side chains of LPS, using glycosyl transferase and hydrolase domains present in the very large multienzyme surface protein which is 3-fold induced in the metaproteome enriched in attached OP3 cells. The utilization of surface polysaccharides is not limited to epibiotic (ultramicro)bacteria. Cellulosomes of clostridia have been described as extracellular, surface-associated “highly efficient nanomachines” that evolved to perfect depolymerization multienzyme complexes for plant cell wall complex carbohydrates (43). The degradation of LPS polysaccharides by epibionts like “*Nanosynbacter lyticus*” and OP3 LiM has many understudied ecological consequences. For example, in humans, the highly antigenic lipopolysaccharide core of pathogenic bacteria is camouflaged by polysaccharide side chains, and the degradation of these sugars by epibionts exposes the core lipid to the immune system. In addition, lipopolysaccharides have not been used as the substrate to isolate anaerobic microorganisms. Future studies will reveal whether lipopolysaccharide-fermenting ultramicrobacteria can be isolated on membrane particles.

Beyond carbohydrate-active enzymes, OP3 LiM has a range of degradative enzyme domains in its very large multienzyme surface protein. This attack protein may open an entry to the cytoplasm of the prey, as is visible on thin-section electron micrographs. Unusual proteins encoded in the genome and highly expressed proteins suggested a draft of the predation process. To paralyze the prey's metabolism, noncoding RNA molecules and excreted ATPases may block essential metabolic steps and deplete the host's energy charge. The abundant 18-strand beta-barrel pore in the outer membrane has the size to allow import and export macromolecules. Several PiIT ATPases may act in the energy-driven uptake of macromolecules, probably through a stator in the inner membrane. The MreB fragment, a bacterial actin usually absent in coccal cells but present in OP3 LiM, may serve as the foothold and cytoplasmic endpoint of a transport mechanism. An energy-economical depolymerization of nucleic acids is ensured by a nucleotidyltransferase.

This study provides a closed genome of high quality of the OP3 LiM population in a limonene-degrading methanogenic enrichment culture. The smallest inoculum was 1 μ L in a dilution series, potentially presenting a population bottleneck. The 16S rRNA gene sequence indicated a large phylogenetic distance from the next related species. The phylogenetic probe OP3-565 showed the morphology of the identified cells and guided the physical enrichment of OP3 LiM cells and correlation of a conspicuous morphotype in electron micrographs to OP3 LiM. Metaproteomes provided insights to infer the putative metabolism. Based on these observations, we propose naming the OP3 LiM cells “*Candidatus Velamenicoccus archaeovor*us” gen. nov., sp. nov.

Taxonomic note. We deposited the genome of OP3 LiM as “*Candidatus Vampirococcus archaeovor*us” strain LiM on 22 January 2019 at NCBI ([CP019384.1](https://doi.org/10.1093/nar/47/1/CP019384)). Moreira et al. (17) reported in 2021 an anaerobic predatory organism of anoxygenic phototrophic bacteria that fits the original description of the genus *Vampirococcus* better than OP3 LiM. The 16S rRNA gene identity of these two organisms is 67.8%. In consequence, we have updated the NCBI database file of [CP019384](https://doi.org/10.1093/nar/47/1/CP019384) and changed the name to “*Candidatus Velamenicoccus archaeovor*us” strain LiM.

Description of “*Candidatus Velamenicoccus*” gen. nov. *Candidatus Velamenicoccus* (O.L. n. *velamen*, covering; N.L. masc. n. *coccus* [from Gr. masc. n. *kokkos*, grain, seed], berry, coccus; N.L. masc. n. *Velamenicoccus*, a coccus with a covering).

Members of the genus “*Ca. Velamenicoccus*” are obligately anaerobic, chemoheterotrophic bacteria. The only representative so far is the species “*Ca. Velamenicoccus archaeovor*us,” a predatory bacterium. The genus “*Ca. Velamenicoccus*” belongs to a lineage within the candidate division OP3 or phylum “*Candidatus Omnitrophica*.”

Description of “*Candidatus Velamenicoccus archaeovor*us” sp. nov. *Candidatus Velamenicoccus archaeovor*us (Gr. masc. adj. *archaios* [Latin transliteration *archaeos*], ancient; L. v. *voro*, to eat, to devour; N.L. masc. adj. *archaeovor*us, archaea [ancient microorganisms] devouring).

The species is represented by the phylotype and strain LiM and its genome (GenBank number [CP019384](https://doi.org/10.1093/nar/47/1/CP019384)). The genome of 1.97 Mb has a GC content of 52.9%. OP3 LiM was highly enriched in a limonene-degrading methanogenic enrichment culture. Cells are coccoid. They are ultramicrobacteria 0.2 μm in diameter and occur free living and attached to other microorganisms. The bacterium is maintained in slowly growing *Methanosaeta*-rich methanogenic enrichment cultures in freshwater medium with low concentrations of limonene as the carbon source at 28°C. It can be visualized by the FISH probe OP3-565.

MATERIALS AND METHODS

Cultivation of methanogenic enrichment cultures. The methanogenic enrichment culture originated from a wastewater sample taken in 1997 (20). From a dilution-to-extinction series prepared in 1999, a culture that had been inoculated with 1 μL enrichment culture became the origin of all cultures investigated in this study. Twelve lineages were established in 2005 and maintained with an annual transfer of 10% (vol/vol) inoculum. The cultures contained 300 mL freshwater methanogenic medium including 2 mM acetate and 1 mM cysteine, 30 mL 2,2,4,6,8,8-heptamethylnonane (HMN), and 1.5 mL of *R*-(+)-limonene in 500-mL borosilicate bottles (19, 20). OP3 LiM-specific PCR and CARD-FISH were performed to select lineages for the experiments.

Cell separation by Percoll density gradient centrifugation. Cell biomass from 100 mL of enrichment culture was pelleted at 15,500 × *g* for 10 min in a Beckman 70.1 Ti ultracentrifuge rotor (Beckman, Palo Alto, CA). The pellet was resuspended in a mixture of 45 mL Percoll (GE Healthcare, Freiburg, Germany) and 5 mL 1.5 M NaCl. Portions of a 10-mL suspension were centrifuged at 36,680 × *g* for 60 min in the aforementioned rotor. Gradient fractions of 1 mL were assayed for the presence of OP3 LiM cells by applying OP3 LiM-specific PCR with a template dilution series and CARD-FISH with the probe OP3-565. Fractions enriched in OP3 LiM cells had densities of 1.05 ± 0.05 g/cm³ and were combined from several separations. The enriched fractions were further concentrated with a second gradient centrifugation using the aforementioned conditions. OP3 LiM

cells were macroscopically visible as a band at a density of 1.05 g/cm^3 and were collected for an OP3 LiM-enriched metagenome and electron microscopy.

Cell separation by differential centrifugation. For the separation of large cells and aggregates from small cells in enrichment cultures, a Beckman SW28 ultracentrifuge rotor was used at $7,600 \text{ rpm}$ ($7,643 \times g$) for 20 min, corresponding to a sedimentation coefficient of 10,000S. The pellet (10kS cells) was resuspended in 1 mL 10 mM Tris–1 mM EDTA, pH 8.0 (TE). The supernatant was centrifuged at $27,000 \text{ rpm}$ ($96,467 \times g$) for 160 min, corresponding to a sedimentation coefficient of 100S. The pellet of each tube was resuspended in 0.5 mL TE, and a 10,000S pellet of aggregated cells and a few large cells (100S aggregates) was obtained at $12,400 \text{ rpm}$ ($16,331 \times g$, 10,000 S) for 3 min. Cell suspensions (10kS cells and 100S aggregates) and the supernatant of the last centrifugation (100S cells) were stored at -80°C . Alternatively, 100S cells were pelleted using a Beckmann Ti 50.2 ultracentrifuge rotor at $184,048 \times g$ for 55 min (100S cell pellet).

Transmission electron microscopy. For negative-staining transmission electron microscopy, bacterial cultures were adsorbed onto carbon film, washed in 20 mM Tris HCl–1 mM EDTA, pH 6.9, and stained for 1 min with 4% (wt/vol) aqueous uranyl acetate (44). After transfer onto copper grids and air drying, samples were examined in a Zeiss EM902A transmission electron microscope (TEM) (Zeiss, Oberkochen, Germany) operated at 80 kV at calibrated magnifications. Cells were analyzed using the program MeasureIT (Olympus Soft Imaging System GmbH, Münster, Germany).

Thin-section TEM. Bacterial cultures were centrifuged in 1.5-mL portions at $16,000 \times g$ for 5 min or, in later experiments, at $2,500 \times g$ for 5 min (at 28°C) and frozen in their medium together with hexadecane or resuspended with a few microliters of 20% Ficoll in medium before high-pressure freezing on an HPM010 (Abra Fluid AG, Switzerland). The planchettes were freeze substituted in an AFS2 machine (Leica Microsystems, Wetzlar, Germany). Ficoll-containing samples were freeze substituted with 0.1% uranyl acetate as described by Ronchi et al. (45). Hexadecane-containing samples were freeze substituted in 1% osmium tetroxide and 0.1% uranyl acetate as according to Cohen et al. (46). In separate experiments, EPON infiltration was performed in steps, as follows. Samples were rinsed four times at 0°C for 15 min and then infiltrated with 30% (vol/vol) Epon in acetone for 3 h. The temperature was raised to 10°C , and cells were infiltrated with 50% (vol/vol) Epon in acetone mix for 3 h. The temperature was raised to room temperature and cells were infiltrated with a 70% (vol/vol) Epon-acetone mix overnight. Cells were then infiltrated with 100% Epon, and the solution was exchanged after 4 h and left overnight. After a final exchange in the morning, the cells were placed in the oven at 60°C for 48 h. Thin sections were cut at 70 nm on a diamond knife (Diatome, Switzerland), floated onto Formvar-coated slot grids (G2500PD; Plano GmbH, Wetzlar, Germany), poststained with uranyl acetate and Reynolds lead citrate, and imaged on either a JEOL 2100 Plus (200-kV) or a CM120 Philips Biotwin (120-kV) electron microscope.

Scanning electron microscopy. Cells fixed with formaldehyde (1.3% [wt/vol] in $1\times$ phosphate-buffered saline, pH 7.4 [PBS]) were spotted on silicon wafers (5 by 7 mm; Plano, Wetzlar, Germany) and incubated for 60 min at room temperature. Then cells were dehydrated in an ethanol series (30%, 50%, 70%, 80%, and 96% [vol/vol], each for 10 min) and critical point dried (Leica EM CPD 300; Leica, Vienna, Austria). Secondary electron micrographs were obtained with a Quanta FEG 250 (FEI, Hillsboro, OR, USA).

Extraction of nucleic acids. OP3 LiM-enriched fractions were extracted for genomic DNA according to Martín-Platero et al. (47). After in-house quality control by spectroscopy and agarose gel

electrophoresis, sample quality control by capillary electrophoresis and sequencing was performed by the Max Planck Genome Center, Cologne, Germany (<http://mpgc.mpipz.mpg.de/home/>). For PCR analyses, DNA was extracted from biomass of 1-mL samples of enrichment culture using the FastDNA spin kit for soil (MP Biomedicals, Santa Ana, CA, USA) according to the manufacturer's instructions.

OP3 LiM-specific PCR. OP3 LiM was detected using as the template 1.0 ng extracted genomic DNA or 1.0 μ L of culture with an optical density at 600 nm (OD_{600}) of 0.2 to 0.5 with freeze-thaw-fractured cells (48) together with a 1.7 μ M concentration each of primers OP3-565F and OP3-1481R (Table 1), 15.0 μ L 2 \times GoTaq master mix (Promega, Madison, WI, USA), and 13 μ L water. The PCR protocol was 4 min at 94°C, 30 cycles of 94°C for 1 min, 62°C for 1 min, and 72°C for 3 min, and finally 72°C for 10 min. After amplicon analysis by separation in a 1% (wt/vol) agarose gel and ethidium bromide staining, sequences were obtained from dideoxynucleotide-terminated oligonucleotides. The sequencing reaction was performed for 60 cycles with an initial denaturing step of 20 s at 96°C, a denaturation of 10 s at 96°C, 5 s at 62°C, and 4 min at 62°C. The products were purified by molecular sieve chromatography using Sephadex G50 Superfine (GE Healthcare Life Sciences, Freiburg, Germany) and were separated on an ABI Prism 3130 XL genetic analyzer (Applied Biosystems, Foster City, CA, USA).

TABLE 1 PCR primers used in this study (Table view)

Primer	Sequence (5'–3')	Position	Description	Reference
OP3-565F	GGGTGTAAAGGCAGGTA	608–626 ^a	OP3-specific 16S rRNA gene forward primer	This study
OP3-1481R	TACGACTTAGCGCCAGTC	1525–1543 ^a	OP3-specific 16S rRNA gene reverse primer	This study
8-27F	AGAGTTTGATCTGGCTCAG	8–27 ^b	Bacterial universal 16S rRNA gene forward primer	71
907R	CCGTC AATTCMTTGAGTTT	907–926 ^b	Bacterial universal 16S rRNA gene reverse primer	72

^a OP3 LiM 16 rRNA location.

^b *E. coli* 16S rRNA location.

Fluorescence *in situ* hybridization of 16S rRNA. For cell identity visualization, we used the catalyzed reporter deposition-fluorescence *in situ* hybridization (CARD-FISH) technique (49). For brighter signals of OP3 LiM cells, clone sequences (GenBank accession numbers FN646451.1, FN646447.1, FN646441.1, FN646440.1, and FN646435.1) were used to manually design helper oligonucleotides, two for each adjacent side of probe OP3-565 (Table 2). They were used at the probe concentration. One mL of enrichment cultures was fixed with formaldehyde (1.3% [wt/vol] in 1 \times PBS) for 60 min at room temperature. Fixed cultures were filtered on 0.2- μ m isopore membrane filters (Millipore, Darmstadt, Germany) and washed three times with 15 mL 1 \times PBS (pH 7.4). After air drying, permeabilization was performed for 60 min at 37°C by lysozyme treatment (10 mg/mL). Probes (Table 3) were hybridized for 160 min at 46°C. After staining with DAPI (1 μ g/mL), cells were visualized using an epifluorescence microscope (Nikon Eclipse 50i; Nikon, Tokyo, Japan) or a confocal laser scanning microscope (CLSM; Zeiss LSM 780, Zeiss, Oberkochen, Germany). The visualization of two populations required two separate CARD-FISH experiments with two probes and differently labeled tyramides for the first and second signal amplification, respectively. Horseradish peroxidase present at the first probe was inactivated after the signal amplification by incubation of the filter with 3% H₂O₂ for 10 min at room temperature. After rinsing with 1 L water and air drying, filters were stored overnight at –20°C, and the second CARD-FISH experiment was performed the next day. Cells were counterstained with 1 μ g/mL DAPI and optionally 5 μ g/mL Nile red in water. To visualize spacer plugs,

the sample was washed a few seconds in water and then quickly in pure ethanol. Cells were embedded in Citifluor-Vectashield (4:1 [vol/vol]).

TABLE 2 Helpers^a (nonlabeled) used and designed in this study (Table view)

Helper	Sequence (5'–3')
H548-A	AATAAATCCGAGTAACGC
H548-C	AATCAATCCGAGTAACGC
H583-TC	CTCCCCACTTGT CAGGCCGCC
H583-CT	CCTCCCCACTTGT CAGGCCGCC

^a All helpers were used in a mix.

TABLE 3 HRP labeled oligonucleotide probes used in this study^a (Table view)

Probe name	Probe sequence (5'–3')	Position	Target group	FA concn (%) ^b	Reference
EUB338 I ^c	GCTGCCTCCCGTAGGAGT	338–355	Most <i>Bacteria</i>	35	73
EUB338 II ^c	GCAGCCACCCGTAGGTGT	338–355	<i>Planctomycetales</i>	35	74
EUB338 III ^c	GCTGCCACCCGTAGGTGT	338–355	<i>Verrucomicrobiales</i>	35	74
ARCH-915	GTGCTCCCCCGCCAATTCCT	915–934	<i>Archaea</i>	35	75
OP3-565	TACCTGCCCTTTACACCC	608–626 ^d	Candidate OP3 LiM	30	19

^a Alternatively, ARCH-915 was labeled with carboxyfluorescein at the bases shown in italics: *GUG CTC CCC CGC CAA TTC CT*.

^b Formamide (FA) concentration in the hybridization buffer.

^c Used in a mix.

^d OP3 LiM 16S rRNA location.

LIVE/DEAD staining of cells. Enrichment samples of 1 mL were centrifuged for 15 min at 4,000 × *g*. The cell pellet was resuspended in 1 mL of anoxic 0.85% (wt/vol) NaCl and mixed with 3 μL of the LIVE/DEAD dye mix (LIVE/DEAD BacLight viability kit [Invitrogen]) yielding 0.03 μmol/mL propidium iodide and 0.005 μmol/mL SYTO 9. After 15 min of incubation in the dark, 5 μL was pipetted on a glass slide (ground edges, 90°; Scientific Menzel-Gläser, Thermo Fisher, Schwerte, Germany) and covered with a coverslip (Menzel-Gläser Cover slips number 1.5). The edges of the coverslip were sealed with clear nail polish. The slides were directly imaged under the CLSM (Zeiss LSM 780) using differential interference contrast to increase contrast in the bright-field image. Propidium iodide was excited with the 561-nm laser line and detected in the window from 570 to 668 nm. For SYTO 9, the 488-nm laser line and the detection window from 499 to 552 nm were used. Laser intensities were adjusted for each sample to avoid overexposure.

SR-SIM. Substrate incubation samples were visualized on a Zeiss ELYRA PS.1 (Carl Zeiss, Jena, Germany) using 561-, 488-, and 405-nm lasers and BP 573-613, BP 502-538, and BP 420-480+LP 750 optical filters. Z-stack images were taken with a Plan-Apochromat 63×/1.4 numerical aperture oil objective and processed with the software ZEN (Carl Zeiss, Jena, Germany). Superresolution structured illumination microscopy (SR-SIM) images are taken by exciting the sample using nonuniform wide-field illumination. The laser light passes through an optical grating, generating a striped sinusoidal interference pattern. This pattern then combines with the sample information originating from structures below the diffraction limit to generate moiré fringes. The image was

detected by an electron-multiplying charge-coupled device (EMCCD) camera and contains high-spatial-frequency sample information shifted to a lower-spatial-frequency band that is transmitted through the objective. Mathematical reconstructions from raw image slices then allow reconstruction of a high-resolution image with doubled resolution in the *x-y* plane (50).

Genome assembly from an OP3 LiM-enriched metagenome. DNA of an OP3 LiM-enriched sample from the second Percoll gradient was sequenced by 454 Titanium pyrosequencing technology (450-bp reads; 454 GS FLX; Roche, Basel, Switzerland) as well as by MiSeq technology (2 × 250-bp reads; Illumina, San Diego, CA, USA) by the Max Planck Genome Center, Cologne, Germany (<http://mpgc.mpipz.mpg.de/home/>). 454 pyrosequencing yielded 491,907 reads. An assembly of 454 reads with Newbler v. 2.3 (51) resulted in a metagenome of 5,779 contigs with 16,026,544 bp. Analysis of tetranucleotide frequencies using JSpecies (52) indicated that five of the six largest contigs covering 1,650,908 bp were part of the OP3 LiM genome. For further analyses, the raw reads of the 454 sequencing were processed in mothur 1.29.1 (53), resulting in 426,697 quality-controlled reads. A MiSeq read set of the same biological sample was used to finish the genome. A total of 9,888,618 paired-end reads were quality-controlled using FastQC (www.bioinformatics.babraham.ac.uk/projects/fastqc/). Raw reads were processed with dynamic trimming with SolexaQA v.2.2. (54) and normalization with Khmer 1.0 (55). The assembly by SPAdes 3.1.0 (56) contained 28,618 contigs with 61,364,565 bp. The assembly was inspected using QUILT v2.3 (57).

Contigs of the Newbler assembly were binned using Metawatt-2.1 (58). Twenty-two contigs of two *Planctomycetales* bins with 1,949,258 bp were selected as targets of a mapping of processed 454 and MiSeq reads within Geneious R8 (Biomatters, Auckland, New Zealand). In addition, MiSeq raw reads were processed with tools of the BMAP package (version 32.27; <http://sourceforge.net/projects/bbmap/>). Contigs of the two bins were extended by read mapping using 454 and MiSeq quality-controlled reads and then *de novo* assembled by Geneious R8 (Biomatters, Auckland, New Zealand). The assembly was improved by 14 rounds of read mapping with BMAP (version 32.27), with assembly of the mapping reads in SPAdes 3.5 taking the actual assembly as trusted and binning of contigs with Metawatt 2.1. When the improvement in the assembly became zero, the MiSeq read data set was trimmed with a different stringency and then used for the next mapping. Almost all bin information was assembled in one contig. Finally, assemblies were based on the Newbler or the first Geneious assembly as the trusted assembly to correct errors introduced in the assembly process. The comparison of the final assemblies revealed a linear presentation of a circular genome with different start points.

Manual visual inspection of mapping results of 454 and MiSeq reads obtained with Geneious R8 and with the Burrows-Wheeler Aligner (BWA) algorithm (Sequencher 5.3; Gene Codes, Ann Arbor, MI, USA) revealed questionable regions in the genome. These regions were repetitive elements which were identified using REPuter (59) and dot plot visualization in Geneious R8 (Biomatters, Auckland, New Zealand). Flanking sequences of repetitive elements were manually identified in the visualization of sequence diversity of mapped reads. This indicated a false-positive assembly of two or more repetitive elements into one in the contig. Sequence comparison using dot plot visualization identified repetitive elements of the draft genome also in two small contigs of the 454 assembly assigned by Metawatt 2.1 to the OP3 LiM bin. Both contigs were *de novo* assembled from 454 reads that mapped to the Newbler contigs using Sequencher 5.3 and Geneious R8 in several mapping rounds to verify

and extent the Newbler assembly for these two contigs. *In silico* read walking from one flanking sequence to the other site across the repetitive elements failed due to read length shortage (repetitive elements were longer than 454 reads) and additional repetitive elements within the genetic content between the repetitive elements.

To clarify the physical sequence order around the triple repetitive element with a second repetitive element within the region, combinatorial PCRs were performed with primers developed with Primer3 v.4.0.0 and located on the flanking sites of repetitive elements. PCR conditions were 4 min at 94°C, 41 cycles of 94°C for 1 min, 55°C for 1 min and 72°C for 4 min, and finally 72°C for 10 min. Amplicons were purified using a PCR purification kit or from agarose gels using a gel extraction kit (both from Qiagen, Hilden, Germany) and sequenced as described above. Sequences of amplicons were obtained using 60 cycles at 96°C for 10 s and 58°C for 5 s with ramping of 1°C per s, and 60°C for 4 min using BigDye Terminator chemistry (Applied Biosystems, Foster City, CA, USA) and analyzed on a 3130xl genetic analyzer (Applied Biosystems). Based on *in silico* read mappings and *in vitro* amplicons and their sequences, the genetic content was assembled between the repetitive elements and integrated into the large contig, thus providing a closed genome of OP3 LiM. The OP3 LiM genome was verified by mapping processed reads onto the OP3 LiM genome in Geneious R9 and visual inspection.

Genome annotation. The OP3 LiM genome was annotated using several pipelines and manual annotations. The NCBI Prokaryotic Genome Annotation (60) was refined using JCoast (61) and Geneious with results of an in-house annotation based on GenDB (62), Rapid Annotations using Subsystems Technology (RAST) (63), and online resources of NCBI. CRISPR structures were searched using CRISPRFinder (64). Artemis release 16.0.0 (65) was applied for the nucleotide composition. RNAMmer 1.2 was used to predict 5S, 16S, and 23S rRNA genes in the genome sequence (66). The number of tRNA genes was identified using ARAGON v1.2.38 (67).

Linearization at *ori*. The OP3 LiM genome was circularized and linearized at the origin of replication (*ori*), with the start codon of the chromosomal replication initiator protein DnaA gene as base 1. The GC skew (68, 69) and the pattern of ORF orientation supported this decision, as analyzed with GenSkew (<http://genskew.csb.univie.ac.at>) and ORF prediction programs. However, the genome contains a homopolymer of 21 guanines as a potential telomere.

Metaproteomic analysis. From two cultures representing biological replicates, cell pellets were obtained by differential centrifugation (10kS cells, 100S aggregates, and 100S cells). They were resuspended in TE buffer (10 mM Tris-HCl, 10 mM EDTA [pH 7.5], containing cOmplete protease inhibitor [Roche]) and extracted by sonication (3 times for 30 s each). The protein content was determined by photometric measurement of the absorption at 595 nm using Nanoquant (Carl Roth, Karlsruhe, Germany). To increase the protein concentration, necessary for direct loading on a gel, samples of 10kS cells and 100S aggregates were concentrated in a vacuum centrifuge. Proteins of the samples were separated by size by using SDS-PAGE. Each gel lane was sliced into 10 equal pieces, and the proteins were digested with trypsin. Peptides were eluted in an ultrasonic bath for 15 min, concentrated, and finally purified with ZipTips with C₁₈ resin (Millipore, Billerica, MA, USA). The peptide mix was separated on a nano-high-performance liquid chromatograph (nano-HPLC) (Easy-nLCII HPLC system; Thermo Fisher Scientific, Dreieich, Germany) and analyzed by tandem mass spectrometry (MS/MS) in an LTQ Orbitrap Velos mass spectrometer (Thermo Fisher Scientific) (70). For protein identification, tandem mass spectra were extracted using the Sorcerer-SEQUEST

platform, version 3.5 (Sage-N Research, Milpitas, CA), searching the MS/MS data against a metaproteome database generated from the metagenomes and common laboratory contaminants. Search parameters were a parent ion tolerance of 10 ppm, fragment ion mass tolerance of 1.00 Da, and oxidation of methionine (15.99 Da) as variable modification (maximum of three modifications per peptide). MS/MS-based peptide and protein identifications were validated with Scaffold V4.4.8 (Proteome Software, Portland, OR). Peptide false discovery rates (FDRs) were set to 1%, and protein FDRs were set to 5% throughout all experiments. Quantification of each protein was performed as total spectral counts (TSC) for each protein, because the molecular weight of ORFs varied largely.

Data availability. The genome of OP3 LiM was deposited under number [CP019384](https://doi.org/10.1093/ncbi/CP019384) in GenBank. The mass spectrometry proteomics data were deposited with the ProteomeXchange Consortium via the PRIDE partner repository with the data set identifiers [PXD025008](https://doi.org/10.6019/PXD025008) and <https://doi.org/10.6019/PXD025008>.

SUPPLEMENTARY MATERIAL

Supplemental material is available online only.

Supplemental file 1

Fig. S1 to S19 and Table S1. Download [aem.02407-21-s0001.pdf](#), PDF file, 8.2 MB

ACKNOWLEDGMENTS

We thank the Max Planck-Genome-center Cologne (<http://mpgc.mpiiz.mpg.de/home/>) for performing NGS sequencing.

This study was funded by the Max Planck Society. Jana Kizina and Almud Lonsing are members of the International Max Planck Research School of Marine Microbiology (MarMic).

Concept: J.H., investigation: J.K., S.F.A.J., G.A.M., A.L., C.P., A.K., R.S.-M., K.S., J.H., data curation: J.K., S.M., J.H., resources and supervision: R.S.-M., E.R., S.L., M.R., T.S., J.H., visualization and writing original draft: J.K., J.H., writing - review and editing: all authors.

REFERENCES

1. Bada JL. 2013. New insights into prebiotic chemistry from Stanley Miller's spark discharge experiments. *Chem Soc Rev* 42:2186–2196. [Crossref](#). [PubMed](#). [Web of Science](#).
2. Robertson MP, Joyce GF. 2012. The origins of the RNA world. *Cold Spring Harb Perspect Biol* 4:a003608. [Crossref](#). [PubMed](#). [Web of Science](#).
3. Martin WF, Sousa FL, Lane N. 2014. Evolution. Energy at life's origin. *Science* 344:1092–1093. [Crossref](#). [PubMed](#). [Web of Science](#).
4. Bengtson S. 2002. Origins and early evolution of predation, p 289–317. In Kowalewski M, Kelley PH (ed), *The fossil record of predation. The Paleontological Society papers 8*. The Paleontological Society, McLean, VA. [Crossref](#).
5. Jurkevitch E. 2007. A brief history of short bacteria: a chronicle of *Bdellovibrio* (and like organisms) research, p 1–9. In Jurkevitch E (ed), *Predatory prokaryotes*. Springer, Berlin, Germany. [Crossref](#).
6. Pérez J, Moraleda-Muñoz A, Marcos-Torres FJ, Muñoz-Dorado J. 2016. Bacterial predation: 75 years and counting! *Environ Microbiol* 18:766–779. [Crossref](#). [PubMed](#). [Web of Science](#).
7. Jurkevitch E. 2007b. Predatory behaviors in bacteria-diversity and transitions. *Microbe* 2:67–73. [Crossref](#).
8. Soo RM, Woodcroft BJ, Parks DH, Tyson GW, Hugenholtz P. 2015. Back from the dead; the curious tale of the predatory cyanobacterium *Vampirovibrio chlorellavorus*. *PeerJ* 3:e968. [Crossref](#). [PubMed](#).
9. Tudor JJ, McCann MP. 2007. Genomic analysis and molecular biology of predatory prokaryotes, p 153–189. In Jurkevitch E (ed), *Predatory prokaryotes*. Springer, Berlin, Germany. [Crossref](#).

10. Jurkevitch E, Davidov Y. 2006. Phylogenetic Diversity and Evolution of Predatory Prokaryotes. P 11–56. In Jurkevitch E (ed), *Predatory prokaryotes*. Springer, Berlin, Germany.
11. Mahmoud KK, Koval SF. 2010. Characterization of type IV pili in the life cycle of the predator bacterium *Bdellovibrio*. *Microbiology (Reading)* 156:1040–1051. [Crossref](#). [PubMed](#). [Web of Science](#).
12. Wang Z, Kadouri DE, Wu M. 2011. Genomic insights into an obligate epibiotic bacterial predator: *Micavibrio aeruginosavorus* ARL-13. *BMC Genomics* 12:453. [Crossref](#). [PubMed](#). [Web of Science](#).
13. Koval SF, Hynes SH, Flannagan RS, Pasternak Z, Davidov Y, Jurkevitch E. 2013. *Bdellovibrio exovor*us sp. nov., a novel predator of *Caulobacter crescentus*. *Int J Syst Evol Microbiol* 63:146–151. [Crossref](#). [PubMed](#). [Web of Science](#).
14. Ganuza E, Sellers CE, Bennett BW, Lyons EM, Carney LT. 2016. A novel treatment protects *Chlorella* at commercial scale from the predatory bacterium *Vampirovibrio chlorellavor*us. *Front Microbiol* 7:848. [Crossref](#). [PubMed](#). [Web of Science](#).
15. Bedree JK, Bor B, Cen L, Edlund A, Lux R, McLean JS, Shi W, He X. 2018. Quorum sensing modulates the epibiotic-parasitic relationship between *Actinomyces odontolyticus* and its *Saccharibacteria* epibiont, a *Nanosynbacter lyticus* strain, TM7x. *Front Microbiol* 9:2049. [Crossref](#). [PubMed](#). [Web of Science](#).
16. Guerrero R, Pedros-Alio C, Esteve I, Mas J, Chase D, Margulis L. 1986. Predatory prokaryotes: predation and primary consumption evolved in bacteria. *Proc Natl Acad Sci USA* 83:2138–2142. [Crossref](#). [PubMed](#). [Web of Science](#).
17. Moreira D, Zivanovic Y, López-Archilla AI, Iniesto M, López-García P. 2021. Reductive evolution and unique predatory mode in the CPR bacterium *Vampirococcus lugosii*. *Nature Comm* 12:2454. [Crossref](#). [PubMed](#).
18. Rinke C, Schwientek P, Sczyrba A, Ivanova NN, Anderson IJ, Cheng JF, Darling A, Malfatti S, Swan BK, Gies EA, Dodsworth JA, Hedlund BP, Tsiamis G, Sievert SM, Liu WT, Eisen JA, Hallam SJ, Kyrpides NC, Stepanauskas R, Rubin EM, Hugenholtz P, Woyke T. 2013. Insights into the phylogeny and coding potential of microbial dark matter. *Nature* 499:431–436. [Crossref](#). [PubMed](#). [Web of Science](#).
19. Rotaru AE, Schauer R, Probian C, Mussmann M, Harder J. 2012. Visualization of candidate division OP3 cocci in limonene-degrading methanogenic cultures. *J Microbiol Biotechnol* 22:457–461. [Crossref](#). [PubMed](#). [Web of Science](#).
20. Harder J, Foss S. 1999. Anaerobic formation of the aromatic hydrocarbon p-cymene from monoterpenes by methanogenic enrichment cultures. *Geomicrobiol J* 16:295–305. [Crossref](#). [Web of Science](#).
21. Raskin L, Stromley JM, Rittmann BE, Stahl DA. 1994. Group-specific 16S rRNA hybridization probes to describe natural communities of methanogens. *Appl Environ Microbiol* 60:1232–1240. [Crossref](#). [PubMed](#). [Web of Science](#).
22. Patel GB, Sprott GD, Humphrey RW, Beveridge TJ. 1986. Comparative analyses of the sheath structures of *Methanothrix concilii* GP6 and *Methanospirillum hungatei* strains GP1 and JF1. *Can J Microbiol* 32:623–631. [Crossref](#). [Web of Science](#).
23. Kubota K, Imachi H, Kawakami S, Nakamura K, Harada H, Ohashi A. 2008. Evaluation of enzymatic cell treatments for application of CARD-FISH to methanogens. *J Microbiol Methods* 72:54–59. [Crossref](#). [PubMed](#). [Web of Science](#).
24. Yamaguchi T, Kawakami S, Hatamoto M, Imachi H, Takahashi M, Araki N, Yamaguchi T, Kubota K. 2015. In situ DNA-hybridization chain reaction (HCR): a facilitated in situ HCR system for the detection of environmental microorganisms. *Environ Microbiol* 17:2532–2541. [Crossref](#). [PubMed](#). [Web of Science](#).
25. Mikucki JA, Liu Y, Delwiche M, Colwell FS, Boone DR. 2003. Isolation of a methanogen from deep marine sediments that contain methane hydrates, and description of *Methanoculleus submarinus* sp. nov. *Appl Environ Microbiol* 69:3311–3316. [Crossref](#). [PubMed](#). [Web of Science](#).
26. Fuchs BM, Glöckner FO, Wulf J, Amann R. 2000. Unlabeled helper oligonucleotides increase the in situ accessibility to 16S rRNA of fluorescently labeled oligonucleotide probes. *Appl Environ Microbiol* 66:3603–3607. [Crossref](#). [PubMed](#). [Web of Science](#).

27. Wolański M, Donczew R, Zawilak-Pawlik A, Zakrzewska-Czerwińska J. 2014. oriC-encoded instructions for the initiation of bacterial chromosome replication. *Front Microbiol* 5:735. [Crossref](#). [PubMed](#). [Web of Science](#).
28. Frank AC, Lobry JR. 1999. Asymmetric substitution patterns: a review of possible underlying mutational or selective mechanisms. *Gene* 238:65–77. [Crossref](#). [PubMed](#). [Web of Science](#).
29. Haugen P, Bhattacharya D, Palmer JD, Turner S, Lewis LA, Pryer KM. 2007. Cyanobacterial ribosomal RNA genes with multiple, endonuclease-encoding group I introns. *BMC Evol Biol* 7:159. [Crossref](#). [PubMed](#). [Web of Science](#).
30. Nesbø CL, Doolittle WF. 2003. Active self-splicing group I introns in 23S rRNA genes of hyperthermophilic bacteria, derived from introns in eukaryotic organelles. *Proc Natl Acad Sci USA* 100:10806–10811. [Crossref](#). [PubMed](#). [Web of Science](#).
31. Brown CT, Hug LA, Thomas BC, Sharon I, Castelle CJ, Singh A, Wilkins MJ, Wrighton KC, Williams KH, Banfield JF. 2015. Unusual biology across a group comprising more than 15% of domain *Bacteria*. *Nature* 523:208–211. [Crossref](#). [PubMed](#). [Web of Science](#).
32. Raghavan R, Miller SR, Hicks LD, Minnick MF. 2007. The unusual 23S rRNA gene of *Coxiella burnetii*: two self-splicing group I introns flank a 34-base-pair exon, and one element lacks the canonical omegaG. *J Bacteriol* 189:6572–6579. [Crossref](#). [PubMed](#). [Web of Science](#).
33. Flood BE, Fliss P, Jones DS, Dick GJ, Jain S, Kaster AK, Winkel M, Mußmann M, Bailey J. 2016. Single-cell (meta-)genomics of a dimorphic *Candidatus Thiomargarita nelsonii* reveals genomic plasticity. *Front Microbiol* 7:603. [Crossref](#). [PubMed](#). [Web of Science](#).
34. Horvath P, Barrangou R. 2010. CRISPR/Cas, the immune system of bacteria and archaea. *Science* 327:167–170. [Crossref](#). [PubMed](#). [Web of Science](#).
35. Mao X, Zhang H, Yin Y, Xu Y. 2012. The percentage of bacterial genes on leading versus lagging strands is influenced by multiple balancing forces. *Nucleic Acids Res* 40:8210–8218. [Crossref](#). [PubMed](#). [Web of Science](#).
36. Chen X, Zhang J. 2013. Why are genes encoded on the lagging strand of the bacterial genome? *Genome Biol Evol* 5:2436–2439. [Crossref](#). [PubMed](#). [Web of Science](#).
37. Veyron S, Peyroche G, Cherfils J. 2018. FIC proteins: from bacteria to humans and back again. *Pathog Dis* 76:fty012. [Crossref](#).
38. Harms A, Stanger FV, Scheu PD, de Jong IG, Goepfert A, Glatter T, Gerdes K, Schirmer T, Dehio C. 2015. Adenylation of gyrase and Topo IV by FicT toxins disrupts bacterial DNA topology. *Cell Rep* 12:1497–1507. [Crossref](#). [PubMed](#). [Web of Science](#).
39. Smith KS, Ingram-Smith C. 2007. *Methanosaeta*, the forgotten methanogen? *Trends Microbiol* 15:150–155. [Crossref](#). [PubMed](#). [Web of Science](#).
40. Chien IC, Meschke JS, Gough HL, Ferguson JF. 2013. Characterization of persistent virus-like particles in two acetate-fed methanogenic reactors. *PLoS One* 8:e81040. [Crossref](#). [PubMed](#). [Web of Science](#).
41. Bor B, McLean JS, Foster KR, Cen L, To TT, Serrato-Guillen A, Dewhirst FE, Shi W, He X. 2018. Rapid evolution of decreased host susceptibility drives a stable relationship between ultrasmall parasite TM7x and its bacterial host. *Proc Natl Acad Sci USA* 115:12277–12282. [Crossref](#). [PubMed](#). [Web of Science](#).
42. Bor B, Collins AJ, Murugkar PP, Balasubramanian S, To TT, Hendrickson EL, Bedree JK, Bidlack FB, Johnston CD, Shi W, McLean JS, He X, Dewhirst FE. 2020. Insights obtained by culturing *Saccharibacteria* with their bacterial hosts. *J Dent Res* 99:685–694. [Crossref](#). [PubMed](#). [Web of Science](#).
43. Fontes CM, Gilbert HJ. 2010. Cellulosomes: highly efficient nanomachines designed to deconstruct plant cell wall complex carbohydrates. *Annu Rev Biochem* 79:655–681. [Crossref](#). [PubMed](#). [Web of Science](#).
44. Valentine RC, Shapiro BM, Stadtman ER. 1968. Regulation of glutamine synthetase. XII. Electron microscopy of the enzyme from *Escherichia coli*. *Biochemistry* 7:2143–2152. [Crossref](#). [PubMed](#). [Web of Science](#).
45. Ronchi P, Mizzon G, Machado P, D’Imprima E, Best BT, Cassella L, Schnorrenberg S, Motero MG, Jechlinger M, Ephrussi A, Leptin M, Mahamid J, Schwab Y. 2021. High-precision targeting workflow for volume electron microscopy. *J Cell Biol* 220:e202104069. [Crossref](#). [PubMed](#). [Web of Science](#).

46. Cohen M, Santarella R, Wiesel N, Mattaj I, Gruenbaum Y. 2008. Electron microscopy of lamin and the nuclear lamina in *Caenorhabditis elegans*. *Methods Cell Biol* 88:411–429. [Crossref](#). [PubMed](#). [Web of Science](#).
47. Martín-Platero AM, Valdivia E, Maqueda M, Martínez-Bueno M. 2007. Fast, convenient, and economical method for isolating genomic DNA from lactic acid bacteria using a modification of the protein “salting-out” procedure. *Anal Biochem* 366:102–104. [Crossref](#). [PubMed](#). [Web of Science](#).
48. Hahnke RL, Harder J. 2013. Phylogenetic diversity of *Flavobacteria* isolated from the North Sea on solid media. *Syst Appl Microbiol* 36:497–504. [Crossref](#). [PubMed](#). [Web of Science](#).
49. Pernthaler A, Pernthaler J, Amann R. 2004. Sensitive multi-color fluorescence in situ hybridization for the identification of environmental microorganisms, p 711–726. In Kowalchuk G, de Bruijn FJ, Head IM, Akkermans DL, van Elsas JD (ed), *Molecular microbial Ecology Manual*, 2nd ed, vol 1. Kluwer Academic Publishers, Dordrecht, Boston, London.
50. Schermelleh L, Heintzmann R, Leonhardt H. 2010. A guide to super-resolution fluorescence microscopy. *J Cell Biol* 190:165–175. [Crossref](#). [PubMed](#). [Web of Science](#).
51. Margulies M, Egholm M, Altman WE, Attiya S, Bader JS, Bemben LA, Berka J, Braverman MS, Chen YJ, Chen Z, Dewell SB, Du L, Fierro JM, Gomes XV, Godwin BC, He W, Helgesen S, Ho CH, Irzyk GP, Jando SC, Alenquer ML, Jarvie TP, Jirage KB, Kim JB, Knight JR, Lanza JR, Leamon JH, Lefkowitz SM, Lei M, Li J, Lohman KL, Lu H, Makhijani VB, McDade KE, McKenna MP, Myers EW, Nickerson E, Nobile JR, Plant R, Puc BP, Ronan MT, Roth GT, Sarkis GJ, Simons JF, Simpson JW, Srinivasan M, Tartaro KR, Tomasz A, Vogt KA, Volkmer GA, Wang SH, Wang Y, Weiner MP, Yu P, Begley RF, Rothberg JM. 2005. Genome sequencing in microfabricated high-density picolitre reactors. *Nature* 437:376–380. [Crossref](#). [PubMed](#). [Web of Science](#).
52. Richter M, Rosselló-Móra R. 2009. Shifting the genomic gold standard for the prokaryotic species definition. *Proc Natl Acad Sci USA* 106:19126–19131. [Crossref](#). [PubMed](#). [Web of Science](#).
53. Schloss PD, Westcott SL, Ryabin T, Hall JR, Hartmann M, Hollister EB, Lesniewski RA, Oakley BB, Parks DH, Robinson CJ, Sahl JW, Stres B, Thallinger GG, Van Horn DJ, Weber CF. 2009. Introducing mothur: open-source, platform-independent, community-supported software for describing and comparing microbial communities. *Appl Environ Microbiol* 75:7537–7541. [Crossref](#). [PubMed](#). [Web of Science](#).
54. Cox MP, Peterson DA, Biggs PJ. 2010. SolexaQA: at-a-glance quality assessment of Illumina second-generation sequencing data. *BMC Bioinformatics* 11:485. [Crossref](#). [PubMed](#). [Web of Science](#).
55. Crusoe MR, Alameldin HF, Awad S, Boucher E, Caldwell A, Cartwright R, Charbonneau A, Constantinides B, Edverson G, Fay S, Fenton J, Fenzl T, Fish J, Garcia-Gutierrez L, Garland P, Gluck J, González I, Guermont S, Guo J, Gupta A, Herr JR, Howe A, Hyer A, Härpfer A, Irber L, Kidd R, Lin D, Lippi J, Mansour T, McANulty P, McDonald E, Mizzi J, Murray KD, Nahum JR, Nanlohy K, Nederbragt AJ, Ortiz-Zuazaga H, Ory J, Pell J, Pepe-Ranney C, Russ ZN, Schwarz E, Scott C, Seaman J, Sievert S, Simpson J, Skennerton CT, Spencer J, Srinivasan R, Standage D, et al. 2015. The khmer software package: enabling efficient sequence analysis. *F1000Res* 4:900. [Crossref](#). [PubMed](#).
56. Bankevich A, Nurk S, Antipov D, Gurevich AA, Dvorkin M, Kulikov AS, Lesin VM, Nikolenko SI, Pham S, Prjibelski AD, Pyshkin AV, Sirotkin AV, Vyahhi N, Tesler G, Alekseyev MA, Pevzner PA. 2012. SPAdes: a new genome assembly algorithm and its applications to single-cell sequencing. *J Comput Biol* 19:455–477. [Crossref](#). [PubMed](#). [Web of Science](#).
57. Gurevich A, Saveliev V, Vyahhi N, Tesler G. 2013. QUASt: quality assessment tool for genome assemblies. *Bioinformatics* 29:1072–1075. [Crossref](#). [PubMed](#). [Web of Science](#).
58. Strous M, Kraft B, Bisdorf R, Tegetmeyer HE. 2012. The binning of metagenomic contigs for microbial physiology of mixed cultures. *Front Microbiol* 3:410. [Crossref](#). [PubMed](#). [Web of Science](#).
59. Kurtz S, Choudhuri JV, Ohlebusch E, Schleiermacher C, Stoye J, Giegerich R. 2001. REPuter: the manifold applications of repeat analysis on a genomic scale. *Nucleic Acids Res* 29:4633–4642. [Crossref](#). [PubMed](#). [Web of Science](#).

60. Angiuoli SV, Gussman A, Klimke W, Cochrane G, Field D, Garrity G, Kodira CD, Kyrpides N, Madupu R, Markowitz V, Tatusova T, Thomson N, White O. 2008. Toward an online repository of Standard Operating Procedures (SOPs) for (meta)genomic annotation. *Omics* 12:137–141. [Crossref](#). [PubMed](#). [Web of Science](#).
61. Richter M, Lombardot T, Kostadinov I, Kottmann R, Duhaime MB, Peplies J, Glöckner FO. 2008. JCoast—a biologist-centric software tool for data mining and comparison of prokaryotic (meta)genomes. *BMC Bioinformatics* 9:177. [Crossref](#). [PubMed](#). [Web of Science](#).
62. Meyer F, Goesmann A, McHardy AC, Bartels D, Bekel T, Clausen J, Kalinowski J, Linke B, Rupp O, Giegerich R, Pühler A. 2003. GenDB—an open source genome annotation system for prokaryote genomes. *Nucleic Acids Res* 31:2187–2195. [Crossref](#). [PubMed](#). [Web of Science](#).
63. Aziz RK, Bartels D, Best AA, DeJongh M, Disz T, Edwards RA, Formsma K, Gerdes S, Glass EM, Kubal M, Meyer F, Olsen GJ, Olson R, Osterman AL, Overbeek RA, McNeil LK, Paarmann D, Paczian T, Parrello B, Pusch GD, Reich C, Stevens R, Vassieva O, Vonstein V, Wilke A, Zagnitko O. 2008. The RAST server: rapid annotations using subsystems technology. *BMC Genomics* 9:75. [Crossref](#). [PubMed](#). [Web of Science](#).
64. Grissa I, Vergnaud G, Pourcel C. 2007. CRISPRFinder: a web tool to identify clustered regularly interspaced short palindromic repeats. *Nucleic Acids Res* 35:W52–57. [Crossref](#). [PubMed](#). [Web of Science](#).
65. Rutherford K, Parkhill J, Crook J, Horsnell T, Rice P, Rajandream MA, Barrell B. 2000. Artemis: sequence visualization and annotation. *Bioinformatics* 16:944–945. [Crossref](#). [PubMed](#). [Web of Science](#).
66. Lagesen K, Hallin P, Rødland EA, Staerfeldt HH, Rognes T, Ussery DW. 2007. RNAmmer: consistent and rapid annotation of ribosomal RNA genes. *Nucleic Acids Res* 35:3100–3108. [Crossref](#). [PubMed](#). [Web of Science](#).
67. Laslett D, Canback B. 2004. ARAGORN, a program to detect tRNA genes and tmRNA genes in nucleotide sequences. *Nucleic Acids Res* 32:11–16. [Crossref](#). [PubMed](#). [Web of Science](#).
68. Bao Q, Tian Y, Li W, Xu Z, Xuan Z, Hu S, Dong W, Yang J, Chen Y, Xue Y, Xu Y, Lai X, Huang L, Dong X, Ma Y, Ling L, Tan H, Chen R, Wang J, Yu J, Yang H. 2002. A complete sequence of the *T. tengcongensis* genome. *Genome Res* 12:689–700. [Crossref](#). [PubMed](#). [Web of Science](#).
69. Necşulea A, Lobry JR. 2007. A new method for assessing the effect of replication on DNA base composition asymmetry. *Mol Biol Evol* 24:2169–2179. [Crossref](#). [PubMed](#). [Web of Science](#).
70. Heinz E, Williams TA, Nakjang S, Noël CJ, Swan DC, Goldberg AV, Harris SR, Weinmaier T, Markert S, Becher D, Bernhardt J, Dagan T, Hacker C, Lucocq JM, Schweder T, Rattei T, Hall N, Hirt RP, Embley TM. 2012. The genome of the obligate intracellular parasite *Trachipleistophora hominis*: new insights into microsporidian genome dynamics and reductive evolution. *PLoS Pathog* 8:e1002979. [Crossref](#). [PubMed](#). [Web of Science](#).
71. Lane DJ, Pace B, Olsen GJ, Stahl DA, Sogin ML, Pace NR. 1985. Rapid determination of 16S ribosomal RNA sequences for phylogenetic analyses. *Proc Natl Acad Sci USA* 82:6955–6959. [Crossref](#). [PubMed](#). [Web of Science](#).
72. Lane DJ. 1991. 16S/23S rRNA sequencing, p 115–175. In Stackebrandt E, Goodfellow M (ed), *Nucleic acid techniques in bacterial systematics*. John Wiley and Sons, New York, NY.
73. Amann RI, Krumholz L, Stahl DA. 1990. Fluorescent-oligonucleotide probing of whole cells for determinative, phylogenetic, and environmental studies in microbiology. *J Bacteriol* 172:762–770. [Crossref](#). [PubMed](#). [Web of Science](#).
74. Daims H, Brühl A, Amann R, Schleifer KH, Wagner M. 1999. The domain-specific probe EUB338 is insufficient for the detection of all *Bacteria*: development and evaluation of a more comprehensive probe set. *Syst Appl Microbiol* 22:434–444. [Crossref](#). [PubMed](#). [Web of Science](#).
75. Stahl DA, Amann R. 1991. Development and application of nucleic acid probes in bacterial systematics, p 205–248. In Stackebrandt E, Goodfellow M (ed), *Nucleic acid techniques in bacterial systematics*. John Wiley and Sons Ltd., New York, NY.

Supplemental material: figures S1 – S19 and table S1

Methanosaeta* and *Candidatus Velamenicoccus archaeovorus

Jana Kizina^a, Sebastian F. A. Jordan^a, Gerrit Alexander Martens^a, Almud Lonsing^a,
Christina Probian^a, Androniki Kolovou^b, Rachel Santarella-Mellwig^b, Erhard Rhiel^c,
Sten Littmann^a, Stephanie Markert^d, Kurt Stüber^e, Michael Richter^a, Thomas Schweder^d
and Jens Harder^a#

^aMax Planck Institute for Marine Microbiology, Bremen, Germany

^bElectron Microscopy Core Facility, EMBL Heidelberg, Heidelberg, Germany

^cInstitute for Chemistry and Biology of the Marine Environment, Carl von Ossietzky
University of Oldenburg, Oldenburg, Germany

^dDepartment of Pharmaceutical Biotechnology, Institute for Pharmacy, University of
Greifswald, Germany

^eMax Planck-Genome-centre Cologne, Cologne, Germany

Address correspondence to Jens Harder: jharder@mpi-bremen.de



Fig S1 The fine structure of a sheath in a transmission electron micrograph (see Fig 1 for overview image). In this part of a resuspended cell pellet of the methanogenic limonene enrichment culture, we assigned the sheath to *Methanosaeta*, based on the spacer disc at the end of the filament and the presence and structure of the sheath. The cells inside the sheath may have been shrunk during the TEM preparation. The small black cells with a grey cape were assigned to the ultramicrobacterium OP3 LiM. Bar corresponds to 250 nm.

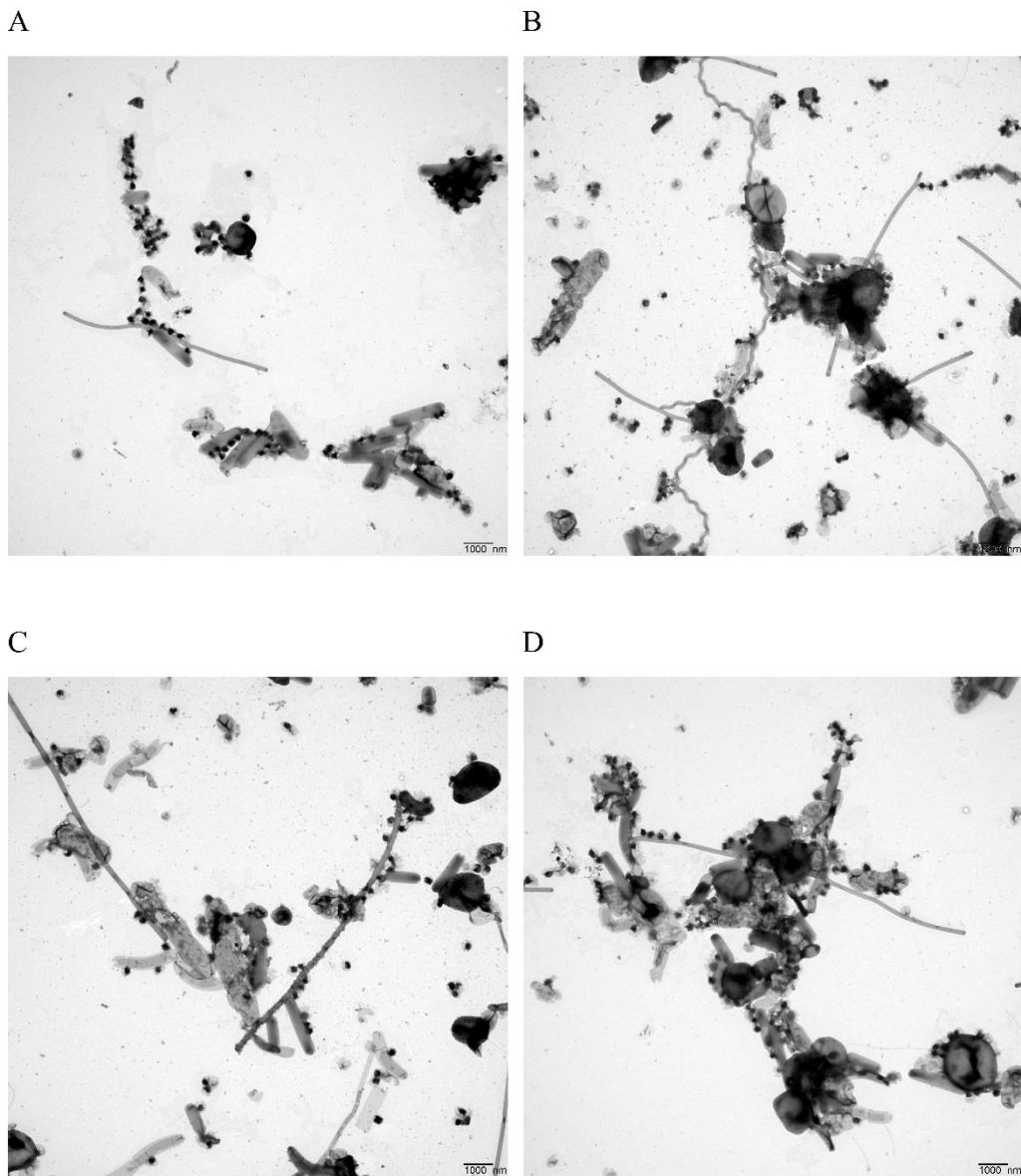


Fig S2 Morphological diversity in the methanogenic limonene enrichment culture as seen in transmission electron micrographs of a resuspended cell pellet. The cells may have been shrunk by the TEM preparation. The small black cells with a grey cape were assigned to the ultramicrobacterium OP3 LiM. Bar corresponds to 1 μm .

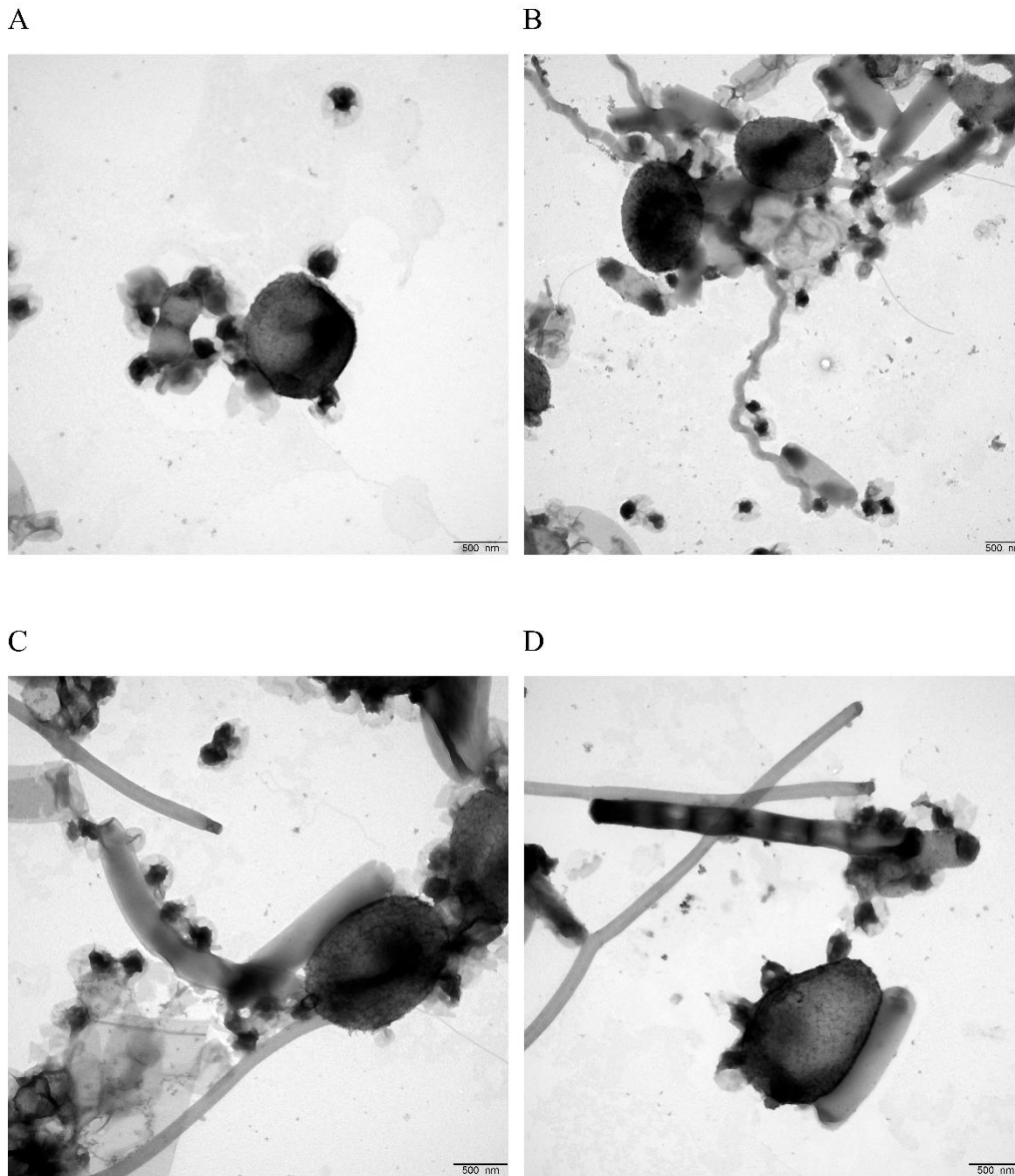


Fig S3 Morphological diversity in the methanogenic limonene enrichment culture as seen in transmission electron micrographs of a resuspended cell pellet. The cells may have been shrunk by the TEM preparation. The small black cells with a grey cape were assigned to the ultramicrobacterium OP3 LiM. Bar corresponds to 500 nm.

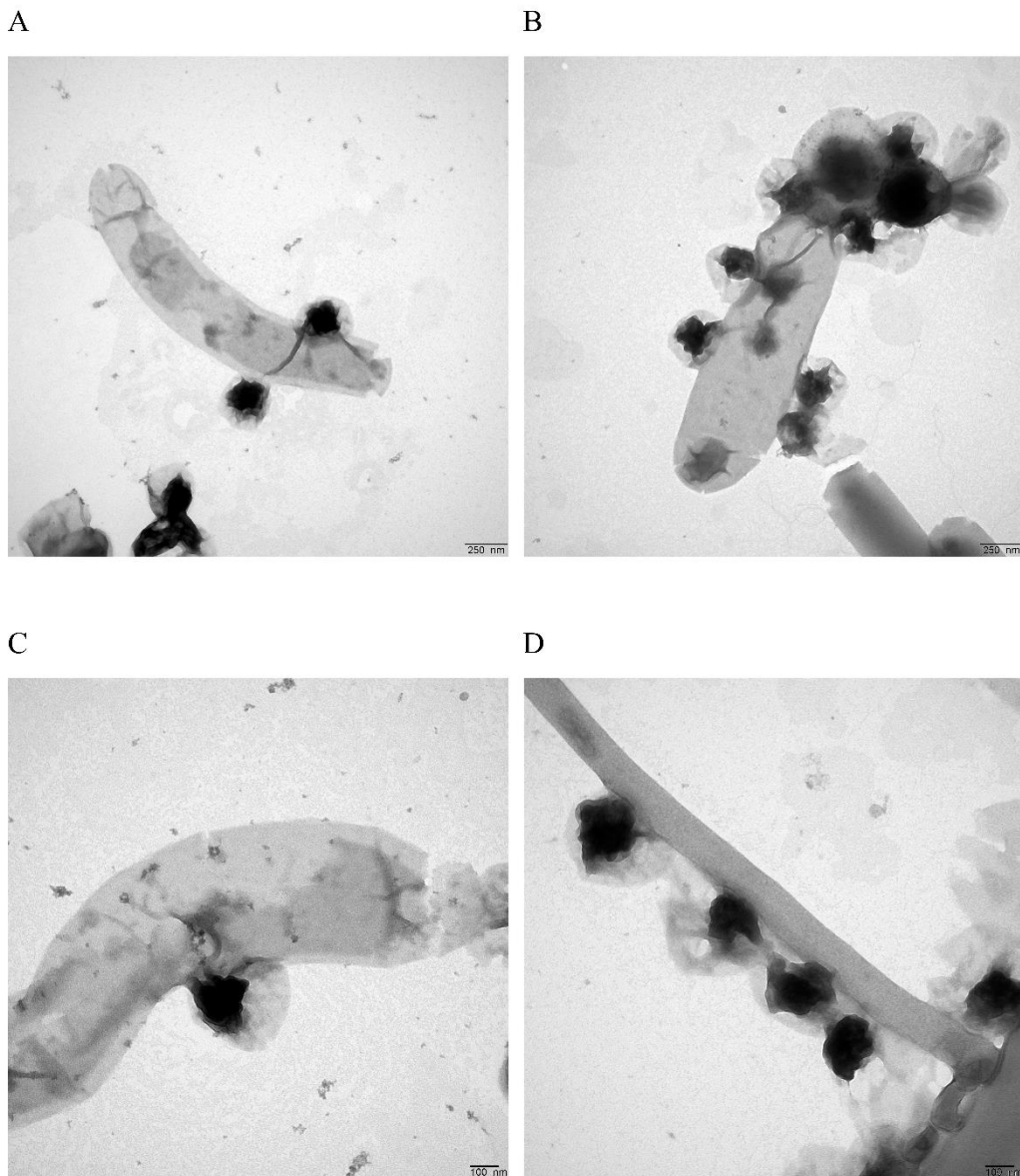


Fig S4 Morphological diversity in the methanogenic limonene enrichment culture as seen in transmission electron micrographs of a resuspended cell pellet. The cells may have been shrunk by the TEM preparation. The small black cells with a grey cape were assigned to the ultramicrobacterium OP3 LiM. Bar corresponds to 250 (A and B) and 100 nm (C and D).

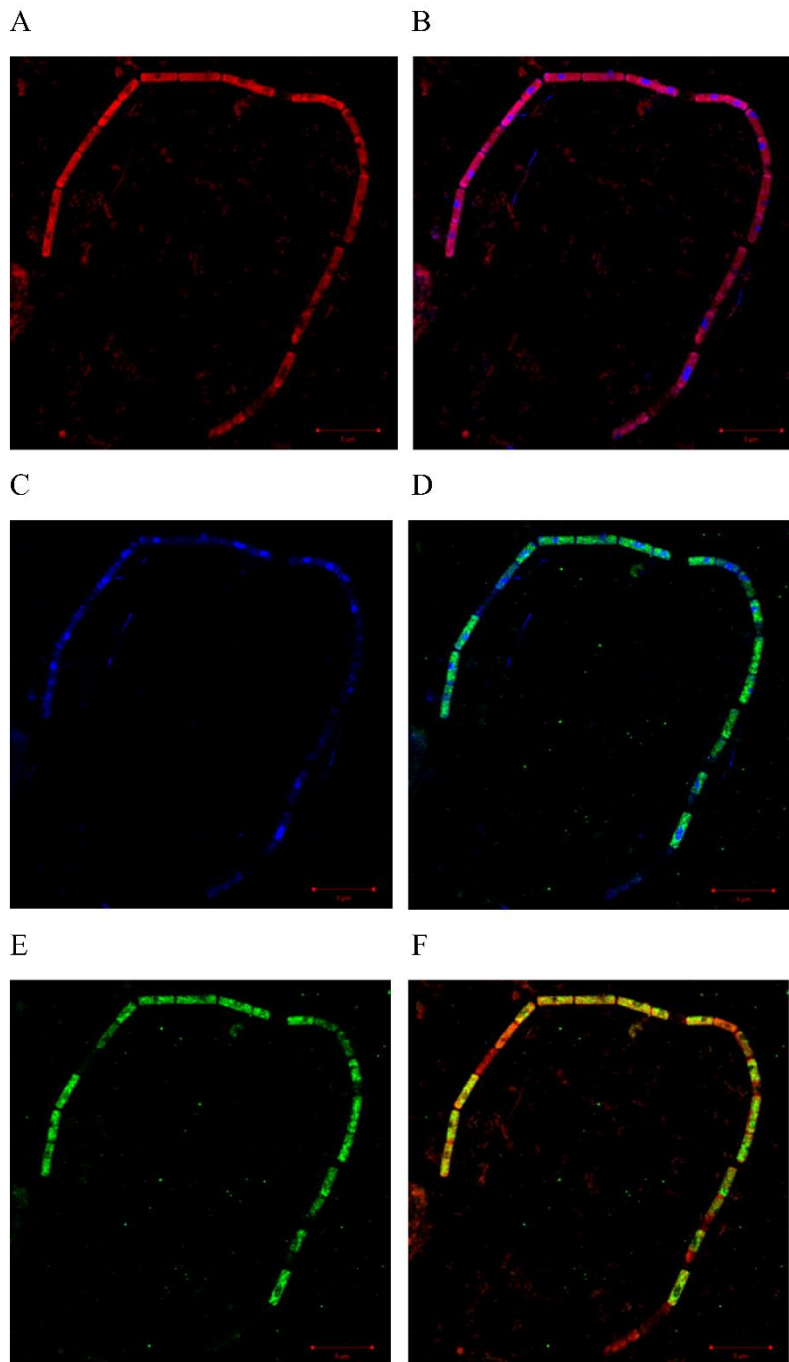


Fig S5 Structured illumination microscopy images of a filament assigned to *Methanosaeta*. Staining was achieved using Nile Red (red) for lipids (A, B, F), probe Arch915 labelled with four 6-carboxyfluorescein molecules (green) for rRNA (D, E, F), DAPI (blue) for DNA B, C, D). Bar corresponds to 5 µm. See Fig 3 for DIC image.

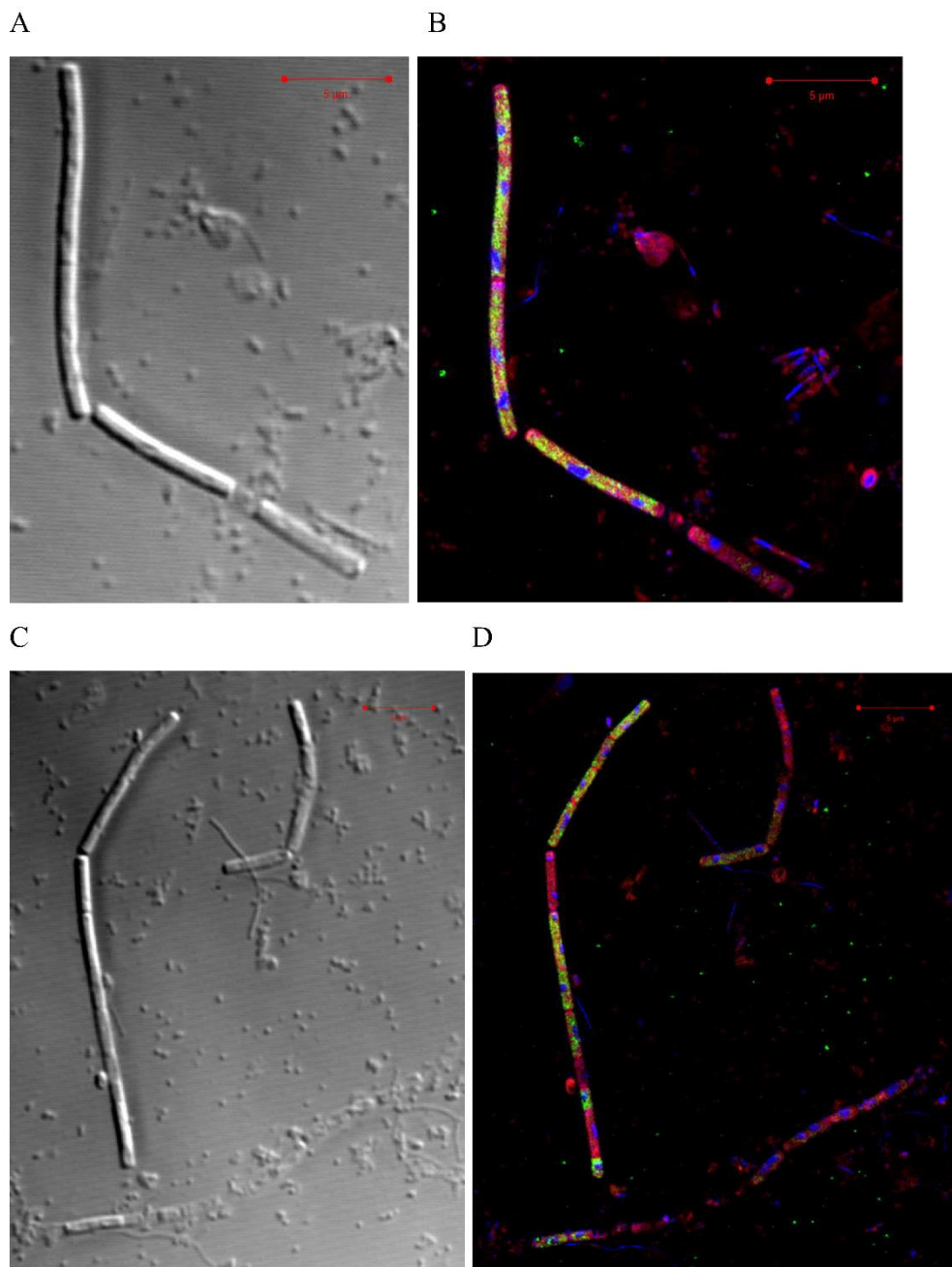


Fig S6 DIC (A, C) and SR-SIM images (B, D) of the limonene enrichment culture. Staining was achieved using Nile Red (red) for lipids, probe Arch915 labelled with four 6-carboxyfluorescein molecules (green) for rRNA, DAPI (blue) for DNA. Bar corresponds to 5 µm.

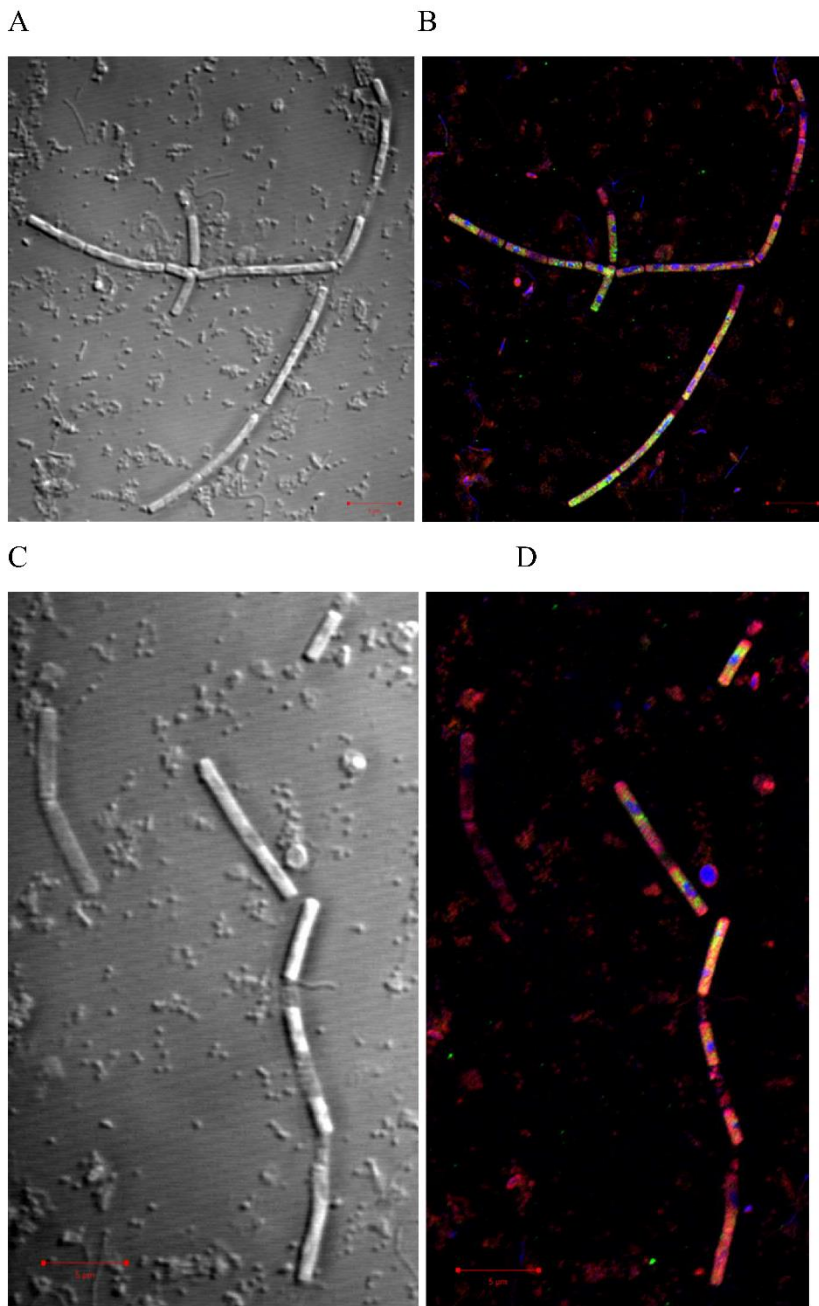


Fig S7 DIC and SR-SIM images (A, C and B, D, respectively) of the limonene enrichment culture. Staining was achieved using Nile Red (red) for lipids, probe Arch915 labelled with four 6-carboxyfluorescein molecules (green) for rRNA, DAPI (blue) for DNA. Bar corresponds to 5 μm .

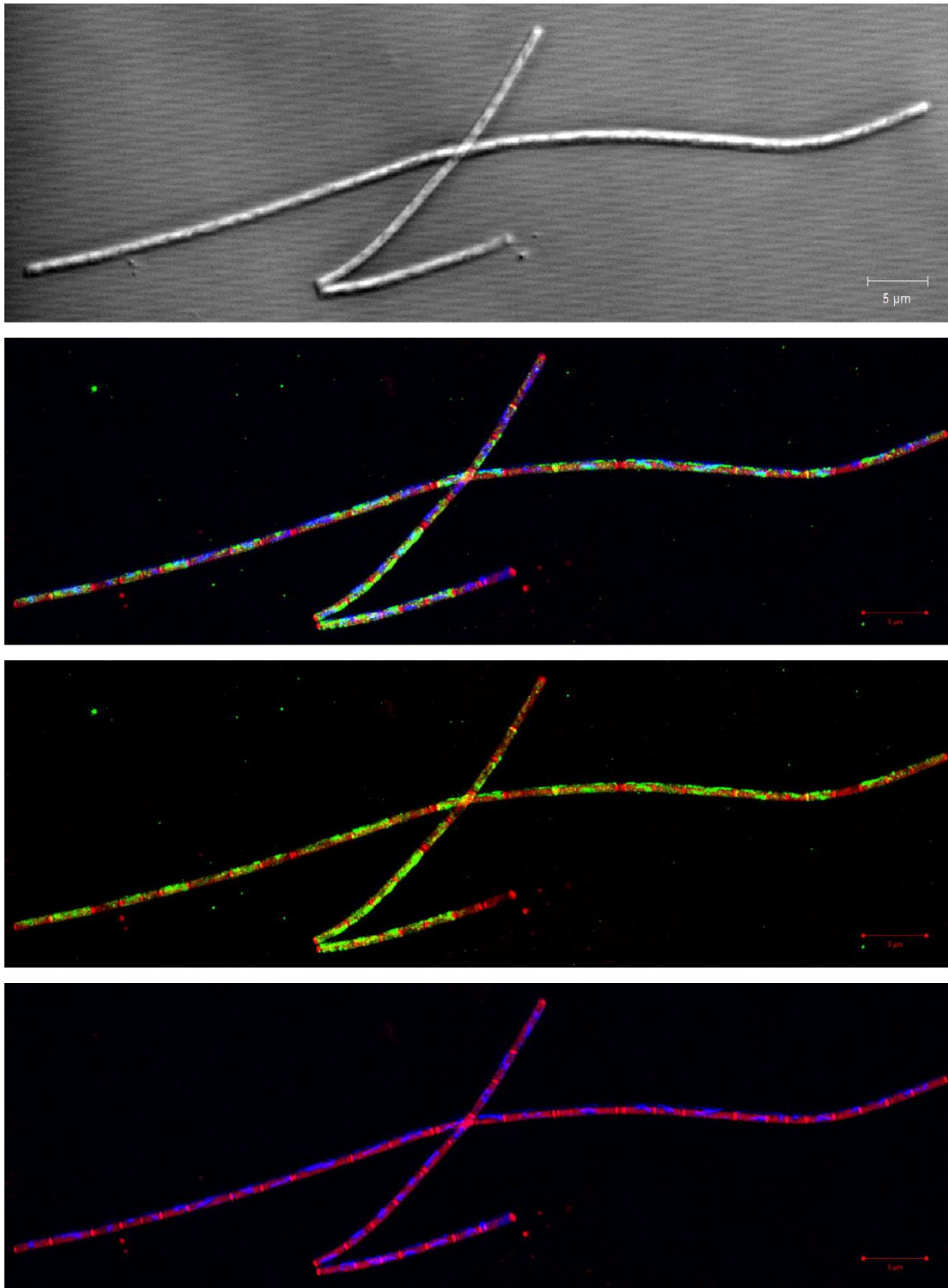


Fig S8 DIC and SR-SIM overlay images of a stationary culture of *Methanosaeta concilii* GP6^T (DSM3671). Staining was achieved using Nile Red (red) for lipids, probe Arch915 labelled with four 6-carboxyfluorescein molecules (green) for rRNA, DAPI (blue) for DNA. Bar corresponds to 5 µm. Overstaining by Nile Red resulted in the

visibility of spacer plug-like structures. Pairs of these structures were annotated as future filament lysis locations, with a DNA-depleted microcell between them resulting from unequal cell division.

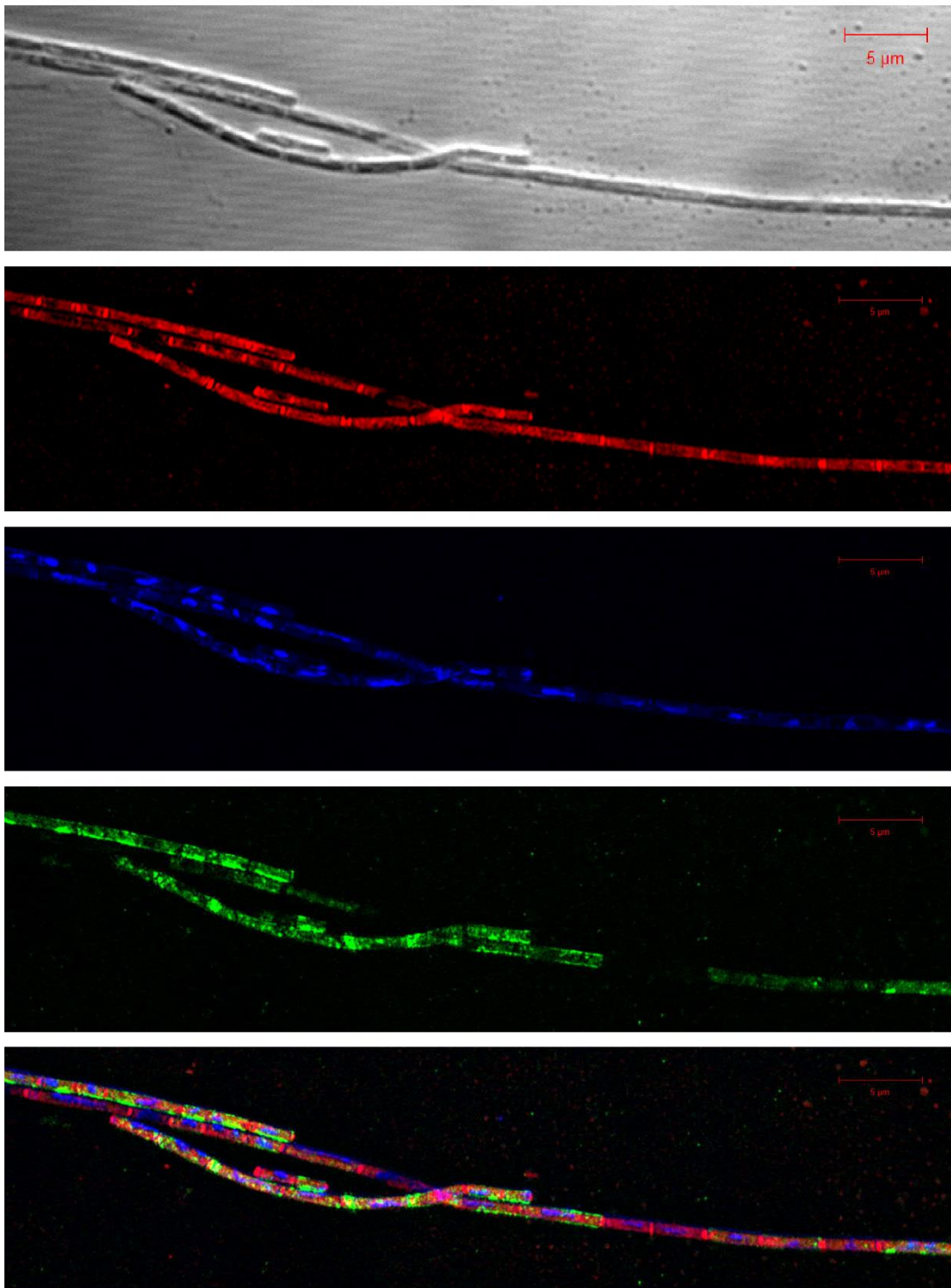


Fig S9 DIC and SR-SIM overlay images of a stationary culture of *Methanosaeta concilii* GP6^T (DSM3671). Staining was achieved using Nile Red (red) for lipids, probe Arch915 labelled with four 6-carboxyfluorescein molecules (green) for rRNA, DAPI

(blue) for DNA. Bar corresponds to 5 μm . Nile Red overstaining made spacer plug-like structures visible.

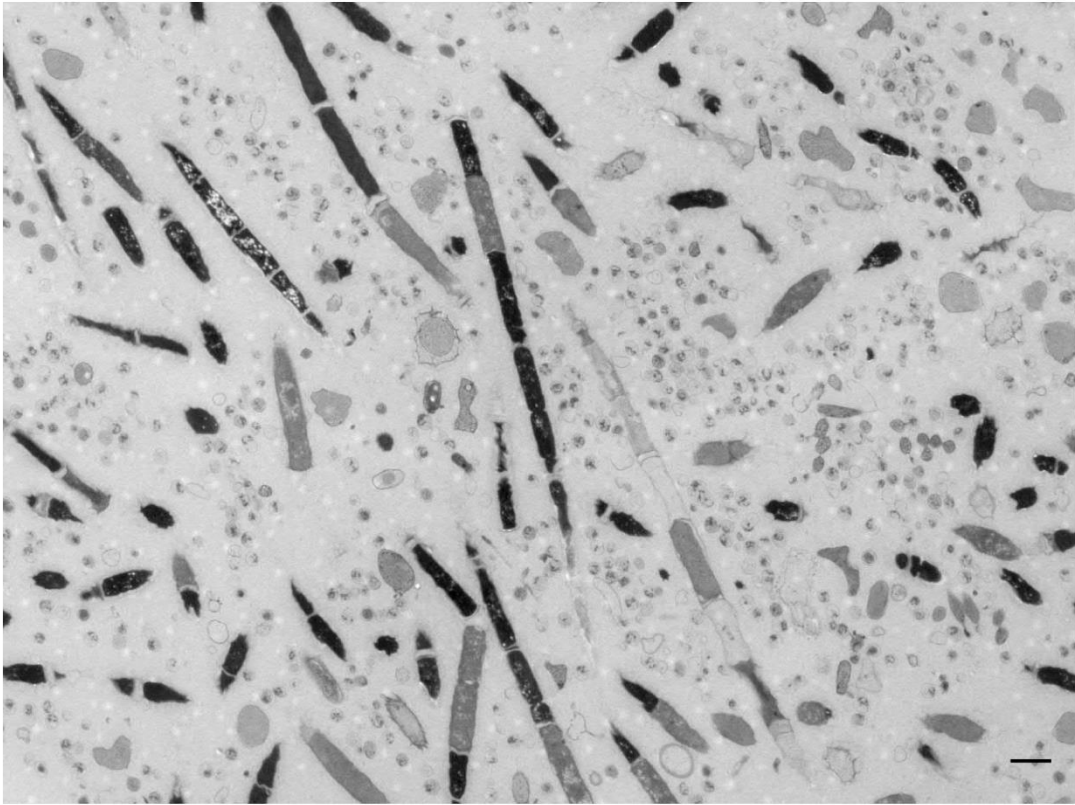


Fig S10 Thin section TEM image of methanogenic limonene enrichment culture. We assigned the longitudinal sectioned filaments to *Methanosaeta*. The intensity of the cell stain was expected to correlate with the availability of acidic groups inside the cell. *Methanosaeta* cells stained differently. The interpretation is not clear. Either healthy cells stain black, including that a grey stain or the absence of a stain represents sick or dead cell within the filament, or the staining of healthy cells was hampered by the cellular structures, including that black or unstained cells represent sick and dead cells. A subpopulation of *Methanosaeta* cells contained unstained structures of unknown origin which were present in grey and in black cells. Bar corresponds to 1 μm .

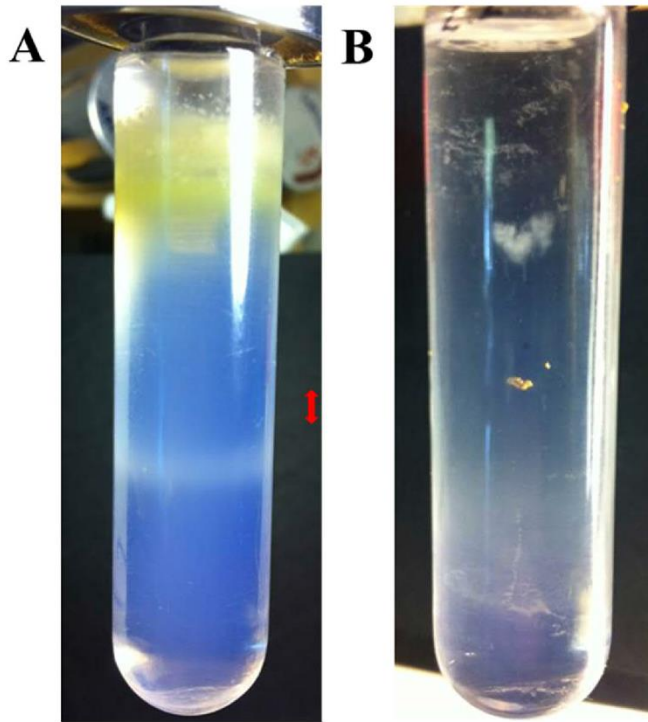


Fig S11 First (A) and second (B) Percoll gradient after cell separation by centrifugation. (A) OP3 LiM cells were most enriched above the clearly visible bottom band (marked by red range). (B) showed a diffusive layer in which 80% of all cells were OP3 LiM cells according to CARD-FISH analysis.

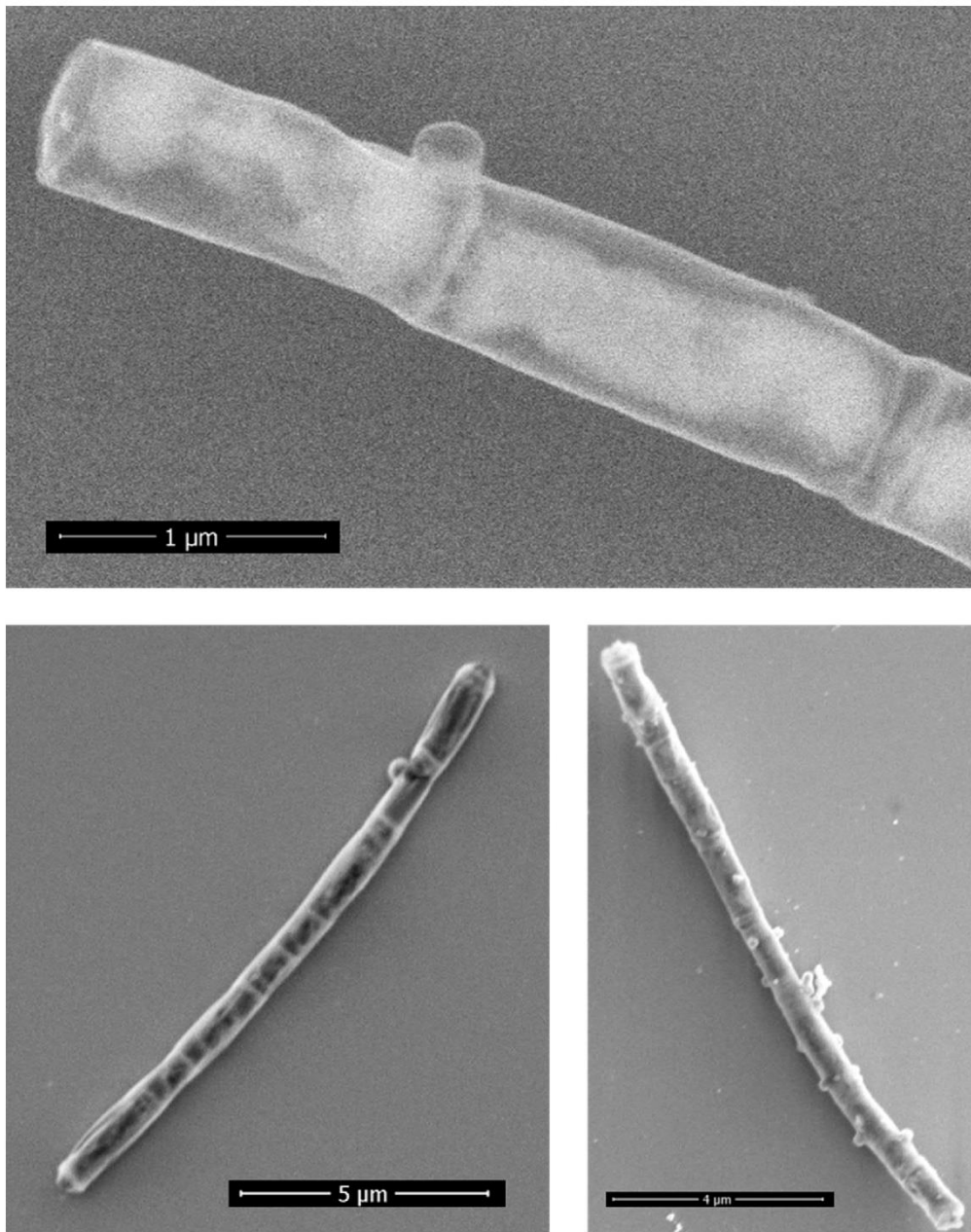


Fig S12 SEM micrographs of *Methanosaeta* filaments with attached ultramicrobial cells. Bar indicates 1, 5 and 4 μm, respectively.

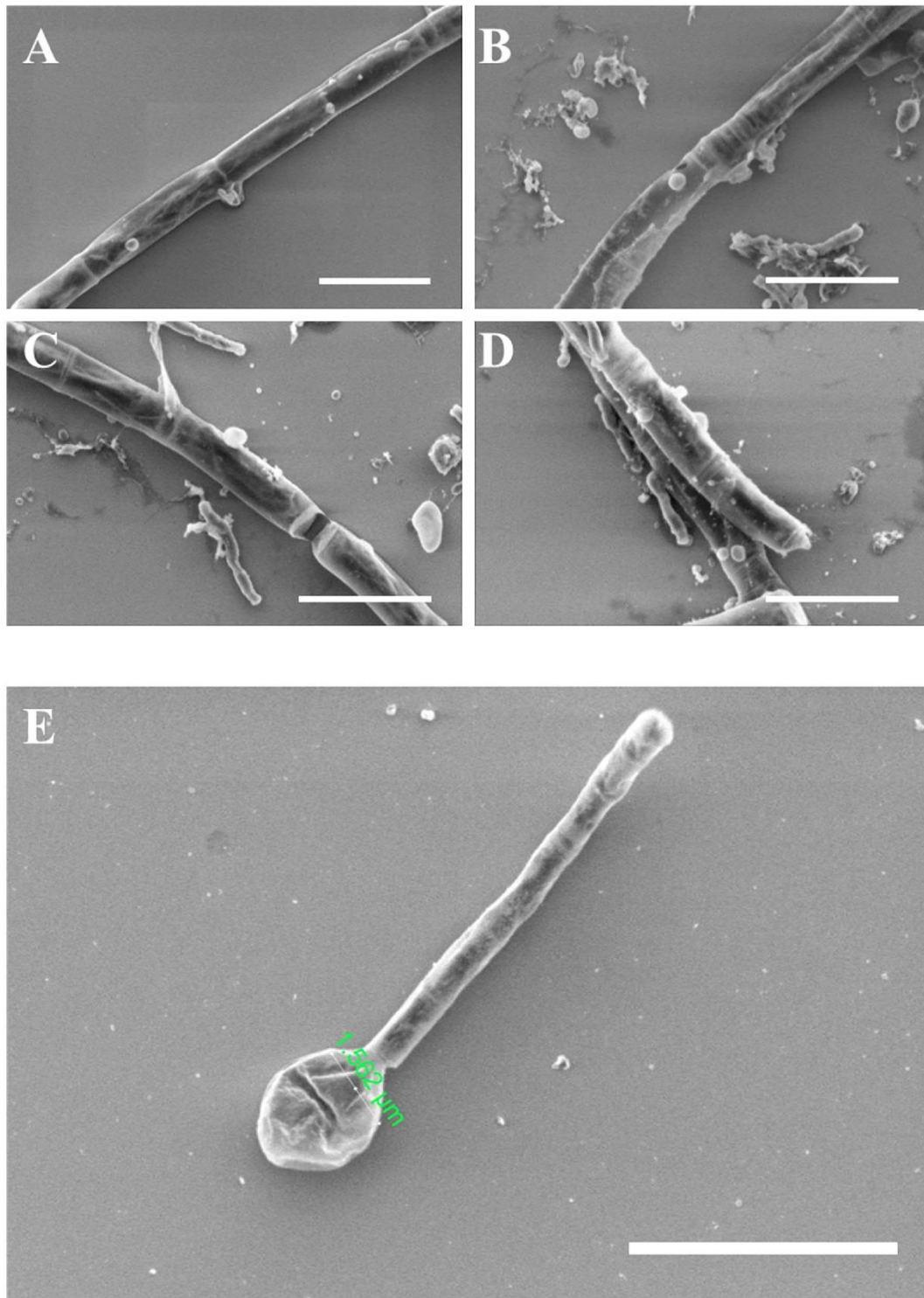
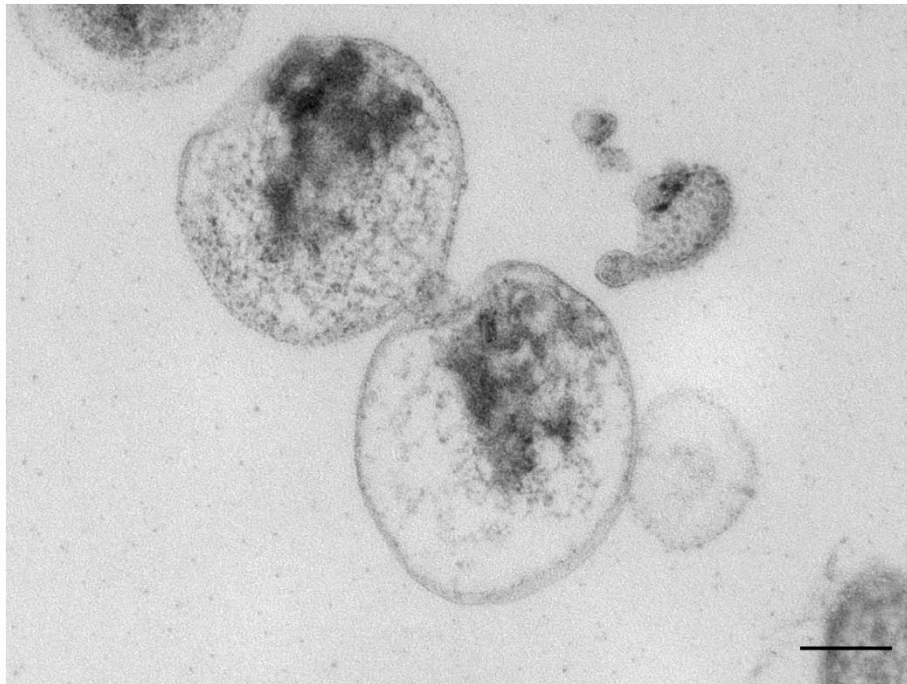


Fig S13 SEM micrographs of *Methanosaeta* and attached cells including a *Methanosaeta* filament with two visibly deformed cells. Scale bar: 2 μm (A-D) and 5 μm (E).

A



B

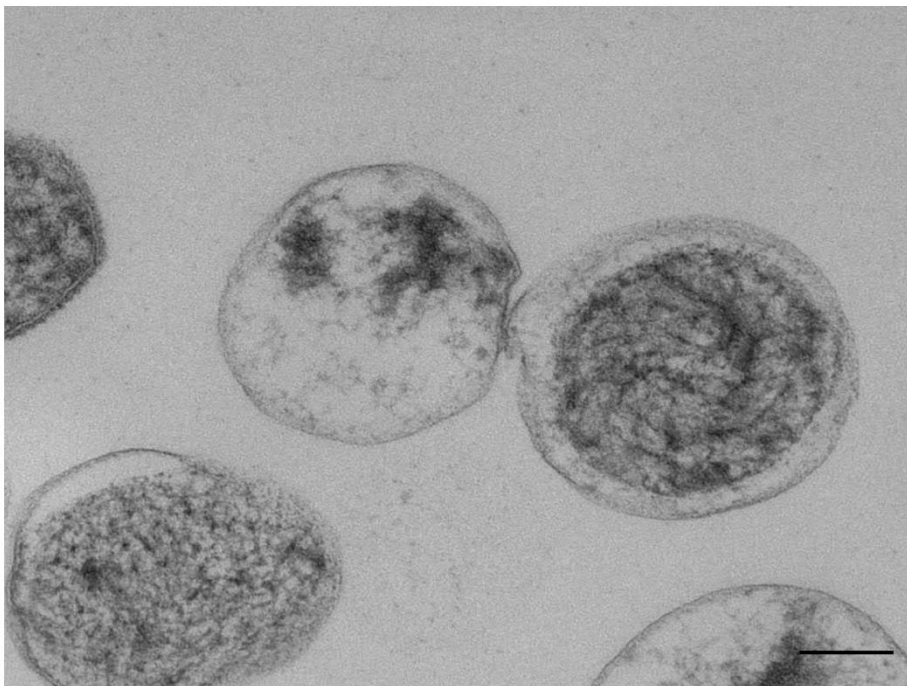
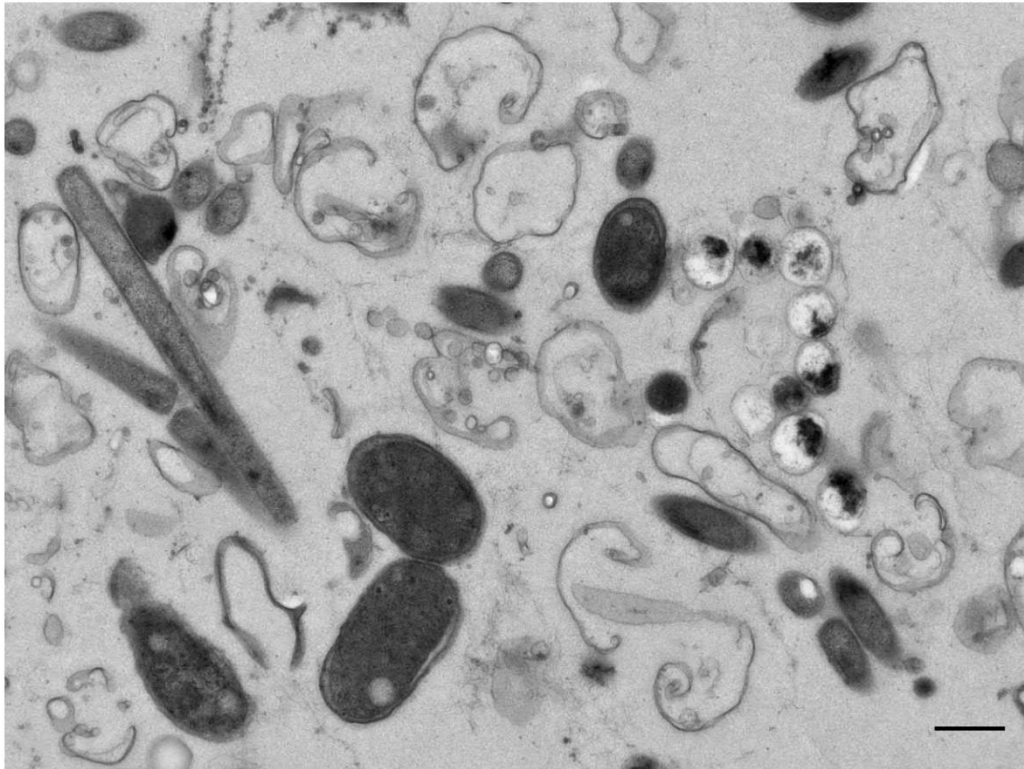


Fig S14 Thin section TEM images of the methanogenic limonene enrichment culture showing contacts between ultramicrobacteria. Bar corresponds to 200 nm.

A



B

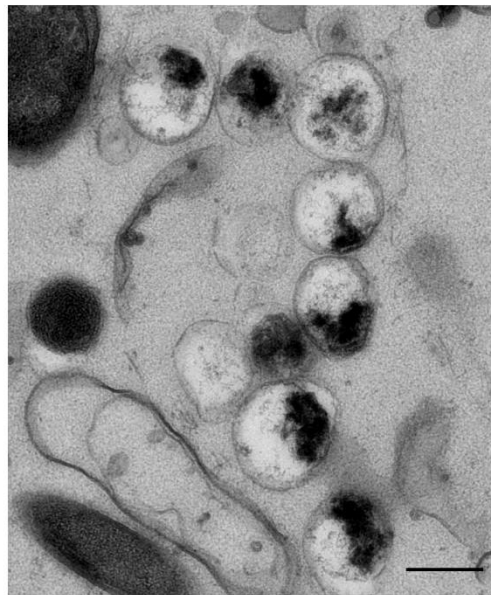


Fig S15 Aged enrichment cultures showed in thin section TEM images a large number of membrane remains of cells (A). The morphological type of ultramicrobacteria assigned to OP3 LiM cells was present as intact cells (A, enlarged in B). Bar corresponds to 500 nm (A) or 300 nm (B).

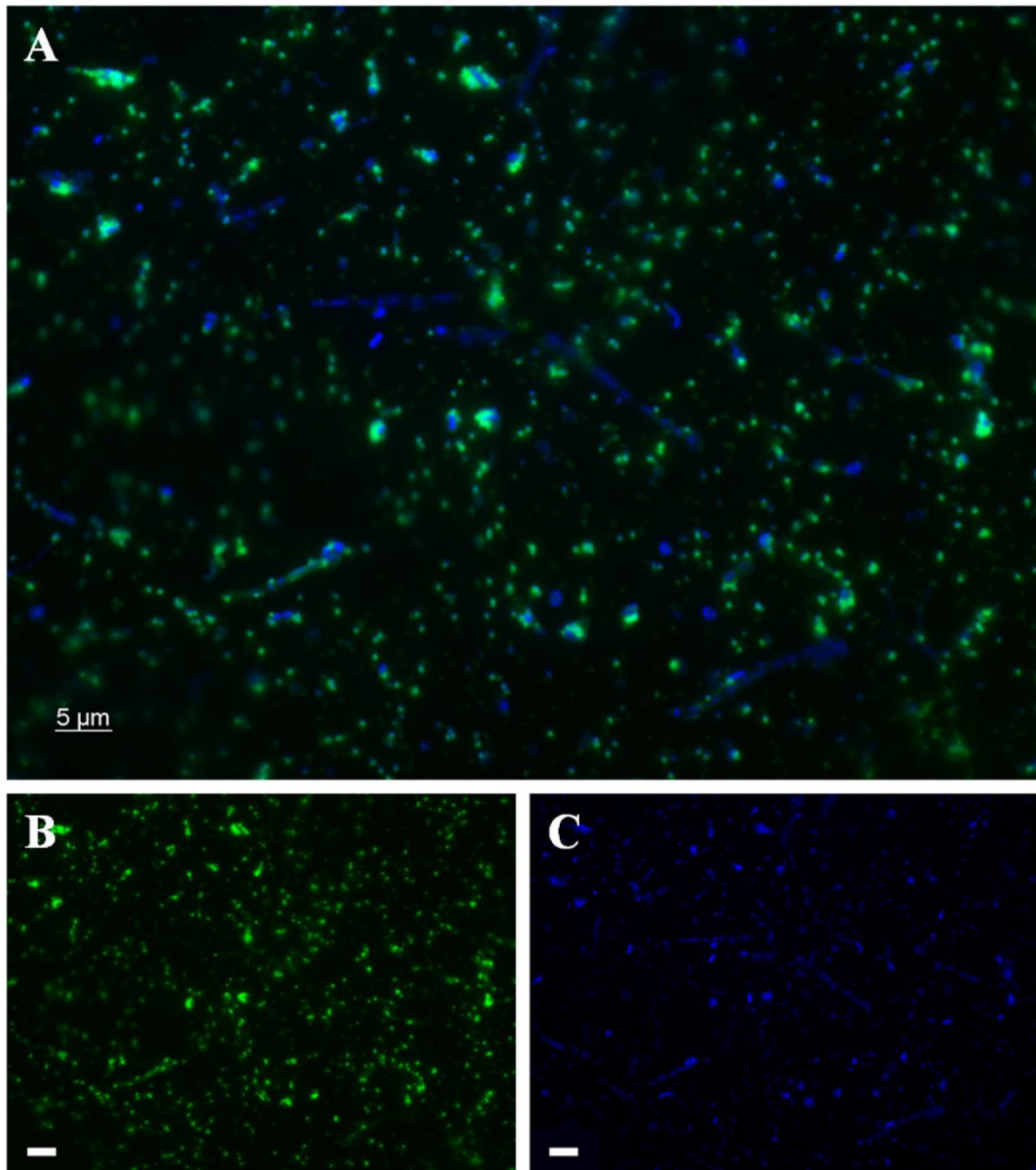


Fig S16 Detection of OP3 LiM cells by CARD-FISH. Cells were detected using the probe OP3-565 with helper oligonucleotides (green, A and B) and DAPI staining (blue, A and C). Scale bars, 5 μm .

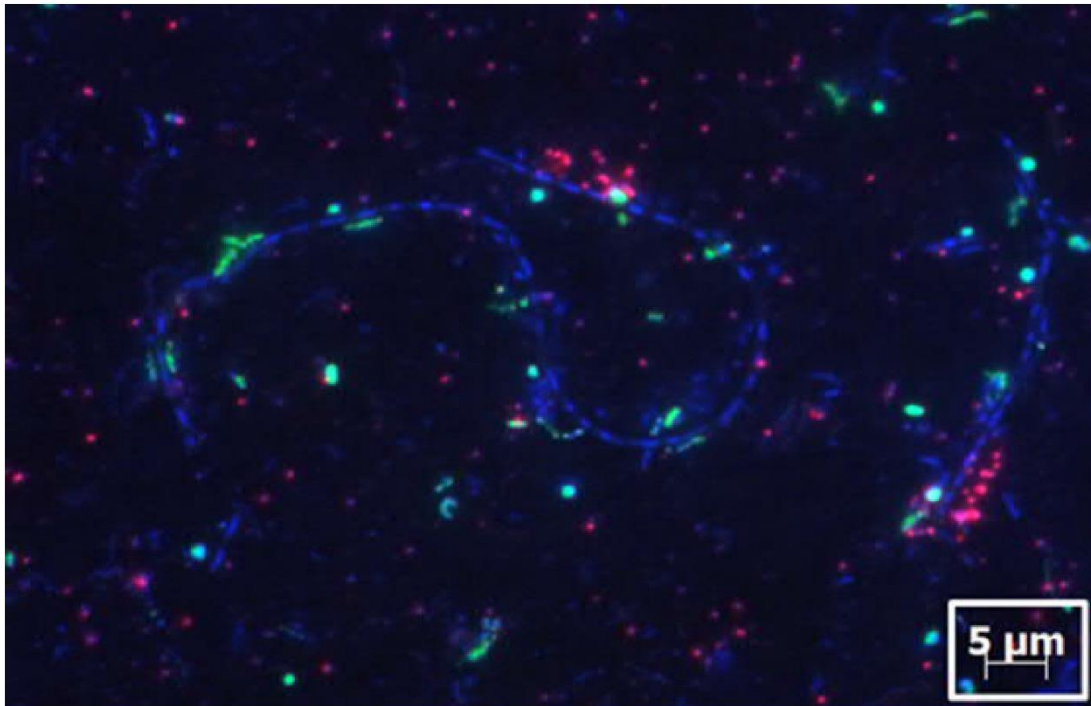


Fig S17 Detection of *Bacteria* in a methanogenic enrichment culture. DNA is stained with DAPI (blue), the 16S rRNA probe mix EUB338 I-III revealed bacterial cells without OP3 LiM (green) and the OP3 LiM-specific probe OP3-565 detected OP3 LiM cells (red) in CARD-FISH double hybridization experiments. Scale bar, 5 μm.

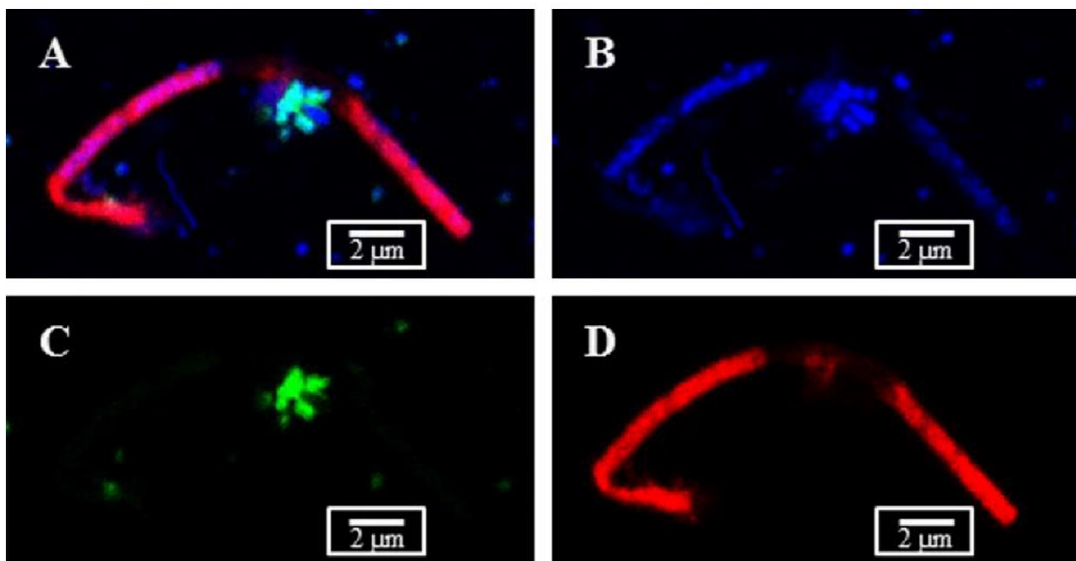


Fig S18 Detection of *Archaea* and OP3 LiM in a methanogenic enrichment culture in CLSM images. Overlay (A) and individual signals of DNA (B), OP3 LiM (C) and *Archaea* (D) were obtained by DAPI staining and double CARD-FISH with probes OP3-565 together with helper oligonucleotides and ARCH-915, respectively. Scale bar represents 2 μm.

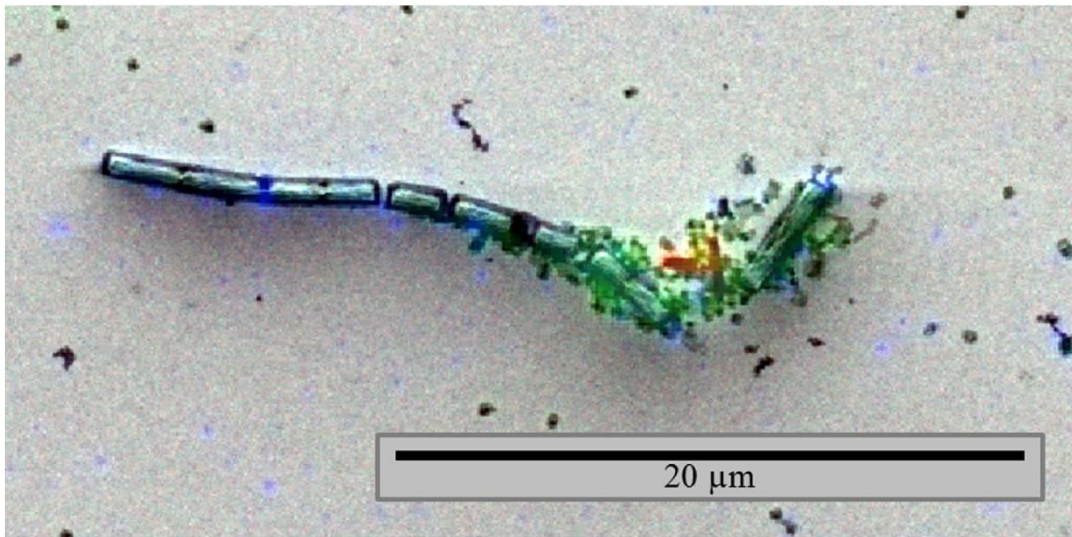


Fig. S19 Correlative fluorescence and electron microscopy image of the detection of OP3 LiM and *Archaea*. A Quanta FEG 250 equipped with the DELMIC SECOM platform was used to visualize a double CARD-FISH sample of the limonene methanogenic enrichment culture. OP3 LiM rRNA was detected with OP3-595 (red, Alexa-594 as label on the tyramide). Archaeal cells were detected with ARCH915 (green; Alexa-488). DNA was stained by DAPI (blue).

Formaldehyde-fixed cells (1.3% v/v for 1 h at RT) of the limonene methanogenic enrichment cultures were centrifuged at 13,000 rpm for 10 min. The pellet was washed twice in 500 μl 1 x PBS, twice in 200 μl MilliQ water and resuspended in 300 μl MilliQ water. The suspension was diluted with MilliQ water and 80 μl were placed on an indium oxide coated cover glass (DELMIC B.V., JA Delft, The Netherlands). After air drying CARD-FISH was performed.

Supplement table 1: Absolute detection level of individual proteins (total spectral counts (TSC)) of OP3 LiM in size-fractions of the microbial community. Two biological replicates were analysed. SD standard deviation. The mass spectrometry proteomics data are available via the PRIDE partner repository with the dataset identifier PXD025008 and 10.6019/PXD025008.

Protein annotation	NCBI locus_tag	Molecular Weight [kDa]	TSC (10kS-aggregates)	SD of TSC (10kS-aggregates)	TSC (100S-aggregates)	SD of TSC (100S-aggregates)	TSC (100S-cells)	SD of TSC (100S-cells)
18-strand beta-barrel outer membrane pore	BU251_04490	53	2718	384	1076	247	2133	443
co-chaperone GroES	BU251_04735	58	1019	151	546	14	1226	50
GPI-membrane anchored actin binding protein with KELCH repeats	BU251_08370	499	1132	448	503	286	940	197
translation elongation factor Tu	BU251_03080	44	760	6	543	34	961	24
pyruvate phosphate dikinase	BU251_04585	101	723	192	410	29	883	117
DNA-directed RNA polymerase subunit beta	BU251_03055	128	640	175	427	18	722	57
DNA-directed RNA polymerase subunit beta'	BU251_03060	152	507	212	401	11	638	9
hypothetical protein	BU251_02745	326	465	187	156	136	466	82
polyribonucleotide nucleotidyltransferase	BU251_05955	77	396	16	210	14	406	11
peptidyl-prolyl cis-trans isomerase	BU251_00050	35	339	14	200	25	356	75
Clp protease, ATP-binding subunit	BU251_06185	91	352	37	161	12	361	6
very large multi-enzyme surface protein	BU251_00165	4382	647	132	100	40	127	9
ATPase for pilus assembly/DNA-transfer CpaF	BU251_00065	82	320	57	174	8	358	11
flagellar motor protein MotD	BU251_01115	24	322	48	126	11	328	54
glutamine synthetase type III	BU251_06845	83	312	93	144	14	315	4
transmembrane receptor with multiple PAS sensor domains and histidine kinase domain	BU251_04855	170	271	81	204	8	292	6
outer membrane protein assembly factor BamA	BU251_00665	87	344	66	142	15	262	11
carbamoyl phosphate synthase large subunit	BU251_01710	118	258	118	139	1	334	17
phosphoglycerate dehydrogenase	BU251_01095	57	283	48	145	7	299	19
glyceraldehyde-3-phosphate dehydrogenase	BU251_04240	36	256	10	170	16	298	70
transcription termination factor Rho	BU251_01315	47	258	8	154	6	289	37
expressed DUF4139 protein, secreted	BU251_00900	63	329	71	87	17	229	11
porA	BU251_00535	43	220	41	128	16	291	27
chromosome segregation protein SMC, archaeal type	BU251_00240	45	254	15	133	17	246	25
bifunctional preprotein translocase subunit SecD/SecE	BU251_05045	80	230	74	146	42	235	11
molecular chaperone DnaK	BU251_02930	68	225	34	116	16	261	16
hypothetical protein	BU251_00210	43	128	11	155	8	310	4
DNA-directed RNA polymerase subunit alpha	BU251_03225	37	214	33	108	6	255	24
putative exosortase-affiliated protein	BU251_00715	13	241	21	79	1	251	61
chromosome segregation protein SMC, common bacterial type	BU251_04010	124	181	64	114	24	227	11
multidrug efflux pump subunit AcrB	BU251_07465	116	172	57	141	23	198	15
pyruvate:ferredoxin oxidoreductase, gamma subunit	BU251_00525	20	177	18	109	30	224	10
actin-like protein MreB	BU251_05665	39	197	15	91	6	218	34
expressed membrane protein	BU251_07630	28	217	3	88	8	181	10
phosphoribosylpyrophosphate synthetase	BU251_00625	35	187	31	89	18	199	43
translation initiation factor IF-2	BU251_01185	79	177	59	79	35	215	47
translation elongation factor EF-G, a GTPase	BU251_08040	71	172	46	93	16	204	22
ketol-acid reductoisomerase	BU251_00930	36	174	23	88	1	204	16
fructose-bisphosphate aldolase class-II	BU251_00910	49	170	38	85	9	204	52
DNA topoisomerase type II (gyrase) subunit B	BU251_00020	91	150	45	99	16	205	32
GMP synthase (glutamine amidotransferase II)	BU251_04725	62	156	76	89	31	205	27
glutamate dehydrogenase	BU251_01730	49	173	11	94	12	179	19
S-adenosylmethionine synthetase	BU251_00890	42	170	27	88	4	186	13
dihydroxy-acid dehydratase	BU251_00915	59	202	77	63	1	179	26
phosphoglycerate kinase	BU251_04065	43	159	11	69	1	192	4
periplasmatic-binding component of ABC transportsystem specific for sn-glycerol-3-phosphate	BU251_00190	51	161	0	82	6	177	13

23

24

gliding motility-associated ABC transporter substrate-binding protein GidG	BU251_08675	94	156	40	98	11	165	7
class V aminotransferase	BU251_01090	41	152	40	89	16	175	25
proton-translocating pyrophosphatase	BU251_00290	72	152	35	105	18	155	11
hypothetical protein	BU251_01495	124	145	65	90	16	170	41
carB	BU251_08275	120	148	39	78	6	170	10
secretion system protein with secretin and TonB domains	BU251_00060	58	177	30	62	2	156	8
bacterioferritin	BU251_00560	18	172	32	61	11	161	35
30S ribosomal protein S2	BU251_00830	33	155	11	63	1	175	14
DNA gyrase subunit A	BU251_00030	91	145	49	75	16	165	3
recombinase RecA	BU251_05105	39	143	1	64	15	175	11
expressed membrane protein	BU251_00135	78	153	32	78	13	140	18
secreted protein containing glycoside hydrolase family 2, TIM barrel domain	BU251_00205	94	169	46	42	18	158	8
ATP-dependent Zn proteases FtsH	BU251_04905	69	148	35	61	10	159	11
PilT, pilus retraction ATPase	BU251_03245	39	142	17	63	9	161	34
type I site-specific deoxyribonuclease, HsdR family	BU251_02610	117	132	41	68	11	157	16
alkaline phosphatase	BU251_06895	46	144	6	53	9	155	26
secretion system ATPase related to VirB11/CpaF	BU251_06755	52	130	31	65	3	155	40
short chain acyl-CoA dehydrogenase	BU251_04025	42	159	4	55	4	136	7
3-deoxy-7-phosphoheptulonate synthase	BU251_03910	37	140	27	70	7	137	13
phosphodiesterase with HD domain of ribonucleaseY	BU251_09405	59	140	39	59	5	145	8
LPS assembly outer membrane protein LptD	BU251_02985	80	186	65	45	11	112	21
phosphoenolpyruvate-protein kinase (PTS system EI component in bacteria)	BU251_00880	65	144	45	46	6	148	1
ATP-dependent protease/chaperone ClpB	BU251_02910	97	129	54	58	1	151	1
cellobiose phosphorylase	BU251_00185	93	134	68	63	20	138	19
alanyl-tRNA synthetase	BU251_05095	96	114	44	65	10	152	12
inosine 5-monophosphate dehydrogenase	BU251_04730	41	128	9	63	5	139	21

25

type II secretory pathway ATPase GspE/PulE or type IV pilus assembly pathway ATPase PilB	BU251_05920	43	123	27	60	14	146	4
protein with domain of unknown function (DUF4912)	BU251_01125	40	136	15	51	7	141	36
6-phosphofructokinase	BU251_03785	37	134	1	55	2	138	8
sensory signal transduction histidine kinase with PAS domain	BU251_02305	114	123	74	61	13	140	23
ribonuclease E or G	BU251_05565	57	133	71	52	9	136	14
transcription termination control protein NusA	BU251_01180	43	135	18	42	1	134	17
S-adenosyl-L-homocysteine hydrolase	BU251_00895	46	102	20	53	4	150	11
hypothetical protein	BU251_02695	93	141	47	42	15	121	11
GMP synthase [glutamine-hydrolyzing], ATP pyrophosphatase subunit B	BU251_01715	36	126	21	46	5	131	31
isoleucine--tRNA synthetase	BU251_05700	105	108	46	54	13	136	19
ATP phosphoribosyltransferase	BU251_00580	33	129	29	46	8	122	28
signal transducer with PAS sensor, GGDEF (diguanylate cyclase) and metal dependent phosphohydrolase domains	BU251_01490	82	116	37	55	12	118	10
pillin related to pseudopilin PulG	BU251_05600	15	138	25	39	9	113	8
bifunctional UDP-3-O-[3-hydroxymyristoyl] N-acetylglucosamine deacetylase/[3R]-hydroxymyristoyl-[acyl-carrier-protein] dehydratase	BU251_00680	48	111	10	57	4	120	18
2-isopropylmalate synthase	BU251_00935	56	112	33	50	6	124	5
hypothetical protein	BU251_01680	53	105	48	44	18	131	4
flagellar motor protein MotB	BU251_07995	23	127	1	43	2	108	16
signal transducer with HAMP signal, GAF transducer and GGDEF (diguanylate cyclase) domains	BU251_05635	51	106	17	52	1	113	35
ATP-dependent Zn-protease	BU251_07530	62	107	33	41	8	121	9
type II secretory pathway component GspD/PulD (secretin)	BU251_05575	57	174	111	31	23	63	5
type II secretory pathway component GspD/PulD (secretin) with additional STN domain	BU251_05205	56	131	41	40	9	96	20

26

phosphohexomutase superfamily	BU251_05530	54	100	1	56	0	110	13
phosphopyruvate hydratase	BU251_07525	46	58	11	61	8	144	24
30S ribosomal protein S4	BU251_03220	24	114	22	40	8	106	10
short-chain dehydrogenase/reductase (SDR)	BU251_07850	28	106	2	41	13	113	12
DNA polymerase I	BU251_01300	97	93	42	48	9	115	24
glycine--tRNA ligase	BU251_04590	59	104	30	42	4	109	7
LemA/RetS hybrid sensor kinase-response regulator protein	BU251_06830	21	101	12	50	9	104	23
hypothetical protein	BU251_08610	49	139	56	45	4	70	13
glycogen branching enzyme, GH-57-type, archaeal	BU251_01130	63	124	45	32	1	97	16
phenylalanyl-tRNA synthetase beta chain	BU251_09435	75	103	23	38	5	112	19
His-Xaa-Ser system radical SAM maturase HxsB	BU251_04195	55	95	30	47	16	109	21
carbohydrate-selective porin	BU251_08565	50	132	10	38	13	80	13
acetolactate synthase large subunit	BU251_00920	59	99	23	43	3	107	7
homocitrate synthase	BU251_01105	46	98	9	37	2	112	4
PilT-related nuclelease/ATPase with N-terminal membrane domain	BU251_06155	37	114	21	37	8	93	12
cell division protein FtsK	BU251_05950	84	102	45	35	10	105	10
Lipid A export ATP-binding/permease protein MsbA	BU251_00825	68	95	45	53	15	94	0
outer membrane chaperone Skp (OmpH)	BU251_00670	21	88	15	34	11	119	1
hypothetical protein	BU251_08110	90	98	30	40	1	101	4
type II secretory protein GspE	BU251_08460	64	113	25	31	0	94	12
translation elongation factor G	BU251_03075	78	85	23	42	4	110	2
uptake hydrogenase large subunit	BU251_07790	52	100	28	39	9	98	13
rubrerythrin	BU251_08710	22	83	8	36	7	117	8
membrane-associated protein with peptidoglycan-binding domains OmpA and LysM	BU251_03585	37	94	5	38	0	100	25
homoserine dehydrogenase	BU251_04810	46	88	37	32	8	111	6
short-chain dehydrogenase/reductase (classical (c) SDR)	BU251_07470	28	98	0	36	10	96	4
pyruvate:ferredoxin oxidoreductase, beta subunit	BU251_00540	31	94	15	35	14	100	3

27

protein with TMH, SAF and Flp pilus assembly protein RcpC domains	BU251_00055	34	100	18	44	8	82	14
hypothetical protein	BU251_04665	51	93	28	25	1	108	18
preprotein translocase subunit SecA	BU251_07115	107	78	29	44	4	103	3
filamentation induced by cAMP protein Fic	BU251_02690	41	96	34	34	9	94	21
DNA-binding transcriptional regulator, LacI/PurRfamily	BU251_04575	32	95	8	34	8	95	5
citramalate synthase	BU251_04780	59	91	30	28	8	103	0
cation transporter	BU251_07400	121	85	26	45	8	91	3
phosphoribosylformylglycinamide synthase II	BU251_07255	107	76	42	47	13	98	22
glyA	BU251_07210	46	92	5	40	3	88	1
lipoprotein-anchoring transpeptidase ErfK/SrfK	BU251_07955	30	88	8	46	4	85	1
replicative DNA helicase	BU251_00660	51	85	22	38	7	94	33
hypothetical protein	BU251_00140	19	91	13	43	4	81	7
type II/IV secretion system ATP hydrolase TadaA/VirB11/CpaF, TadaA subfamily	BU251_01160	48	75	10	34	0	106	26
acetyl-CoA carboxylase biotin carboxylase subunit	BU251_03800	51	86	12	30	6	99	24
sensor histidine kinase Walk	BU251_06285	84	69	30	33	0	113	18
hypothetical protein	BU251_00620	44	87	6	33	2	95	18
hypothetical protein	BU251_02750	32	89	1	41	7	81	15
DNA topoisomerase I	BU251_00965	76	80	24	31	4	100	6
hypothetical protein	BU251_03520	35	100	19	30	5	77	28
type II restriction endonuclease, TdeIII	BU251_09035	32	90	15	34	6	80	13
TonB-dependent receptor	BU251_03850	74	103	41	25	28	73	5
D-alanine/D-glutamate or branched-chain amino acid aminotransferase	BU251_06200	32	97	16	34	12	70	24
nicotinamidase/pyrazinamidase	BU251_09520	31	99	23	27	9	76	1
type II secretion system protein E	BU251_05620	63	78	25	29	3	93	0
ABC-type transporter Mla maintaining outer membrane lipid asymmetry, periplasmic component MlaD	BU251_06165	23	90	4	31	1	76	23

28

type II and III secretion system protein	BU251_08410	58	101	11	26	1	70	12
valine--tRNA ligase	BU251_04775	106	67	30	34	1	95	15
dihydrodipicolinate synthase	BU251_05300	31	56	21	45	1	95	10
type II secretion system protein E	BU251_05625	64	79	29	28	3	88	6
adenylosuccinate lyase	BU251_07260	48	82	34	27	4	84	8
type II secretion system protein E	BU251_05200	63	73	28	27	1	91	11
glycogen synthase, ADP-glucose transglucosylase	BU251_01305	54	80	32	27	5	82	17
twitching motility protein PIIT	BU251_07125	41	76	20	28	4	82	11
acetate kinase	BU251_02105	41	71	16	33	10	80	1
transcriptional regulator	BU251_00245	37	74	22	29	9	82	29
secreted protein containing molecular chaperone lbpA, HSP20 family	BU251_06465	23	72	14	21	6	90	6
xanthine dehydrogenase, molybdenum binding subunit	BU251_01525	84	65	16	36	4	80	6
membrane or secreted protein containing peptidoglycan-binding Lysin subgroup domain (LysM)	BU251_04695	19	72	12	30	7	79	12
hypothetical protein	BU251_03790	22	85	2	23	4	73	4
transcriptional regulator	BU251_03940	27	80	18	25	1	75	1
dolichyl-phosphooligosaccharide-protein glycotransferase	BU251_01235	99	73	39	32	2	75	6
peptidyl-prolyl cis-trans isomerase	BU251_00575	59	80	8	26	4	71	4
phosphoenolpyruvate carboxykinase	BU251_07675	65	73	25	28	1	75	7
serine--tRNA ligase	BU251_06770	48	73	15	23	3	79	20
putative chromosome-partitioning protein ParB	BU251_02435	38	78	17	21	0	75	1
periplasmic solute-binding protein of ABC transporter	BU251_08595	28	23	0	33	13	117	6
glycerol-3-phosphate ABC transporter ATP-bindingprotein	BU251_02960	41	74	11	24	2	74	4
aspartate-semialdehyde dehydrogenase	BU251_05370	37	71	1	27	2	75	4
GDP-mannose 4,6-dehydratase	BU251_04420	44	64	8	28	1	77	24
phosphate regulon sensor protein PhoR (SphS)	BU251_08525	56	60	8	31	3	78	13

29

type II/IV secretion system ATP hydrolase TadA/VirB11/CpaF, TadA subfamily	BU251_01170	48	61	8	28	1	79	0
lysine--tRNA ligase	BU251_06215	56	71	19	25	3	72	6
macrolide export ATP-binding/permease protein MacB	BU251_01860	72	61	12	37	6	69	4
D,D-heptose 7-phosphate kinase	BU251_02245	37	73	26	21	5	73	16
chorismate mutase	BU251_03920	40	64	12	26	5	77	12
phosphoenolpyruvate synthase	BU251_02360	98	60	31	30	6	75	6
sugar dehydratase	BU251_04395	37	69	19	24	4	72	20
aspartyl-tRNA synthetase	BU251_04700	54	59	31	30	11	74	35
glutamine--fructose-6-phosphate transaminase (isomerizing)	BU251_08340	66	60	37	26	7	76	7
protein containing glucosylceramidase domain	BU251_00225	82	65	47	21	8	76	24
phosphoribosylformylglycinamide cyclo-ligase	BU251_07235	36	72	16	19	6	70	16
SOS ribosomal protein L5	BU251_03150	22	67	10	17	4	76	19
hypothetical protein	BU251_00220	61	71	12	22	0	65	8
threonine--tRNA ligase	BU251_09480	65	64	33	21	4	73	5
hypothetical protein	BU251_05255	95	71	19	23	5	64	13
D-beta-D-heptose 7-phosphate kinase	BU251_01045	37	59	18	27	6	71	13
excinuclease ABC subunit A	BU251_05260	106	53	13	33	4	70	2
cell division trigger factor	BU251_01080	34	59	1	23	1	74	5
aspartate kinase	BU251_04785	44	58	7	27	8	70	11
hypothetical protein	BU251_08365	31	69	4	28	4	58	12
pilus assembly protein PilC	BU251_05610	43	67	8	28	1	59	7
acetylglutamate kinase	BU251_05335	31	71	11	23	2	60	3
cell division protein FtsH	BU251_09340	57	55	15	27	2	67	28
UDP-glucose 6-dehydrogenase	BU251_02990	49	57	12	22	1	70	4
secreted protein containing Sulphatase-modifyingfactor domain	BU251_02725	55	55	29	30	2	61	32
type IV pili twitching motility protein PIIT	BU251_08545	37	62	6	22	4	62	16
serine O-acetyltransferase	BU251_06140	25	61	16	18	3	65	8

30

signal transduction histidine kinase	BU251_02005	68	67	22	17	6	60	11
3-isopropylmalate dehydrogenase	BU251_05380	38	55	14	29	10	59	8
hypothetical protein	BU251_07160	37	67	10	23	7	53	0
argininosuccinate synthase	BU251_05320	44	57	11	18	0	68	9
DNA polymerase III beta subunit	BU251_00010	42	57	1	23	4	62	9
ATP-dependent Clp protease proteolytic subunit	BU251_01085	22	53	14	25	6	63	13
nucleoside-diphosphate sugar epimerase	BU251_04330	39	53	12	25	11	62	4
expressed secreted protein of unknown function	BU251_07650	18	61	9	14	2	65	4
30S ribosomal protein S7	BU251_03070	18	59	3	18	1	61	2
phosphate acetyltransferase	BU251_02100	34	55	3	23	3	59	1
hypothetical protein	BU251_05220	60	61	25	18	1	59	1
LL-diaminopimelate aminotransferase	BU251_05290	43	57	4	19	4	61	13
hypothetical protein	BU251_07480	59	83	37	10	8	45	6
hypothetical protein	BU251_07935	27	36	5	19	1	81	21
adenine phosphoribosyltransferase	BU251_04570	19	56	16	17	4	62	2
bifunctional phosphoribosylaminoimidazolecarboxamide formyltransferase/IMP cyclohydrolase	BU251_08030	55	58	34	15	0	61	4
hypothetical protein	BU251_08680	34	60	13	20	5	54	13
signal peptidase I	BU251_03890	23	59	18	17	4	58	8
hypothetical protein	BU251_08450	56	62	23	16	4	55	1
carboxyl-terminal protease	BU251_05005	48	59	28	20	1	54	18
hypothetical protein	BU251_07980	82	48	18	25	6	59	6
chorismate synthase	BU251_00950	37	57	12	21	3	54	6
50S ribosomal protein L1	BU251_03040	24	59	13	15	9	58	6
glucose-6-phosphate isomerase	BU251_06840	55	53	18	20	4	59	19
50S ribosomal protein L17	BU251_03230	21	61	17	16	4	53	2
isocitrate dehydrogenase	BU251_03255	39	54	6	21	1	55	1
3-oxoacyl-[acyl-carrier-protein] reductase	BU251_05405	26	54	11	17	0	59	15

31

secreted protein containing peptidoglycan binding-like domain	BU251_07685	14	60	9	24	5	46	1
protein with tandem amino-acid binding ACT domain	BU251_03865	14	31	4	21	5	77	18
V-type ATP synthase subunit A	BU251_08230	64	62	36	16	1	51	6
prolyl-tRNA synthetase, bacterial type	BU251_01165	47	50	8	20	4	57	4
ATP-dependent RNA helicase RHE	BU251_03485	47	50	23	14	11	63	15
UDP-N-acetyl-D-glucosamine dehydrogenase	BU251_00395	47	56	10	16	1	54	4
nucleoside diphosphate kinase	BU251_09525	19	50	1	20	4	56	6
hypothetical protein	BU251_02845	85	52	24	15	6	58	7
hypothetical protein	BU251_09550	55	58	27	19	3	47	10
aspartate aminotransferase	BU251_03235	43	50	3	20	4	54	1
HlyD family secretion protein	BU251_00740	31	58	1	20	4	46	1
lipid A deacylase PagL	BU251_06260	22	63	3	15	1	45	4
hypothetical protein	BU251_04295	64	51	24	18	9	53	3
LSU ribosomal protein L25	BU251_00630	25	51	11	22	4	49	12
glycogen phosphorylase	BU251_02170	76	50	23	15	7	56	18
histidinol-phosphate transaminase	BU251_03915	41	47	6	19	4	55	11
aminotransferase	BU251_04305	41	50	21	17	6	54	3
glutamyl-tRNA synthetase and Glutamyl-tRNA(Gln) synthetase	BU251_09360	49	48	8	20	1	53	11
rubredoxin-oxygen (nitric oxide) oxidoreductase	BU251_00555	44	42	10	22	6	56	7
50S ribosomal protein L4	BU251_03095	24	53	9	18	3	49	3
hypothetical protein	BU251_04285	64	51	25	17	6	53	5
hypothetical protein	BU251_07835	57	51	16	22	9	47	18
50S ribosomal protein L10	BU251_03045	20	54	7	12	4	53	10
AAA family ATPase	BU251_08745	67	47	21	15	2	58	19
PTS fructose/mannitol-specific subunit IIA [OP3 merged]	BU251_00865	17	52	8	17	3	49	13
hypothetical protein	BU251_05915	22	47	1	19	1	52	14
glutamate racemase	BU251_01600	29	52	10	14	4	51	2

32

triose-phosphate isomerase	BU251_04060	28	14	0	30	6	73	8
hypothetical protein	BU251_08385	85	50	13	18	9	50	2
UDP-2,3-diacetylglucosamine pyrophosphatase	BU251_00690	31	53	6	17	4	47	8
30S ribosomal protein S5	BU251_03175	18	50	2	17	1	50	3
histidyl-tRNA synthetase	BU251_04705	41	42	12	21	2	54	6
hypothetical protein	BU251_08735	39	53	8	14	0	48	9
arabinose 5-phosphate isomerase	BU251_01060	34	50	16	14	6	50	14
RND multidrug efflux transporter	BU251_02070	112	37	19	24	6	54	4
two-component sensor histidine kinase	BU251_02800	40	48	20	15	1	51	23
hypothetical protein	BU251_08750	46	46	8	19	0	48	9
30S ribosomal protein S1	BU251_05515	58	49	30	14	8	50	8
tRNA nucleotidyltransferase, CC-adding	BU251_07185	56	38	29	18	7	56	1
50S ribosomal protein L6	BU251_03165	20	48	11	13	4	50	3
lipid hydroperoxide peroxidase	BU251_03560	30	49	1	19	6	44	11
hypothetical protein	BU251_04365	44	45	16	14	8	53	6
asnC family transcriptional regulator	BU251_06855	19	53	4	14	2	45	8
phosphate regulon transcriptional regulatory protein PhoB (SphR)	BU251_08530	28	42	4	15	6	55	1
histidine kinase	BU251_06870	26	48	1	15	6	48	8
transmembrane and TPR repeat-containing protein	BU251_04045	88	51	29	15	4	44	1
DNA polymerase III	BU251_04620	64	45	28	10	4	54	18
aminopeptidase	BU251_05970	42	43	4	16	2	51	1
hypothetical protein	BU251_07155	20	56	4	13	0	40	8
ADP-glyceromanno-heptose 6-epimerase	BU251_00975	37	48	17	17	3	44	8
response regulator receiver modulated metal dependent phosphohydrolase	BU251_05615	42	46	7	16	2	47	4
mannosylglucosyl-3-phosphoglycerate phosphatase	BU251_01100	67	57	36	14	1	38	1
3-isopropylmalate dehydratase large subunit	BU251_05390	46	47	6	17	1	44	3
glycosyltransferase	BU251_01000	77	38	25	14	8	56	5
50S ribosomal protein L2	BU251_03105	31	41	7	16	8	50	6

33

argininosuccinate lyase	BU251_05315	50	45	9	14	1	48	11
gpsA	BU251_04830	36	49	21	13	9	45	8
O-acetylhomoserine aminocarboxypropyltransferase	BU251_06620	47	37	14	21	1	48	6
hypothetical protein	BU251_09455	24	43	6	19	3	45	11
hypothetical protein	BU251_05595	40	50	6	16	6	41	2
type I restriction-modification system subunit M	BU251_02600	59	42	10	17	1	47	4
NAD-dependent dehydratase	BU251_02995	35	48	16	15	1	42	1
UDP-3-O-[3-hydroxymyristoyl] glucosamine N-acyltransferase	BU251_00675	37	41	8	13	4	52	18
response regulator PleD	BU251_04005	45	34	2	16	1	56	1
imidazoleglycerol-phosphate dehydratase	BU251_05085	23	46	6	15	3	44	4
hypothetical protein	BU251_01645	98	42	16	14	2	49	6
Single-stranded DNA-binding protein	BU251_00645	17	39	9	16	1	49	25
adenylosuccinate synthetase	BU251_01110	47	41	21	18	2	45	6
hypothetical protein	BU251_01430	16	52	12	12	2	41	12
transcription termination/anti-termination protein NusG	BU251_03030	20	42	3	17	1	45	0
phosphoribosylamine--glycine ligase	BU251_07230	45	39	5	15	5	51	2
30S ribosomal protein S3	BU251_03120	27	49	11	13	1	42	15
phosphate starvation-inducible protein PhoH, predicted ATPase	BU251_04690	36	38	4	20	4	46	21
hypothetical protein	BU251_07005	17	47	0	12	1	44	6
sensory box histidine kinase/response regulator	BU251_00720	72	44	23	15	5	44	6
UMP kinase	BU251_00840	26	41	4	13	4	49	21
hypothetical protein	BU251_01400	30	47	7	12	8	44	11
histidinol dehydrogenase	BU251_05090	45	39	18	20	1	43	11
threonine synthase	BU251_04805	38	38	11	20	7	44	2
methylthioadenosine phosphorylase	BU251_05705	31	48	15	15	2	39	4
trehalose-6-phosphate synthase	BU251_03545	57	45	25	13	4	43	12
NGG1p interacting factor NIF3	BU251_03995	37	49	6	14	1	38	7

34

galactose-1-phosphate uridylyltransferase	BU251_04250	39	41	13	17	0	42	11
transcriptional regulator	BU251_05945	31	47	12	14	3	39	11
hypothetical protein	BU251_03445	38	46	5	16	1	38	1
phosphoheptose isomerase	BU251_03780	21	47	1	14	4	39	0
leucine--tRNA ligase	BU251_07605	93	43	26	12	0	45	2
DNA ligase	BU251_05855	74	40	28	13	6	47	11
hypothetical protein	BU251_07920	37	35	6	17	4	47	13
RNA polymerase sigma factor RpoD	BU251_03985	60	36	24	15	4	48	15
hypothetical protein	BU251_05850	20	41	13	12	4	45	7
hypothetical protein	BU251_05190	15	41	6	17	0	40	8
electron transfer flavoprotein subunit alpha	BU251_04015	43	37	7	15	3	45	4
3-dehydroquinase synthase	BU251_03705	39	38	13	14	8	45	10
hypothetical protein	BU251_08615	22	49	8	9	1	39	6
hypothetical protein	BU251_00870	46	41	4	17	4	38	1
6-phosphofructokinase	BU251_00905	35	41	11	16	1	40	15
2-dehydropantoate 2-reductase	BU251_04670	33	45	11	11	7	40	4
UDP-N-acetylglucosamine 1-carboxyvinyltransferase	BU251_01335	46	38	11	13	1	45	1
excinuclease ABC subunit B	BU251_04755	84	33	4	19	1	43	1
phosphate transport system regulatory protein PhoU	BU251_06640	26	43	9	13	6	40	1
hypothetical protein	BU251_08670	39	43	6	14	4	39	10
rubrerythrin-2	BU251_08345	20	38	9	14	4	43	13
bifunctional ornithine acetyltransferase/N-acetylglutamate synthase	BU251_05340	42	48	1	7	0	39	7
tyrosine--tRNA ligase	BU251_05510	47	42	10	12	1	41	4
hypothetical protein	BU251_09505	20	36	2	16	3	43	2
LPS export ABC transporter ATP-binding protein	BU251_01075	27	40	3	16	5	38	1
anthranilate synthase component I	BU251_05060	55	39	23	12	1	43	1
tryptophan synthase subunit beta	BU251_06020	41	39	20	12	3	42	4

35

saccharopine dehydrogenase	BU251_06225	46	31	11	17	4	45	11
thymidylate synthase (FAD)	BU251_07300	28	42	8	13	4	39	2
hypothetical protein	BU251_02705	57	38	53	15	21	41	57
expressed secreted protein of unknown function	BU251_01035	12	39	9	15	6	39	12
hypothetical protein	BU251_01295	30	42	5	8	2	43	10
glutamate-5-semialdehyde dehydrogenase	BU251_05545	45	43	18	16	5	34	5
acetyl-CoA synthase corrinoid activation protein	BU251_04960	69	33	7	13	5	46	13
ornithine carbamoyltransferase	BU251_05325	36	37	10	17	4	38	1
N-acetyl-gamma-glutamyl-phosphate reductase	BU251_05345	38	37	13	12	2	43	4
type II secretion system protein F	BU251_08455	45	39	13	15	1	37	8
hypothetical protein	BU251_09530	34	40	10	13	7	38	11
GDP-mannose 4,6-dehydratase	BU251_02255	37	39	12	15	3	37	1
peptide deformylase	BU251_05075	29	37	2	14	8	40	8
hypothetical protein	BU251_07885	46	39	11	12	1	40	2
hypothetical protein	BU251_07455	23	38	4	16	4	36	4
1-deoxy-D-xylulose-5-phosphate synthase	BU251_09385	68	39	20	11	5	40	1
hypothetical protein	BU251_06370	61	47	24	2	2	41	6
FOF1 ATP synthase subunit beta	BU251_08470	53	34	8	11	1	45	11
amidophosphoribosyltransferase	BU251_03495	54	35	10	18	1	36	4
ribose 1,5-bisphosphate isomerase	BU251_06360	27	36	0	12	2	41	10
hypothetical protein	BU251_00285	44	40	11	15	1	34	13
thioredoxin	BU251_00490	31	35	8	16	0	38	2
aspartate aminotransferase family protein	BU251_05330	44	39	5	12	5	38	13
CTP synthetase	BU251_02160	59	40	20	10	6	38	5
hypothetical protein	BU251_04495	23	35	4	19	2	34	4
NAD+ synthase	BU251_06865	61	39	22	10	1	39	5
hypothetical protein	BU251_08605	55	35	25	12	6	41	18
phenylalanine--tRNA ligase subunit alpha	BU251_09440	39	32	5	14	1	41	0
4-alpha-glucanotransferase	BU251_04235	70	38	19	14	0	34	0

36

ATP synthase FO sector subunit b	BU251_08485	28	37	1	16	4	33	2
outer membrane efflux protein	BU251_00735	62	41	16	13	1	31	5
hypothetical protein	BU251_05585	23	38	2	11	6	36	5
deoxyguanosinetriphosphate triphosphohydrolase	BU251_03730	47	34	1	11	4	39	2
DNA recombination protein RmuC-like protein	BU251_00475	45	35	9	13	0	35	11
electron transfer flavoprotein subunit beta	BU251_04020	28	33	3	17	6	33	3
pantothenate kinase type II, CoaX-like	BU251_04745	28	40	11	14	9	30	1
diadenylate cyclase spyDAC	BU251_04895	29	34	6	13	3	36	3
arginine--tRNA ligase	BU251_06000	63	35	23	8	4	40	1
hypothetical protein	BU251_08420	32	42	18	9	5	32	14
translation initiation factor IF-3	BU251_09475	18	35	4	10	2	38	4
[Fe] hydrogenase large subunit	BU251_02020	68	36	23	10	0	36	14
hypothetical protein	BU251_04835	76	32	25	12	7	38	1
DNA-binding protein similar to integration host factor beta subunit	BU251_04715	10	32	2	17	6	33	6
orotate phosphoribosyltransferase	BU251_06055	21	33	6	15	5	34	6
protein arginine kinase	BU251_06190	40	32	8	10	3	39	3
50S ribosomal protein L3	BU251_03090	22	30	6	10	8	41	3
cysteine synthase A	BU251_06615	33	25	10	15	6	41	13
type II secretion system protein GspG	BU251_08440	14	36	13	8	1	36	4
glycosyl transferase	BU251_04300	31	38	9	11	1	31	7
hypothetical protein	BU251_06665	30	14	0	20	7	46	1
hypothetical protein	BU251_07475	34	36	1	9	2	36	2
hypothetical protein	BU251_08095	73	34	19	10	1	36	10
hypothetical protein	BU251_08400	53	33	1	14	0	33	5
hypothetical protein	BU251_03725	41	30	18	12	4	38	1
mannose-1-phosphate guanylyltransferase/mannose-6-phosphate isomerase	BU251_04425	55	36	26	8	1	36	14
hypothetical protein	BU251_07910	55	31	17	14	3	34	18

37

ATP-dependent RNA helicase DBP9	BU251_05130	56	40	23	8	4	31	1
hypothetical protein	BU251_00550	39	35	16	13	1	30	4
ribosome recycling factor	BU251_00845	21	35	5	11	4	33	11
hypothetical protein	BU251_05215	22	32	4	14	1	32	13
hypothetical protein	BU251_07965	26	30	1	12	0	36	4
nucleoside-diphosphate-sugar pyrophosphorylase	BU251_00885	30	30	11	15	5	33	10
preprotein translocase subunit SecE	BU251_04055	15	37	10	9	3	32	1
SMC-Scp complex subunit ScpB	BU251_03925	27	31	8	11	4	36	6
transaldolase	BU251_06585	42	35	0	14	2	29	5
UDP-N-acetylglucosamine 3-dehydrogenase	BU251_00820	36	34	18	11	7	32	16
lipid A biosynthesis lauroyl acyltransferase	BU251_00995	71	33	25	10	4	34	4
hypothetical protein	BU251_01120	13	36	4	10	1	31	12
ribokinase	BU251_01555	33	33	4	9	2	36	1
chemotaxis protein MotB	BU251_01565	25	31	6	12	2	35	9
nicotinate-nucleotide diphosphorylase (carboxylating)	BU251_04770	33	37	16	9	6	31	6
hypothetical protein	BU251_07105	45	34	4	9	1	34	1
transketolase, C-terminal section	BU251_00595	33	31	11	13	1	32	7
ATP-dependent DNA helicase UvrD/PcrA	BU251_06335	124	24	26	12	4	41	2
NADH-cytochrome b5 reductase 1	BU251_04630	27	30	7	16	3	29	1
hypothetical protein	BU251_05125	26	24	0	25	16	27	1
threonylcarbamoyl-AMP synthase	BU251_07220	38	29	2	13	1	34	13
transketolase, N-terminal section	BU251_00590	31	36	6	7	4	32	1
Type III restriction-modification system methylation subunit	BU251_02840	65	37	23	6	1	31	1
adenylate kinase	BU251_03190	24	30	4	13	1	32	4
malate dehydrogenase	BU251_03260	31	29	7	13	4	33	1
hypothetical protein	BU251_03645	23	13	3	20	3	41	4
hypothetical protein	BU251_07410	51	34	1	13	3	27	1

38

acetyl-CoA carboxylase carboxyltransferase subunit alpha	BU251_07950	36	36	6	8	1	31	11
hypothetical protein	BU251_01505	77	34	18	5	2	35	1
efflux transporter, RND family, MFP subunit	BU251_01855	35	35	5	8	4	31	3
Holliday junction DNA helicase RuvA	BU251_03930	25	35	3	10	1	29	5
hypothetical protein	BU251_06910	40	30	6	13	0	31	2
V-type ATP synthase subunit E	BU251_08240	22	34	2	11	4	30	4
hypothetical protein	BU251_08395	25	30	3	11	4	33	8
ribonucleoside-diphosphate reductase, adenosylcobalamin-dependent	BU251_07050	85	27	16	11	1	35	4
hypothetical protein	BU251_01740	32	34	4	13	4	26	3
glucose-1-phosphate adenylyltransferase	BU251_04035	46	31	11	12	1	31	1
hypothetical protein	BU251_04985	19	31	25	6	0	36	15
6,7-dimethyl-8-ribityllumazine synthase	BU251_05895	17	29	1	10	3	34	2
glycosyl transferase family 2	BU251_00420	42	37	21	10	1	26	1
hypothetical protein	BU251_06255	17	32	4	10	4	31	1
magnesium-translocating P-type ATPase	BU251_06040	95	28	6	9	4	35	10
1D-myo-inositol 2-acetamido-2-deoxy-alpha-D-glucopyranoside deacetylase	BU251_06345	54	43	23	10	1	19	3
hypothetical protein	BU251_07940	58	19	7	23	2	30	0
bifunctional homocysteine S-methyltransferase/methylenetetrahydrofolate reductase	BU251_00970	67	30	10	9	6	32	2
DNA repair protein RadA	BU251_06160	51	23	11	13	6	36	4
hypothetical protein	BU251_05150	59	32	6	7	5	32	1
crotonobetainyl-CoA--carnitine CoA-transferase	BU251_02215	28	32	9	9	5	30	5
galactose-1-phosphate uridylyltransferase	BU251_04245	39	28	6	6	1	36	4
primosomal protein N'	BU251_04260	72	30	28	6	4	33	3
phosphoribosylaminoimidazolesuccinocarboxamide synthase	BU251_07670	33	31	4	11	0	27	4
hypothetical protein	BU251_08070	36	30	9	11	5	29	6

39

acetyl-CoA synthase corrinoid iron-sulfur protein, large subunit	BU251_04965	49	31	10	8	1	30	7
diaminopimelate decarboxylase	BU251_05310	47	32	12	10	1	27	10
cell division/cell wall cluster transcriptional repressor MirZ	BU251_05800	17	25	7	11	5	33	0
cysteine--tRNA ligase	BU251_06135	54	32	23	5	7	32	21
hypothetical protein	BU251_02950	48	22	1	11	0	35	1
hypothetical protein	BU251_03535	26	34	4	11	0	24	6
RNA polymerase sigma factor RpoD	BU251_04580	32	27	8	11	2	31	2
hypothetical protein	BU251_07625	53	30	1	10	7	28	3
hypothetical protein	BU251_07845	25	16	4	6	1	47	9
hypothetical protein	BU251_04220	16	32	1	8	2	29	6
ATP-dependent DNA helicase RecG	BU251_05470	79	26	23	8	6	35	4
3-phosphoshikimate 1-carboxyvinyltransferase	BU251_05670	47	27	8	12	6	29	4
hypothetical protein	BU251_06560	38	31	8	8	2	30	9
hypothetical protein	BU251_05265	9	23	10	14	3	30	11
3-isopropylmalate dehydratase	BU251_03250	19	33	7	7	0	27	1
cochaperone GroES (HSP10)	BU251_04740	11	19	1	12	4	37	2
hypothetical protein	BU251_07840	34	26	0	11	1	30	1
translation elongation factor Ts	BU251_00835	20	34	1	5	1	28	6
NADP-reducing hydrogenase HdnC	BU251_02015	65	34	14	7	6	25	1
hypothetical protein	BU251_06400	14	29	4	10	2	28	6
hypothetical protein	BU251_04430	25	28	4	9	1	30	8
pantoate--beta-alanine ligase	BU251_05275	32	31	4	9	1	26	0
ribose-phosphate 3-epimerase	BU251_05535	24	27	7	7	1	32	2
V-type ATP synthase subunit B	BU251_08225	48	28	7	12	5	26	3
hypothetical protein	BU251_07860	18	26	1	11	2	29	0
formate--tetrahydrofolate ligase	BU251_01265	60	25	9	13	2	28	4
ribonuclease TTHA0252	BU251_01455	55	25	21	8	4	32	0
hypothetical protein	BU251_04485	61	29	11	10	1	26	5

40

16S rRNA (cytidine[1402]-2'-O)-methyltransferase	BU251_08335	24	26	1	11	3	28	0
phosphate-acyl-ACP acyltransferase	BU251_05420	36	25	3	8	1	31	3
transporter, putative	BU251_01905	94	27	6	11	5	26	4
elongation factor 4	BU251_03895	67	24	16	8	7	32	15
UDP-N-acetylmuramoylalanyl-D-glutamyl-2, 6-diaminopimelate-D-alanyl-D-alanine ligase	BU251_05770	51	24	11	7	3	33	8
GTP cyclohydrolase I FoIE	BU251_08015	21	31	1	2	2	31	14
lipopolysaccharide heptosyltransferase II	BU251_01020	77	19	9	10	4	35	5
nicotinate phosphoribosyltransferase	BU251_03565	60	25	11	12	4	27	8
4-hydroxy-tetrahydrodipicolinate reductase	BU251_05295	30	26	3	11	4	27	13
magnesium and cobalt transport protein CorA	BU251_06405	37	31	4	11	1	22	4
phosphopantothoenylcysteine decarboxylase / Phosphopantothoenylcysteine synthetase	BU251_04800	24	25	4	13	6	26	4
elongation factor P	BU251_03810	21	24	4	11	0	27	7
NADH-ubiquinone oxidoreductase chain 49kDa	BU251_04515	47	22	6	11	6	29	9
hypothetical protein	BU251_05820	35	30	17	8	3	24	8
acetolactate synthase small subunit	BU251_00925	18	31	3	9	1	22	4
thioredoxin-disulfide reductase	BU251_02120	34	29	4	7	3	25	13
cytidylate kinase	BU251_03900	25	23	5	10	3	29	6
hypothetical protein	BU251_07040	29	28	5	9	0	25	1
3-methyl-2-oxobutanoate dehydrogenase subunit VorB	BU251_09375	39	25	11	11	3	25	7
outer membrane vitamin B12 receptor BtuB	BU251_00360	68	38	21	5	0	18	1
expressed secreted protein of unknown function	BU251_01030	9	29	7	11	1	21	0
4-hydroxy-3-methylbut-2-en-1-yl diphosphate synthase	BU251_01155	39	23	8	9	2	30	1
glucose-1-phosphate adenyllyltransferase	BU251_06835	47	22	8	11	3	28	6
3-deoxy-8-phosphooctulonate synthase	BU251_01055	30	28	8	8	4	25	2
acetyl-CoA carboxylase subunit beta	BU251_02955	31	28	3	5	3	27	3
hypothetical protein	BU251_04445	28	28	4	8	0	25	1
hypothetical protein	BU251_04685	51	27	7	8	1	26	2

41

glutaminyl-tRNA synthase (glutamine-hydrolyzing)subunit B	BU251_05010	53	28	16	7	0	25	4
hypothetical protein	BU251_08805	23	28	1	7	1	26	4
hypothetical protein	BU251_01790	25	26	2	8	2	27	6
serine phosphatase RsbU, regulator of sigma subunit	BU251_02685	29	23	6	10	4	27	6
transcriptional regulator	BU251_03330	20	32	2	2	3	26	2
hypothetical protein	BU251_01365	33	29	4	9	1	21	0
chromosomal replication initiator protein DnaA	BU251_00005	50	23	16	8	5	28	18
hypothetical protein	BU251_01690	39	14	0	11	6	34	8
hypothetical protein	BU251_02370	30	31	0	7	1	21	4
hypothetical protein	BU251_03955	27	31	11	3	1	25	2
2-oxoglutarate oxidoreductase, alpha subunit	BU251_06595	63	25	11	5	2	29	10
amidophosphoribosyltransferase	BU251_07240	53	22	8	8	1	28	10
hypothetical protein	BU251_08375	19	30	3	5	1	24	6
aminoacyl-tRNA hydrolase	BU251_00635	21	24	1	6	0	28	6
50S ribosomal protein L19	BU251_01250	13	23	1	9	1	26	3
lipoprotein-releasing system transmembrane protein LolC	BU251_06210	43	24	12	8	1	27	13
mechanosensitive ion channel protein MscS	BU251_03550	39	27	8	10	3	20	13
3-oxoacyl-ACP synthase	BU251_05415	36	27	6	9	3	22	1
metallophosphoesterase	BU251_09400	30	26	6	6	0	26	1
ABC transporter ATP-binding protein	BU251_02290	27	26	3	6	4	25	1
homoserine O-acetyltransferase	BU251_06610	43	26	8	5	0	26	0
30S ribosomal protein S11	BU251_03215	13	25	5	6	4	26	9
acetyl-CoA decarboxylase/synthase complex subunit delta	BU251_04950	35	24	5	6	3	27	9
hypothetical protein	BU251_07715	19	29	4	0	0	28	2
3-hydroxyacyl-[acyl-carrier-protein] dehydratase, FabZ form	BU251_03300	16	28	0	8	5	20	1
L-Aspartate dehydrogenase	BU251_05270	28	26	7	6	3	24	1

42

hypothetical protein	BU251_06705	43	21	5	9	0	26	2
glutamate 5-kinase	BU251_05550	38	18	6	9	1	28	8
electron transport complex protein RnfG	BU251_07565	21	24	8	4	6	27	7
polysaccharide biosynthesis protein vipB/tviC	BU251_00400	35	27	8	5	3	22	4
hypothetical protein	BU251_01540	15	21	1	8	1	26	5
outer membrane lipoprotein-sorting protein	BU251_01910	30	25	4	8	2	22	5
ADP-heptose--lipooligosaccharide heptosyltransferase II	BU251_04280	42	24	7	7	1	23	7
redox-sensing transcriptional repressor Rex	BU251_02000	24	23	8	9	4	23	2
hypothetical protein	BU251_03905	30	24	7	9	0	21	2
cell division protein FtsA	BU251_05720	46	20	4	7	6	27	1
hypothetical protein	BU251_09120	24	17	6	9	0	28	6
acyl-[acyl-carrier-protein]--UDP-N-acetylglucosamine O-acyltransferase	BU251_00685	27	22	12	8	1	24	8
type II secretion system protein F	BU251_05195	46	23	8	5	3	25	10
beta-ketoacyl-[acyl-carrier-protein] synthase II	BU251_05395	44	1	1	10	0	42	6
DNA-binding response regulator	BU251_06710	26	23	0	7	4	24	4
aromatic amino acid aminotransferase	BU251_06850	43	22	4	6	1	26	8
hydroxyacid dehydrogenase	BU251_07285	39	26	8	7	4	21	1
preprotein translocase subunit YajC	BU251_05050	10	27	5	6	1	20	1
lipid-A-disaccharide synthase	BU251_00815	43	25	16	5	1	23	2
ABC transporter ATP-binding protein	BU251_06175	28	20	1	12	1	21	3
hypothetical protein	BU251_06330	108	19	18	6	2	28	2
hypothetical protein	BU251_07075	16	27	3	7	1	18	4
hypothetical protein	BU251_07945	21	24	6	6	1	23	8
putative ABC transporter ATP-binding protein YdiF	BU251_01950	73	19	15	4	5	30	4
molecular chaperone DnaJ	BU251_02935	42	21	8	6	0	25	2
50S ribosomal protein L22	BU251_03115	15	21	6	9	4	23	5
phosphomethylpyrimidine synthase	BU251_03760	48	19	6	8	1	25	8
hypothetical protein	BU251_00385	23	23	7	8	2	21	1

43

carbon monoxide dehydrogenase medium chain	BU251_01515	30	22	4	6	4	24	4
type II 3-dehydroquinone dehydratase	BU251_03820	16	22	1	9	2	21	3
leuD	BU251_05385	19	18	1	9	1	24	0
cell division protein FtsI [Peptidoglycan synthetase]	BU251_05780	63	20	6	6	0	25	4
hypothetical protein	BU251_08405	28	26	1	7	4	19	4
hypothetical protein	BU251_00275	46	17	11	14	8	20	4
3-deoxy-D-manno-octulosonate cytidyltransferase	BU251_01050	28	23	7	5	0	23	2
N-acetylmuramoyl-L-alanine amidase	BU251_01595	39	22	4	8	2	21	0
GDP-L-fucose synthetase	BU251_02230	36	22	7	8	2	21	3
transcriptional regulatory protein	BU251_03600	15	20	1	7	6	24	1
FOF1 ATP synthase subunit alpha	BU251_08480	55	19	8	11	1	21	8
sensory box/GGDEF family protein	BU251_08500	33	22	7	9	2	20	0
gatA	BU251_05015	52	22	11	9	4	20	8
lipoprotein-releasing system ATP-binding protein LolD	BU251_06205	25	21	3	6	1	23	6
hypothetical protein	BU251_05000	33	27	13	4	2	20	1
hypothetical protein	BU251_06025	29	21	4	10	1	19	1
monogalactosyldiacylglycerol synthase	BU251_04265	41	22	9	6	3	22	1
glycine dehydrogenase (aminomethyl-transferring)	BU251_07320	49	20	7	6	2	24	5
hypothetical protein	BU251_07895	36	23	8	5	1	22	8
methylenetetrahydrofolate dehydrogenase (NADP+) / methenyltetrahydrofolate cyclohydrolase	BU251_01275	32	22	10	7	3	20	8
hypothetical protein	BU251_01500	21	20	1	5	3	24	1
superoxide reductase	BU251_01930	15	24	1	6	1	20	1
hypothetical protein	BU251_02385	26	17	2	9	5	24	1
50S ribosomal protein L7/L12	BU251_03050	13	17	1	10	1	22	4
DNA-binding response regulator	BU251_08535	36	17	1	8	1	25	6
hypothetical protein	BU251_01390	11	20	3	9	1	19	4
Na(+)/H(+) antiporter NhaA 2	BU251_01820	27	21	11	4	0	24	1

44

30S ribosomal protein S8	BU251_03160	15	21	1	8	2	20	1
acetyl-CoA carboxylase, biotin carboxyl carrier protein	BU251_03805	18	20	3	7	2	22	4
diaminopimelate epimerase	BU251_05305	30	23	4	8	6	18	0
[acyl-carrier-protein] S-malonyltransferase	BU251_05410	32	25	1	4	0	20	1
3-oxoacyl-[acyl-carrier-protein] reductase	BU251_03290	28	21	6	5	3	22	2
glutamate synthase (NADPH), homotetrameric	BU251_01725	51	15	2	6	1	27	2
Holliday junction branch migration DNA helicase RuvB	BU251_01200	38	19	2	6	1	23	8
YggS family pyridoxal phosphate enzyme	BU251_05715	26	15	3	11	2	21	6
hypothetical protein	BU251_07930	36	21	6	4	6	22	4
anthranilate/aminodeoxychorismate synthase component II	BU251_01340	21	20	6	5	0	21	4
hypothetical protein	BU251_03625	12	17	4	8	4	22	3
dihydropteroate synthase	BU251_04900	43	19	10	3	0	24	16
nickel responsive regulator	BU251_07710	16	20	0	9	2	18	2
hypothetical protein	BU251_07915	53	6	1	9	3	31	0
hypothetical protein	BU251_02485	26	19	4	7	1	20	4
5-methyltetrahydrofolate:corrinoid iron-sulfur protein methyltransferase	BU251_04945	29	21	4	6	3	19	6
2'-5' RNA ligase	BU251_05910	21	23	4	3	0	20	7
tryptophan synthase subunit alpha	BU251_06015	29	21	4	8	2	18	4
S-methyl-S-thioribose-1-phosphate isomerase	BU251_06100	39	21	8	6	4	19	1
hydroxylamine reductase	BU251_06240	47	16	1	9	3	21	6
hypothetical protein	BU251_08715	12	23	0	5	1	18	2
cell division protein FtsQ	BU251_05725	32	21	2	7	3	18	2
methionine-tRNA ligase	BU251_06115	57	21	9	4	6	21	8
DNA polymerase III subunits gamma and tau	BU251_06525	61	14	9	3	1	29	4
Phosphoenolpyruvate carboxykinase (ATP)	BU251_07535	29	22	1	8	0	16	2
hypothetical protein	BU251_00250	39	21	8	4	2	20	4
outer membrane efflux protein	BU251_01850	46	21	4	4	1	20	7

45

hypothetical protein	BU251_06010	49	21	8	6	4	18	7
hypothetical protein	BU251_08415	23	19	12	3	4	23	6
hypothetical protein	BU251_08725	39	17	8	8	2	20	6
hypothetical protein	BU251_00695	10	12	7	10	4	23	13
glycosyl transferase	BU251_01005	40	17	8	7	3	20	3
3-isopropylmalate dehydratase large subunit	BU251_03240	45	18	7	8	1	18	3
DNA polymerase	BU251_04655	26	17	1	7	5	21	8
peptide chain release factor 1	BU251_01325	40	14	2	4	3	26	0
30S ribosomal protein S13	BU251_03210	15	19	0	7	2	18	1
hypothetical protein	BU251_06290	42	20	1	6	2	18	3
hypothetical protein	BU251_07430	25	19	4	7	4	19	1
RIP metalloprotease RseP	BU251_01150	40	19	5	7	4	18	5
hypothetical protein	BU251_01840	28	19	5	6	1	19	0
radical SAM domain heme biosynthesis protein	BU251_05185	37	16	6	8	4	19	1
hypothetical protein	BU251_07190	16	19	1	6	1	18	3
glycine cleavage system protein T	BU251_07310	40	20	4	4	1	20	1
methionine biosynthesis protein MetW	BU251_06605	23	19	3	6	1	18	6
hypothetical protein	BU251_08360	30	21	1	4	0	18	1
undecaprenyl-diphosphatase	BU251_03640	28	22	8	6	5	15	7
lipopolysaccharide biosynthesis protein RfbH	BU251_04390	50	21	15	4	2	18	0
4-hydroxythreonine-4-phosphate dehydrogenase PdxA	BU251_05030	34	19	8	4	1	20	1
imidazole glycerol phosphate synthase subunit HisH	BU251_05080	23	20	1	6	5	17	4
response regulator MprA	BU251_06275	34	26	2	2	2	15	8
tryptophanyl-tRNA synthetase	BU251_01190	37	19	1	7	1	16	1
thioredoxin	BU251_02095	12	16	4	9	1	17	1
hypothetical protein	BU251_03355	36	18	4	5	4	20	8
NAD(P)H-hydrate dehydratase	BU251_04880	30	17	2	9	2	17	4
tRNA lysidine(34) synthetase TIS	BU251_04910	38	18	12	6	2	19	4

46

hypothetical protein	BU251_07085	37	18	6	5	1	19	1
50S ribosomal protein L9	BU251_00655	17	17	1	7	1	18	1
endo-1,4-beta-xylanase C	BU251_02260	49	16	6	4	3	22	2
DNA-binding response regulator	BU251_02795	26	18	1	7	3	16	3
hypothetical protein	BU251_02890	30	17	0	7	1	18	5
hypothetical protein	BU251_04230	18	19	6	7	3	15	1
hypothetical protein	BU251_05590	21	19	3	5	0	17	1
hypothetical protein	BU251_06725	34	19	1	5	2	18	3
glycine dehydrogenase (aminomethyl-transferring)	BU251_07325	53	19	10	5	1	18	5
phosphoribosylglycinamide formyltransferase	BU251_07540	23	20	6	6	3	16	8
hypothetical protein	BU251_04440	41	14	3	6	2	21	6
16S rRNA (cytosine(1402)-N(4))-methyltransferase	BU251_05790	34	17	8	4	1	20	5
bifunctional pyr operon transcriptional regulator/uracil phosphoribosyltransferase	BU251_08320	20	19	1	2	0	20	13
thiol:disulfide oxidoreductase related to ResA	BU251_00515	19	18	1	7	1	16	5
peptidoglycan D,D-transpeptidase MrdA	BU251_05650	62	20	8	4	2	17	1
ribosome-binding factor A	BU251_05825	16	17	1	5	2	19	1
hypothetical protein	BU251_08425	32	20	5	7	1	14	2
hypothetical protein	BU251_00605	16	20	2	4	0	16	3
UDP-glucose 4-epimerase	BU251_02205	37	17	4	8	2	15	6
hypothetical protein	BU251_03480	18	17	7	3	4	20	4
rod shape-determining protein MreC	BU251_05660	29	17	2	6	1	18	2
hypothetical protein	BU251_01285	31	17	1	6	1	17	2
hypothetical protein	BU251_02980	37	18	10	3	4	19	1
type IV fimbrial assembly, ATPase PilB	BU251_03970	34	19	8	6	2	15	8
hypothetical protein	BU251_05965	26	18	3	4	1	17	7
hypothetical protein	BU251_06425	20	18	0	0	0	21	7
ADP-ribose pyrophosphatase	BU251_07180	30	19	8	6	2	15	5
hypothetical protein	BU251_07880	44	14	3	1	1	24	7

47

hypothetical protein	BU251_08585	27	17	4	4	3	19	6
hypothetical protein	BU251_09055	18	18	1	3	4	19	6
iron-sulfur cluster-binding protein	BU251_02155	42	17	11	6	0	16	1
50S ribosomal protein L15	BU251_03180	17	16	1	7	0	16	4
peptidoglycan-N-acetylglucosamine deacetylase	BU251_06355	29	18	11	6	1	15	7
peptidylprolyl isomerase	BU251_00430	20	14	2	8	1	17	5
hypothetical protein	BU251_03395	105	20	2	0	0	19	1
Rossmann fold protein, TIGR00730 family	BU251_07250	24	15	4	8	2	16	1
glycosyl transferase family 1	BU251_03530	45	15	5	7	4	17	5
hypothetical protein	BU251_03815	39	19	4	2	3	17	2
glucose-1-phosphate cytidyllyltransferase	BU251_04380	32	18	7	4	1	16	4
GDP-fucose synthetase	BU251_04415	37	15	4	6	1	17	10
DNA polymerase III subunit alpha	BU251_04720	129	12	17	6	8	20	10
dTMP kinase	BU251_06130	25	15	4	7	0	16	5
orotidine 5'-phosphate decarboxylase	BU251_08245	26	16	1	5	1	17	3
ferrous iron transporter B	BU251_00305	64	14	13	9	1	15	2
hypothetical protein	BU251_05035	35	17	1	6	1	15	0
hypothetical protein	BU251_05120	30	7	1	6	1	24	3
dihydroorotate dehydrogenase B catalytic subunit	BU251_08250	33	15	6	5	6	18	8
pyruvate kinase	BU251_01215	38	17	6	4	2	17	4
HIT family hydrolase	BU251_03735	18	15	4	4	2	19	1
hypothetical protein	BU251_04435	32	20	5	3	4	15	6
putative nickel insertion protein	BU251_04970	42	17	3	0	0	20	1
50S ribosomal protein L13	BU251_05355	19	16	4	6	0	15	1
protein containing Prepilin-type cleavage/methylation	BU251_05475	14	16	2	6	1	15	4
carboxynorspermidine decarboxylase	BU251_06230	41	14	11	6	0	16	6
phosphopantetheine adenylyltransferase	BU251_06265	21	18	4	5	1	14	1
hypothetical protein	BU251_07405	47	11	8	4	0	21	4
hypothetical protein	BU251_01025	8	15	1	8	1	13	1

48

ferredoxin-NADP reductase	BU251_01720	30	15	6	5	1	17	1
tRNA (adenosine(37)-N6)-dimethylallyltransferaseMiaA	BU251_02895	39	19	8	4	1	13	3
hypothetical protein	BU251_07095	51	20	1	2	3	14	13
Type IV pilin PIIA	BU251_03450	15	19	1	4	1	13	3
hypothetical protein	BU251_04225	16	18	4	4	1	14	2
nicotinate-nucleotide adenyllyltransferase	BU251_05540	22	17	2	3	1	16	0
hypothetical protein	BU251_01915	50	20	1	1	1	14	1
hypothetical protein	BU251_04315	53	16	7	3	4	16	3
response regulator	BU251_04845	14	17	4	3	4	15	1
GTPase Era	BU251_08080	33	15	4	5	2	16	5
hypothetical protein	BU251_08290	48	19	8	4	2	12	0
AAA family ATPase	BU251_09425	55	12	7	4	1	19	8
hypothetical protein	BU251_00440	29	15	1	4	1	16	1
hypothetical protein	BU251_01475	22	14	6	5	0	15	7
glycosyltransferase	BU251_04385	37	13	6	6	0	15	4
protein containing Prepilin-type cleavage/methylation apolipoprotein N-acyltransferase	BU251_07110	56	17	5	5	7	13	1
hypothetical protein	BU251_07700	26	17	2	4	5	14	0
dihydroliipoamide dehydrogenase of 2-oxoglutaratedehydrogenase	BU251_03265	46	14	4	4	2	17	1
undecaprenyl-phosphate glucose phosphotransferase	BU251_04400	55	17	10	6	1	11	6
LPS export ABC transporter permease LptG	BU251_06070	42	16	3	5	1	13	1
stage 0 sporulation protein	BU251_06120	29	17	5	4	1	14	1
hypothetical protein	BU251_01665	40	14	2	4	1	16	2
hypothetical protein	BU251_02325	12	14	1	7	2	13	3
cofactor-independent phosphoglycerate mutase	BU251_04795	44	10	3	3	4	20	6
pyridoxine 5'-phosphate synthase	BU251_04890	27	16	5	2	2	16	4
transcription-repair coupling factor	BU251_05040	71	12	17	0	0	21	3

49

sulfur reduction protein DsrE	BU251_08695	12	15	1	6	5	13	0
3-deoxy-D-manno-octulosonate 8-phosphate phosphatase	BU251_01065	21	17	4	4	0	12	6
hypothetical protein	BU251_04635	37	15	6	4	2	14	8
30S ribosomal protein S9	BU251_05350	15	13	2	6	3	14	0
pyrroline-5-carboxylate reductase	BU251_05710	30	14	1	6	5	14	4
thiamine-phosphate kinase	BU251_06095	33	16	4	3	1	14	1
2-C-methyl-D-erythritol 2,4-cyclodiphosphate synthase	BU251_06145	18	15	1	6	2	13	2
hypothetical protein	BU251_07060	24	14	4	8	2	12	1
hypothetical protein	BU251_07485	24	11	2	6	3	16	1
hypothetical protein	BU251_00260	29	4	1	4	1	24	3
glycosyltransferase	BU251_02200	25	14	4	4	1	15	6
2-oxoacid ferredoxin oxidoreductase beta subunit	BU251_06590	31	13	2	5	1	15	1
hypothetical protein	BU251_07035	32	19	9	3	0	11	1
hypothetical protein	BU251_09100	28	13	5	1	1	19	16
hypothetical protein	BU251_01445	19	15	1	5	0	12	2
cation-binding protein	BU251_01940	21	16	8	3	1	13	10
two-component hybrid sensor and regulator	BU251_02040	14	8	1	5	4	20	4
glycosyl transferase	BU251_04405	47	14	5	4	1	15	6
hypothetical protein	BU251_07680	17	13	2	6	2	14	1
hypothetical protein	BU251_08185	52	18	7	0	0	14	2
carbamoyl phosphate synthase small subunit	BU251_08280	41	10	0	5	0	17	6
hypothetical protein	BU251_02180	13	13	18	4	6	15	4
hypothetical protein	BU251_03795	13	16	4	1	1	15	2
thiamine biosynthesis lipoprotein ApbE	BU251_04615	35	14	4	4	3	14	2
V-type ATP synthase subunit I	BU251_08210	65	12	10	6	8	14	1
quinolinate synthase	BU251_00375	34	12	1	4	1	15	3
tetraacyldisaccharide 4'-kinase	BU251_00990	39	16	8	4	3	11	7
indole-3-glycerol phosphate synthase	BU251_01350	29	13	4	4	2	15	1

50

adenosine monophosphate-protein transferase SoFic	BU251_01735	43	15	10	3	1	13	0
hypothetical protein	BU251_06935	29	14	1	4	1	13	0
radical SAM protein	BU251_00570	32	13	1	2	2	16	0
shikimate dehydrogenase	BU251_00940	28	15	4	3	1	13	1
geranylgeranyl diphosphate synthase	BU251_01585	40	16	8	1	1	13	6
hypothetical protein	BU251_02330	28	13	4	4	0	13	6
rRNA small subunit methyltransferase E	BU251_05140	27	13	2	4	1	13	7
50S ribosomal protein L21	BU251_05560	11	13	0	6	4	11	3
hypothetical protein	BU251_08010	15	16	1	5	1	9	0
4-(cytidine 5'-diphospho)-2-C-methyl-D-erythritol kinase	BU251_00600	34	13	6	4	1	12	1
thiamin-phosphate pyrophosphorylase	BU251_03770	17	13	1	4	1	12	3
hypothetical protein	BU251_05165	18	13	5	4	2	13	1
methylenetetrahydrofolate dehydrogenase	BU251_00215	51	15	0	2	2	12	1
hypothetical protein	BU251_02700	14	14	5	3	1	13	2
hypothetical protein	BU251_08520	44	11	3	4	1	14	1
site-specific tyrosine recombinase XerD	BU251_05840	34	15	8	3	1	11	6
transketolase, C-terminal section	BU251_02225	33	9	3	6	1	13	3
3-oxoacyl-[acyl-carrier-protein] synthase, KASII	BU251_03305	45	11	2	3	4	15	1
30S ribosomal protein S16	BU251_05055	9	10	1	7	0	11	4
ATP-dependent permease AU51	BU251_06085	20	12	3	3	1	13	1
hypothetical protein	BU251_08555	10	15	4	3	1	10	2
2-oxoglutarate ferredoxin oxidoreductase subunitgamma	BU251_09365	18	16	1	0	0	12	1
hypothetical protein	BU251_00280	36	12	1	4	1	12	1
anthranilate phosphoribosyltransferase	BU251_01345	36	12	4	1	1	14	4
30S ribosomal protein S12	BU251_03065	14	10	1	4	2	14	4
hypothetical protein	BU251_04925	26	12	6	2	2	14	9
type IV pilin PilA	BU251_06880	19	12	1	3	0	12	3

51

5-(carboxyamino)imidazole ribonucleotide mutase	BU251_07225	17	13	1	5	1	10	0
electron transport complex protein RnfC	BU251_07570	47	9	1	0	0	18	1
molybdopterin biosynthesis protein MoeA	BU251_07780	39	13	11	2	2	13	6
guanylate kinase	BU251_09515	23	12	7	3	1	12	1
ABC transporter	BU251_00455	35	10	2	4	5	14	2
NADP-reducing hydrogenase HdnA	BU251_02025	18	10	3	4	2	13	0
riboflavin biosynthesis protein RibF	BU251_05810	36	12	5	4	1	12	2
pstB	BU251_06650	28	12	1	3	1	12	1
hypothetical protein	BU251_07865	30	9	1	4	1	14	4
hypothetical protein	BU251_07975	14	10	1	6	1	11	1
hypothetical protein	BU251_09135	21	14	0	2	2	11	0
2-oxoglutarate oxidoreductase	BU251_09370	26	14	4	2	2	12	4
peptidase	BU251_01945	37	14	1	2	3	11	6
hypothetical protein	BU251_04185	29	9	6	5	2	13	1
alanine racemase	BU251_06075	42	12	8	2	3	13	1
2-C-methyl-D-erythritol 4-phosphate cytidyltransferase	BU251_06150	26	11	1	4	3	12	6
sporulation initiation inhibitor Soj	BU251_09535	30	10	5	5	3	12	5
undecaprenyl pyrophosphate phosphatase	BU251_06960	17	9	13	0	0	17	1
hypothetical protein	BU251_08445	20	13	1	3	1	10	0
hypothetical protein	BU251_09460	24	13	2	0	0	13	1
xanthine and CO dehydrogenases maturation factor, XdhC/CoxF family	BU251_01535	27	12	1	2	3	12	1
hypothetical protein	BU251_03000	25	11	0	5	1	10	1
30S ribosomal protein S17	BU251_03135	11	9	1	4	1	12	4
protein containing Type IV pilus assembly PilZ domain	BU251_03685	15	12	2	4	1	10	3
hypothetical protein	BU251_08035	26	9	1	2	3	14	8
protein containing Prepilin-type cleavage/methylation	BU251_08430	21	14	3	3	1	8	1

52

UDP-N-acetylmuramoyl-L-alanyl-D-glutamate--2, 6-diaminopimelate ligase	BU251_05775	53	9	8	3	1	13	4
tRNA (adenosine(37)-N6)-threonylcarbamoyltransferase complex dimerization subunit type 1 TsaB	BU251_06080	25	9	3	5	4	11	1
hypothetical protein	BU251_08075	40	12	8	1	1	12	1
hypothetical protein	BU251_09020	18	13	4	2	2	11	2
ribonuclease PH	BU251_01635	26	12	1	2	3	11	2
hypothetical protein	BU251_02190	45	14	2	0	0	11	8
CO dehydrogenase accessory protein CooC (nickel insertion)	BU251_04955	27	11	1	1	1	13	6
lipopolysaccharide export system permease protein LptF	BU251_06065	41	10	5	6	8	9	1
nitroreductase	BU251_06940	24	3	1	5	0	17	6
23S rRNA (adenine(2503)-C(2))-methyltransferase	BU251_07760	39	11	4	4	1	10	1
transmembrane and TPR repeat-containing protein	BU251_07990	80	12	4	1	1	12	4
V-type ATP synthase subunit D	BU251_08220	24	9	3	3	4	13	5
tRNA (guanosine(37)-N1)-methyltransferase TrmD	BU251_01245	27	9	4	3	1	12	1
dephospho-CoA kinase	BU251_01310	26	9	1	3	1	12	0
hypothetical protein	BU251_01560	18	12	1	2	2	11	1
NADPH-dependent 7-cyano-7-deazaguanine reductaseQueF	BU251_01620	15	11	1	3	1	11	2
NAD(P)H-hydrate epimerase	BU251_02315	23	11	2	3	4	11	1
acetoacetyl-CoA synthetase / long-chain-fatty-acid-CoA ligase	BU251_03635	58	9	7	4	6	11	2
hypothetical protein	BU251_04470	18	8	2	3	1	13	3
[NiFe] hydrogenase metallocenter assembly protein HypE	BU251_07815	36	13	5	2	2	10	2
hypothetical protein	BU251_08140	11	8	3	4	1	12	0
DNA helicase UvrD	BU251_08300	46	8	3	3	0	13	2
hypothetical protein	BU251_08380	11	10	1	4	0	10	1
lipase family protein	BU251_09030	31	12	1	4	1	8	1

53

hypothetical protein	BU251_09105	38	14	4	0	0	10	2
non-canonical purine NTP pyrophosphatase, RdgB/HAM1 family	BU251_01640	22	11	1	3	1	10	1
alpha-D-GlcNAc alpha-1,2-L-rhamnosyltransferase	BU251_01875	42	10	4	2	2	12	2
hypothetical protein	BU251_02850	24	10	1	3	1	11	1
ADP-heptose synthase / D-glycero-beta-D-manno-heptose 7-phosphate kinase	BU251_03775	17	9	4	4	1	10	1
GTP-binding protein Obg	BU251_05555	35	10	5	3	1	11	1
hypothetical protein	BU251_01225	42	11	9	3	1	9	4
hypothetical protein	BU251_01920	17	9	1	4	2	11	1
30S ribosomal protein S10	BU251_03085	11	13	3	0	0	10	2
trehalose-phosphatase	BU251_03540	30	9	6	2	3	12	2
D-glycero-D-manno-heptose 1-phosphate guanosyltransferase	BU251_04270	26	10	0	2	0	11	4
hypothetical protein	BU251_06250	27	9	6	4	1	10	0
hypothetical protein	BU251_07055	81	10	5	3	4	11	4
hypothetical protein	BU251_08265	28	12	4	2	3	9	3
CO dehydrogenase/acetyl-CoA synthase, acetyl-CoA synthase subunit	BU251_01290	79	7	10	1	1	14	8
hypothetical protein	BU251_02055	26	10	2	4	3	9	4
GNAT family N-acetyltransferase	BU251_04030	17	11	1	3	4	9	2
imidazole glycerol phosphate synthase subunit HisF	BU251_05070	31	9	3	1	1	12	1
16S rRNA (guanine(966)-N(2))-methyltransferase RsmD	BU251_05440	20	11	2	2	2	10	1
ribosomal large subunit pseudouridine synthase D	BU251_05675	34	10	4	3	0	10	2
fumarate hydratase	BU251_01625	30	12	3	1	1	9	2
UDP-glucose 4-epimerase	BU251_02235	37	10	5	3	1	10	1
hypothetical protein	BU251_03005	41	10	8	2	3	10	3
hypothetical protein	BU251_07870	15	11	2	6	2	6	1
phosphate transport system regulatory protein PhoU	BU251_09090	23	12	1	0	0	10	8

54

FAD-dependent pyridine nucleotide-disulfide oxidoreductase	BU251_01590	42	7	4	1	1	13	1
hypothetical protein	BU251_02165	14	9	2	4	1	9	1
crossover junction endodeoxyribonuclease RuvC	BU251_03935	18	12	1	2	2	8	1
RNA-binding transcriptional accessory protein	BU251_05975	81	8	11	2	2	12	3
hypothetical protein	BU251_06995	35	10	1	2	3	9	0
N-acetylmuramoyl-L-alanine amidase	BU251_07290	29	9	1	2	2	11	1
thiol:disulfide oxidoreductase related to ResA	BU251_07350	18	9	1	3	1	9	0
shikimate kinase I	BU251_00955	19	12	1	0	0	9	1
50S ribosomal protein L16	BU251_03125	16	9	1	3	0	9	1
hypothetical protein	BU251_05455	26	9	1	3	1	9	1
hypothetical protein	BU251_05640	64	9	9	1	1	11	3
hypothetical protein	BU251_06775	21	11	4	2	3	8	1
hypothetical protein	BU251_08705	13	10	3	2	3	9	1
iron-sulfur cluster assembly protein SufB	BU251_02295	40	8	0	3	0	9	0
ribosome-recycling factor	BU251_02440	60	13	11	0	0	7	3
50S ribosomal protein L18	BU251_03170	10	8	1	6	1	7	1
hypothetical protein	BU251_04345	39	11	4	1	1	9	8
hypothetical protein	BU251_05445	45	12	2	0	0	9	2
protein containing General secretory system II	BU251_05630	17	9	0	2	3	9	1
ATP-dependent RNA helicase DOB1	BU251_06270	73	10	6	2	2	9	0
DNA repair protein	BU251_06990	21	9	0	1	1	10	3
site-specific DNA-methyltransferase	BU251_08755	38	11	8	0	0	10	8
hypothetical protein	BU251_00985	49	8	2	2	2	11	2
phosphoribosylanthranilate isomerase	BU251_01355	23	11	4	1	1	8	0
endonuclease III	BU251_01440	25	8	3	2	3	10	1
gamma-D-glutamyl-L-lysine dipeptidyl-peptidase	BU251_01700	36	9	0	1	1	10	2
3-oxoadipate enol-lactonase	BU251_02420	33	11	1	2	0	7	1
NADH-ubiquinone oxidoreductase chain C	BU251_04520	18	10	2	5	1	6	2
hypothetical protein	BU251_05250	26	7	2	0	0	13	6

55

rubrerythrin family protein	BU251_06435	32	12	1	1	1	7	1
thioredoxin reductase	BU251_00500	32	10	1	1	1	9	2
lipopolysaccharide heptosyltransferase II	BU251_01015	40	10	4	3	1	7	2
transcription elongation factor GreA	BU251_01570	18	9	2	2	2	9	4
TPR Domain containing protein	BU251_02415	21	10	1	0	0	9	4
hypothetical protein	BU251_04080	81	6	8	2	2	12	2
biotin-[acetyl-CoA-carboxylase] ligase	BU251_04750	29	10	1	2	2	8	1
UDP-N-acetylmuramoylalanine--D-glutamate ligase	BU251_05760	46	6	4	2	0	11	1
hypothetical protein	BU251_07000	18	8	3	2	2	10	1
exodeoxyribonuclease VII large subunit	BU251_09395	43	10	1	0	0	9	4
50S ribosomal protein L20	BU251_09465	14	8	1	2	3	9	6
hypothetical protein	BU251_02065	26	10	6	2	2	7	3
hypothetical protein	BU251_07495	33	11	4	0	0	8	1
cysteine desulfurase NifS	BU251_00355	43	9	3	0	0	9	4
nucleotide exchange factor GrpE	BU251_02925	23	6	0	3	4	10	5
3-oxoacyl-[acyl-carrier-protein] synthase, KASII	BU251_03285	46	9	2	2	0	8	2
hypothetical protein	BU251_05480	14	6	8	5	1	8	3
[NiFe] hydrogenase metallocenter assembly protein HypD	BU251_07820	40	6	3	0	0	12	1
hypothetical protein	BU251_02425	46	7	1	1	1	10	3
hypothetical protein	BU251_04000	21	10	2	3	4	5	7
hypothetical protein	BU251_04375	37	7	4	3	4	8	3
hypothetical protein	BU251_08175	27	9	1	0	0	9	1
hypothetical protein	BU251_00080	59	10	8	0	0	7	3
protease	BU251_00340	19	7	3	3	0	7	1
polysaccharide deacetylase	BU251_01010	30	9	1	1	1	7	0
1-deoxy-D-xylulose-5-phosphate reductoisomerase	BU251_01145	42	8	1	0	0	9	1
formiminotetrahydrofolate cyclodeaminase	BU251_01270	18	9	1	0	0	9	1
hypothetical protein	BU251_01425	42	9	5	3	1	6	3
type I methionyl aminopeptidase	BU251_03195	27	7	3	1	1	9	3

56

hypothetical protein	BU251_03630	15	8	0	3	1	6	0
general secretion pathway protein B	BU251_03700	16	9	1	0	0	8	0
pantetheine-phosphate adenyllyltransferase	BU251_05435	18	7	1	3	4	8	1
UDP-N-acetylglucosamine-N-acetylmuramyl- (pentapeptide) pyrophosphoryl-undecaprenol N- acetylglucosamine transferase	BU251_05750	39	7	2	3	1	8	4
phosphate ABC transporter ATP-binding protein	BU251_06645	28	5	3	0	0	12	4
ADP-ribose pyrophosphatase	BU251_07270	20	7	3	2	0	8	1
glycine cleavage system protein H	BU251_07315	14	6	1	3	1	8	0
hypothetical protein	BU251_08435	15	8	3	1	1	8	3
diguanylate cyclase/phosphodiesterase (GGDEF & EAL domains) with PAS/PAC sensor(s)	BU251_02970	39	8	3	1	1	8	6
50S ribosomal protein L14	BU251_03140	13	8	2	3	1	6	3
hypothetical protein	BU251_03595	27	8	0	2	2	7	4
TPR domain protein	BU251_05745	31	6	4	3	1	9	4
hypothetical protein	BU251_05845	21	8	1	2	3	7	2
hypothetical protein	BU251_06110	28	9	2	0	0	8	0
glycoside hydrolase	BU251_06450	43	7	6	2	0	8	2
hypothetical protein	BU251_07345	28	9	5	3	0	5	1
electron transport complex protein RnfD	BU251_07560	34	8	6	1	1	8	1
DNA-binding response regulator	BU251_09110	26	10	3	0	0	7	2
hypothetical protein	BU251_01755	17	8	1	0	0	9	4
GGDEF/response regulator receiver domain protein	BU251_02965	49	8	4	0	0	9	6
hypothetical protein	BU251_03750	25	8	3	3	1	5	1
hypothetical protein	BU251_04360	46	7	4	0	0	9	1
hypothetical protein	BU251_06485	44	6	0	1	1	9	1
hypothetical protein	BU251_02740	10	6	0	2	2	8	4
GTPase HflX	BU251_02900	47	10	1	0	0	6	1
50S ribosomal protein L11	BU251_03035	15	9	4	0	0	7	1
septum formation protein Maf	BU251_06045	49	9	4	0	0	7	1

57

ferric siderophore transport system, biopolymer transport protein ExbB	BU251_07595	22	9	1	0	0	7	2
6-carboxytetrahydropterin synthase QueD	BU251_01605	15	7	4	2	2	7	1
heavy metal translocating P-type ATPase	BU251_01785	82	4	5	3	4	9	8
GNAT family N-acetyltransferase	BU251_02570	18	7	4	4	1	5	0
hypothetical protein	BU251_03570	36	8	0	0	0	7	0
thymidylate synthase	BU251_04840	29	6	1	3	1	7	2
hypothetical protein	BU251_04850	15	9	3	0	0	6	1
hypothetical protein	BU251_06950	22	8	0	0	0	7	0
hypothetical protein	BU251_06980	26	6	0	0	0	9	0
hypothetical protein	BU251_07175	22	6	2	0	0	10	5
hypothetical protein	BU251_09355	32	8	1	2	0	5	1
response regulator	BU251_01435	15	6	1	1	1	8	1
hypothetical protein	BU251_02250	30	8	4	0	0	7	0
putative metallophosphoesterase	BU251_03740	18	5	1	4	1	7	1
nucleotidyl transferase	BU251_04325	27	9	1	0	0	6	0
pyridoxamine 5'-phosphate oxidase	BU251_05460	16	7	2	0	0	8	1
ribosome silencing factor	BU251_06005	13	6	3	2	0	7	2
hypothetical protein	BU251_07010	35	7	3	1	1	7	1
hypothetical protein	BU251_08180	24	8	6	3	1	4	6
cupin	BU251_09130	12	8	1	1	1	6	2
sugar ABC transporter permease	BU251_00195	40	9	5	2	3	4	5
hypothetical protein	BU251_00450	18	7	2	0	0	8	2
outer membrane efflux protein precursor	BU251_02060	50	8	2	1	1	6	2
metal-dependent hydrolase of the beta-lactamase superfamily	BU251_02140	29	8	4	1	1	6	1
hypothetical protein	BU251_04165	42	7	5	2	2	6	4
hypothetical protein	BU251_04465	14	5	0	1	1	8	6
secreted protein containing Prepilin-type cleavage/methylation	BU251_05170	15	6	3	3	1	6	1

58

hypothetical protein	BU251_05690	15	6	2	0	0	9	1
transmembrane and TPR repeat-containing protein	BU251_06315	32	7	1	2	2	6	1
cupin	BU251_07280	11	6	1	1	1	8	1
hypothetical protein	BU251_07890	24	7	3	0	0	7	1
phosphate regulon transcriptional regulatory protein PhoB (SphR)	BU251_02310	14	5	1	3	0	6	0
hypothetical protein	BU251_03610	26	8	2	0	0	6	1
ATP-dependent DNA helicase UvrD/PcrA, proteobacterial protein	BU251_04660	78	6	8	0	0	8	0
biopolymer transport protein ExbD	BU251_05245	14	8	1	1	1	5	1
hypothetical protein	BU251_06220	14	8	0	0	0	6	1
hypothetical protein	BU251_07415	55	6	5	0	0	8	0
RNA polymerase sigma factor RpoE	BU251_07660	21	6	3	0	0	8	2
hypothetical protein	BU251_07765	17	8	1	1	1	5	1
TrpB-like pyridoxal-phosphate dependent enzyme	BU251_08145	50	8	2	0	0	6	4
tRNA (NG-threonylcarbamoyladenosine(37)-N6)-methyltransferase TrmO	BU251_08690	19	7	2	1	1	6	1
30S ribosomal protein S6	BU251_00640	12	7	1	1	1	5	4
hypothetical protein	BU251_00700	27	5	3	0	0	8	4
preprotein translocase subunit SecY	BU251_03185	50	7	4	0	0	6	1
hypothetical protein	BU251_04320	74	5	1	2	3	7	2
RNA-splicing ligase RtcB	BU251_06780	53	4	5	0	0	10	2
secondary thiamine-phosphate synthase enzyme	BU251_08135	16	7	0	0	0	6	0
5,10-methylenetetrahydrofolate reductase	BU251_01280	28	7	2	2	0	4	0
hypothetical protein	BU251_01480	9	6	1	3	1	5	1
phosphatidylethanolamine-binding protein	BU251_02175	22	1	1	3	0	9	1
sigma-54 dependent transcriptional regulator/response regulator	BU251_02755	23	7	3	0	0	6	1
hypothetical protein	BU251_03365	31	5	3	0	0	8	2
D-alanyl-D-alanine carboxypeptidase	BU251_04605	30	7	3	1	1	5	2
hypothetical protein	BU251_04650	17	5	1	3	1	6	1

59

transmembrane and TPR repeat-containing protein CG5038	BU251_04940	75	6	4	3	4	5	4
protein containing Prepilin-type cleavage/methylation	BU251_05175	19	9	1	0	0	4	5
polyphosphoinositide phosphatase	BU251_05210	18	7	1	0	0	6	1
hypothetical protein	BU251_05450	8	6	1	1	1	6	0
hypothetical protein	BU251_07090	19	7	1	1	1	5	0
tRNA uridine-5-carboxymethylaminomethyl(34) synthesis GTPase MnmE	BU251_08020	52	7	4	0	0	6	1
hypothetical protein	BU251_01205	22	6	4	2	2	5	0
50S ribosomal protein L23	BU251_03100	11	6	2	1	1	6	1
50S ribosomal protein L24	BU251_03145	12	5	1	0	0	7	1
hypothetical protein	BU251_03580	40	6	1	0	0	6	1
hypothetical protein	BU251_03590	16	5	2	0	0	8	2
hypothetical protein	BU251_04340	38	5	4	1	1	7	6
riboflavin synthase bacterial/eukaryotic	BU251_05875	21	8	5	0	0	5	2
30S ribosomal protein S15	BU251_05960	10	5	0	2	3	5	0
UDP-3-O-acetyl-N-acetylglucosamine deacetylase	BU251_06375	23	6	1	0	0	7	5
archaease	BU251_06785	13	7	1	2	2	4	0
hypothetical protein	BU251_07015	17	5	1	3	1	5	1
cold-shock protein	BU251_01795	7	3	0	5	6	4	6
hypothetical protein	BU251_01885	63	7	9	1	1	4	3
hypothetical protein	BU251_04140	43	6	1	0	0	6	1
methionyl-tRNA formyltransferase	BU251_04255	34	4	6	1	1	7	2
rRNA small subunit methyltransferase A	BU251_05025	31	5	1	0	0	7	1
transcription antitermination factor NusB	BU251_05890	16	6	1	0	0	6	3
hypothetical protein	BU251_06765	23	6	1	0	0	6	1
hypothetical protein	BU251_08685	34	4	1	1	1	7	1
tyrosine recombinase XerC	BU251_00980	33	6	4	0	0	5	1
hypothetical protein	BU251_03310	43	7	6	0	0	4	0

60

nitroreductase family protein	BU251_03980	19	6	0	1	1	4	1
hypothetical protein	BU251_04150	46	6	4	0	0	6	4
hypothetical protein	BU251_04160	41	5	4	1	1	5	1
hypothetical protein	BU251_04480	22	6	1	0	0	6	4
tRNA (adenosine(37)-N6)- threonylcarbamoyltransferase complex ATPase subunit type 1 TsaE TPR repeat	BU251_00710	74	5	1	3	1	3	0
fumarate hydratase	BU251_01630	20	6	1	0	0	5	1
dTDP-4-dehydrohamnose 3,5-epimerase	BU251_02240	20	2	3	1	1	8	1
phosphoribosyl transferase domain protein	BU251_02270	25	6	0	0	0	5	1
hypothetical protein	BU251_02400	40	5	2	0	0	6	4
hypothetical protein	BU251_03835	35	6	5	0	0	5	1
excinuclease ABC subunit C	BU251_07135	48	6	4	0	0	5	3
transcriptional regulator NrdR	BU251_07195	18	6	1	0	0	5	2
peptide deformylase	BU251_02115	16	4	1	2	3	5	1
hypothetical protein	BU251_02395	26	5	0	0	0	5	1
AmmeMemoRadISam system protein B	BU251_03830	29	7	5	0	0	4	1
cytidine deaminase	BU251_07205	17	5	1	1	1	4	0
ribose 5-phosphate isomerase B	BU251_07215	17	4	5	0	0	7	2
hypothetical protein	BU251_08060	28	5	4	0	0	6	2
hypothetical protein	BU251_00145	8	3	4	2	3	5	3
multidrug ABC transporter ATP-binding protein	BU251_00750	35	5	1	0	0	5	0
hypothetical protein	BU251_01890	56	6	3	0	0	4	1
purine nucleoside phosphorylase	BU251_02135	17	6	1	0	0	4	1
hypothetical protein	BU251_03845	28	4	1	0	0	6	5
protein-arginine kinase activator protein	BU251_06195	20	6	2	0	0	4	6
hypothetical protein	BU251_06530	15	7	1	0	0	3	0
hypothetical protein	BU251_07900	20	5	0	0	0	5	1
flavin reductase	BU251_08165	20	5	1	0	0	5	1

61

Fe-S cluster assembly scaffold protein NifU	BU251_09415	14	6	1	0	0	4	0
pyruvate:ferredoxin oxidoreductase, delta subunit	BU251_00530	9	2	3	0	0	7	3
hypothetical protein	BU251_00785	68	7	2	0	0	3	1
NADP-reducing hydrogenase HdnB	BU251_02010	13	4	1	1	1	5	1
transketolase, N-terminal section	BU251_02220	31	4	3	0	0	5	0
SAM-dependent methyltransferase	BU251_02375	31	5	1	1	1	4	1
K ⁺ /H ⁺ antiporter	BU251_02790	53	4	1	0	0	6	2
hypothetical protein	BU251_04110	40	4	3	2	2	4	5
hypothetical protein	BU251_04355	60	5	7	1	1	3	4
hypothetical protein	BU251_07725	22	0	0	9	4	0	0
hypothetical protein	BU251_09450	55	6	8	0	0	3	4
phosphoglycolate phosphatase	BU251_00265	26	2	3	1	1	6	4
tRNA preQ1(34) S-adenosylmethionine ribosyltransferase-isomerase QueA	BU251_01230	39	4	5	0	0	5	3
CTP:molybdopterin cytidyltransferase	BU251_01530	22	5	0	0	0	4	1
cold-shock protein	BU251_03490	7	0	0	0	0	9	12
hypothetical protein	BU251_03665	9	3	1	1	1	5	1
hypothetical protein	BU251_03680	22	5	1	0	0	4	1
hypothetical protein	BU251_04990	17	6	0	0	0	3	4
phosphoribosyl-AMP cyclohydrolase	BU251_05065	14	5	1	0	0	4	0
ribosomal protein S12p methylthiotransferase	BU251_05940	50	4	2	2	3	3	1
4-hydroxy-3-methylbut-2-enyl diphosphate reductase	BU251_06320	31	4	5	0	0	5	7
RNA pseudouridine synthase	BU251_06490	28	4	0	0	0	5	6
recombination protein RecR	BU251_06520	22	3	1	0	0	6	1
hypothetical protein	BU251_06695	47	9	1	0	0	0	0
A/G-specific adenine glycosylase	BU251_06715	40	7	7	0	0	2	2
phosphoribosylformylglycinamide synthase I	BU251_07245	29	6	2	0	0	3	4
V-type ATP synthase subunit K	BU251_08205	15	7	3	0	0	2	2
tRNA guanosine(34) transglycosylase Tgt	BU251_01240	42	3	0	0	0	5	1

62

membrane fusion component of tripartite multidrug resistance system	BU251_01770	32	4	1	1	1	4	1
MBL fold metallo-hydrolase	BU251_02285	32	4	2	0	0	5	1
hypothetical protein	BU251_04290	37	5	7	0	0	3	0
hypothetical protein	BU251_04980	12	5	1	0	0	3	1
hypothetical protein	BU251_06050	44	3	0	1	1	4	0
hypothetical protein	BU251_07985	58	7	1	0	0	1	1
5-formyltetrahydrofolate cyclo-ligase	BU251_09410	22	4	1	1	1	4	1
hypothetical protein	BU251_00115	14	3	4	0	0	5	1
glycosyl transferase, family 2	BU251_00790	27	4	1	0	0	4	0
undecaprenyl diphosphate synthase	BU251_01135	28	3	0	1	1	4	1
hypothetical protein	BU251_01575	18	3	4	0	0	5	1
hypothetical protein	BU251_02735	8	3	1	1	1	4	1
hypothetical protein	BU251_03745	64	6	8	0	0	2	3
secondary thiamine-phosphate synthase enzyme	BU251_04640	17	5	2	0	0	3	0
hypothetical protein	BU251_06885	26	4	5	0	0	4	0
[NiFe] hydrogenase metalcenter assembly protein HypF	BU251_07830	43	5	6	0	0	3	4
hypothetical protein	BU251_08160	20	4	1	0	0	4	1
O-acetyl-ADP-ribose deacetylase	BU251_00810	18	4	1	0	0	3	0
ribosome maturation factor RimP	BU251_01175	18	4	1	0	0	4	1
hypothetical protein	BU251_02365	26	2	0	1	1	4	1
proline iminopeptidase	BU251_03505	32	4	6	0	0	3	1
hypothetical protein	BU251_03870	24	4	1	0	0	4	1
hypothetical protein	BU251_03875	13	3	4	2	2	3	1
DNA polymerase III subunit delta'	BU251_06125	35	4	5	0	0	4	5
3-oxoacyl-[acyl-carrier protein] reductase	BU251_06470	28	4	2	0	0	4	1
hypothetical protein	BU251_06720	37	4	1	0	0	4	2
hypothetical protein	BU251_06900	34	4	0	0	0	3	4
hypothetical protein	BU251_06915	18	4	1	0	0	3	0

63

peptide chain release factor 2	BU251_07120	41	4	2	1	1	3	4
magnesium transporter	BU251_09075	51	3	1	0	0	5	6
TPR domain protein	BU251_00585	36	4	0	0	0	3	1
ABC-type multidrug transport system, permease component	BU251_00755	43	4	1	0	0	3	4
RNA polymerase sigma factor RpoE	BU251_01465	23	5	1	0	0	2	0
hypothetical protein	BU251_01825	15	4	1	0	0	3	1
K ⁺ /H ⁺ antiporter	BU251_02335	53	3	4	0	0	4	1
3-methyl-2-oxobutanoate hydroxymethyltransferase	BU251_04595	29	3	0	0	0	4	1
undecaprenyl phosphate-alpha-4-amino-4-deoxy-L-arabinose arabinosyl transferase	BU251_06350	64	4	5	1	1	2	3
hypothetical protein	BU251_06955	21	2	2	0	0	5	0
hypothetical protein	BU251_07730	49	4	1	0	0	3	1
hypothetical protein	BU251_08270	30	1	1	0	0	6	1
hypothetical protein	BU251_08655	46	4	2	1	1	2	3
ferrous iron transport protein A	BU251_00300	9	3	1	0	0	3	0
hypothetical protein	BU251_02630	9	0	0	0	0	6	3
hypothetical protein	BU251_02825	10	3	0	0	0	3	1
hypothetical protein	BU251_04175	48	3	4	0	0	4	1
regulatory protein RecX	BU251_05100	18	3	1	0	0	3	1
hypothetical protein	BU251_05835	50	2	3	0	0	4	0
hypothetical protein	BU251_06325	47	2	3	0	0	4	0
hypothetical protein	BU251_06625	31	5	3	0	0	1	1
hypothetical protein	BU251_07150	26	4	1	0	0	3	1
hypothetical protein	BU251_08085	57	3	4	0	0	4	2
dihydroorotase	BU251_08310	46	2	3	0	0	4	0
ATP synthase gamma chain	BU251_08475	36	3	1	0	0	3	1
hypothetical protein	BU251_09050	27	5	2	0	0	2	2
hypothetical protein	BU251_09095	20	3	1	0	0	4	5

64

hypothetical protein	BU251_01040	34	2	2	0	0	4	1
molybdenum ABC transporter ATP-binding protein	BU251_01695	31	2	2	0	0	4	3
MFS transporter	BU251_01995	43	2	3	3	4	1	1
retron-type RNA-directed DNA polymerase	BU251_02720	45	3	4	0	0	3	4
hypothetical protein	BU251_03605	26	3	4	0	0	3	1
tRNA pseudouridine(55) synthase TruB	BU251_05815	26	2	0	0	0	4	2
metal-dependent phosphoesterases (PHP family)	BU251_05885	31	3	4	0	0	3	1
phosphate ABC transporter permease subunit PstC	BU251_06660	32	2	2	1	1	4	1
hypothetical protein	BU251_07755	15	2	2	0	0	4	1
hypothetical protein	BU251_08350	20	2	2	0	0	4	1
hypothetical protein	BU251_09420	23	2	0	0	0	4	1
transcriptional regulator	BU251_00335	17	2	2	0	0	4	1
1-acyl-sn-glycerol-3-phosphate acyltransferase	BU251_00445	29	2	2	0	0	4	1
hypothetical protein	BU251_00725	14	3	0	0	0	2	0
DipZ protein	BU251_01450	24	4	2	0	0	2	2
hypothetical protein	BU251_01750	20	2	2	0	0	4	1
hypothetical protein	BU251_02090	43	4	1	0	0	2	2
DNA-binding response regulator KdpE	BU251_02300	14	3	1	0	0	2	3
hypothetical protein	BU251_02575	10	3	0	0	0	2	0
transcription elongation factor GreAB	BU251_02830	16	2	2	0	0	4	2
hypothetical protein	BU251_02905	10	2	3	0	0	3	0
hypothetical protein	BU251_04040	22	4	5	0	0	2	2
phenylacetate-coenzyme A ligase	BU251_04100	36	3	4	0	0	2	3
D-erythrose-4-phosphate dehydrogenase	BU251_05230	18	3	4	0	0	3	1
tRNA pseudouridine(38-40) synthase TruA	BU251_05365	30	3	0	0	0	2	0
hypothetical protein	BU251_05375	18	4	2	0	0	2	2
50S ribosomal protein L32	BU251_05425	9	4	1	0	0	2	2
UDP-N-acetylenolpyruvoylglucosamine reductase	BU251_05735	33	3	4	0	0	3	1
cold-shock protein	BU251_06340	7	0	0	0	0	5	1

65

tRNA-specific adenosine deaminase	BU251_06905	18	2	3	0	0	3	0
hypothetical protein	BU251_06925	14	3	1	0	0	3	1
tRNA dihydrouridine synthase DusB	BU251_07165	36	3	4	0	0	3	1
hypothetical protein	BU251_07460	11	3	1	0	0	2	0
hypothetical protein	BU251_08050	37	3	4	0	0	2	3
polysaccharide deacetylase	BU251_00415	32	1	1	1	1	3	1
aspartate 1-decarboxylase	BU251_02110	13	1	1	0	0	4	1
hypothetical protein	BU251_04205	13	2	2	0	0	3	0
hypothetical protein	BU251_06930	25	0	0	0	0	5	2
hypothetical protein	BU251_06965	18	2	3	0	0	3	1
hypothetical protein	BU251_07505	21	4	2	0	0	1	1
hypothetical protein	BU251_07925	17	2	3	0	0	3	1
permease	BU251_08260	29	3	1	0	0	2	3
Type II/IV secretion system protein TadC, associated with Flp pilus assembly	BU251_00155	34	2	2	0	0	3	4
hypothetical protein	BU251_00470	14	4	0	0	0	0	0
SsrA-binding protein	BU251_01410	18	3	1	0	0	1	1
acetyltransferase	BU251_02355	26	2	0	0	0	2	0
hypothetical protein	BU251_03825	13	2	2	0	0	3	4
CDP-diacylglycerol-glycerol-3-phosphate 3-phosphatidyltransferase	BU251_05935	22	0	0	0	0	4	0
beta-phosphoglucomutase	BU251_06790	24	1	1	0	0	3	0
biopolymer transport protein ExbD/ToIR	BU251_07590	16	4	1	0	0	0	0
formate hydrogenlyase subunit 5	BU251_08625	59	2	2	0	0	3	4
hypothetical protein	BU251_09350	25	1	1	0	0	3	1
hypothetical protein	BU251_09445	18	4	6	0	0	0	0
ABC transporter permease	BU251_00760	43	1	1	3	4	0	0
protein containing Type IV pilus assembly PilZ domain	BU251_00795	15	2	2	0	0	2	3
putative deoxyribonuclease Yjiv	BU251_02265	28	2	2	0	0	2	0

66

phosphate transport system regulatory protein PhoU	BU251_02805	23	1	1	0	0	3	1
heat-inducible transcription repressor HrcA	BU251_02920	29	2	2	0	0	2	3
tRNA pseudouridine 13 synthase	BU251_03670	44	2	2	0	0	2	3
AmmeMemoRadISam system radical SAM enzyme	BU251_05155	43	2	2	0	0	2	3
hypothetical protein	BU251_05360	59	2	2	0	0	2	3
riboflavin biosynthesis protein RibD	BU251_05880	40	3	4	0	0	1	1
damage-inducible protein CinA	BU251_05905	16	1	1	0	0	3	1
large-conductance mechanosensitive channel	BU251_06495	10	3	1	0	0	1	1
phosphoribosyltransferase	BU251_06820	23	2	3	0	0	2	2
hypothetical protein	BU251_07740	28	3	1	0	0	1	1
hypothetical protein	BU251_00025	30	2	2	0	0	2	2
hypothetical protein	BU251_00075	14	2	0	0	0	1	1
hypothetical protein	BU251_00410	41	2	3	0	0	1	1
HesA/MoeB/ThiF family protein	BU251_02275	27	1	1	0	0	2	3
30S ribosomal protein S19	BU251_03110	11	2	0	0	0	1	1
hypothetical protein	BU251_04600	31	2	2	0	0	2	2
hypothetical protein	BU251_05115	110	2	2	0	0	2	2
1-acyl-sn-glycerol-3-phosphate acyltransferase	BU251_05525	22	2	0	0	0	1	1
cation efflux protein	BU251_06300	34	2	3	0	0	1	1
glutamate synthase large subunit	BU251_06815	167	0	0	0	0	3	4
hypothetical protein	BU251_06970	26	3	0	0	0	0	0
hypothetical protein	BU251_07065	51	2	3	0	0	1	1
electron transport complex protein RnfB	BU251_07545	29	2	3	0	0	1	1
hypothetical protein	BU251_07960	16	2	3	1	1	0	0
hypothetical protein	BU251_08000	10	1	1	0	0	2	0
hypothetical protein	BU251_08125	28	1	1	0	0	2	0
hypothetical protein	BU251_09060	7	2	2	0	0	2	2
hypothetical protein	BU251_00040	10	1	1	0	0	2	2
sugar ABC transporter permease	BU251_00200	35	0	0	0	0	3	4

67

tRNA (5-methylaminomethyl-2-thiouridylylate)-methyltransferase	BU251_00310	40	1	1	0	0	2	2
30S ribosomal protein S18	BU251_00650	13	1	1	0	0	2	2
hypothetical protein	BU251_03575	18	3	1	0	0	0	0
thiamine-phosphate diphosphorylase	BU251_03765	21	0	0	0	0	3	1
acyl-phosphate glycerol 3-phosphate acyltransferase	BU251_03880	24	3	4	0	0	0	0
hypothetical protein	BU251_04155	43	1	1	0	0	2	2
Na ⁺ H ⁺ antiporter subunit A	BU251_04550	67	2	2	0	0	1	1
phosphate ABC transporter, permease protein PstA	BU251_06655	31	2	2	0	0	1	1
hypothetical protein	BU251_06975	27	2	2	0	0	1	1
hypothetical protein	BU251_08130	20	2	2	0	0	1	1
hypothetical protein	BU251_08760	55	1	1	0	0	2	2
catabolite repression HPr-like protein Crh	BU251_00875	11	0	0	0	0	2	0
tRNA(Cytosine32)-2-thiocytidine synthetase	BU251_01485	29	2	3	0	0	0	0
[FeFe]-hydrogenase maturation protein HydF	BU251_02050	45	1	1	0	0	1	1
four helix bundle protein	BU251_02730	15	1	1	0	0	1	1
CDP-diacylglycerol-glycerol-3-phosphate 3-phosphatidyltransferase	BU251_03885	22	1	1	0	0	1	1
hypothetical protein	BU251_04075	38	1	1	0	0	1	1
hypothetical protein	BU251_04190	10	0	0	0	0	2	3
hypothetical protein	BU251_04310	48	2	3	0	0	0	0
hypothetical protein	BU251_05580	14	2	0	0	0	0	0
hypothetical protein	BU251_06245	27	2	3	0	0	0	0
hypothetical protein	BU251_06430	18	1	1	0	0	1	1
hypothetical protein	BU251_06455	17	2	0	0	0	0	0
UDP-N-acetylmuramate--L-alanine ligase	BU251_06500	11	1	1	0	0	1	1
hypothetical protein	BU251_06985	17	1	1	0	0	1	1
hypothetical protein	BU251_08330	59	1	1	1	1	0	0
hypothetical protein	BU251_08795	26	2	0	0	0	0	0

68

DNA-directed RNA polymerase subunit omega	BU251_09510	8	1	1	0	0	1	1
type II secretion system protein F	BU251_00150	31	2	2	0	0	0	0
hypothetical protein	BU251_01360	11	2	2	0	0	0	0
RND family efflux transporter MFP subunit	BU251_03420	32	0	0	0	0	2	2
ATP phosphoribosyltransferase regulatory subunit	BU251_03500	7	2	2	0	0	0	0
hypothetical protein	BU251_03715	10	0	0	0	0	2	2
hypothetical protein	BU251_04085	38	0	0	0	0	2	2
hypothetical protein	BU251_04120	42	2	2	0	0	0	0
hypothetical protein	BU251_04135	38	0	0	0	0	2	2
hypothetical protein	BU251_04460	21	0	0	0	0	2	2
NADH-ubiquinone oxidoreductase chain H	BU251_04510	35	0	0	0	0	2	2
hypothetical protein	BU251_05235	7	2	2	0	0	0	0
nicotinamidase	BU251_06795	24	2	2	0	0	0	0
hypothetical protein	BU251_08025	52	0	0	0	0	2	2
endonuclease	BU251_08090	25	0	0	0	0	2	2
hypothetical protein	BU251_08790	114	0	0	0	0	2	2
hypothetical protein	BU251_00435	20	0	0	0	0	1	1
periplasmic divalent cation tolerance protein CutA	BU251_00505	12	0	0	0	0	1	1
anti-sigma-W factor RsiW	BU251_01470	20	0	0	0	0	1	1
xanthine dehydrogenase iron-sulfur subunit	BU251_01520	17	0	0	0	0	1	1
hypothetical protein	BU251_01895	12	1	1	0	0	0	0
transcriptional regulator	BU251_01900	9	0	0	0	0	1	1
hypothetical protein	BU251_02280	11	1	1	0	0	0	0
hypothetical protein	BU251_02390	13	0	0	0	0	1	1
gamma-glutamylcyclotransferase	BU251_02450	17	1	1	0	0	0	0
hypothetical protein	BU251_03345	25	0	0	0	0	1	1
hypothetical protein	BU251_03720	22	0	0	0	0	1	1
hypothetical protein	BU251_04200	13	1	1	0	0	0	0
hypothetical protein	BU251_04410	45	1	1	0	0	0	0

69

hydrogenase-4 component B / Formate	BU251_04545	51	1	1	0	0	0	0
hydrogenlyase subunit 3	BU251_04995	35	0	0	0	0	1	1
tRNA (adenosine(37)-NG)-								
threonylcarbamoyltransferase complex transferase								
subunit TsaD	BU251_05240	22	1	1	0	0	0	0
hypothetical protein	BU251_05240	22	1	1	0	0	0	0
diadenosine tetraphosphatase	BU251_05860	27	0	0	0	0	1	1
hypothetical protein	BU251_05980	10	0	0	0	0	1	1
hypothetical protein	BU251_05985	13	1	1	0	0	0	0
ABC-type transport system involved in resistanceto	BU251_06180	28	0	0	0	0	1	1
organic solvents, permease component								
gltD	BU251_06810	51	0	0	0	0	1	1
hypothetical protein	BU251_07425	17	0	0	0	0	1	1
hypothetical protein	BU251_08100	44	0	0	0	0	1	1
hypothetical protein	BU251_08235	20	1	1	0	0	0	0
formate hydrogenlyase subunit 7	BU251_08620	19	0	0	0	0	1	1
hypothetical protein	BU251_08770	29	0	0	0	0	1	1
K ⁺ /H ⁺ antiporter	BU251_09115	53	0	0	1	1	0	0
hypothetical protein	BU251_09125	14	1	1	0	0	0	0

70

Chapter 3:
**Anaerobic limonene metabolism in a methanogenic enrichment
involves a glycine radical enzyme**

Almud Lonsing¹, Gerrit Alexander Martens¹, Anastasia Resteu¹, Jana Kizina¹, Jens Harder¹

Manuscript in preparation

¹Max Planck Institute for Marine Microbiology, Bremen, Germany.

Corresponding author: Prof. Dr. Jens Harder

Contribution of the Ph.D. candidate in % of the total work:

Experimental concept and design – 40%

Experimental work/acquisition of experimental data – 80%

Data analysis and interpretation – 80%

Preparation of figures and tables – 90%

Drafting of the manuscript – 90%

Keywords: limonene, monoterpene degradation, *Syntrophobacteraceae*, benzylsuccinate synthase, glycyl radical enzyme, methanogenic enrichment culture, *Candidatus* Velamenicoccus archaeovorans, hydrocarbon activation, fumarate addition

Abstract

Limonene, a monocyclic monoterpene, is present in human environments, food, and as a citrus scent in cosmetics and cleaning products. It enters wastewater and is also anoxically transformed and degraded in methanogenic digester towers. In this study, metagenomics, metatranscriptomic, and metaproteomic datasets of a stable limonene-degrading methanogenic enrichment culture were used to characterize the microbial community and identify the degradation pathway of limonene. Thirty-two metagenome-assembled genomes characterized a complex community of bacteria and methanogenic archaea dominated by *Candidatus Velamenicoccus archaeovorus* as the top predator. The presence of several fermenting bacteria suggested a recycling of biomass in a microbial loop. Only one hydrocarbon-activating enzyme system was expressed, a member of the alkyl- and arylsuccinate synthase family, which are glycine radical enzymes adding fumarate to the hydrocarbons. The limonenylsuccinate synthase gene encodes a modified substrate binding pocket with two smaller amino acids, suggesting an adaptation for the larger structure of limonene. A *Syntrophobacteraceae* genome comprises a limonenylsuccinate synthase operon, a ring cleavage operon, and genes for the final mineralization to acetate, hydrogen, and formate. Nearly all genes for this degradation pathway were highly transcribed and expressed, demonstrating a catalytic role for glycine radical enzymes in methanogenic systems degrading limonene.

3.1 Introduction

One per mille of carbon annually fixed by plants is released into the atmosphere as monoterpenes (Messina *et al.*, 2016). In soil, monoterpenes are released from litter and as root exudates and inhibit the oxidation of ammonium. As carbon and energy source for bacteria, they stimulate nitrogen fixation in the rhizosphere (White, 1991; Smolander *et al.*, 2006). Monoterpenes enter the anthropogenic carbon cycle as food, e.g., orange juice, and as fragrances in numerous consumer products. Although a part of it evaporates during anthropogenic processes, a substantial amount of monoterpenes reaches wastewater treatment facilities. The influent may contain up to 20 µg of limonene per liter, and a broad survey in Sweden in 2004 detected 0.15 to 0.87 µg limonene per g dry weight of sludge (Potter *et al.*, 2014). Clarification in wastewater treatment plants uses the biology of aerobic as well as denitrifying and methanogenic microbial communities for the transformation of monoterpenes.

Monoterpenes are isoprene-derived hydrocarbons with carbon-carbon double bonds (alkenes). Hydrophobicity and the absence of functional groups contribute to a resistance to biodegradation. Molecular oxygen is used by many aerobic bacteria to initiate the functionalization of the alkenes (Marmulla and Harder, 2014). In the absence of oxygen, the alkene may be hydroxylated in the allylic position. The nitrate-reducing *Castellaniella defragrans* isomerizes monoterpenes (Harder and Marmulla, 2017; Puentes-Cala *et al.*, 2018a) and then uses a limonene dehydrogenase to obtain with perillyl alcohol a functionalized compound suitable for further transformations (Puentes-Cala *et al.*, 2018b). Most probable number determinations have shown that for every 140 denitrifying cells in activated sludge, one is capable of growing on monoterpenes. With a population density of three million cells per milliliter, this population indicates a significant potential for reducing the concentration of monoterpenes in wastewater treatment ponds (Harder, 2000). Methanogenic communities present in anaerobic digesters also consumed monoterpenes (Harder and Foss, 1999). The study revealed a biological production of *p*-cymene (*p*-isopropylmethylbenzene), which is a major contaminant in biogas. Stable methanogenic enrichment cultures were established on limonene and maintained (Rotaru *et al.*, 2012; Kizina *et al.*, 2022). Diversity analysis based on the 16S rRNA gene showed that methane production from settled activated sludge involved a community of both Bacteria and Archaea. Additionally, the limonene-degrading enrichment culture surprised with the abundant presence of cells affiliating to a so-far-uncultured phylotype OP3 originally named after Obsidian Pool in Yellowstone, USA (Rotaru *et al.*, 2012), and now classified as Omnitrophota (Perez-Molphe-Montoya *et al.*, 2022; Seymour *et al.*, 2023). The

phylotype LiM (for Limonene-based Methanogenesis) was characterized as a predatory ultramicrobacterium that was found both free-living and attached to other cells and was named *Candidatus (Cand.) Velamenicoccus archaeovorans* (Kizina *et al.*, 2022). As a predatory bacterium, it represents a trophic level above the limonene-degrading community of bacteria and archaea. In this study, we aimed at the identification of the degradation pathway of limonene in the stable methanogenic enrichment culture. So far, the anaerobic degradation pathways for alkanes and aromatic compounds have been investigated in methanogenic communities but not for monoterpenes.

Methane, the smallest alkane, diffuses in marine sediments upwards into the sulfate-reducing zone, where it is oxidized by consortia of archaeal anaerobic methane oxidizers and sulfate-reducing bacteria at psychrophilic as well as at thermophilic conditions (Wegener *et al.*, 2022). The methane-activating key enzyme in Archaea is the methyl-coenzyme M reductase (Mcr), acting as a dehydrogenase on methane (Shima *et al.*, 2012). The enzyme is also present in freshwater methane-oxidizing species, i.e., in *Cand. Methanoperedens* that uses nitrate as an electron acceptor (Kurth *et al.*, 2019; Berger *et al.*, 2021; Lemaire and Wagner, 2022).

Alkyl-coenzyme M reductases (ACR) are related to MCRs. They also use the heterodisulfide CoM-S-S-CoB as cosubstrate to oxidize a variety of alkanes to thioethers. Archaea performing these dehydrogenations are *Cand. Argoarchaeum ethanivorans* (ethane), *Cand. Ethanoperedens* and *Cand. Syntrophoarchaeum* (propane, butane) (Wegener *et al.*, 2022). Pentane to tetradecane are substrates for *Cand. Alkanophaga* (Zehnle *et al.*, 2023). Also, hexadecane and other long-chain alkanes are utilized by ACR-containing Archaea, i.e., by *Cand. Methanoliparia* originating from an oil tank active in brackish water (Laso-Pérez *et al.*, 2019). So far, all ACR-utilizing organisms have been isolated and cultured in marine or hypersaline media (Wegener *et al.*, 2022).

An alternative mechanism to the dehydrogenases is the addition of a carbon-hydrogen bond to fumarate by glycine radical enzymes, better known as succinate synthases. Alkyl- and benzylsuccinate synthases (Ass and Bss) have been characterized from denitrifying strains (Heider *et al.*, 2016). The glycine radical is introduced posttranslationally by an activase cleaving S-adenosyl-methionine into 5-deoxyadenosine and methionine (Frey *et al.*, 1994). The large protein A or alpha subunit of succinate synthases is widely used as marker gene for the activation of hydrocarbons. The presence of *ass* or *bss* genes and alkyl- or benzylsuccinates as metabolites together indicates an anaerobic degradation via glycine radical enzymes.

The first methanogenic alkane enrichment was established on penta- and hexadecane (Zengler *et al.*, 1999), and a careful reanalysis of the sequencing datasets showed the presence of *assA* genes in *Smithella*-single cell genomes (Tan *et al.*, 2014). *Smithella propionica* was the first strain in the genus that belongs to *Syntrophaceae*. Partial *assA* genes were also detected in methanogenic enrichment cultures for n-alkanes from five to fifty carbon atoms. Mesophilic cultures from oil sands tailing ponds utilized 2-iso-hexane and -heptane. Partial *assA* sequences and the detection of alkylsuccinates suggested a succinate synthase as the first enzyme of the degradation pathway in a novel *Peptococcaceae* species (GenBank: KFD41815.1) (Abu Laban *et al.*, 2015; Siddique *et al.*, 2020). Production water from a low-temperature oil field developed a methanogenic community within a year on an equimolar mixture of nonane to dodecane (J. H. Ji *et al.*, 2020). The presence of C9 to C12 alkylsuccinates and partial *assA* genes of *Smithella*-related bacteria indicated the enzymatic activation by fumarate addition. Related bacteria were also present with *assA* genes in methanogenic enrichment cultures on tri- and tetradecane and on C16 to C20 n-alkanes, whereby the inocula originated from Xinjiang Oil reservoir production water (J. H. Ji *et al.*, 2020; J.-H. Ji *et al.*, 2020). *Smithella*-related species catalyzed the degradation of long-chain n-paraffins (C28-C50) in a marine methanogenic enrichment from a polychlorinated biphenyls-contaminated sediment in San Diego Bay (Wawrik *et al.*, 2016). In contrast to these mesophilic *Syntrophaceae*, actinobacteria were in a methanogenic culture from production water from an oil reservoir able to degrade even-numbered n-paraffins (C21-C30) under thermophilic conditions (55°C), with *Cand. Syntrophihaticia* encoding a complete *ass* operon and several hydrogenases and formate dehydrogenase (Liu *et al.*, 2020). These studies documented that the anaerobic degradation of alkanes by microorganisms in methanogenic communities can involve thioether-forming dehydrogenases such as MCR and ACR or alkylsuccinate synthases.

Aromatic compounds as substrates for methanogenic cultures were early established (Grbić-Galić and Vogel, 1987), but the enzymes that activate the aromatic compounds were characterized preferentially in denitrifying bacteria. Besides glycine radical enzymes, oxygen-independent hydroxylation and carboxylation were identified as activation reactions. Ethylbenzene dehydrogenase uses a molybdenum cofactor to introduce the hydroxyl group (Hagel *et al.*, 2022). This mechanism differs from the aforementioned limonene dehydrogenase. Carboxylation has been established for phenol, benzene, and naphthalene and was suggested for phenanthrene (Atashgahi *et al.*, 2018; Himmelberg *et al.*, 2018; Tomei *et al.*, 2021; Heker *et al.*, 2023). The enzymes were classified in the superfamily of 3-octaprenyl-4-hydroxybenzoate decarboxylase (UbiD family decarboxylase) and uses as cofactor prenylated

flavin mononucleotide provided by a flavin prenyltransferase (UbiX). Benzene degradation in methanogenic cultures has been established for decades, still, the activating enzymes in the deltaproteobacterial candidate Sva0485 have not been identified yet (Toth *et al.*, 2021). A clostridial putative naphthalene carboxylase operon was also retrieved from a one-year-old methanogenic enrichment culture (Toth *et al.*, 2018). In a comparative study, genes encoding alkyl- and benzylsuccinate synthases (*bssA*, *assA*, *nmsA* – naphthylmethylsuccinate synthase) were detected in enrichment cultures on short-chain alkanes, toluene, and naphthalene (Tan *et al.*, 2015). Genes for carboxylation (*abcA*, *ancA*) were only retrieved from the short-chain alkane and naphthalene enrichments. The downstream pathways for the fumarate adducts were present in the metagenomes, comprising the beta-oxidation of benzylsuccinate (*bbs*), the benzyl-CoA reductase (*bcr*) and the benzoate anaerobic metabolism (*bam*, class II Bcr).

This overview of the degradation of alkanes and aromatic compounds by methanogenic communities suggests several candidate enzymes for methanogenic limonene degradation. In this study, we have analyzed metagenomes for the presence of these genes. Although candidate genes were identified, metatranscriptomic and metaproteomic analyses argued that limonene degradation is initiated by a limonenylsuccinate synthase in a *Syntrophobacteraceae* bacterium and results in acetate, hydrogen, and formate.

3.2 Material and Methods

Cultivation

The methanogenic enrichment culture on limonene (1-methyl-4-(1-methylethenyl)-cyclohex-1-ene) was initiated in 1997 and, one microliter of the enrichment culture was the basis for the current microbial diversity in 12 lineages in culture with one transfer per year since 2005 (Harder and Foss, 1999; Rotaru *et al.*, 2012; Kizina *et al.*, 2022). The cultures contained 300 mL freshwater methanogenic media, including 2 mM acetate and 1 mM cysteine, 30 mL 2,2,4,6,8,8-heptamethylnonane (HMN), and 1.5 mL of *R*-(+)-limonene in 500 mL borosilicate bottles, were inoculated with 10% vol/vol and were incubated at 28°C with a rotary shaking of 60 rpm. Culture MM-324 (inoculated March 15, 2014) served as inoculum for five parallel cultures that grew from January 31, 2015, until mid-October 2015. Gas formation until the end of September 2015 accumulated to 455 till 665 ml per culture and the biomass was harvested in October/November 2015 (Kizina, 2017).

DNA and RNA extractions and sequencing

Differential centrifugation was performed to separate the cells in the culture by size. A SW28 ultracentrifuge rotor (Beckman, Palo Alto, CA) settled 10kS cells at 7,600 rpm (7,643x g) for 20 min (10,000 S). The pellet was resuspended in 1 ml 10 mM Tris, 1 mM EDTA, pH 8.0 (TE). Centrifugation was continued with the supernatant at 27,000 rpm (96,467x g) for 160 min (100S). The resulting pellet was resuspended in 0.5 ml TE and centrifuged at 12,400 rpm (16,331x g) for 3 min (10,000 S). The resulting pellet of 100S aggregates contained, besides cell clumps, few single large cells, whereas the 100S cells in the supernatant were dominated by small, free-living cells. RNA and DNA were extracted using the RNA PowerSoil® Total RNA Isolation Kit and the RNA PowerSoil® DNA Elution Accessory Kit (Qiagen, Hilden, Germany). NGS sequencing was performed by the Max Planck-Genome-centre Cologne, Germany (<https://mpgc.mpipz.mpg.de/home/>). RNA was analyzed using Illumina HiSeq (150 bp reads) (Illumina, San Diego, CA, SRA accession SRR30230645 Bioproject PRJNA1129558). Illumina HiSeq was also applied for DNA (2 x 250 bp reads, SRA accessions SRR29791532, SRR29888306, and SRR29898063, Bioproject PRJNA1129558). Large DNA reads representing the 10kS cells were obtained with HiFi reads on a PacBio Sequel (Pacific Biosciences, Menlo Park, CA, SRA accession SRR29679608, Bioproject PRJNA1129558).

Bioinformatic analyses

HiFi reads were assembled using the Flye de novo assembler v2.9 (Kolmogorov *et al.*, 2020). The contigs were error-corrected using Inspector (Chen *et al.*, 2021), resulting in 1041 contigs with an N50 of 597098 bp. Binning of these contigs used MaxBin v2.2.7 (Wu *et al.*, 2015), Metabat2 v2.15 (Kang *et al.*, 2015), and Vamb v3.0.2 (Nissen *et al.*, 2021). DAS Tool v1.1.3 (Sieber *et al.*, 2018) calculated a set of 32 optimal non-redundant bins. Bin quality was estimated using CheckM v1.1.3 (Parks *et al.*, 2015). The bin LiM6 assigned to *Syntrophobacteraceae* had a strain heterogeneity of 42.89, an indication of the presence of closely related strains.

Bin quality was improved by mapping Illumina reads of 10kS cells onto the contig LiM6 with 99% identity using bbmap v38.93 (<https://sourceforge.net/projects/bbmap/>) and reassembly of mapped reads together with LiM6 contigs that were set as “-trusted-contigs” using the assembler SPAdes v3.15.3 (Bankevich *et al.*, 2012). Contigs larger than 1500 bp were used as a starting point for an additional three rounds of assembly improvement. The refinement reduced contaminations and removed the strain heterogeneity in bin LiM6. Bins were taxonomically classified using GTDB-Tk v1.7.0 (Chaumeil *et al.*, 2019) based on the Genome

Database Taxonomy. The bin representing *Candidatus Velamenicoccus archaeovor* had eight contigs and covered 98.46% of the closed genomes (GenBank CP019384, see supporting table 1 sheet “genome comparison”) using D-GENIES (Cabanettes and Klopp, 2018). We used the closed genome for read mapping. The relative abundance of each bin within the community was calculated using CoverM v0.6.1 (Woodcroft, 2021). Proteins encoded in the 10kS cell metagenome were annotated using Bakta v1.4.2 with its database v3.1 for bacterial genomes (Schwengers *et al.*, 2021) and Prokka v1.14.6 for the methanogenic archaeal MAGs (Seemann, 2014). Genes of interest were also annotated using the NCBI Conserved Domain Database (Marchler-Bauer *et al.*, 2011). Genome annotations of *Syntrophobacteraceae* LiM6 and *Syntrophobacter fumaroxidans* (Genbank GCA_000014965.1) were analyzed with DIAMOND in sensitive mode (Buchfink *et al.*, 2021).

Metatranscriptomic analyses started with SortMeRNA v4.3.4 (Kopylova *et al.*, 2012) to remove rRNA encoding reads. The non-rRNA reads were trimmed with quality 30 using bbdduk within bbmap v38.93 and then mapped with Bowtie2 v2.4.4 (Langmead and Salzberg, 2012) onto the genes of proteins present in the 10kS cell metagenome. Transcripts per million (TPM) were calculated in R using the total mapped reads to a gene.

The analyses to detect enzymes known to be involved in limonene degradation were conducted using DIAMOND v2.0.15.153 BLASTP analyses (Buchfink *et al.*, 2021) for limonene hydroxylase (AAC25032.1), the limonene dehydrogenase subunit A (CDM25290.1), the limonene-1,2-epoxide hydrolase (CAA77012.1), the limonene 1,2-monooxygenase (CAC20855.1), the methyl-coenzyme M reductase type I (WP_013296337.1, WP_013296341.1, WP_013296338.1) (PDB ID: 5A0Y_1, 5A0Y_2, 5A0Y_3), the methanogenic methyl-coenzyme M reductase (WP_084174107.1, WP_042686194.1, WP_042686201.1) (PDB ID: 7NKG), the methyl-coenzyme M reductase type II (WP_013296302.1, WP_013296305.1, WP_013296303.1) (PDB ID: 5A8R_1, 5A8R_2, 5A8R_3), the alkyl-coenzyme M reductases (RZB32666.1), the phenylphosphate carboxylase (WP_245880909.1), the benzene carboxylase (ADJ94002.1) and the ethylbenzene hydroxylase (WP_011237152.1). Enzymes with an E-value below E-40 and a bitscore above 100 were considered similar. These enzymes were examined in more detail with respect to their TPM, presence in the metaproteome, and their annotations were additionally checked against the Conserved Domain Database (CDD) (Marchler-Bauer *et al.*, 2011) and in NCBI Blast and ExpASY searches (Artimo *et al.*, 2012; Sayers *et al.*, 2022). To phylogenetically place the highly expressed succinate synthase alpha subunit in relation to other characterized alkyl- and arylsuccinate synthases, the amino acid sequence was aligned using the ClustalW algorithm

within the MEGA11 software (version 11.0.13) with a gap opening penalty of 10 and a gap extension penalty of 0.1 (Thompson *et al.*, 1994; Tamura *et al.*, 2021). The alignment was used to calculate a maximum-likelihood tree with a bootstrap value of 1000 using FastTree (version 2.1.11) (Price *et al.*, 2009). The amino acid sequences of the proteins were retrieved from GenBank depositions (Selesi *et al.*, 2010; Wöhlbrand *et al.*, 2013; Strijkstra *et al.*, 2014; Heider *et al.*, 2016).

For the metaproteomic analysis, the proteins of 10kS cells, 100S aggregates, and 100S cells were extracted and analyzed by MS/MS in an LTQ Orbitrap Velos mass spectrometer (Thermo Fisher Scientific, Waltham, Mass.) (Kizina, 2017). The mass spectrometry proteomics data were retrieved from ProteomeCentral with the identifier PXD025008 (<https://doi.org/10.6019/PXD025008>). The mass spectrometric data were analyzed with MaxQuant 1.6.16 (Sinitcyn *et al.*, 2018) using genes present in the 10kS cell metagenome and laboratory contaminations and quantified in relative intensity-based absolute quantification (riBAQ) values (Krey *et al.*, 2014).

3.3 Results

3.3.1 Cell fractionation and – omic datasets

The methanogenic community was harvested differentially by cell size. Most cells and cell aggregates were collected in the 10,000S pellet (10kS cells), very small cells were collected in 100S aggregate and 100S cell fractions (Kizina, 2017). Analysis of 10kS cells by an assembly of PacBio Hifi reads resulted in 32 metagenome-assembled genomes (MAGs) representing a coverage of 98.73% (Table). 4 MAGs were closed genomes, 12 MAGs had 2 to 9 contigs, and further 10 MAGs had 11 to 21 contigs. 27 MAGs were over 80% complete, according to CheckM, and only 6 MAGs showed a strain heterogeneity. 30 MAGs had a completeness of at least 66%, and 29 MAGs had a quality above 50 (Supplementary Table 2). The 32 MAGs were also representative of Illumina-based metagenomes of 100S aggregates and 100S cells, covering at least 97.74% of all reads (Supplementary Table 1). The activity of the community was analyzed by metatranscriptomes and metaproteomes. In the large 10kS cell metatranscriptome, 716 of 123474 coding sequences (cds) present in the assembly had a TPM value of ≥ 50 and 10725 cds a TPM value of ≥ 5 . The metaproteomes detected 5634 cds.

3.3.2 Community members of the limonene-degrading culture

The PacBio metagenome contained 55 individual 16S rRNA gene sequences, of which 50 were assigned to 31 MAGs. Only the incomplete MAG LiM31, a *Bacteroidales*, had no 16S rRNA gene. The unassigned 16S rRNA genes affiliated to *Synergistota* (three sequences), *Sphaerochaetaceae*, and *Methanobacteriota*.

Candidatus Velamenicoccus archaeovor

Cand. Velamenicoccus archaeovor was the member of the microbial community with the highest abundance in the metagenome, with 68.70% of the PacBio reads mapping onto the genome. In the 10kS-cell fraction, 63% of the Illumina reads mapped to the predator genome. A higher abundance was observed for the small cell metagenomes: 83% of the reads from the 100S-aggregates metagenome and 89% of reads of the 100S cell metagenome mapped to the

Cand. Velamenicoccus archaeovor genome. This increase in abundance of the ultramicrobial, predatory *Cand. Velamenicoccus archaeovor* coincided with a depletion of large cells and cell aggregates and an increase of small cells in the 100S cell fractions, according to TEM graphs (Kizina, 2017). The metatranscriptomic abundance of *Cand. Velamenicoccus archaeovor* was lower. 20.0%, 31.5%, and 11.6% of the RNA reads of 10kS cells, 100S aggregates and 100S cells, respectively, were mapped onto the *Cand. Velamenicoccus archaeovor* genome. However, one-third of these values were recognized as contribution of the intron-containing 23S rRNA gene, which had not been removed by the digital filtering of rRNA reads using SortMeRNA, likely due to the presence of an intron within the gene. Overall, the ratio of RNA to DNA in the cell fractions reflected the small cell size of *Cand. Velamenicoccus archaeovor*. A decrease in transcripts in the 100S cells coincided with low numbers of ribosomes in free-living cells in comparison to epibiontic cells, according to CARD-FISH images (Kizina, 2017). The metaproteome identified 1280 proteins out of 1884 annotated coding sequences of *Cand. Velamenicoccus archaeovor* representing 48.7, 15.2, and 51.5% of the protein quantity (riBAQ) in 10ks cells, 100S aggregates, and 100S cells, respectively.

Methanogenic archaea

Five MAGs were identified as archaeal *Methanobacteriota*. Besides a *Methanotherix soehngeni* MAG (LiM3), two MAGs matched species according to GTDB taxonomy: *Methanoculleus* sp001896715 (LiM2) and *Methanospirillum* sp012729995 (LiM14). Novel species candidates were *Methanoregula* LiM1 and *Methanofastidius* LiM21, which presented below 0.5% of the -omic quantities. Proteins of *Methanoculleus* presented 9.7% of detected proteins (359 proteins) in the metaproteome, whereas *Methanoregula* and *Methanotherix* hold 1.0 and 1.3 % riBAQ (306 and 851 proteins), respectively. DNA and RNA contributions were 5.2, 8.2 and 3.7 % and 11.1, 4.1 and 8.1 % for *Methanoculleus*, *Methanoregula* and *Methanotherix*, respectively. This suggested that *Methanoculleus* may perform the majority of hydrogenotrophic and formate-based methanogenesis, whereas filamentous *Methanotherix* acted as the acetoclastic methanogen.

Bacteria

The phylum *Desulfobacterota* was present with seven MAGs (Table1). Based on protein quantity, *Aminidesulfobivrio aminophilus* LiM27 was the second most abundant bacterium, with 14.5% riBAQ originating from 174 proteins. It surprised with a low mRNA content of

0.13%. A second *Aminidesulfovibro* sp., LM12, was detected, with 99 proteins and with 0.9% riBAQ (Table1). *Syntrophales* were represented by *Smithellaceae* gen. nov. LM17 and by LM5 and LM10, which affiliate with two novel families. LM5 and LM17 showed with 2.1 and 1.4 % riBAQ a significant biomass contribution. LiM6 was affiliated as a novel genus to *Syntrophobacteraceae*. It comprised 9.3% riBAQ accumulated of 554 proteins and 9.8% TPM (Table1). The abundance of their proteins suggested *Aminidesulfovibro aminophilus* LM27 and *Syntrophobacteraceae* LiM6 as organisms that may catalyze the limonene degradation.

Bacteroidota accounted for 7.5 % of the total riBAQ. *Lentimicrobium* LiM4 contributed 3.7 % riBAQ and 7.74 % TPM, and *Mangrovibacterium* LiM18 and LiM22 made up 2.4% riBAQ, also with a higher TPM contribution of 11.8%. *Meliobacteraceae* LiM26 of the order *Ignavibacteria* hold 1.1 % riBAQ (Table1).

Individual MAGs of other phyla contributed little to the metaproteome. *Anaerolineaceae* was present with four species, LiM 8, 9, 11, and 29. *Spirochaetota* was present with three MAGs, and a range of phyla were only present with one MAG: *Acidobacteriota*, *Actinobacteriota*, *Eremiobacterota*, *Planctomycetota*, *Synergistota*, and *Thermotogota*. The order *Thermotogales* was represented by *Mesotoga* sp002305955 (LiM7), comprising 3.84% of the TPM and 0.072% riBAQ of the metaproteome with 123 detected CDS. Abundant *Anaerolineales* were *Anaerolineaceae* UBA4781 (LiM8) comprising 1.47% TPM and 33 detected CDS representing 0.052% riBAQ and *Anaerolineaceae* sp002436085 (LiM11) showing 0.74% TPM and 34 detected CDS accounting for 0.024% riBAQ. Altogether, the *Mesotoga* sp002305955 (LiM7) and *Anaerolinaceae* MAGs comprised relatively low abundances in the metatranscriptome as well as the metaproteome (Table1).

Chapter 3 - Anaerobic limonene metabolism involves a glycine radical enzyme

Table 1: Overview of MAGs showing their assigned taxonomy (GTDB-tk), their relative abundance in the PacBio metagenome, their percentage of TPM assigned to each MAG, their percent of detected proteins out of all detected proteins, and the percentage of riBAQ for each MAG in the metaproteome

name	taxonomy	PacBio relative abundance [%]	% TPM	% of detected CDS	% riBAQ
LiM	s__Vampirococcus archaeovorus	68.7	19.95	22.72	48.764
LiM1	g__Methanoregula	8.15	4.08	5.43	1.011
LiM2	s__Methanoculleus	5.21	11.07	6.37	9.713
LiM3	s__Methanotherix soehngeni	3.68	8.13	15.1	1.330
LiM4	s__Lentimicrobium	2.02	7.74	4.51	3.709
LiM5	o__Syntrophales	1.39	3.4	3.75	2.075
LiM6	f__Syntrophobacteraceae	1.25	9.81	9.83	9.292
LiM7	g__Mesotoga	1.04	3.84	2.18	0.083
LiM8	f__Anaerolineaceae	0.94	1.47	0.59	0.054
LiM9	f__Anaerolineaceae	0.83	0.31	0.59	0.177
LiM10	o__Syntrophales	0.57	1.43	0.89	0.223
LiM11	f__Anaerolineaceae	0.56	0.74	0.6	0.020
LiM12	s__Aminidesulfovibrio	0.55	1.08	1.76	0.878
LiM13	o__Treponematales	0.53	0.97	0.83	0.257
LiM14	s__Methanospirillum	0.35	0.31	1.38	0.221
LiM15	c__Coriobacteriia	0.31	0.44	0.12	0.003
LiM16	f__Aminobacteriaceae	0.31	1.26	2.63	0.104
LiM17	f__Smithellaceae	0.28	0.5	1.3	1.396
LiM18	g__Mangrovibacterium	0.22	5.38	1.85	1.457
LiM19	c__Xenobia	0.22	0.28	0.6	0.166
LiM20	p__Spirochaetota	0.18	0.66	0.94	0.686
LiM21	g__Methanofastidiosum	0.18	0.08	0.05	0.000
LiM22	g__Mangrovibacterium	0.16	6.36	2.2	0.934
LiM23	p__Desulfobacterota	0.16	0.02	0.46	0.040
LiM24	f__Thermoanaerobaculaceae	0.15	1.56	0.92	0.031
LiM25	c__Ignavibacteria	0.15	0.05	0.8	0.344
LiM26	f__Melioribacteraceae	0.12	0.27	2.15	1.128
LiM27	s__Aminidesulfovibrio aminophilus	0.12	0.13	3.09	14.460
LiM28	f__Thermoguttaceae	0.12	0.17	1.03	0.158
LiM29	f__Anaerolineaceae	0.11	0.37	0.14	0.000
LiM30	s__Sphaerochaeta	0.1	0.11	0.32	0.025
LiM31	o__Bacteroidales	0.09	0.11	0.12	0.001
unmapped	N/A	1.27	7.91	4.74	1.262

3.3.3 Evidence for hydrocarbon metabolism in the - omic datasets

Candidate genes for hydrocarbon-activating enzymes were identified using characterized enzymes as references in protein similarity searches with DIAMOND in BLASTp mode. The superfamily that includes benzene and phenol carboxylases was present with three genes for quinone biosynthesis annotated as 3-octaprenyl-4-hydroxybenzoate decarboxylase in the MAGs LiM13, LiM6, and LiM20. The expression level of these enzymes was below 10 TPM, and the proteins were not detected in the metaproteome. The same observation was made for molybdopterin-binding proteins of unknown function that had the highest, but overall low similarity to the ethylbenzene hydroxylase. One protein related to limonene hydroxylase was detected in the metaproteome but annotated as a transcriptional regulator. Three genes were related to limonene dehydrogenase, but not transcribed and expressed. Genes for the aerobic oxidation with molecular oxygen, limonene-1,2-monooxygenase, and limonene-1,2-epoxide hydrolase were not present in the metagenome.

Alkyl-coenzyme M dehydrogenase-related proteins were annotated as six methyl-coenzyme M reductases present in *Methanotherx soehngeni* LiM 3, *Methanofastidiosum* sp. LiM21, *Methanospirillum* sp. LiM14, *Methanoregula* sp. LiM1, and twice in *Methanoculleus* LiM2. All large MCR alpha subunits were detected in the metaproteomes. Coding sequences related to benzylsuccinate synthases were also frequently detected. However, annotation varied from pyruvate formate lyase superfamily to benzylsuccinate synthase. Only one of five gene variants of the succinate synthase was highly transcribed and, with a value of 174 TPM, it was among the top 1% of proteins in the metatranscriptome.

The protein annotated as the alpha subunit of the benzylsuccinate synthase was the 1398 most abundant protein in the metaproteome of the 10kS fraction. The benzylsuccinate synthase alpha subunit gene was located on a contig assigned to MAG LiM6 belonging to the family *Syntrophobacteraceae*. This finding, together with the absence of alternative candidate enzymes for limonene activation, initiated a detailed genome analysis of MAG LiM6 to identify the limonene degradation pathway and metabolism in this bacterium.

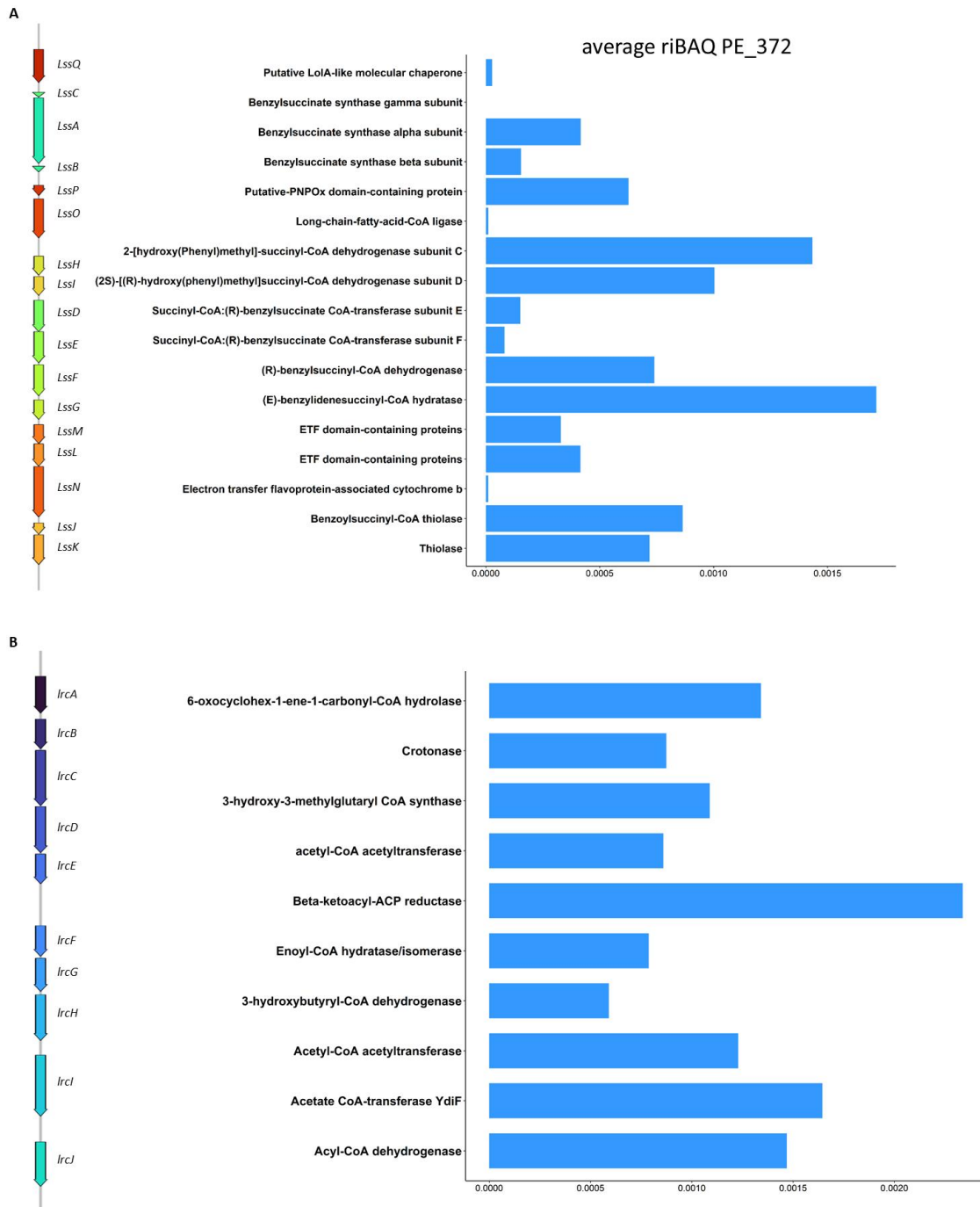
3.3.4 Limonene degradation in *Syntrophobacteraceae* LiM6

Figure 1: Two operons from LiM6 (*Syntrophobacteraceae*) involved in limonene degradation. The first operon, *lssQCABPOHIDEFGMLNJK* (A), includes limonenylsuccinate synthase, and the second operon, *ircABCDEFGHJIJ* (B), encodes 4-isopropenyl-2-oxocyclohexane-1-carbonyl-CoA hydrolase, which is involved in the ring opening. The bar charts display average riBAQ values from 10kS, 100S aggregates and 100 S fractions (PE_372), with protein names from Bakta annotation and CDD annotation.

Hydrocarbon activation by glycy radical enzymes needs accessory enzymes to introduce the glycine radical by S-adenosylmethionine (SAM) cleavage. The genome of LiM6 encodes an expressed chaperon BssE (LMHFDP_382590) for the activation of a glycy radical enzyme by s-adenosylmethionine and several radical SAM proteins, of which transcripts were detected, but none were detected in the metaproteome. The *bssE* gene is located outside of the limonenylsuccinate synthase (*lss*) operon *lssQCABPOHIDEFGMLNJK* (LMHFDP_300800-300880) (Figure 1 A, Supplementary Table 3 and 4).

Transcriptomic and proteomic abundances of proteins suggest the following pathway: limonene degradation is initiated by limonen-7-ylsuccinate synthase (p-mentha-1,8-dien-7-yl-succinate synthase) LssABC (Figure 2). The enzyme adds limonene to fumarate. In the next step, the succinyl-CoA:limonen-7-ylsuccinate CoA-transferase LssDE activates a carboxylic acid and the substrate molecule is enlarged by the coenzyme A moiety providing a handle for highly specific substrate binding. A first beta-oxidation follows, including dehydrogenation, hydration, oxidation and thiolitic cleavage. The limonene-7-yl-succinyl-CoA dehydrogenase LssF introduces the double bond in the alpha-beta position. Electron transfer flavoproteins LssLM are likely the physiological electron acceptors. The limonene-7-ylidene-succinyl-CoA hydratase LssG adds water to the enoyl-CoA ester. The alcohol is oxidized by the NAD⁺-reducing 2-(7-hydroxylimonen-7-yl)-succinyl-CoA dehydrogenase LssHI. The last reaction in the beta oxidation is a thiolitic cleavage of the beta-oxo-acylCoA ester by 2-(7-oxolimonen-7-yl)-succinyl-CoA thiolase LssJK (Figure 2). One product is succinyl-CoA, which is reused by LssDE. The second product is perillyl-CoA, which is the educt for the enzymes encoded in the limonene ring cleavage (*lrc*) operon (Figure 1 B). The *lss* operon encodes also some proteins with a presumably accessory function. The electron transfer flavoprotein-associated cytochrome b and iron-sulfur cluster-binding oxidoreductase LssN may be involved in electron transfers, eventually bifurcation. The acyl-CoA synthetase LssO may replenish intermediate pools of CoA esters. Two expressed proteins lack a clear function in this initial limonene degradation pathway: LssP is a protein with putative pyridoxamine 5'-phosphate oxidase function and LssQ is a putative LolA-like molecular chaperone for lipoproteins or other hydrophobic compounds. Of these proteins, only the small gamma subunit LssC was not detected in the metaproteome (Figure 1 A), but a value of 183 TPM in the metatranscriptome indicated relevance for the enzyme (Supplementary Table 3 and 4). Although located on a different contig, the expression of the membrane proteins HflC and HflK suggested that the activation reaction may be anchored as multiprotein complex on the membrane.

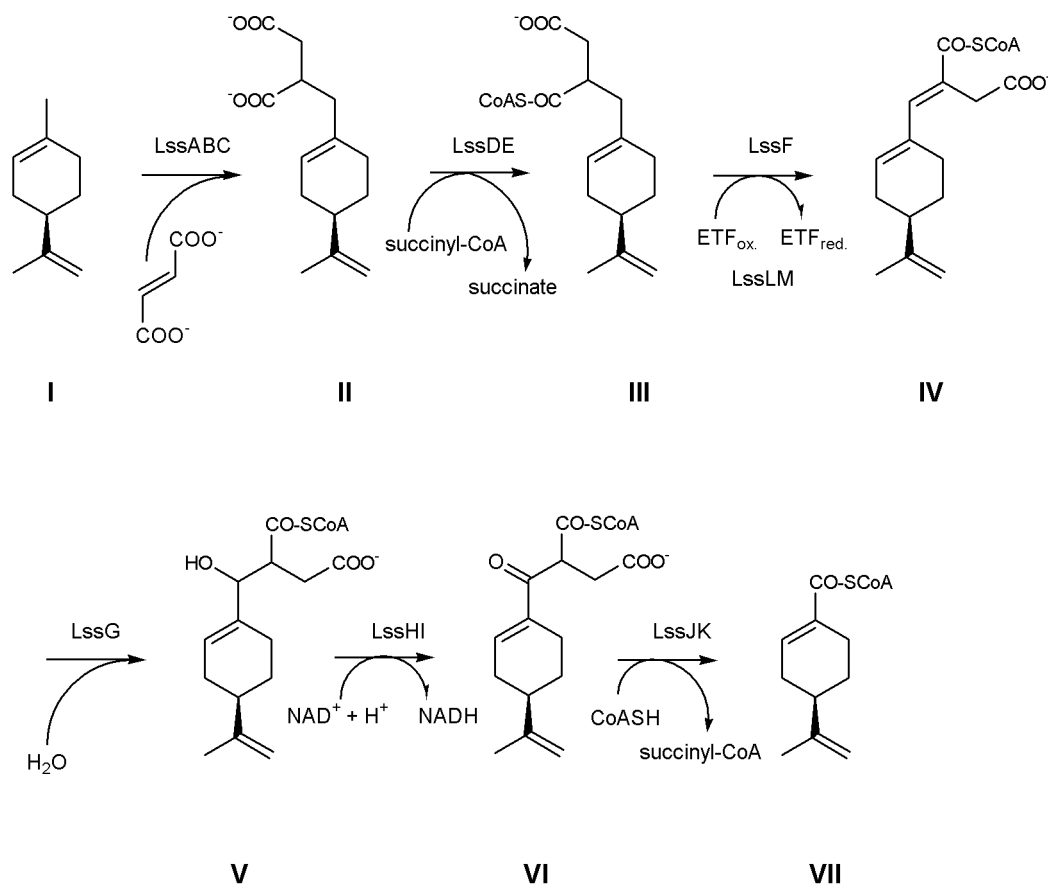


Figure 2: Limonene transformation to perillyl-CoA. I: limonene, II: limonen-7-ylsuccinate, III: limonen-7-ylsuccinate, IV: limonen-7-ylidene-succinyl-CoA, V: 2-(7-hydroxylimonen-7-yl)-succinyl-CoA, VI: 2-(7-oxolimonen-7-yl)-succinyl-CoA, VII: perillyl-CoA.

The limonene ring cleavage (*lrc*, LMHFDP_382595-382510) operon encodes BssE and enzymes for the cyclohexene ring opening and beta oxidations (Figure 1 B). Perillyl-CoA hydration may occur by one of the enoyl-CoA hydratases LrcB or LrcF (Figure 3). Physiological electron acceptors may be a second pair of electron transfer flavoproteins EtfAB (LMHFDP_246400 and 246395), encoded outside of the *lrc* operon. The oxidation of the alcohol 4-isoprenyl-2-hydroxy-cyclohexane-1-carbonyl-CoA may be catalyzed by a dehydrogenase: LrcE, LrcG, or LrcK. The 4-isopropenyl-2-oxo-cyclohexane-1-carbonyl-CoA hydrolase LrcA catalyzes the ring opening to 5-isopropenyl-pimeloyl-CoA. A beta oxidation with an acyl-CoA dehydrogenase LrcJ, an enoyl-CoA hydratase (LrcB or LrcF), a dehydrogenase (LrcE, LrcG or LrcK) and a thiolase (LrcD or LrcH) results in 3-isopropenyl-glutaryl-CoA (Figure 3). 3-modified glutaryl-CoA can be functionalized by acyl-CoA

dehydrogenase and enoyl-CoA hydratase to a tertiary alcohol, in this case, 3-hydroxy-3-isopropenyl-glutaryl-CoA. Besides the aforementioned acyl-CoA dehydrogenases, an annotated glutaryl-CoA dehydrogenase and the adjacent encoded hypothetical protein were detected in the metaproteome and are candidate enzymes for this metabolic step (LMHFDP_341120 and 341110). LrcC, an enzyme of the 3-hydroxy-3-methyl-glutarylCoA synthase superfamily, cleaves at the tertiary carbon of 3-hydroxy-3-isopropenyl-glutaryl-CoA into acetyl-CoA and 4-methyl-3-oxo-pent-4-enoyl-CoA, the endproduct of this operon (Figure 3). Expressed proteins include also an acyl-CoA:acetate CoA transferase LrcI and two proteins with domains of unknown functions LrcL (DUF4125) and LrcM (DUF4037). Adjacent operons include an expressed beta-ketoacyl-acyl carrier protein reductase (LMHFDP_382495) and a tripartite transmembrane efflux transporter MepABC (monoterpene-associated efflux proteins).

presence of the methylmalonyl-CoA mutase small and large subunits (LMHFDP_355185 and 355180). The importance of the B₁₂-enzyme is highlighted by the expression of a cobalt ABC transporter (LMHFDP_206225-206235), a B₁₂-binding protein (LMHFDP_365805), and a corrinoid adenosyltransferase (LMHFDP_205380). Succinyl-CoA hydrolysis can be linked to energy conservation and a succinate dehydrogenase replenishes fumarate, the cosubstrate of the first reaction (LMHFDP_10730-10740). After hydration, an oxaloacetate-decarboxylating malate dehydrogenase produces pyruvate (LMHFDP_533100). Pyruvate oxidation can be catalyzed by a 2-oxoacid:ferredoxin reductase (LMHFDP_626730 + LMHFDP_626735 or LMHFDP_365485 + LMHFDP_365490), providing reducing equivalents at a very negative electrochemical potential for proton or carbonate reduction. Acetyl-CoA transformation to acetate enables energy conservation, i.e. by an acetyl-CoA synthetase (LMHFDP_60925).

Syntrophobacteraceae are well known for multiple hydrogenases (Hase) and formate (FDH) dehydrogenases, i.e. *S. fumaroxidans* has 8 Hases and 6 FDHs (Mollaei *et al.*, 2021). Strain LiM expresses a locus (LMHFDP_383075-383110) encoding two FDH alpha subunits, one FDH beta subunit, a NiFe Hase alpha subunit, a NiFe Hase 20 kDa subunit (NADH acceptor oxidoreductase), a ferredoxin reductase and a ferredoxin. A second loci (LMHFDP_382905-382950) codes for expressed two FDH subunits, a NADH:acceptor oxidoreductase and Hase maturation factors. HypB, a nickel incorporation factor for Hases, was expressed (LMHFDP_365535). Another NiFe Hase loci (LMHFDP_206220-206205) is located beside the aforementioned cobalt ABC transporter. A group of redox-active proteins (LMHFDP_246610-246625) was also expressed and annotations involved NAD-reducing Hase subunit HndC. These proteins are involved in the electron transfers from ferredoxin, NADH, and ETFs and likely involve bifurcation and reverse electron transport. Expressed energy conservation proteins include pyrophosphate-energized proton pump (LMHFDP_560695) and ATP synthase.

The proteome included also enzymes of the C₃-part of glycolysis and additional acyl-CoA dehydrogenases, enoyl-CoA hydratases, dehydrogenases and thiolases. Biosynthetic enzymes were 3-isopropyl-malate dehydratase and acetolactate synthase.

3.3.5 Phylogenetic analysis of the benzylsuccinate synthase alpha subunit

The phylogenetic analysis placed the highly expressed benzylsuccinate synthase alpha subunit together with desulfobacterial enzymes from *Desulfosarcina widdellii* and *Desulfosarcina ovata* in a new clade within the branch of naphthyl/benzyl-succinate synthases and separate from alkylsuccinate synthases (Figure 4). The new clade is more related to the toluene clade than to the *p*-cresol and the naphthalene clades and comprises proteins from *Deltaproteobacteria*. This is not exclusive, *Deltaproteobacteria* encode also alkyl and benzylsuccinate synthases in other clades.

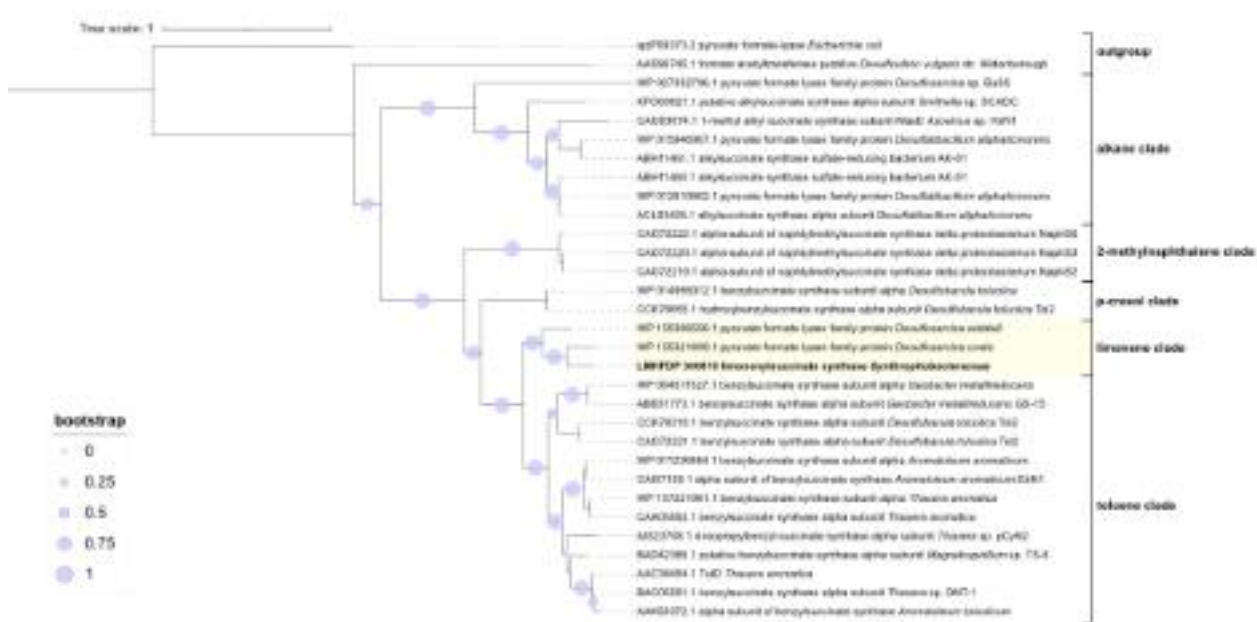


Figure 4 Maximum likelihood phylogenetic tree displaying the relationship of the limonenylysuccinate synthase alpha subunit with other glycine radical enzymes. The tree was generated using a ClustalW protein alignment. The scale bar indicates the number of substitutions per site. Sequences of glycine radical enzymes were obtained from literature (Grundmann *et al.*, 2008; Selesi *et al.*, 2010; Wöhlbrand *et al.*, 2013; Strijkstra *et al.*, 2014; Heider *et al.*, 2016).

3.4 Discussion

An interdependent community of syntrophic bacteria and methanogenic archaea usually performs hydrocarbon degradation in methanogenic habitats. Syntrophic bacteria ferment hydrocarbons into formate, hydrogen, acetate, and carbon dioxide. These catabolic reactions are thermodynamically unfavorable, only the constant consumption of the products by methanogens enables the continuous activity of syntrophic bacteria (Dolfing *et al.*, 2007). The presence of the acetoclastic *Methanothrix* - also validly published as *Methanosaeta* under the International Code of Nomenclature of Prokaryotes - in the limonene-degrading culture was previously established by *in situ* hybridization (Kizina, 2017) and is expected because the genus is known as the oligotrophic acetoclastic methanogen with the highest affinity for acetate (Jetten *et al.*, 1992). Hydrogen, carbon dioxide, and formate are substrates for a range of methanogenic genera, including *Methanoculleus*, *Methanoregula*, and *Methanospirillum* (Demirel and Scherer, 2008; Thauer *et al.*, 2008; Bräuer *et al.*, 2015; Joshi *et al.*, 2018). The latter are also known to oxidize alcohols (Widdel *et al.*, 1988; Kurth *et al.*, 2020). Methylthiol reduction in combination with hydrogen oxidation seems to be the restricted metabolism of *Methanofastidiosum* (Nobu *et al.*, 2016). Its presence in the community is an indication of an active sulfur metabolism in the enrichment culture. The archaeal community in the limonene-degrading enrichment culture is overall similar to other methanogenic communities in mesophilic hydrocarbon enrichments (Embree *et al.*, 2015; Tan *et al.*, 2015; Fowler *et al.*, 2016). In contrast, the bacterial community was larger than expected, with 27 established metagenome-derived genomes. A hydrocarbon-transforming syntroph was expected, accompanied by a few taxa that may provide essential amino acids or equilibrate concentrations of hydrogen and carbon dioxide to formate to the same thermodynamic energy level. The enrichment culture on limonene has an additional trophic level represented by the predating *Cand. Velamenicoccus archaeovor*. This trophic level reduces the biomass of the first trophic layer of syntrophs and methanogens. The huge abundance in - omic datasets, together with earlier electron microscopic observations (Kizina, 2017; Kizina *et al.*, 2022), suggest that *Cand. Velamenicoccus archaeovor* did not attack and digest itself. The predation process is most likely accompanied by a release of prey cytosolic compounds. Also, membrane remains of dead cells were observed in electron micrographs (Kizina, 2017; Kizina *et al.*, 2022). This organic matter can support a catabolic mineralization in a microbial loop in which fermenters, syntrophs, and methanogens use the organic matter to form new biomass (Azam *et al.*, 1983).

This process contributes to the microbial diversity in the enrichment culture. Fermentation may be performed by *Lentimicrobiaceae*, *Ignavibacteria*, and *Anaerolineaceae* (Schink, 1997; Sun *et al.*, 2016; McIlroy *et al.*, 2017; Bei *et al.*, 2021) and by *Aminidesulfobivrio aminophilus*, an established amino acid fermenter initially isolated and validly described as *Desulfobivrio aminophilus* (Baena *et al.*, 1998). *Desulfobivrio* are incomplete oxidizing sulfate-reducing bacteria producing acetate and can grow in syntrophic physiology (Bryant *et al.*, 1977; Muyzer and Stams, 2008). The fermentation of amino acids yields also propionate, butyrate, and eventually alcohols which are substrates for specialists in syntrophic catabolism. Species affiliating to *Syntrophobacteraceae*, *Smithellaceae*, and other *Syntrophales* were present in the community. Of these, only LiM6 affiliating to *Syntrophobacteraceae* featured a limonene degradation pathway. A key to the stable maintenance of this complex community is a slow nutrient-limited growth rate of the methanogenic community allowing the predatory bacteria to thrive at least at equal growth rates. Not only hydrocarbons but also short-chain fatty acids can support syntrophic and methanogenic microorganisms (Jackson *et al.*, 1999; McInerney *et al.*, 2008) which may, facilitate the coexistence of predatory bacteria, potentially including members of the candidate phylum Omnitrochota.

The metatranscriptomic and metaproteomic observations of this study revealed the expression of two operons from the genome of *Syntrophobacteraceae* LiM6 for the degradation of limonene into smaller organic molecules. Key enzymes are limonenylsuccinate synthase and 4-isopropenyl-2-oxo-cyclohexane-1-carbonyl-CoA hydrolase which activate limonene and catalyze the ring opening. The annotation was based on high amino acid similarity to well-characterized enzymes. For the large alpha subunit, the limonenylsuccinate synthase shares 61% identical amino acids with the benzyl succinate synthase alpha subunit of *Thauera aromatica*. The differences are sufficient to form together with uncharacterized proteins from other Deltaproteobacteria a phylogenetic clade that is separated from the toluene and *p*-cresol clades of benzylsuccinate synthases. Amino acids involved in fumarate binding and catalysis by benzylsuccinate synthase are conserved in the limonenylsuccinate synthase. Some differences are located in the substrate binding pocket of the active site as defined by crystal structures of benzylsuccinate synthase with toluene (Funk *et al.*, 2015; Heider *et al.*, 2016). In the substrate binding pocket, Ile-384 (nomenclature of the benzylsuccinate synthase from *Thauera aromatica*) blocks the bottom of the pocket opposite to the 4-position of the benzyl ring. It is replaced by smaller amino acids in succinate synthases with a large substituent in the 4-position: a threonine in limonenylsuccinate synthase and a valine in 4-isopropylbenzylsuccinate synthase of pCyN2. Deltaproteobacterial glycine radical proteins within the limonene

clade (i.e. WP_155306556.1 from *Desulfosarcina widdellii* and WP_155321809.1 from *Desulfosarcina ovata*) have with alanine a different amino acid on this position and may not transform limonene. Substitution of Ile-617 by valine enlarges the substrate cavity. The mutant Val-617 (I617V) lifts a steric hindrance and enables the activation of m-xylene by Bss (Salii *et al.*, 2021). Limonenylsuccinate synthase and naphthylmethylsuccinate synthase have a histidine at this position which is also smaller than isoleucine and in the size range of valine. The space requirement of limonene is due to its structure: unlike aromatic compounds, limonene is not a planar ring.

This study has revealed a novel pathway for the degradation of limonene. The limonene dehydrogenase of the denitrifying *Castellaniella defragrans* requires an electron acceptor with a positive reduction potential (Puentes-Cala *et al.*, 2018a, 2018b). This system may not be operable in microorganisms living in an anoxic, highly reduced environment, i.e. sulfate-reducing or methanogenic systems. In contrast, the oxygen-sensitive glycine radical enzymes are present in facultative and obligate anaerobes involved in a range of anoxic reactions. Our findings revealed an additional use of the radical enzyme by nature, the anaerobic functionalization of limonene to limonenylsuccinate.

Acknowledgment

We thank the Max Planck-Genome-center Cologne (<http://mpgc.mpipz.mpg.de/home/>) for performing NGS sequencing.

This study was funded by the Max Planck Society. Jana Kizina and Almud Lonsing are members of the International Max Planck Research School of Marine Microbiology (MarMic).

References

- Abu Laban, N., Dao, A., Semple, K., and Foght, J. (2015) Biodegradation of C7 and C8 isoalkanes under methanogenic conditions. *Environmental Microbiology* **17**: (12) 4898–4915.
- Artimo, P., Jonnalagedda, M., Arnold, K., Baratin, D., Csardi, G., de Castro, E., et al. (2012) ExPASy: SIB bioinformatics resource portal. *Nucleic acids research* **40**: (Web Server issue) W597-603.
- Atashgahi, S., Hornung, B., van der Waals, M.J., da Rocha, U.N., Hugenholtz, F., Nijssse, B., et al. (2018) A benzene-degrading nitrate-reducing microbial consortium displays aerobic and anaerobic benzene degradation pathways. *Scientific Reports* **8**: (1) 4490.
- Azam, F., Fenchel, T., Field, J., Gray, J.S., Meyer, L., and Thingstad, T.F. (1983) The ecological role of water-column microbes in the sea. *Marine Ecology Progress Series* **10**: (3) 257–263.
- Baena, S., Fardeau, M.L., Labat, M., Ollivier, B., Garcia, J.L., and Patel, B.K.C. (1998) *Desulfovibrio aminophilus* sp. nov., a novel amino acid degrading and sulfate reducing bacterium from an anaerobic dairy wastewater lagoon. *Systematic and Applied Microbiology* **21**: (4) 498–504.
- Bankevich, A., Nurk, S., Antipov, D., Gurevich, A.A., Dvorkin, M., Kulikov, A.S., et al. (2012) SPAdes: a new genome assembly algorithm and its applications to single-cell sequencing. *Journal of computational biology* **19**: (5) 455–77.
- Bei, Q., Peng, J., and Liesack, W. (2021) Shedding light on the functional role of the *Ignavibacteria* in Italian rice field soil: a meta-genomic/transcriptomic analysis. *Soil Biology and Biochemistry* **163**: 108444.
- Berger, S., Cabrera-Orefice, A., Jetten, M.S.M., Brandt, U., and Welte, C.U. (2021) Investigation of central energy metabolism-related protein complexes of ANME-2d methanotrophic archaea by complexome profiling. *Biochimica et Biophysica Acta (BBA) - Bioenergetics* **1862**: (1) 148308.
- Bräuer, S., Cadillo-Quiroz, H., Kyrpides, N., Woyke, T., Goodwin, L., Detter, C., et al. (2015) Genome of *Methanoregula boonei* 6A8 reveals adaptations to oligotrophic peatland environments. *Microbiology* **161**: (8) 1572–1581.

- Bryant, M.P., Campbell, L.L., Reddy, C.A., and Crabill, M.R. (1977) Growth of *Desulfovibrio* in lactate or ethanol media low in sulfate in association with H₂-utilizing methanogenic bacteria. *Applied and Environmental Microbiology* **33**: (5) 1162–1169.
- Buchfink, B., Reuter, K., and Drost, H.-G. (2021) Sensitive protein alignments at tree-of-life scale using DIAMOND. *Nature Methods* **18**: (4) 366–368.
- Cabanettes, F. and Klopp, C. (2018) D-GENIES: dot plot large genomes in an interactive, efficient and simple way. *PeerJ* **6**: e4958.
- Chaumeil, P.-A., Mussig, A.J., Hugenholtz, P., and Parks, D.H. (2019) GTDB-Tk: a toolkit to classify genomes with the Genome Taxonomy Database. *Bioinformatics* **36**: (6) 1925–1927.
- Chen, Y., Zhang, Y., Wang, A.Y., Gao, M., and Chong, Z. (2021) Accurate long-read de novo assembly evaluation with Inspector. *Genome Biology* **22**: (1) 312.
- Demirel, B. and Scherer, P. (2008) The roles of acetotrophic and hydrogenotrophic methanogens during anaerobic conversion of biomass to methane: a review. *Reviews in Environmental Science and Bio/Technology* **7**: (2) 173–190.
- Dolfing, J., Larter, S.R., and Head, I.M. (2007) Thermodynamic constraints on methanogenic crude oil biodegradation. *The ISME Journal* **2**: (4) 442–452.
- Embree, M., Liu, J.K., Al-Bassam, M.M., and Zengler, K. (2015) Networks of energetic and metabolic interactions define dynamics in microbial communities. *Proceedings of the National Academy of Sciences* **112**: (50) 15450–15455.
- Fowler, S.J., Toth, C.R.A., and Gieg, L.M. (2016) Community structure in methanogenic enrichments provides insight into syntrophic interactions in hydrocarbon-impacted environments. *Frontiers in Microbiology* **7**: 562.
- Frey, M., Rothe, M., Wagner, A.F., and Knappe, J. (1994) Adenosylmethionine-dependent synthesis of the glycy radical in pyruvate formate-lyase by abstraction of the glycine C-2 pro-S hydrogen atom. Studies of [2H]glycine-substituted enzyme and peptides homologous to the glycine 734 site. *Journal of Biological Chemistry* **269**: (17) 12432–12437.
- Funk, M.A., Marsh, E.N.G., and Drennan, C.L. (2015) Substrate-bound structures of benzylsuccinate synthase reveal how toluene is activated in anaerobic hydrocarbon degradation. *The Journal of Biological Chemistry* **290**: (37) 22398–22408.
- Grbić-Galić, D. and Vogel, T.M. (1987) Transformation of toluene and benzene by mixed methanogenic cultures. *Applied and Environmental Microbiology* **53**: (2) 254–260.

- Grundmann, O., Behrends, A., Rabus, R., Amann, J., Halder, T., Heider, J., and Widdel, F. (2008) Genes encoding the candidate enzyme for anaerobic activation of n-alkanes in the denitrifying bacterium, strain HxN1. *Environmental Microbiology* **10**: (2) 376–385.
- Hagel, C., Blaum, B., Friedrich, T., and Heider, J. (2022) Characterisation of the redox centers of ethylbenzene dehydrogenase. *Journal of Biological Inorganic Chemistry* **27**: (1) 143–154.
- Harder, J. (2000) Anaerobic utilization of essential oils by denitrifying bacteria. *Biodegradation* **11**: (1) 55–63.
- Harder, J. and Foss, S. (1999) Anaerobic formation of the aromatic hydrocarbon p-cymene from monoterpenes by methanogenic enrichment cultures. *Geomicrobiology Journal* **16**: (4) 295–305.
- Harder, J. and Marmulla, R. (2017) Catabolic pathways and enzymes involved in the anaerobic degradation of terpenes. In *Anaerobic Utilization of Hydrocarbons, Oils, and Lipids*. Boll, M. (ed). Cham: Springer International Publishing, pp. 1–15.
- Heider, J., Szaleniec, M., Martins, B.M., Seyhan, D., Buckel, W., and Golding, B.T. (2016) Structure and function of benzylsuccinate synthase and related fumarate-adding glyceryl radical enzymes. *Microbial Physiology* **26**: (1–3) 29–44.
- Heker, I., Haberhauer, G., and Meckenstock, R.U. (2023) Naphthalene carboxylation in the sulfate-reducing enrichment culture N47 is proposed to proceed via 1,3-dipolar cycloaddition to the cofactor prenylated flavin mononucleotide. *Applied and Environmental Microbiology* **89**: (3) e01927-22.
- Himmelberg, A.M., Bröls, T., Farmani, Z., Weyrauch, P., Barthel, G., Schrader, W., and Meckenstock, R.U. (2018) Anaerobic degradation of phenanthrene by a sulfate-reducing enrichment culture. *Environmental Microbiology* **20**: (10) 3589–3600.
- Jackson, B.E., Bhupathiraju, V.K., Tanner, R.S., Woese, C.R., and McInerney, M.J. (1999) *Syntrophus aciditrophicus* sp. nov., a new anaerobic bacterium that degrades fatty acids and benzoate in syntrophic association with hydrogen-using microorganisms. *Archives of Microbiology* **171**: (2) 107–114.
- Jetten, M.S.M., Stams, A.J.M., and Zehnder, A.J.B. (1992) Methanogenesis from acetate: a comparison of the acetate metabolism in *Methanothrix soehngenii* and *Methanosarcina* spp. *FEMS Microbiology Reviews* **8**: (3–4) 181–197.
- Ji, J.-H., Liu, Y.-F., Zhou, L., Irfan, M., Mbadinga, S.M., Pan, P., et al. (2020) Methanogenic biodegradation of C13 and C14 n-alkanes activated by addition to fumarate. *International Biodeterioration & Biodegradation* **153**: 104994.

- Ji, J.H., Zhou, L., Mbadanga, S.M., Irfan, M., Liu, Y.F., Pan, P., et al. (2020) Methanogenic biodegradation of C₉ to C_{12n}-alkanes initiated by *Smithella* via fumarate addition mechanism. *AMB Express* **10**: (1) 23.
- Joshi, A., Lanjekar, V., Dhakephalkar, P.K., and Dagar, S.S. (2018) Cultivation of multiple genera of hydrogenotrophic methanogens from different environmental niches. *Anaerobe* **50**: 64–68.
- Kang, D.D., Froula, J., Egan, R., and Wang, Z. (2015) MetaBAT, an efficient tool for accurately reconstructing single genomes from complex microbial communities. *PeerJ* **3**: e1165.
- Kizina, J. (2017) Insights into the biology of Candidate Division OP3 LiM populations. *PhD Thesis*.
- Kizina, J., Jordan Sebastian, F.A., Martens Gerrit, A., Lonsing, A., Probian, C., Kolovou, A., et al. (2022) *Methanosaeta* and “*Candidatus Velamenicoccus archaeovor*us”. *Applied and Environmental Microbiology* **88**: (7) e02407-21.
- Kolmogorov, M., Bickhart, D.M., Behsaz, B., Gurevich, A., Rayko, M., Shin, S.B., et al. (2020) metaFlye: scalable long-read metagenome assembly using repeat graphs. *Nature Methods* **17**: (11) 1103–1110.
- Kopylova, E., Noé, L., and Touzet, H. (2012) SortMeRNA: fast and accurate filtering of ribosomal RNAs in metatranscriptomic data. *Bioinformatics* **28**: (24) 3211–7.
- Krey, J.F., Wilmarth, P.A., Shin, J.B., Klimek, J., Sherman, N.E., Jeffery, E.D., et al. (2014) Accurate label-free protein quantitation with high- and low-resolution mass spectrometers. *Journal of proteome research* **13**: (2) 1034–1044.
- Kurth, J.M., Op den Camp, H.J.M., and Welte, C.U. (2020) Several ways one goal - methanogenesis from unconventional substrates. *Applied Microbiology and Biotechnology* **104**: (16) 6839–6854.
- Kurth, J.M., Smit, N.T., Berger, S., Schouten, S., Jetten, M.S.M., and Welte, C.U. (2019) Anaerobic methanotrophic archaea of the ANME-2d clade feature lipid composition that differs from other ANME archaea. *FEMS Microbiology Ecology* **95**: (7) fiz082.
- Langmead, B. and Salzberg, S.L. (2012) Fast gapped-read alignment with Bowtie 2. *Nature methods* **9**: (4) 357–9.
- Laso-Pérez, R., Hahn, C., van Vliet Daan, M., Tegetmeyer Halina, E., Schubotz, F., Smit Nadine, T., et al. (2019) Anaerobic degradation of non-methane alkanes by “*Candidatus Methanoliparia*” in hydrocarbon seeps of the Gulf of Mexico. *mBio* **10**: (4) 10.1128/mbio.01814-19.

- Lemaire, O.N. and Wagner, T. (2022) A structural view of alkyl-coenzyme M reductases, the first step of alkane anaerobic oxidation catalyzed by Archaea. *Biochemistry* **61**: (10) 805–821.
- Liu, Y.-F., Chen, J., Liu, Z.-L., Shou, L.-B., Lin, D.-D., Zhou, L., et al. (2020) Anaerobic degradation of paraffins by thermophilic actinobacteria under methanogenic conditions. *Environmental Science & Technology* **54**: (17) 10610–10620.
- Marchler-Bauer, A., Lu, S., Anderson, J.B., Chitsaz, F., Derbyshire, M.K., DeWeese-Scott, C., et al. (2011) CDD: a Conserved Domain Database for the functional annotation of proteins. *Nucleic acids research* **39**: (Database issue) D225-9.
- Marmulla, R. and Harder, J. (2014) Microbial monoterpene transformations - a review. *Frontiers in Microbiology* **5**: 346.
- McIlroy, S.J., Kirkegaard, R.H., Dueholm, M.S., Fernando, E., Karst, S.M., Albertsen, M., and Nielsen, P.H. (2017) Culture-independent analyses reveal novel *Anaerolineaceae* as abundant primary fermenters in anaerobic digesters treating waste activated sludge. *Frontiers in Microbiology* **8**: 1134.
- McInerney, M.J., Struchtemeyer, C.G., Sieber, J., Mouttaki, H., Stams, A.J.M., Schink, B., et al. (2008) Physiology, ecology, phylogeny, and genomics of microorganisms capable of syntrophic metabolism. *Annals of the New York Academy of Sciences* **1125**: (1) 58–72.
- Messina, P., Lathière, J., Sindelarova, K., Vuichard, N., Granier, C., Ghattas, J., et al. (2016) Global biogenic volatile organic compound emissions in the ORCHIDEE and MEGAN models and sensitivity to key parameters. *Atmospheric Chemistry Physics* **16**: (22) 14169–14202.
- Mollaie, M., Suarez-Diez, M., Sedano-Nunez, V.T., Boeren, S., Stams, A.J.M., and Plugge, C.M. (2021) Proteomic analysis of a syntrophic coculture of *Syntrophobacter fumaroxidans* MPOBT and *Geobacter sulfurreducens* PCAT. *Frontiers in Microbiology* **12**: 708911.
- Muyzer, G. and Stams, A.J.M. (2008) The ecology and biotechnology of sulphate-reducing bacteria. *Nature Reviews Microbiology* **6**: (6) 441–454.
- Nissen, J.N., Johansen, J., Allesøe, R.L., Sønderby, C.K., Armenteros, J.J.A., Grønbech, C.H., et al. (2021) Improved metagenome binning and assembly using deep variational autoencoders. *Nature Biotechnology* **39**: (5) 555–560.
- Nobu, M.K., Narihiro, T., Kuroda, K., Mei, R., and Liu, W.-T. (2016) Chasing the elusive Euryarchaeota class WSA2: genomes reveal a uniquely fastidious methyl-reducing methanogen. *The ISME Journal* **10**: (10) 2478–2487.

- Parks, D.H., Imelfort, M., Skennerton, C.T., Hugenholtz, P., and Tyson, G.W. (2015) CheckM: assessing the quality of microbial genomes recovered from isolates, single cells, and metagenomes. *Genome research* **25**: (7) 1043–55.
- Perez-Molphe-Montoya, E., Küsel, K., and Overholt, W.A. (2022) Redefining the phylogenetic and metabolic diversity of phylum Omnitrophota. *Environmental Microbiology* **24**: (11) 5437–5449.
- Potter, A., Andersson, J., Sjöblom, A., Junedahl, E., Cousins, A., Brorström-Lundén, E., and Cato, I. (2014) Results from the swedish national screening programme 2004 - subreport 3: limonene.
- Price, M.N., Dehal, P.S., and Arkin, A.P. (2009) FastTree: computing large minimum evolution trees with profiles instead of a distance matrix. *Molecular biology and evolution* **26**: (7) 1641–50.
- Puentes-Cala, E., Liebeke, M., Markert, S., and Harder, J. (2018a) Anaerobic degradation of bicyclic monoterpenes in *Castellaniella defragrans*. *Metabolites* **8**: (1) 12.
- Puentes-Cala, E., Liebeke, M., Markert, S., and Harder, J. (2018b) Limonene dehydrogenase hydroxylates the allylic methyl group of cyclic monoterpenes in the anaerobic terpene degradation by *Castellaniella defragrans*. *Journal of Biological Chemistry* **293**: (24) 9520–9529.
- Rotaru, A.E., Schauer, R., Probian, C., Mussmann, M., and Harder, J. (2012) Visualization of candidate division OP3 cocci in limonene-degrading methanogenic cultures. *Journal of Microbiology and Biotechnology* **22**: (4) 457–61.
- Salii, I., Szaleniec, M., Zein, A.A., Seyhan, D., Sekuła, A., Schühle, K., et al. (2021) Determinants for substrate recognition in the glyceryl radical enzyme benzylsuccinate synthase revealed by targeted mutagenesis. *ACS Catalysis* **11**: (6) 3361–3370.
- Sayers, E.W., Bolton, E.E., Brister, J.R., Canese, K., Chan, J., Comeau, D.C., et al. (2022) Database resources of the national center for biotechnology information. *Nucleic Acids Research* **50**: (D1) D20-d26.
- Schink, B. (1997) Energetics of syntrophic cooperation in methanogenic degradation. *Microbiology and molecular biology reviews* **61**: (2) 262–80.
- Schwengers, O., Jelonek, L., Dieckmann, M.A., Beyvers, S., Blom, J., and Goesmann, A. (2021) Bakta: rapid and standardized annotation of bacterial genomes via alignment-free sequence identification. *Microbial Genomics* **7**: (11).
- Seemann, T. (2014) Prokka: rapid prokaryotic genome annotation. *Bioinformatics* **30**: (14) 2068–9.

- Selesi, D., Jehmlich, N., von Bergen, M., Schmidt, F., Rattei, T., Tischler, P., et al. (2010) Combined genomic and proteomic approaches identify gene clusters involved in anaerobic 2-methylnaphthalene degradation in the sulfate-reducing enrichment culture N47. *Journal of bacteriology* **192**: (1) 295–306.
- Seymour, C.O., Palmer, M., Becraft, E.D., Stepanauskas, R., Friel, A.D., Schulz, F., et al. (2023) Hyperactive nanobacteria with host-dependent traits pervade Omnitrotophota. *Nature Microbiology* **8**: (4) 727–744.
- Shima, S., Krueger, M., Weinert, T., Demmer, U., Kahnt, J., Thauer, R.K., and Ermler, U. (2012) Structure of a methyl-coenzyme M reductase from Black Sea mats that oxidize methane anaerobically. *Nature* **481**: (7379) 98–101.
- Siddique, T., Semple, K., Li, C., and Foght, J.M. (2020) Methanogenic biodegradation of iso-alkanes and cycloalkanes during long-term incubation with oil sands tailings. *Environmental Pollution* **258**: 113768.
- Sieber, C.M.K., Probst, A.J., Sharrar, A., Thomas, B.C., Hess, M., Tringe, S.G., and Banfield, J.F. (2018) Recovery of genomes from metagenomes via a dereplication, aggregation and scoring strategy. *Nature Microbiology* **3**: (7) 836–843.
- Sinitcyn, P., Tiwary, S., Rudolph, J., Gutenbrunner, P., Wichmann, C., Yilmaz, Ş., et al. (2018) MaxQuant goes Linux. *Nature Methods* **15**: (6) 401.
- Smolander, A., Ketola, R.A., Kotiaho, T., Kanerva, S., Suominen, K., and Kitunen, V. (2006) Volatile monoterpenes in soil atmosphere under birch and conifers: effects on soil N transformations. *Soil Biology and Biochemistry* **38**: (12) 3436–3442.
- Strijkstra, A., Trautwein, K., Jarling, R., Wöhlbrand, L., Dörries, M., Reinhardt, R., et al. (2014) Anaerobic activation of p-cymene in denitrifying betaproteobacteria: methyl group hydroxylation versus addition to fumarate. *Applied and environmental microbiology* **80**: (24) 7592–603.
- Sun, L., Toyonaga, M., Ohashi, A., Turlousse, D.M., Matsuura, N., Meng, X.Y., et al. (2016) *Lentimicrobium saccharophilum* gen. nov., sp. nov., a strictly anaerobic bacterium representing a new family in the phylum *Bacteroidetes*, and proposal of *Lentimicrobiaceae* fam. nov. *International Journal of Systematic and Evolutionary Microbiology* **66**: (7) 2635–2642.
- Tamura, K., Stecher, G., and Kumar, S. (2021) MEGA11: Molecular Evolutionary Genetics Analysis Version 11. *Molecular biology and evolution* **38**: (7) 3022–3027.

- Tan, B., Jane Fowler, S., Laban, N.A., Dong, X., Sensen, C.W., Foght, J., and Gieg, L.M. (2015) Comparative analysis of metagenomes from three methanogenic hydrocarbon-degrading enrichment cultures with 41 environmental samples. *The ISME Journal* **9**: (9) 2028–2045.
- Tan, B., Nesbø, C., and Foght, J. (2014) Re-analysis of omics data indicates *Smithella* may degrade alkanes by addition to fumarate under methanogenic conditions. *The ISME Journal* **8**: (12) 2353–6.
- Thauer, R.K., Kaster, A.-K., Seedorf, H., Buckel, W., and Hedderich, R. (2008) Methanogenic archaea: ecologically relevant differences in energy conservation. *Nature Reviews Microbiology* **6**: (8) 579–591.
- Thompson, J.D., Higgins, D.G., and Gibson, T.J. (1994) CLUSTAL W: improving the sensitivity of progressive multiple sequence alignment through sequence weighting, position-specific gap penalties and weight matrix choice. *Nucleic acids research* **22**: (22) 4673–80.
- Tomei, M.C., Mosca Angelucci, D., Clagnan, E., and Brusetti, L. (2021) Anaerobic biodegradation of phenol in wastewater treatment: achievements and limits. *Applied Microbiology and Biotechnology* **105**: (6) 2195–2224.
- Toth, C.R.A., Berdugo-Clavijo, C., O’Farrell, C.M., Jones, G.M., Sheremet, A., Dunfield, P.F., and Gieg, L.M. (2018) Stable isotope and metagenomic profiling of a methanogenic naphthalene-degrading enrichment culture. *Microorganisms* **6**: (3) 65.
- Toth, C.R.A., Luo, F., Bawa, N., Webb, J., Guo, S., Dworatzek, S., and Edwards, E.A. (2021) Anaerobic benzene biodegradation linked to the growth of highly specific bacterial clades. *Environmental Science & Technology* **55**: (12) 7970–7980.
- Wawrik, B., Marks, C.R., Davidova, I.A., McInerney, M.J., Pruitt, S., Duncan, K.E., et al. (2016) Methanogenic paraffin degradation proceeds via alkane addition to fumarate by ‘*Smithella*’ spp. mediated by a syntrophic coupling with hydrogenotrophic methanogens. *Environmental Microbiology* **18**: (8) 2604–2619.
- Wegener, G., Laso-Pérez, R., Orphan, V.J., and Boetius, A. (2022) Anaerobic degradation of alkanes by marine archaea. *Annual Review of Microbiology* **76**: (1) 553–577.
- White, C.S. (1991) The role of monoterpenes in soil nitrogen cycling processes in ponderosa pine. *Biogeochemistry* **12**: (1) 43–68.
- Widdel, F., Rouvière, P.E., and Wolfe, R.S. (1988) Classification of secondary alcohol-utilizing methanogens including a new thermophilic isolate. *Archives of Microbiology* **150**: (5) 477–481.

- Wöhlbrand, L., Jacob, J.H., Kube, M., Mussmann, M., Jarling, R., Beck, A., et al. (2013) Complete genome, catabolic sub-proteomes and key-metabolites of *Desulfobacula toluolica* Tol2, a marine, aromatic compound-degrading, sulfate-reducing bacterium. *Environmental Microbiology* **15**: (5) 1334–1355.
- Woodcroft, B. (2021) CoverM: program. *Woodcroft: Queensland University*.
- Wu, Y.-W., Simmons, B.A., and Singer, S.W. (2015) MaxBin 2.0: an automated binning algorithm to recover genomes from multiple metagenomic datasets. *Bioinformatics* **32**: (4) 605–607.
- Zehnle, H., Laso-Pérez, R., Lipp, J., Riedel, D., Benito Merino, D., Teske, A., and Wegener, G. (2023) *Candidatus* Alkanophaga archaea from Guaymas Basin hydrothermal vent sediment oxidize petroleum alkanes. *Nature Microbiology* **8**: (7) 1199–1212.
- Zengler, K., Richnow, H.H., Rosselló-Mora, R., Michaelis, W., and Widdel, F. (1999) Methane formation from long-chain alkanes by anaerobic microorganisms. *Nature* **401**: (6750) 266–9.

Supplementary material

Supplementary Table 1: Relative abundance of MAGs for the 10kS (pellet 1), 100S aggregates (pellet2) and 100S cells (pellet3) based on Illumina reads (DNA).

name	pellet 1 Illumina	pellet 2 Illumina	pellet 3 Illumina	taxonomy
	relative abundance [%]	relative abundance [%]	relative abundance [%]	
LiM	63.01	83.02	89.29	s__Vampirococcus archaeovorur
LiM1	7.17	4.43	2.06	g__Methanoregula
LiM2	7.9	0.81	0.44	s__Methanoculleus
LiM3	3.72	1.13	0.55	s__Methanotherix soehngeni
LiM4	1.91	0.99	0.24	s__Lentimicrobium
LiM5	1.76	0.96	0.36	o__Syntrophales
LiM6	1.36	0.72	0.6	f__Syntrophobacteraceae
LiM7	1.36	0.58	0.17	g__Mesotoga
LiM8	0.93	1.24	0.24	f__Anaerolineaceae
LiM9	0.87	0.53	0.07	f__Anaerolineaceae
LiM10	0.67	0.14	0.23	o__Syntrophales
LiM11	0.64	0.37	0.19	f__Anaerolineaceae
LiM12	0.96	0.55	0.23	s__Aminidesulfovibrio
LiM13	0.52	0.23	0.07	o__Treponematales
LiM14	0.32	0.06	0.03	s__Methanospirillum
LiM15	0.45	0.27	0.1	c__Coriobacteria
LiM16	0.41	0.21	0.04	f__Aminobacteriaceae
LiM17	0.21	0.08	0.07	f__Smithellaceae
LiM18	0.17	0.07	0.27	g__Mangrovibacterium
LiM19	0.24	0.07	0.02	c__Xenobia
LiM20	0.19	0.12	0.09	p__Spirochaetota
LiM21	0.13	0.1	0.01	g__Methanofastidiosum
LiM22	0.18	0.09	0.77	g__Mangrovibacterium
LiM23	0.11	0.05	0.01	p__Desulfobacterota
LiM24	0.35	0.2	0.1	f__Thermoanaerobaculaceae
LiM25	0.12	0.14	0.11	c__Ignavibacteria
LiM26	0.21	0.19	0.16	f__Melioribacteraceae
LiM27	0.22	0.18	1.22	s__Aminidesulfovibrio aminophilus
LiM28	0.2	0.05	0.06	f__Thermoguttaceae
LiM29	0.1	0.08	0.03	f__Anaerolineaceae
LiM30	0.08	0.03	0.04	s__Sphaerochaeta
LiM31	0.08	0.05	0.04	o__Bacteroidales
unmapped	3.45	2.26	2.09	

Chapter 3 - Anaerobic limonene metabolism involves a glycine radical enzyme

Supplementary Table 2: CheckM results of PacBio MAGs before refinement. Quality is calculated as “completeness minus 5x contamination”.

name	Completeness	Contamination	Quality	Strain heterogeneity	Genome size (bp)	# contigs
LiM12	100	0	100	0	3075031	1
LiM7	99.84	0	99.84	0	3073864	15
LiM13	98.85	0	98.85	0	2973146	1
LiM3	99.35	0.65	96.1	0	3098710	4
LiM2	98.04	0.65	94.79	0	3026955	1
LiM14	99.02	0.98	94.12	0	3622126	7
LiM17	96.13	0.65	92.88	0	3885460	16
LiM1	99.22	1.31	92.67	0	2781926	2
LiM16	100	1.69	91.55	0	3329856	1
LiM20	96.63	1.12	91.03	0	4543582	9
LiM4	97.67	1.61	89.62	0	5157114	5
LiM18	100	2.42	87.9	42.86	8231951	13
LiM24	90.12	0.85	85.87	0	5252351	19
LiM22	97.85	2.42	85.75	0	8360960	21
LiM19	94.23	1.71	85.68	0	9177046	19
LiM	92.05	1.61	84	0	1993543	8
LiM27	86.7	0.59	83.75	100	3106712	47
LiM25	92.35	1.91	82.8	0	3239037	11
LiM6	97.26	3.39	80.31	42.86	3418264	63
LiM26	84.53	1.12	78.93	50	3483156	43
LiM5	98.06	4.26	76.76	33.33	4092719	2
LiM28	74.9	0	74.9	0	7314612	84
LiM10	83.39	1.94	73.69	0	2768134	15
LiM11	92.73	3.82	73.63	0	2408773	4
LiM15	96.67	5	71.67	0	2848877	3
LiM23	89.16	3.55	71.41	0	3278977	11
LiM8	90.91	4.55	68.16	0	2771241	3
LiM29	69.24	0.91	64.69	0	2443355	21
LiM9	92.73	6.36	60.93	0	2493101	9
LiM30	66.37	3.41	49.32	33.33	2289486	51
LiM31	44.57	0.24	43.37	0	1292300	33
LiM21	43.27	1.49	35.82	0	714189	5

Supplementary Table 3: Genes, including new annotation, predicted to be involved in limonene degradation in LiM6.

locus tag	new annotation	gene
LMHFDP_246395	electron transfer flavoprotein subunit beta	Etf
LMHFDP_246400	electron transfer flavoprotein subunit alpha	Etf
LMHFDP_300765		
LMHFDP_300800	Putative LolA-like molecular chaperone	LssQ
LMHFDP_300805	Limonen-7-ylsuccinate synthase gamma subunit (p-Mentha-1,8-dien-7-yl-succinate synthase)	LssC
LMHFDP_300810	Limonen-7-ylsuccinate synthase alpha subunit (p-Mentha-1,8-dien-7-yl-succinate synthase)	LssA
LMHFDP_300815	Limonen-7-ylsuccinate synthase beta subunit (p-Mentha-1,8-dien-7-yl-succinate synthase)	LssB
LMHFDP_300820	Putative-PNPOx domain-containing protein	LssP
LMHFDP_300825	Acyl-CoA synthetase	LssO
LMHFDP_300830	2-(7-hydroxylimonen-7-yl)-succinyl-CoA dehydrogenase subunit A	LssH
LMHFDP_300835	2-(7-hydroxylimonen-7-yl)-succinyl-CoA dehydrogenase catalytic subunit B	LssI
LMHFDP_300840	Succinyl-CoA:limonen-7-ylsuccinate CoA-transferase subunit A	LssD
LMHFDP_300845	Succinyl-CoA:limonen-7-ylsuccinate CoA-transferase subunit b	LssE
LMHFDP_300850	Limonen-7-yl-succinyl-CoA dehydrogenase	LssF
LMHFDP_300855	Limonen-7-ylidene-succinyl-CoA hydratase	LssG
LMHFDP_300860	Electron transferring flavoprotein subunit B	LssM
LMHFDP_300865	electron transferring flavoprotein subunit A	LssL
LMHFDP_300870	Electron transfer flavoprotein-associated cytochrome b and iron-sulfur cluster-binding oxidoreductase	LssN
LMHFDP_300875	2-(7-oxolimonen-7-yl)-succinyl-CoA thiolase subunit A	LssJ
LMHFDP_300880	2-(7-oxolimonen-7-yl)-succinyl-CoA thiolase subunit B	LssK
LMHFDP_382510	acyl-CoA dehydrogenase forms enoyl(=2-en-acyl)-CoA esters	lrcJ
LMHFDP_382515	CoA transferase / propionylCoA:acetate CoA transferase	lrcI
LMHFDP_382520	thiolase / acyl-CoA acetyltransferase	lrcH
LMHFDP_382525	3-hydroxyacyl-CoA dehydrogenase	lrcG
LMHFDP_382530	enoyl-CoA hydratase/isomerase	lrcF
LMHFDP_382540	5-isopropenyl-3-hydroxy-pimeloyl-CoA //beta-ketoacyl-ACP reductase	lrcE
LMHFDP_382545	5-isopropenyl-3-oxo-pimeloyl-CoA thiolase // thiolase/acetyl-CoA transferase	lrcD
LMHFDP_382550	3-isopropenyl-3-hydroxy-glutarylCoA synthase / 3-hydroxy-3-methyl-glutarylCoA synthase superfamily	lrcC
LMHFDP_382555	3-isopropenyl-glutaconyl-CoA hydratase / crotonase/enoyl-CoA hydratase	lrcB
LMHFDP_382560	4-isopropenyl-2-oxocyclohexane-1-carbonyl-CoA hydrolase	lrcA
LMHFDP_382595	flavin-containing oxidoreductase	
LMHFDP_382605	membrane-fusion subunit of putative monoterpene efflux pump (tripartite transporter)	mepB
LMHFDP_382610	efflux pump subunit of putative monoterpene efflux pump (tripartite transporter)	mepA
LMHFDP_382615	efflux transmembrane channel protein of putative monoterpene efflux pump (tripartite transporter)	mepC
LMHFDP_382915	NADH-acceptor oxidoreductase subunit F (51 kDa) of NiFe hydrogenase	Hox
LMHFDP_560725	isocitrate/3-isopropylmalate dehydrogenase	ipmC
LMHFDP_560730	3-isopropylmalate dehydratase small subunit	ipmB
LMHFDP_560735	3-isopropylmalate dehydratase large subunit	ipmA
LMHFDP_562065	succinat dehydrogenase/fumarate reductase iron-sulfur subunit	SdhABC
LMHFDP_562070	succinat dehydrogenase/fumarate reductase flavoprotein subunit	SdhABC
LMHFDP_562075	succinat dehydrogenase/fumarate reductase cytochrome b subunit	SdhABC

Chapter 3 - Anaerobic limonene metabolism involves a glycine radical enzyme

Supplementary Table 4: Genes predicted to be involved in limonene degradation in LiM6, with corresponding TPM and riBAQ values.

locus tag	gene	TPM 1906C	riBAQ PE1_372	riBAQ PE2_372	riBAQ PE3_372	riBAQ PE1_373	riBAQ PE2_373	riBAQ PE3_373
LMHFDP_246395	Etf	52	1.252E-04	1.365E-03	2.569E-03	3.950E-04	1.149E-04	5.980E-06
LMHFDP_246400	Etf	47	1.924E-05	3.131E-04	7.181E-04	1.153E-05	9.273E-05	5.663E-05
LMHFDP_300765		0	#N/A	#N/A	#N/A	#N/A	#N/A	#N/A
LMHFDP_300800	LssQ	17	0.000E+00	2.164E-05	6.194E-05	#N/A	#N/A	#N/A
LMHFDP_300805	LssC	183	#N/A	#N/A	#N/A	#N/A	#N/A	#N/A
LMHFDP_300810	LssA	174	2.604E-05	2.985E-04	9.259E-04	2.452E-05	8.486E-05	9.991E-05
LMHFDP_300815	LssB	55	0.000E+00	1.103E-04	3.528E-04	4.547E-05	4.279E-05	9.801E-07
LMHFDP_300820	LssP	11	2.285E-04	4.459E-04	1.205E-03	2.309E-05	1.019E-04	1.134E-04
LMHFDP_300825	LssO	12	0.000E+00	1.221E-05	1.992E-05	3.815E-06	9.355E-06	1.671E-05
LMHFDP_300830	LssH	41	2.667E-04	1.294E-03	2.738E-03	5.779E-05	1.285E-04	1.064E-03
LMHFDP_300835	LssI	49	2.077E-04	8.815E-04	1.923E-03	8.903E-05	1.207E-04	1.240E-04
LMHFDP_300840	LssD	21	6.554E-06	1.770E-04	2.723E-04	2.683E-05	4.983E-05	0.000E+00
LMHFDP_300845	LssE	24	0.000E+00	1.355E-04	1.129E-04	1.071E-05	3.348E-05	1.696E-05
LMHFDP_300850	LssF	31	5.805E-05	1.294E-03	8.691E-04	5.147E-05	9.897E-05	1.658E-06
LMHFDP_300855	LssG	35	2.214E-04	1.807E-03	3.114E-03	3.636E-05	1.278E-04	1.527E-04
LMHFDP_300860	LssM	24	3.655E-05	3.661E-04	5.858E-04	4.124E-05	8.564E-05	1.015E-04
LMHFDP_300865	LssL	25	3.520E-05	4.531E-04	7.574E-04	#N/A	#N/A	#N/A
LMHFDP_300870	LssN	22	0.000E+00	8.281E-06	2.383E-05	6.070E-06	8.397E-06	1.549E-06
LMHFDP_300875	LssJ	40	2.390E-04	5.421E-04	1.812E-03	2.078E-05	1.134E-04	8.525E-05
LMHFDP_300880	LssK	73	1.121E-04	7.077E-04	1.335E-03	2.390E-05	8.911E-05	1.524E-04
LMHFDP_382510	lrcJ	42	1.957E-04	9.866E-04	1.180E-03	9.760E-05	9.907E-05	2.984E-05
LMHFDP_382515	lrcI	47	3.569E-05	6.921E-04	1.043E-03	#N/A	#N/A	#N/A
LMHFDP_382520	lrcH	30	6.286E-05	1.489E-03	2.136E-03	#N/A	#N/A	#N/A
LMHFDP_382525	lrcG	15	1.456E-04	1.630E-03	3.159E-03	#N/A	#N/A	#N/A
LMHFDP_382530	lrcF	19	6.075E-05	1.770E-03	2.576E-03	6.458E-05	1.238E-04	1.785E-05
LMHFDP_382540	lrcE	104	3.219E-04	2.271E-03	4.419E-03	#N/A	#N/A	#N/A
LMHFDP_382545	lrcD	57	4.592E-05	9.850E-04	1.549E-03	5.484E-05	1.057E-04	2.639E-04
LMHFDP_382550	lrcC	40	7.371E-05	1.316E-03	1.877E-03	4.653E-05	1.203E-04	8.859E-04
LMHFDP_382555	lrcB	31	1.483E-04	9.600E-04	1.514E-03	1.364E-04	1.061E-04	0.000E+00
LMHFDP_382560	lrcA	194	1.154E-04	1.708E-03	2.201E-03	0.000E+00	1.308E-04	5.488E-07
LMHFDP_382595		9	0.000E+00	2.570E-06	3.741E-06	4.122E-07	1.034E-06	0.000E+00
LMHFDP_382605	mepB	8	3.122E-06	4.404E-05	5.338E-05	1.013E-06	1.502E-05	2.945E-06
LMHFDP_382610	mepA	9	0.000E+00	7.939E-06	7.400E-06	#N/A	#N/A	#N/A
LMHFDP_382615	mepC	4	5.174E-06	2.715E-05	2.678E-05	4.315E-05	3.783E-06	0.000E+00
LMHFDP_382915	Hox	11	0.000E+00	4.507E-06	9.036E-06	#N/A	#N/A	#N/A
LMHFDP_560725	ipmC	5	1.808E-06	3.161E-05	6.266E-05	4.613E-06	5.447E-06	0.000E+00
LMHFDP_560730	ipmB	5	#N/A	#N/A	#N/A	#N/A	#N/A	#N/A
LMHFDP_560735	ipmA	4	#N/A	#N/A	#N/A	6.525E-07	8.229E-07	0.000E+00
LMHFDP_562065	SdhABC	32	6.945E-06	1.192E-04	4.739E-04	1.341E-05	7.587E-05	1.203E-04
LMHFDP_562070	SdhABC	23	8.962E-06	1.626E-04	1.936E-04	#N/A	#N/A	#N/A
LMHFDP_562075	SdhABC	26	5.525E-06	1.141E-04	1.721E-04	3.373E-05	4.875E-05	1.306E-04

Chapter 4:
New insights into the biology of *Candidatus Velamenicoccus*
archaeovor

Almud Lonsing¹, Anastasia Resteu¹, Jana Kizina¹, Jens Harder¹

Manuscript in preparation

¹Max Planck Institute for Marine Microbiology, Bremen, Germany.

Corresponding author: Prof. Dr. Jens Harder

Contribution of the Ph.D. candidate in % of the total work:

Experimental concept and design – 50%

Experimental work/acquisition of experimental data – 50%

Data analysis and interpretation – 80%

Preparation of figures and tables – 90%

Drafting of the manuscript – 80%

Keywords: *Candidatus Velamenicoccus archaeovor*, predatory bacteria, Group I intron, stress response, *Methanotherix*, SR-SIM

Abstract

Candidatus Velamenicoccus archaeovorus is a predatory ultramicrobacterium belonging to the candidate phylum Omnitrophota within the *Planctomycetota–Verrucomicrobiota–Chlamydiota* (PVC) superphylum. This study investigates its predatory “lifestyle” and transcriptomic analyses and high-resolution microscopy. Using catalyzed reporter deposition-fluorescence *in situ* hybridization (CARD-FISH) and structured illumination microscopy (SR-SIM), we visualized interactions between *Candidatus Velamenicoccus archaeovorus* and prey cells, including *Methanothrix* filaments, revealing that attached cells exhibit higher ribosomal content and metabolic activity compared to free-living cells. Transcriptomic data revealed that attached cells express higher levels of proteins involved in type IV pili assembly, secretion systems, and cell wall degradation, supporting active predation. In contrast, free-living cells express higher levels of stress response proteins, indicating dormancy and nutrient limitation. Furthermore, a group I intron was detected in the 23S rRNA of *Candidatus Velamenicoccus archaeovorus* as well as in prey cells, suggesting potential lateral gene transfer during predation. These findings advance our understanding of *Candidatus Velamenicoccus archaeovorus* as a predator, contributing to the study of microbial predation and gene mobility in microbial ecosystems.

4.1 Introduction

Candidatus Velamenicoccus archaeovor is a predatory ultramicrobacterium classified within the candidate phylum Omnitrophota (formerly OP3), which belongs to the *Planctomycetota–Verrucomicrobiota–Chlamydiota* (PVC) superphylum (Rinke *et al.*, 2013; Oren and Garrity, 2021; Kizina *et al.*, 2022). Members of Omnitrophota were first discovered in Obsidian Pool in Yellowstone National Park and have been found to be a group of very small bacteria, predicted to have a host-dependent lifestyle (Hugenholtz *et al.*, 1998; Glöckner *et al.*, 2010; Seymour *et al.*, 2023).

The genomes of predatory bacteria encode proteins, which are involved in their specific predatory mechanism. For epibiotic predators, like *Candidatus Velamenicoccus archaeovor* or *Vampirococcus lugosii*, but also other predatory bacteria, genes expressing proteins involved in finding their prey, attach to them, lyse their cell wall and membrane, and finally uptake their cell contents and degrade them, are often encoded (Moreira *et al.*, 2021; Kizina *et al.*, 2022). These proteins can include various secretion systems, hydrolytic enzymes, and motility-associated proteins. Since the attack occurs outside of the host cell, the proteins involved have to be secreted. Although not an epibiotic predator, *Myxococcus xanthus* hunts in packs and uses a Tad-like secretion system. The Tad-like secretion system belongs to the type IV filament superfamily, which also includes the Type II secretion systems. It is primarily involved in adhesion and essential for initiating contact-dependent prey killing. This system allows the bacterium to form tight interactions with prey cells, leading to their intoxication and subsequent death. After the death of the prey cell, an atypical type III-like secretion system leads to cell lysis (Seef *et al.*, 2021; Thiery *et al.*, 2022; Kaimer *et al.*, 2023). In *Bdellovibrio bacteriovorus*, a well-characterized predatory bacterium, genomic analysis has identified four inner-membrane and five outer-membrane secretion systems (Barabote *et al.*, 2007). Type I and type II secretion systems, as well as the twin-arginine translocation (Tat) system, are involved in transporting proteins, the tat system specifically being involved in translocating proteins across the inner membrane (Barabote *et al.*, 2007; Bratanis *et al.*, 2020). The genome of *Candidatus Velamenicoccus archaeovor* also encodes a Sec secretion system and a type II secretion system (Kizina *et al.*, 2022).

Hydrolytic enzymes like proteases, e.g., elastase-like serine protease, DNases, RNases, and lipases, are encoded in genomes of predatory bacteria like *Bdellovibrio bacteriovorus*. These are not only involved in degrading the prey's cellular components but can also inhibit and

reduce the biofilm formed by other bacteria (Monnappa *et al.*, 2014; Bratanis and Lood, 2019; Bratanis *et al.*, 2020). *Candidatus Velamenicoccus archaeovorus* has a very large multienzyme surface protein with glycosyl transferases, glycosyl hydrolases, and peptidases involved in degrading the prey's cell wall (Kizina *et al.*, 2022). These large proteins are common throughout Omnitrophota (West-Roberts *et al.*, 2023).

To locate and find their prey, predatory bacteria have evolved different strategies. *Myxococcus xanthus* hunts in packs and uses two types of mobility, the “adventurous” and the “social” mobility, in which cells work together. Adventurous motility ensures directional movement of the cells toward prey, while social motility enables cooperation between cells within a swarm, increasing predation efficiency. The protein AglZ, located in the inner membrane, is part of the Agl-Glt machinery and plays an important role in adventurous -motility and gliding motility in general in *Myxococcus xanthus* (Kährström, 2014; Rombouts *et al.*, 2023). *Candidatus Velamenicoccus archaeovorus* is not flagellated, and unlike *Myxococcus xanthus*, it lacks the Agl-Glt machinery essential for gliding motility. Nevertheless, it expresses two stator proteins, MotD and MotB, which are typically associated with flagellar motility. These proteins may also be involved in processes other than motility, such as predation (Kizina *et al.*, 2022).

Bdellovibrio bacteriovorus has been shown to use a set of chimeric fiber proteins from the mosaic adhesive trimer (MAT) superfamily located on its surface. These proteins are trimeric β -helices with diverse terminal domains that resemble bacteriophage tail fibers involved in recognizing and possibly attaching to its prey by creating a stable contact point. Once cell-cell contact is established between predator and prey, *Bdellovibrio bacteriovorus* invades its prey and releases a vesicle into the host's periplasm. Inside this vesicle are MAT proteins and a homolog of CpoB, a protein involved in cell division (Caulton *et al.*, 2024; Du Toit, 2024). The predator produces peptidoglycan hydrolases, which belong to the penicillin-binding protein four family. These peptidoglycan hydrolases modify the prey's peptidoglycan, which leads to the rounding of the prey cell, turning it into the round and osmotically stable “bdelloplast” (Lerner *et al.*, 2012) when two genes coding for N-acetylglucosamine deacetylases involved in modifying the peptidoglycan of the prey were deleted. The prey cells' peptidoglycan structure remained, resulting in a “bacterial ghost”- a hollow shell lacking cytoplasmic content (Lambert *et al.*, 2016). *Candidatus Velamenicoccus archaeovorus* uses a large multienzyme surface protein containing glycosyl hydrolases and peptidases to degrade the polysaccharide components of its prey's outer membrane. After predation, it leaves behind empty cells, as observed in *Methanothrix* (Kizina *et al.*, 2022).

The genome of *Bdellovibrio bacteriovorus* also shows indications for lateral gene transfer and is also shaped by mobile genetic elements (Gophna *et al.*, 2006). Mobile genetic elements are selfish fragments of DNA that can move within and between genomes and are unable to replicate by themselves. These mobile genetic elements include e.g. plasmids, bacteriophages, transposons, and group I and group II introns (Lambowitz and Belfort, 1993; Frost *et al.*, 2005).

Group I introns are self-splicing sequences found in the genomes of bacteria, as well as in the nuclear rDNA genes of protists, fungal mitochondria, and bacteriophages (Reinhold-Hurek and Shub, 1992; Haugen *et al.*, 2005). Group I introns have also been found in archaea by bioinformatics studies based on homology search using covariance models (Nawrocki *et al.*, 2018). They can self-splice from precursor RNA, leaving the exons joined together and avoiding the disruption of the host gene function. On the DNA level, group I introns can transfer to intronless alleles of the same gene via a process called “homing”. A homing endonuclease, a site-specific endonuclease encoded in some group I introns, recognizes the sequence of the intronless gene and cuts a double-strand break in the DNA. The DNA breaks are then repaired by the cell's homologous recombination machinery, which uses the intron-containing version of the gene as a template. There are four families of homing endonucleases, the Cys box, LAGLIDADG, GIY-YIG, and HNH families, which are based on conserved amino acid motifs (Haugen *et al.*, 2005; Hausner *et al.*, 2014). Group I introns were found to be relatively rare in bacteria. *Simikania negevensis* and *Coxiella burnetii*, both of which are obligate intracellular pathogens, are suggested to have obtained group I introns in their 23S rRNA genes from eukaryotic host cells (Nesbø and Doolittle, 2003). In archaea, group I introns were found in multiple archaeal phyla, showing they are broadly distributed across the archaeal domain (Nawrocki *et al.*, 2018). In *Candidatus Velamenicoccus archaeovorus*, a group I intron is inserted into the 23S rRNA, encoding a LAGLIDADG/HNH homing endonuclease. This enzyme facilitates the mobility of the intron by inducing double-strand breaks in intron-less genes, promoting the spread of the intron through homologous recombination (Stoddard, 2005; Kizina, 2017).

In this study, the predatory “lifestyle” of *Candidatus Velamenicoccus archaeovorus* is investigated based on transcriptomic data as well as high-resolution micrographs and FISH. The hypothesis is proposed that the transcriptome of attached *Candidatus Velamenicoccus archaeovorus* cells differs from that of free-living cells, with differentially expressed genes

encoding proteins involved in predatory activity. Of particular interest was confirming a previous finding that the bacterial intron can be detected in archaeal cells.

4.2 Material and Methods

Microscopic techniques

For cell identification and visualization, catalyzed reporter deposition-fluorescence *in situ* hybridization (CARD-FISH) was used (Pernthaler *et al.*, 2004). To increase signal brightness, helper oligonucleotides were designed and added for each side of the HRP-labelled OP3-565 probe. The probes and helper concentrations in the working solution was ~8,4 pmol/μl.

Table 1: Oligonucleotide probes used for CARD-FISH, targeting *Candidatus Velamenicoccus archaeovor*, *Archaea*, and *Bacteria*. Horseradish peroxidase (HRP).

Name	Sequence (5'–3')	Position	Target	FA (%)	Detection	Reference
OP3-565	TACCTGCCCTTT ACACCC	608–626	Candidate OP3 LiM	30	HRP	(Rotaru <i>et al.</i> , 2012)
ARCH-915	GTGCTCCCCCGC CAATTCCT	915–934	<i>Archaea</i>	35	HRP	(Stahl and Amann, 1991)
MX825	TCGCACCGTGGC CGACACCTAGC	825–847	<i>Methanosaetaceae</i>	50	Atto488	(Raskin <i>et al.</i> , 1994)
EUB338 I	GCTGCCTCCCGT AGGAGT	338–355	Most <i>Bacteria</i>	35	HRP	(Amann <i>et al.</i> , 1990)
EUB338 II	GCAGCCACCCGT AGGTGT	338–355	<i>Planctomycetales</i>	35	HRP	(Daims <i>et al.</i> , 1999)
EUB338 III	GCTGCCACCCGT AGGTGT	338–355	<i>Verrucomicrobiales</i>	35	HRP	(Daims <i>et al.</i> , 1999)
H548-A	AATAAATCCGA GTAACGC	590–608	Candidate OP3 LiM (helper)			(Kizina, 2017)
H548-C	AATCAATCCGA GTAACGC	590–608	Candidate OP3 LiM (helper)			(Kizina, 2017)
H583-TC	CTCCCCACTTGT CAGGCCGCC	626–647	Candidate OP3 LiM (helper)			(Kizina, 2017)
H583-CT	CCTCCCCACTTGT CAGGCCGCC	626–647	Candidate OP3 LiM (helper)			(Kizina, 2017)

Table 2: Oligonucleotide probes used for CARD-FISH and FISH, targeting the group I intron in *Candidatus Velamenicoccus archaeovor*.

Name	Sequence (5'–3')	Position	Target	FA (%)	Detection	Reference
hen1-2235	CCGCCAAGTAGTAG	2235–	Intron in 23S rRNA	20	HRP	(Kizina, 2017)
	CCGATT	2255	of OP3 LiM			
hen2-2309	AACTTTCCACGGTG	2309–	Intron in 23S rRNA	20	HRP or Atto594	(Kizina, 2017)
	ACTTGT	2329	of OP3 LiM			
hen3-2538	TCTTTATCATTCTG	2538–	Intron in 23S rRNA	20	HRP	(Kizina, 2017)
	CCAGTTCG	2561	of OP3 LiM			
hen2-rc2309	ACAAGTCACCGTGG	2309–	Reverse complement	20	Atto594	(Kizina, 2017)
	AAAGTT	2329	of hen2			
hen2-2309+4bp	AACTTTCCA C CGG	2309–	hen2 + 4 random	20	Atto594	(Kizina, 2017)
	TGA A CTTGTA	2329	basepairs			

1 mL of the enrichment culture was fixed with formaldehyde (1.3% [wt/vol]) for 60 min at room temperature. 1-5 μ L of fixed culture were dried on a glass slide and, after drying, rinsed carefully with 1 ml of 1 \times PBS (pH 7.4). Cells were embedded in 0.1 % [wt/vol]) low-melting agarose to reduce cell loss during the washing steps. After air drying, permeabilization was performed for 60 min at 37°C by lysozyme treatment (10 mg/mL) and proteinase K treatment (200 μ g/ml) for 2 minutes at room temperature. Endogenous peroxidases were deactivated by incubating the samples in 0.1 M HCl for 1 min at RT, washing them with 1x PBS, and incubating them in 3% H₂O₂ for 10 minutes at RT. The probes were hybridized for 160 minutes at 46 °C, and for CARD-FISH, signals were amplified by incubating with tyramides (Alexa-594 for red and Alexa-488 for green) at 46°C for 45 min in the dark, and then the preparation was stained with DAPI (1 μ g/mL). To visualize lipids and other hydrophobic compounds, the cells were stained with Nile Red (5 μ g/ml) in PCR water for 15 minutes and washed 2x for 10 minutes with MilliQ water. Cells were embedded in Citifluor-Vectashield (4:1 [vol/vol]). Images were obtained with superresolution structured illumination microscopy (SR-SIM) using a confocal laser scanning microscope (CLSM; Zeiss LSM 780, Zeiss, Oberkochen, Germany) or an epifluorescence microscope (Nikon Eclipse 50i; Nikon, Tokyo).

Transcriptome analysis

To separate the aggregated cells containing the attached *Candidatus Velamenicoccus archaeovorus* from the small free-living cells, differential centrifugation was performed using a SW28 ultracentrifuge rotor (Beckman, Palo Alto, CA). The majority of cells were collected in a 10 kS pellet by centrifugation at 7,800 rpm for 20 min (10,000 S). The supernatant was centrifuged again at 27,000 rpm (96,467x g) for 160 min (100S) and should contain all cells present in the supernatant. This pellet was resuspended in 0.5 ml TE and centrifuged at 12,400 rpm (16,331x g) for 3 min (10,000 S), resulting in a pellet of 100S aggregates, which contained the attached *Candidatus Velamenicoccus archaeovorus* cells. The 100S cells in the supernatant contained very small cells, most of which were the free-living *Candidatus Velamenicoccus archaeovorus* cells. RNA and DNA were extracted using the RNA PowerSoil® Total RNA Isolation Kit and the RNA PowerSoil® DNA Elution Accessory Kit (Qiagen, Hilden, Germany). NGS sequencing was performed by the Max Planck-Genome-centre Cologne, Germany (<https://mpgc.mpipz.mpg.de/home/>). RNA was analyzed using Illumina HiSeq (150 bp reads) (Illumina, San Diego, CA, SRA accession SRR30230645 Bioproject PRJNA1129558). Large DNA reads representing the 100kS cells were obtained with HiFi reads on a PacBio Sequel (Pacific Biosciences, Menlo Park, CA, SRA accession SRR29679608, Bioproject PRJNA1129558). For metatranscriptomic analyses, rRNA reads were removed using SortMeRNA v4.3.4 (Kopylova *et al.*, 2012). The non-rRNA reads were trimmed with quality 30 using bbdduk within bbmap v38.93 (<https://sourceforge.net/projects/bbmap/>) and then mapped with Bowtie2 v2.4.4 (Langmead and Salzberg, 2012) onto the genes of proteins annotated using Bakta v1.4.2 with its database v3.1 for bacterial genomes (Schwengers *et al.*, 2021) present in the 10kS cell metagenome. Transcripts per million (TPM) were calculated in R using the total mapped reads to a gene. The fold change was calculated by either dividing the attached cells TPM (100S aggregates) by the free-living cells (100S cells) or the reverse direction.

4.3 Results and Discussion

4.3.1 Structured illumination micrographs show *Candidatus Velamenicoccus archaeovor* attached to morphologically different cells

Catalyzed reporter deposition-fluorescence in situ hybridization (CARD-FISH) targeting *Candidatus Velamenicoccus archaeovor* revealed small coccoid-shaped cells either free-living or attached to other cells in fluorescent microscopic and confocal laser scanning images (Kizina *et al.*, 2022). To obtain additional information, structured illumination microscopy was applied. It provides a spatial resolution of approximately 100 nanometers (Gustafsson, 2005). Using the CARD-FISH probe OP3-565 (Table 1), *Candidatus Velamenicoccus archaeovor* was imaged attaching to morphologically different cells. Large filaments were annotated as *Methanotherix* filaments based on their morphology and hybridization with probe MX825 (Table 1), which targets *Methanosaetaceae* (Figure 1 A, Supplementary figure 3 and 4). *Candidatus Velamenicoccus archaeovor* formed clusters around round-shaped large cells (Figure 1 B and Supplementary figure 1 D-F) and attached also to thin filaments. These long small filaments visible in Figure 1 C (Supplementary figure 1 G-I and Supplementary figure 5) can be assigned to *Anaerolinaceae* based on the morphology of described species and the presence of *Anaerolinaceae* genomes LiM8, LiM9, LiM11 and LiM29 (McIlroy *et al.*, 2017).

Images of the CARD-FISH and DAPI-stained cells display the clusters of small coccoid cells surrounding larger round cells, best seen in Figure 1 B (Supplementary figure 1). The CARD-FISH probe signals (Figure 1 and Supplementary figure 1 B, E, and H) confirm the identity of these small cells as *Candidatus Velamenicoccus archaeovor*. Notably, the CARD-FISH signals of the cells interacting with their prey are brighter compared to the signals from cells not in direct contact with other bacteria, indicating a higher ribosome content. The large cells have been identified as bacteria by bacterial probe EUB338 I-III (Kizina, 2017). A notable feature of the 16S rRNA of *Candidatus Velamenicoccus archaeovor* is a sequence that does not hybridize in FISH experiments with EUB338 I-III (Rotaru *et al.*, 2012). A combination of the CARD-FISH probes Arch-915 and EUB338 I-III (Table 1) amplified with the same tyramide demonstrated in combination with OP3-565 that the large cell surrounded by *Candidatus Velamenicoccus archaeovor* is either an archaeon targeted by Arch-915 or bacteria targeted by EUB338 I-III (Figure 2).

The large number of *Candidatus Velamenicoccus archaeovorus* cells around the central cell suggests that the cells are actively proliferating. Proliferation requires many resources and metabolic activity, which usually results in a larger ribosome content. CARD-FISH signals of the attached cells are brighter than the signals of cells that are not attached. This indicates that attached cells have higher ribosome contents, which could be due to higher metabolic activity. Genome analysis of *Candidatus Velamenicoccus archaeovorus* revealed that it also has the ability to live syntrophic on either hydrogen or formate, indicating that its attachment could also be syntrophic or more complex depending on the environment and available resources (Kizina, 2017).

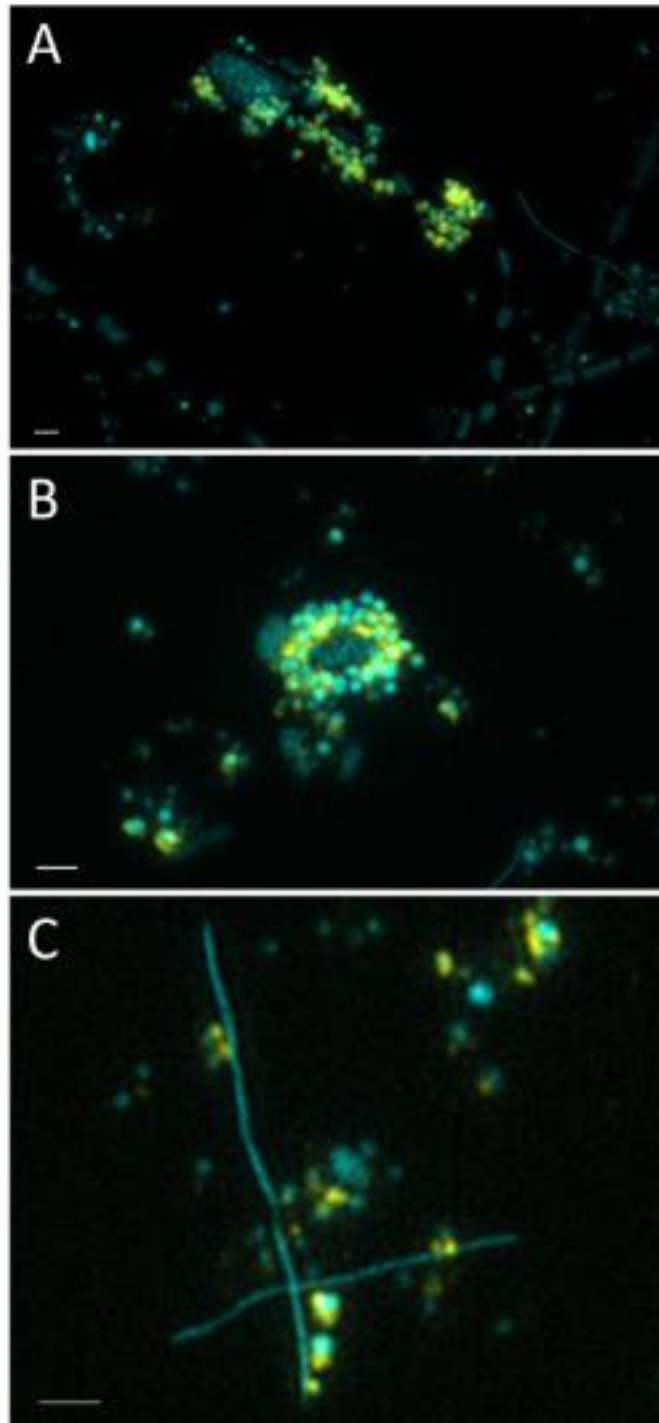


Figure 1: Structured illumination micrograph of the limonene-degrading methanogenic enrichment culture visualized as a maximum intensity projection. This overlay shows signals of OP3-565 CARD-FISH hybridized cells that identify the cells as *Candidatus Velamenicoccus archaeovor* (yellow) and DAPI-stained cells (cyan). The scale bar is 1 µm. Individual channels are shown in Supplementary figure 1.

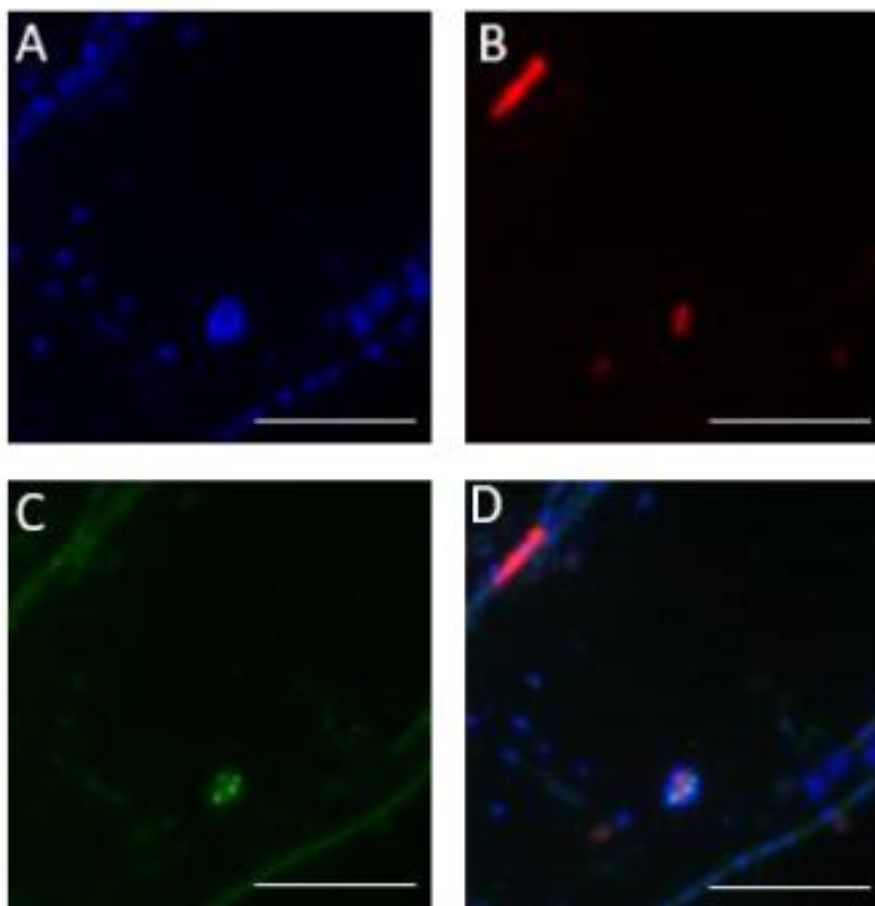


Figure 2: Epifluorescence micrograph of the limonene-degrading methanogenic enrichment culture. (A) DAPI stained cells. (B) Signals of EUBI-III and ARCH-915 CARD-FISH hybridized cells, which identify cells as *Archaea* or *Bacteria*. (C) Signals of OP3-565 CARD-FISH hybridized cells identify the cells as *Candidatus Velamenicoccus archaeovor*. (D) Overlay. The scale bar is 10 μm .

4.3.2 Structured illumination micrographs of Nile Red stained cells

Nile Red is a stain for hydrophobic compounds, including lipids. This stain reveals cell structures, like membranes, in more detail, revealing the cell's volume. Morphologically distinct cells could be visualized, revealing the complexity of the limonene-degrading methanogenic enrichment culture (Supplementary figure 7). *Methanotherix* filaments can be identified by their characteristic morphology, while small coccoid-shaped cells, consistent with the morphology of *Candidatus Velamenicoccus archaeovor*, were also visible and observed attached to other cells (Supplementary figure 5, 6, and 7).

In *Methanotherix*, the filament, as well as the spacer plugs, structures that separate the cells within the filament, are stained by the Nile Red (Figure 3, Supplementary figure 2 and 7), indicating that these structures also contain larger amounts of hydrophobic compounds.

Candidatus Velamenicoccus archaeovorus cells or cell assemblages are attached to the filaments, showing a strong OP3-565 signal and strong Nile Red staining (Figure 3, Supplementary figure 2 and 3). This indicated that these attached cells are metabolically active and predated on the filament. The strong Nile Red staining is indicative of a large amount of lipids or other hydrophobic compounds. More *Candidatus Velamenicoccus archaeovorus* cells with weaker CARD-FISH signals were also attached to the filament and *Candidatus Velamenicoccus archaeovorus* cells in the free-living state were detected with even weaker signals (Figure 3 and Supplementary figure 2 and 3). The Nile Red stained area in the images was larger than the probe-stained area of the cells. This suggested that the tyramide dye was preferentially labeling intracellular targets.

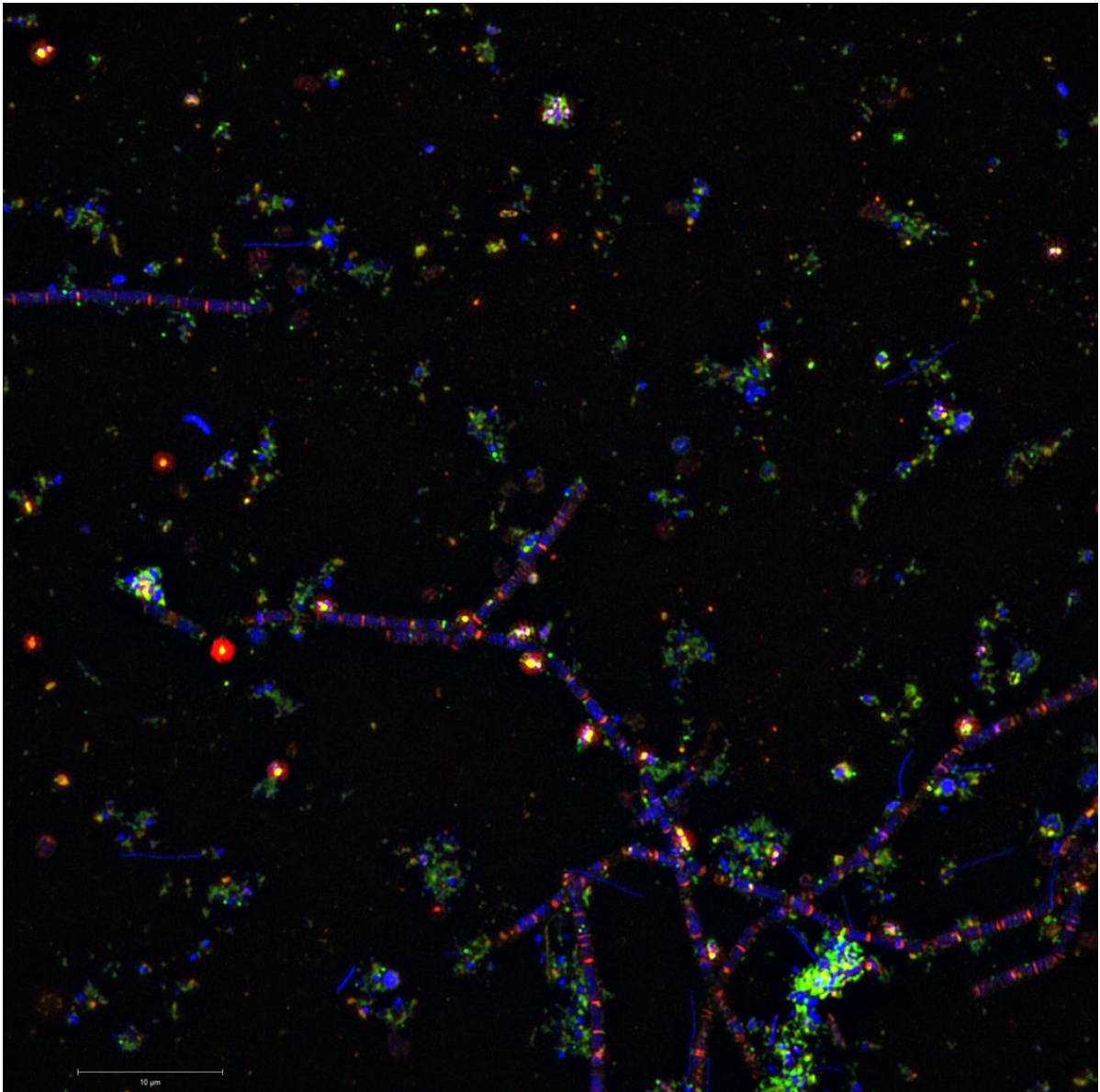


Figure 3: Structured illumination micrograph of the limonene-degrading methanogenic enrichment culture visualized as a maximum intensity projection. Nile Red stained hydrophobic compounds (red), signals of OP3-565 CARD-FISH hybridized cells, which identify the cells as *Candidatus Velamenicoccus archaeovor* (green) and DAPI stained cells (blue), are shown in this overlay. The scale bar is 10 μm. Individual channels are shown in Supplementary figure 2.

4.3.3 SR-SIM images visualize the presence of a predator group I intron molecule in prey cells

The group I intron containing a homing endonuclease inserted in the 23S rRNA of the *Candidatus Velamenicoccus archaeovor* was targeted using CARD-FISH probes hen2-2309 (Figure 4 and Supplementary figure 8) as well as hen1-2235 and hen3-2538 (Figure 6 and Supplementary figure 9). The probe hen2-rc2309 is the reverse complement of the probe hen2-2309 (Table 2), which also targets the intron and was used as a control (Figure 6 and Supplementary figure 9). The probe hen2-2309 (Table 2) showed clear signals within cells of *Methanothrix* filaments as well as in other cells within the community (Figure 4 and Supplementary figure 8). Some of the *Methanothrix* cells with a hen2-2309 had *Candidatus Velamenicoccus archaeovor* cells attached to them. The modified probe, hen2-2309+4bp (Table 2), with four additional random base pairs, was initially not expected to bind the target site. However, signals were still observed in some cells hybridized with hen2-2309+4bp (Figure 5 and Supplementary figure 10). These unexpected signals may result from partial binding of the probe, possibly influenced by secondary structures at the intron target site.

The probes hen1-2235 and hen3-2538 (Table 2) target different regions of the intron and also revealed signals within cells of *Methanothrix* filaments, which have small cells attached to them (Figure 6 and Supplementary figure 9). The cells positive with the hen1-2235 and hen3-2538 signals did not show any signals of the probe hen2-rc2309, the reverse complement of hen2-2309 (Figure 6 and Supplementary figure 9). This indicates that the probes targeting the intron specifically bind to the intron's sequence, and the observed signals are not due to nonspecific binding of the probes. Furthermore, the whole cell is stained, indicating that multiple copies of the intron are present within the cells within *Methanothrix* filaments. Possibly, these are cells to which *Candidatus Velamenicoccus archaeovor* has attached and into which the intron was passed. An alternative explanation is the uptake of extracellular intron RNA, however, this process likely needs energy and intact cells. The occurrence of the intron primarily in dead cells indicated a strong link to the physical interaction during predation.

Mobile genetic elements, specifically group I introns, have been described to have a potential for horizontal gene transfer (Hausner *et al.*, 2014). Spreading into new hosts is a hallmark of introns. During the attack on *Methanothrix* cells by *Candidatus Velamenicoccus archaeovor*, some cells may die, potentially hindering the intron's ability to insert itself into new sequences. However, since the intron can self-splice, it might still find a new host outside the dead cells.

There is no evidence that intron RNA manipulates host metabolism, as it leaves the host's genes functional. Currently, there is no support for a role of introns in the predation mechanism of *Candidatus Velamenicoccus archaeovorus*. Although it is known that non-coding RNA molecules can have an inhibitory action on the prey metabolism by modulating gene expression or interfering with essential metabolic pathways. These non-coding RNA molecules are also present in extracellular RNA (Ghosal *et al.*, 2015; Pita *et al.*, 2020; Tosar *et al.*, 2021; Borniego and Innes, 2023).

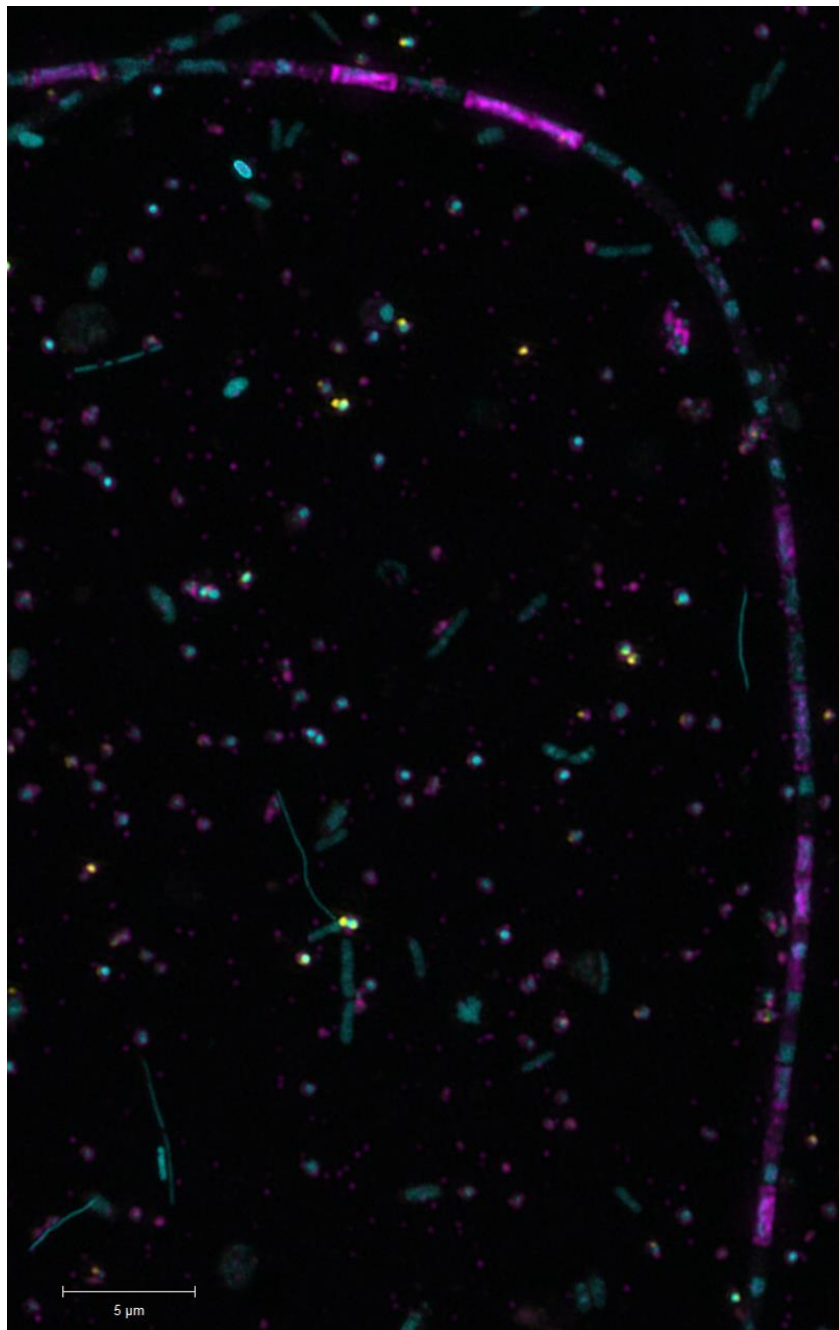


Figure 4: Structured illumination micrograph of the limonene-degrading methanogenic enrichment culture visualized as a maximum intensity projection. Signals of hen2-2309 -CARD-FISH, targeting the group I intron found in the 23S rRNA of OP3 LiM (magenta), signals of OP3-565 CARD-FISH hybridized cells which identify the cells as *Candidatus Velamenicoccus archaeovor* (yellow) and DAPI stained cells (cyan) are shown in this overlay. The scale bar is 5 μm. Individual channels are shown in Supplementary figure 8.

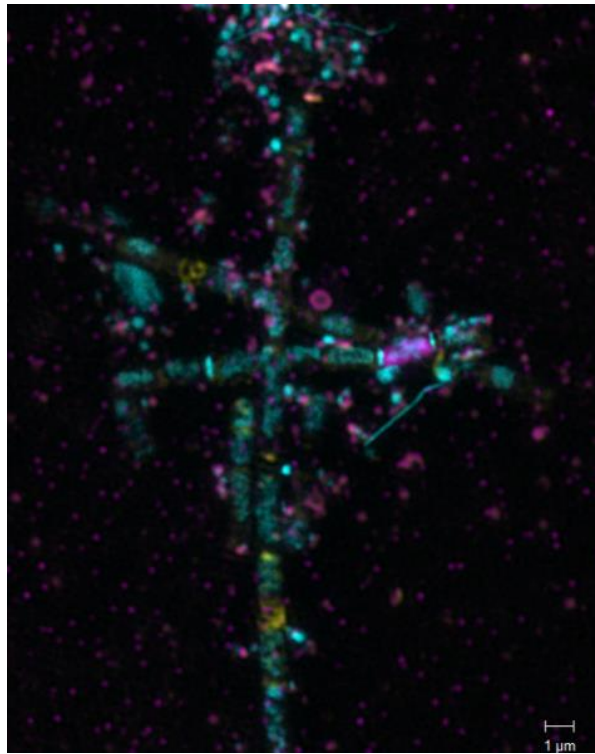


Figure 5: Structured illumination micrograph of the limonene-degrading methanogenic enrichment culture visualized as a maximum intensity projection. Signals of hen2-2309+4bp signals, probe targeting the intron in the 23S rRNA of OP3 LiM with four random basepairs (magenta), signals of EUBI-III and ARCH-915 CARD-FISH hybridized cells, which identify cells as *Archaea* or *Bacteria* (yellow) and DAPI stained cells (cyan) are shown in this overlay. The scale bar is 1 μm. Individual channels are shown in Supplementary figure 10.

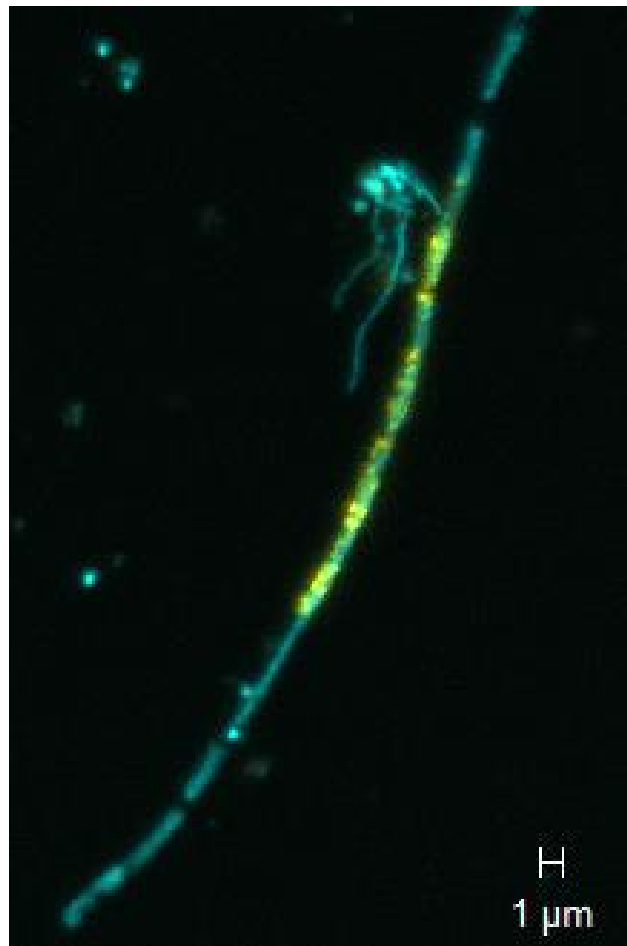


Figure 6: Structured illumination micrograph of the limonene-degrading methanogenic enrichment culture visualized as a maximum intensity projection. Signals of hen1-2235 and hen3-2538 CARD-FISH signals, targeting the intron in the 23S rRNA of OP3 LiM (yellow), hen2-rc2309 reverse complement signal (magenta), and DAPI stained cells (cyan) are shown in this overlay. The scale bar is 1 μm. Individual channels are shown in Supplementary figure 9.

4.3.4 Transcriptomic analysis of *Candidatus Velamenicoccus archaeovorus*

Transcriptomes of the 100kS pellet with large cells and attached small cells and of the 100S cells comprising small cells were compared. The datasets comprised 1239904 (100kS pellet) and 888607 (100S cells) reads, of which 796976 (100kS pellet) and 119346 (100S cells) mapped to protein sequences in the genome of *Candidatus Velamenicoccus archaeovorus*, respectively.

Free-living *Candidatus Velamenicoccus archaeovorus*

The metatranscriptomic analysis of the 100S cells fraction containing *Candidatus Velamenicoccus archaeovorus* in a free-living state revealed 203 proteins with at least two-fold higher expression level. The annotation of these genes revealed a participation in a variety of stress responses.

Nutrient acquisition and membrane transport by beta-barrel outer membrane pore

Free-living *Candidatus Velamenicoccus archaeovorus* cells transcribe an 18-strand beta-barrel outer membrane pore almost 4 times as much as attached cells (Figure 7). These proteins are often substrate-specific. The OpdK protein from *Pseudomonas aeruginosa* has an 18-stranded beta-barrel with a pore size of approximately 8 Å and is involved in the transport of small acids (Biswas *et al.*, 2008). In *Myxococcus xanthus*, an 18-strand β-barrel pore is involved in exporting large molecules like polysaccharides from the periplasm to the extracellular environment, which is important for biofilm formation and surface motility (Schwabe *et al.*, 2022). In *Escherichia coli* and *Pseudomonas aeruginosa*, outer membrane proteins play a crucial role in adhesion to epithelial cells. Similarly, *Acinetobacter baumannii*, a pathogen responsible for pneumonia, uses outer membrane pores to invade epithelial cells (Krishnan and Prasadarao, 2012).

Although these examples involve bacteria that are pathogenic to humans and attack epithelial cells, they illustrate that outer membrane pores can be involved in adhesion and invasion processes. In free-living *Candidatus Velamenicoccus archaeovorus*, the increased transcription of these pores might allow the bacterium to quickly attach to prey bacteria. Additionally, these outer membrane pores could be used to acquire nutrients from the environment, which is important for survival during periods of nutrient deprivation.

The upregulated membrane-associated protein with peptidoglycan-binding domains OmpA and LysM is also involved in nutrient acquisition. The outer membrane protein A (OmpA) is an eight-stranded beta-barrel protein that has been studied in Gram-negative bacteria and is known for its role in multiple processes, including membrane stability, biofilm formation, and infecting other cells (Smith *et al.*, 2007). OmpA stabilizes the bacterial outer membrane by anchoring it to the peptidoglycan layer, resulting in stability of the cell and protection against environmental stresses (Park *et al.*, 2012; Samsudin *et al.*, 2016; Paulsson *et al.*, 2021). In *Acinetobacter baumannii*, OmpA facilitates biofilm formation and is indicated to be a factor involved in adherence to epithelial cells as well as fungal filaments (Gaddy *et al.*, 2009). In *Escherichia coli*, OmpA has been linked to virulence by assisting bacterial adhesion to and invasion of host cells (Wu *et al.*, 2009). These studies show the versatility of OmpA and its upregulation in free-living *Candidatus Velamenicoccus archaeovor* could improve its cell's stability, making it more resistant to environmental stressors. It could also be involved in the attachment to other cells, as free-living cells would be ready to attach quickly to their prey when they find it. In pathogenic bacteria, LysM domains can bind to the peptidoglycan of host cell walls and act as virulence factors, as it has been described in pathogens, e.g., *Staphylococcus aureus* (Buist *et al.*, 2008; Mesnage *et al.*, 2014).

Sensing the environment and stress response

A histidine kinase and an anti-sigma factor antagonist showed a 3- to 4-fold upregulation in expression in free-living *Candidatus Velamenicoccus archaeovor* (Figure 7). Proteins like the histidine kinase are involved in signal transduction, sensing, and adapting to different factors in the environment (Bhate *et al.*, 2015). The anti-sigma factor antagonist plays a role in regulating stress responses in Gram-negative bacteria by modulating the activity of sigma factors, which control the expression of stress-response genes. These antagonists bind to anti-sigma factors, releasing sigma factors to activate transcription of genes (Bastiaansen *et al.*, 2017; Bishop *et al.*, 2017). A protein annotated as sporulation initiation phosphotransferase F and likely shifting cells into dormancy is upregulated in free-living *Candidatus Velamenicoccus archaeovor*. While sporulation is traditionally associated with Gram-positive bacteria, e.g., *Bacillus subtilis*, recent studies indicate that some Gram-negative bacteria are also able to form spore-like structures or enter dormancy when environmental conditions become unfavorable (Stephenson and Hoch, 2002; Tocheva *et al.*, 2013). For example, the predatory Gram-negative bacterium *Myxococcus xanthus* is known to form stress-resistant structures called myxospores in response to nutrient limitation or other environmental stressors (Kaiser *et al.*, 2010). Studies on *Enterobacter* species, *Klebsiella pneumoniae*, and

Escherichia coli have shown that these bacteria can enter a reversible dormant state in response to stress, such as osmotic pressure (Sachidanandham and Yew-Hoong Gin, 2009). Therefore, the upregulation of the sporulation initiation phosphotransferase F in free-living *Candidatus Velamenicoccus archaeovorus* cells may indicate its involvement in a stress response mechanism, potentially contributing to dormancy as a response to a nutrient-limited environment. The anti-sigma factor antagonist, histidine kinase, as well as the sporulation initiation phosphotransferase F upregulated in the free-living cells underline their stressed state, likely due to nutrient limitation.

The upregulation of the site-specific DNA-methyltransferase (Figure 7), suggests that DNA is methylated to protect it, which could be beneficial for survival under nutrient-deprived conditions. Methylation protects bacterial DNA from oxidative stress or other environmental damage, ensuring genetic stability. It can also change the expression pattern of genes, letting the cells adapt to changing environmental conditions (Nye *et al.*, 2020; Ramaloko and Osei Sekyere, 2022).

Ribosome recycling factors, also upregulated in the free-living *Candidatus Velamenicoccus archaeovorus* (Figure 7), are involved in the final step of protein synthesis, known as ribosome recycling. Without ribosome recycling, ribosomes would remain bound to the mRNA and tRNA, wasting cellular resources and energy. By facilitating the release of mRNA and tRNA, ribosome recycling factors help cells avoid unnecessary translation and make ribosomes available for new protein synthesis. Under conditions where nutrients are limited these factors might contribute to the survival of the cells by saving resources (Kiel *et al.*, 2007).

The upregulation of phage tail fiber protein and phage minor tail protein in free-living predatory bacteria may reflect an adaptive response, possibly linked to a secretion system that shares structural similarities with bacteriophage tails. These proteins might assist as attachment molecules in predation. This upregulation may help the bacterium remain prepared for interactions with prey (Pukatzki *et al.*, 2007).

The filamentation induced by cAMP protein Fic, also upregulated in free-living cells, has been shown to be involved in a variety of processes. In pathogenic bacteria, they are involved in disarming host defenses by modifying host GTPases through AMPylation. They are also part of toxin-antitoxin systems, which are stress response mechanisms in bacteria. These systems

allow the bacteria to enter a dormant state to survive harsh conditions (Cruz and Woychik, 2014).

Overall, the upregulated expression of these proteins coincides with the 'dormancy' of free-living cells, as observed by decreased ribosome content in FISH images. These 'free-living' cells are not attached to prey and have to survive in a nutrient-deprived state by reducing their metabolic activity, protecting their DNA as well as sensing their environment for nutrients or prey. They also need to remain prepared to attack prey when it becomes available and transport nutrients into the cell.

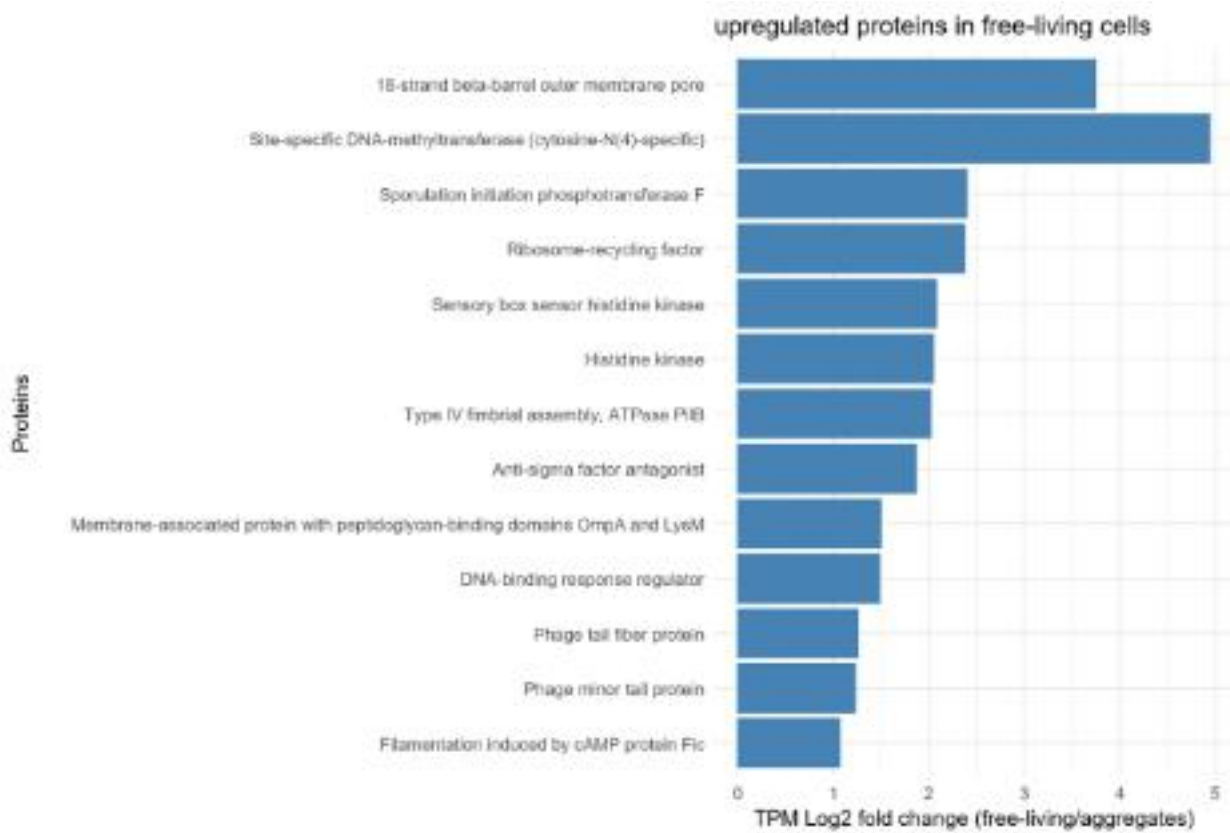


Figure 7: Logarithmic fold change of transcripts per million (TPM) of free-living vs. attached *Candidatus Velamenicoccus archaeovor* cells.

Attached-living *Candidatus Velamenicoccus archaeovor*

Attachment and motility

In *Candidatus Velamenicoccus archaeovor*, the upregulated proteins involved in type IV pili assembly are possibly involved in attaching to the prey, even inserting components into the prey. In *Bdellovibrio bacteriovorus*, type IV pili are involved in predation by helping the predator attach to prey cells (Chanyi and Koval, 2014). Proteins such as PilE, PilM, and PilO are essential for pilus biogenesis, enabling the formation and extension of type IV pili. PilT, an ATPase also present in *Candidatus Velamenicoccus archaeovor*, allows the retraction of pili, enabling bacteria to pull themselves towards their target, a process known as “twitching motility” (Mattick, 2002; Giltner *et al.*, 2010). Prepilin-type cleavage proteins process precursor pilins, preparing them for integration into functional pili (Dupuy and Pugsley, 1994). In pathogens, type IV pili are involved in adhesion to host tissues and can also be cytotoxic, potentially delivering toxins to host cells via secretion systems (Nudleman and Kaiser, 2004).

In attached *Candidatus Velamenicoccus archaeovor*, upregulated proteins annotated as stator proteins MotB, MotD, and the M-ring protein are usually involved in the anchoring the bacterial flagellum to the cell membrane (Figure 8). However, in non-flagellated bacteria, these proteins can also facilitate ion transport and may contribute to other membrane-associated processes, potentially facilitating nutrient transport (Takekawa *et al.*, 2021).

Proteins involved in secretion systems, like alkaline protease secretion protein AprF, type II/IV secretion system ATP hydrolase Tada/VirB11/CpaF, Tada subfamily, type II secretion system core protein G, and protein-export membrane protein SecG are also upregulated in attached *Candidatus Velamenicoccus archaeovor* (Figure 8). In *Pseudomonas aeruginosa*, a secreted protease contributes to its virulence by degrading host proteins (Filloux, 2011). The cell-contact-dependent killing of *Myxococcus xanthus* relies on a combination of Tad- and type III-like protein secretion systems, which work together to induce prey cell death and lysis (Thiery *et al.*, 2022). In attached *Candidatus Velamenicoccus archaeovor* cells, these upregulated proteins could be involved in similar mechanisms, such as attaching to prey as well as secreting compounds into the prey cell.

The 23S rRNA is upregulated in the attached fraction of cells, likely due to the presence of a group I intron within its sequence, which includes a homing endonuclease. This intron-containing sequence was not recognized as rRNA during bioinformatics analysis and, therefore, was not filtered out. The upregulation of the 23S rRNA may be attributed to the presence of the intron in *Candidatus Velamenicoccus archaeovorus* and its prey, e.g., *Methanothrix*, leading to a higher number of transcripts mapping to the 23S rRNA region. These mobile genetic elements are ribozymes capable of excising themselves from RNA transcripts. Given that the intron has been observed in prey cells, it might play a role in a predatory mechanism (Nielsen and Johansen, 2009).

Furthermore, proteins involved in cell wall degradation are upregulated as well in attached *Candidatus Velamenicoccus archaeovorus*, like the N-acetylmuramoyl-L-alanine amidase, polysaccharide deacetylase, Peptidoglycan D,D-transpeptidase MrdA and Peptidoglycan-N-acetylglucosamine deacetylase. The N-acetylmuramoyl-L-alanine amidase is known to be able to degrade peptidoglycan (Vermassen *et al.*, 2019), and the polysaccharide deacetylase ensures the proper localization and function of the Type IV secretion in *Legionella pneumophila* allowing it to manipulate host cellular processes (Boamah *et al.*, 2023). *Bdellovibrio bacteriovorus* uses a peptidoglycan-D,D-transpeptidase, to modify its prey's cell wall, making it stable for it to replicate within the prey's periplasm (Kuru *et al.*, 2017). In *Streptococcus pneumoniae*, the peptidoglycan N-acetylglucosamine deacetylase modifies the bacterial cell wall's peptidoglycan, making it more resistant to lysozyme degradation, a possible defense mechanism of the prey cell (Vollmer and Tomasz, 2002).

Outer membrane efflux proteins are part of a tripartite exporter of cytosolic compounds. Its primary task is to transport any dangerous compound out of the cell (Kumawat *et al.*, 2023). This makes it a potential defense mechanism for *Candidatus Velamenicoccus archaeovorus* (Figure 8). It could also transport compounds to the outside, which harm or weaken the prey. The ABC transporter ATP-binding protein is involved in transporting a variety of substrates, including substrates, across the cytoplasmic membrane (Akhtar and Turner, 2022). These could play a role in acquiring nutrients, which then allows *Candidatus Velamenicoccus archaeovorus* to proliferate.

The cell division protein FtsB is involved in cell division, indicating that *Candidatus Velamenicoccus archaeovorus* has enough nutrients and energy to proliferate, as can also be seen in the SR-SIM images (Figure 1) (Kureisaite-Ciziene *et al.*, 2018).

Overall, the attached *Candidatus Velamenicoccus archaeovor*us cells express proteins involved in cell membrane degradation, proteins for the uptake of nutrients, and proteins indicating proliferation and metabolic activity (Figure 8). This is consistent with SR-SIM images, which show brighter CARD-FISH signals for attached cells, reflecting their higher metabolic activity.

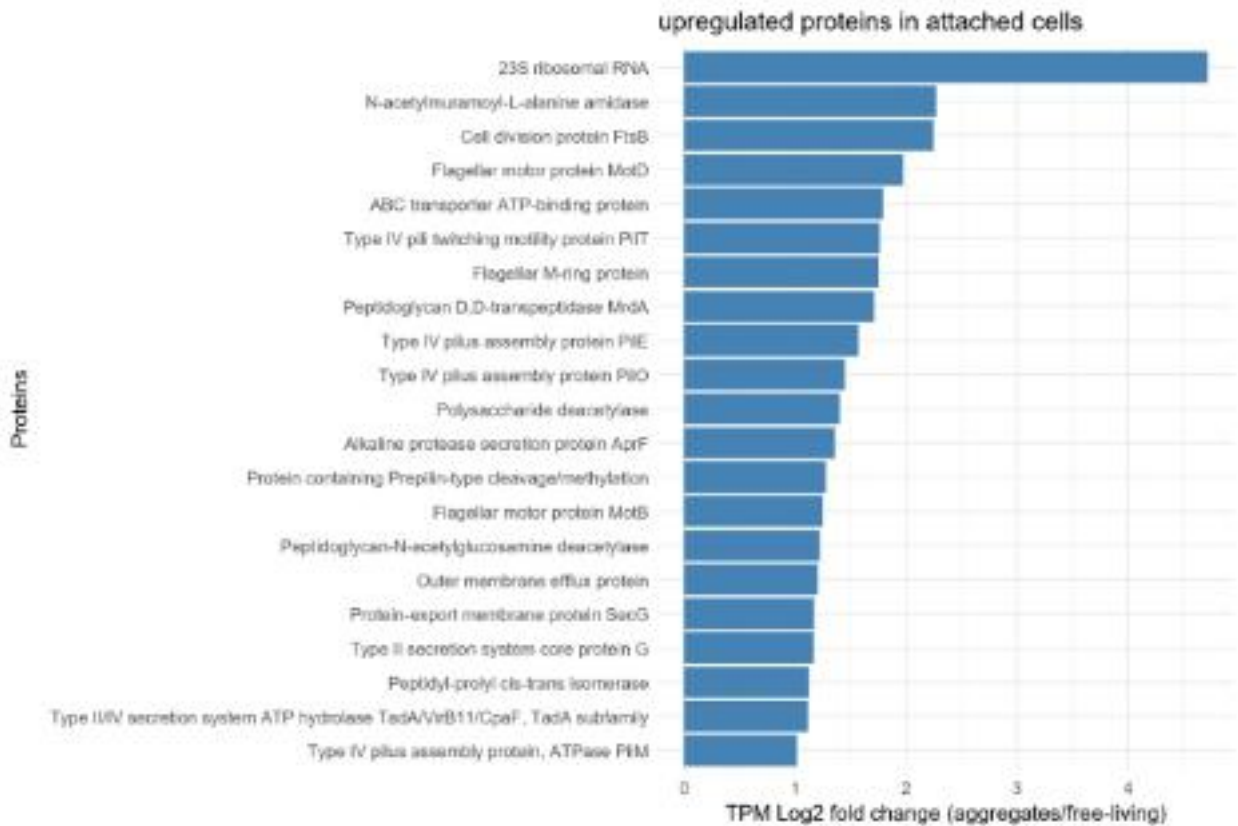


Figure 8: Logarithmic fold change of transcripts per million (TPM) of attached vs. free-living *Candidatus Velamenicoccus archaeovor*us cells

References

- Akhtar, A.A. and Turner, D.P.J. (2022) The role of bacterial ATP-binding cassette (ABC) transporters in pathogenesis and virulence: therapeutic and vaccine potential. *Microbial Pathogenesis* **171**: 105734.
- Amann, R.I., Krumholz, L., and Stahl, D.A. (1990) Fluorescent-oligonucleotide probing of whole cells for determinative, phylogenetic, and environmental studies in microbiology. *Journal of bacteriology* **172**: (2) 762–70.
- Barabote, R.D., Rendulic, S., Schuster, S.C., and Saier, M.H. (2007) Comprehensive analysis of transport proteins encoded within the genome of *Bdellovibrio bacteriovorus*. *Genomics* **90**: (4) 424–446.
- Bastiaansen, K.C., Civantos, C., Bitter, W., and Llamas, M.A. (2017) New insights into the regulation of cell-surface signaling activity acquired from a mutagenesis screen of the *Pseudomonas putida* IutY sigma/anti-sigma factor. *Frontiers in Microbiology* **8**: 747.
- Bhate, M.P., Molnar, K.S., Goulian, M., and DeGrado, W.F. (2015) Signal transduction in histidine kinases: insights from new structures. *Structure* **23**: (6) 981–994.
- Bishop, T.F., Martin, L.W., and Lamont, I.L. (2017) Activation of a cell surface signaling pathway in *Pseudomonas aeruginosa* requires ClpP protease and new sigma factor synthesis. *Frontiers in Microbiology* **8**: 2442.
- Biswas, S., Mohammad, M.M., Movileanu, L., and van den Berg, B. (2008) Crystal structure of the outer membrane protein OpdK from *Pseudomonas aeruginosa*. *Structure* **16**: (7) 1027–1035.
- Boamah, D., Gilmore, M.C., Bourget, S., Ghosh, A., Hossain, M.J., Vogel, J.P., et al. (2023) Peptidoglycan deacetylation controls type IV secretion and the intracellular survival of the bacterial pathogen *Legionella pneumophila*. *Proceedings of the National Academy of Sciences* **120**: (23) e2119658120.
- Borniego, M.L. and Innes, R.W. (2023) Extracellular RNA: mechanisms of secretion and potential functions. *Journal of Experimental Botany* **74**: (7) 2389–2404.
- Bratanis, E., Andersson, T., Lood, R., and Bukowska-Faniband, E. (2020) Biotechnological potential of *Bdellovibrio* and like organisms and their secreted enzymes. *Frontiers in Microbiology* **11**: 662.
- Bratanis, E. and Lood, R. (2019) A novel broad-spectrum elastase-like serine protease from the predatory bacterium *Bdellovibrio bacteriovorus* facilitates elucidation of site-specific IgA glycosylation pattern. *Frontiers in Microbiology* **10**: 971.
- Buist, G., Steen, A., Kok, J., and Kuipers, O.P. (2008) LysM, a widely distributed protein motif for binding to (peptido)glycans. *Molecular Microbiology* **68**: (4) 838–847.

- Caulton, S.G., Lambert, C., Tyson, J., Radford, P., Al-Bayati, A., Greenwood, S., et al. (2024) *Bdellovibrio bacteriovorus* uses chimeric fibre proteins to recognize and invade a broad range of bacterial hosts. *Nature Microbiology* **9**: (1) 214–227.
- Chanyi, R.M. and Koval, S.F. (2014) Role of type IV pili in predation by *Bdellovibrio bacteriovorus*. *PLoS One* **9**: (11) e113404.
- Cruz, J.W. and Woychik, N.A. (2014) Teaching Fido new ModIFICation tricks. *PLoS pathogens* **10**: (9) e1004349.
- Daims, H., Brühl, A., Amann, R., Schleifer, K.H., and Wagner, M. (1999) The domain-specific probe EUB338 is insufficient for the detection of all *Bacteria*: development and evaluation of a more comprehensive probe set. *Systematic and Applied Microbiology* **22**: (3) 434–44.
- Du Toit, A. (2024) *Bdellovibrio bacteriovorus* finds its prey. *Nature Reviews Microbiology* **22**: (3) 119–119.
- Dupuy, B. and Pugsley, A.P. (1994) Type IV prepilin peptidase gene of *Neisseria gonorrhoeae* MS11: presence of a related gene in other piliated and nonpiliated *Neisseria* strains. *Journal of Bacteriology* **176**: (5) 1323–31.
- Filloux, A. (2011) Protein secretion systems in *Pseudomonas aeruginosa*: an essay on diversity, evolution, and function. *Frontiers in Microbiology* **2**: 155.
- Frost, L.S., Leplae, R., Summers, A.O., and Toussaint, A. (2005) Mobile genetic elements: the agents of open source evolution. *Nature Reviews Microbiology* **3**: (9) 722–732.
- Gaddy, J.A., Tomaras, A.P., and Actis, L.A. (2009) The *Acinetobacter baumannii* 19606 OmpA protein plays a role in biofilm formation on abiotic surfaces and in the interaction of this pathogen with eukaryotic cells. *Infection and Immunity* **77**: (8) 3150–3160.
- Ghosal, A., Upadhyaya, B.B., Fritz, J.V., Heintz-Buschart, A., Desai, M.S., Yusuf, D., et al. (2015) The extracellular RNA complement of *Escherichia coli*. *MicrobiologyOpen* **4**: (2) 252–266.
- Giltner, C.L., Habash, M., and Burrows, L.L. (2010) *Pseudomonas aeruginosa* minor pilins are incorporated into type IV pili. *Journal of Molecular Biology* **398**: (3) 444–461.
- Glöckner, J., Kube, M., Shrestha, P.M., Weber, M., Glöckner, F.O., Reinhardt, R., and Liesack, W. (2010) Phylogenetic diversity and metagenomics of candidate division OP3. *Environmental Microbiology* **12**: (5) 1218–1229.
- Gophna, U., Charlebois, R.L., and Doolittle, W.F. (2006) Ancient lateral gene transfer in the evolution of *Bdellovibrio bacteriovorus*. *Trends in Microbiology* **14**: (2) 64–69.

- Gustafsson, M.G.L. (2005) Nonlinear structured-illumination microscopy: wide-field fluorescence imaging with theoretically unlimited resolution. *Proceedings of the National Academy of Sciences* **102**: (37) 13081–13086.
- Haugen, P., Simon, D.M., and Bhattacharya, D. (2005) The natural history of group I introns. *Trends in Genetics* **21**: (2) 111–119.
- Hausner, G., Hafez, M., and Edgell, D.R. (2014) Bacterial group I introns: mobile RNA catalysts. *Mobile DNA* **5**: (1) 8.
- Hugenholtz, P., Pitulle, C., Hershberger Karen, L., and Pace Norman, R. (1998) Novel division level bacterial diversity in a Yellowstone hot spring. *Journal of Bacteriology* **180**: (2) 366–376.
- Kåhrström, C.T. (2014) Multitasking in *Myxococcus*. *Nature Reviews Microbiology* **12**: (2) 73–73.
- Kaimer, C., Weltzer, M.L., and Wall, D. (2023) Two reasons to kill: predation and kin discrimination in myxobacteria. *Microbiology* **169**: (7).
- Kaiser, D., Robinson, M., and Kroos, L. (2010) Myxobacteria, polarity, and multicellular morphogenesis. *Cold Spring Harbor perspectives in biology* **2**: (8) a000380.
- Kiel, M.C., Kaji, H., and Kaji, A. (2007) Ribosome recycling: an essential process of protein synthesis. *Biochemistry and Molecular Biology Education* **35**: (1) 40–44.
- Kizina, J. (2017) Insights into the biology of Candidate Division OP3 LiM populations. *PhD Thesis*.
- Kizina, J., Jordan Sebastian, F.A., Martens Gerrit, A., Lonsing, A., Probian, C., Kolovou, A., et al. (2022) *Methanosaeta* and “*Candidatus Velamenicoccus archaeovor*”. *Applied and Environmental Microbiology* **88**: (7) e02407-21.
- Kopylova, E., Noé, L., and Touzet, H. (2012) SortMeRNA: fast and accurate filtering of ribosomal RNAs in metatranscriptomic data. *Bioinformatics* **28**: (24) 3211–7.
- Krishnan, S. and Prasadarao, N.V. (2012) Outer membrane protein A and OprF: versatile roles in Gram-negative bacterial infections. *The FEBS Journal* **279**: (6) 919–931.
- Kumawat, M., Nabi, B., Daswani, M., Viquar, I., Pal, N., Sharma, P., et al. (2023) Role of bacterial efflux pump proteins in antibiotic resistance across microbial species. *Microbial Pathogenesis* **181**: 106182.
- Kureisaite-Ciziene, D., Varadajan, A., McLaughlin, S.H., Glas, M., Silva, A.M., Luirink, R., et al. (2018) Structural analysis of the interaction between the bacterial cell division proteins FtsQ and FtsB. *mBio* **9**: (5) 10.1128/mbio.01346-18.

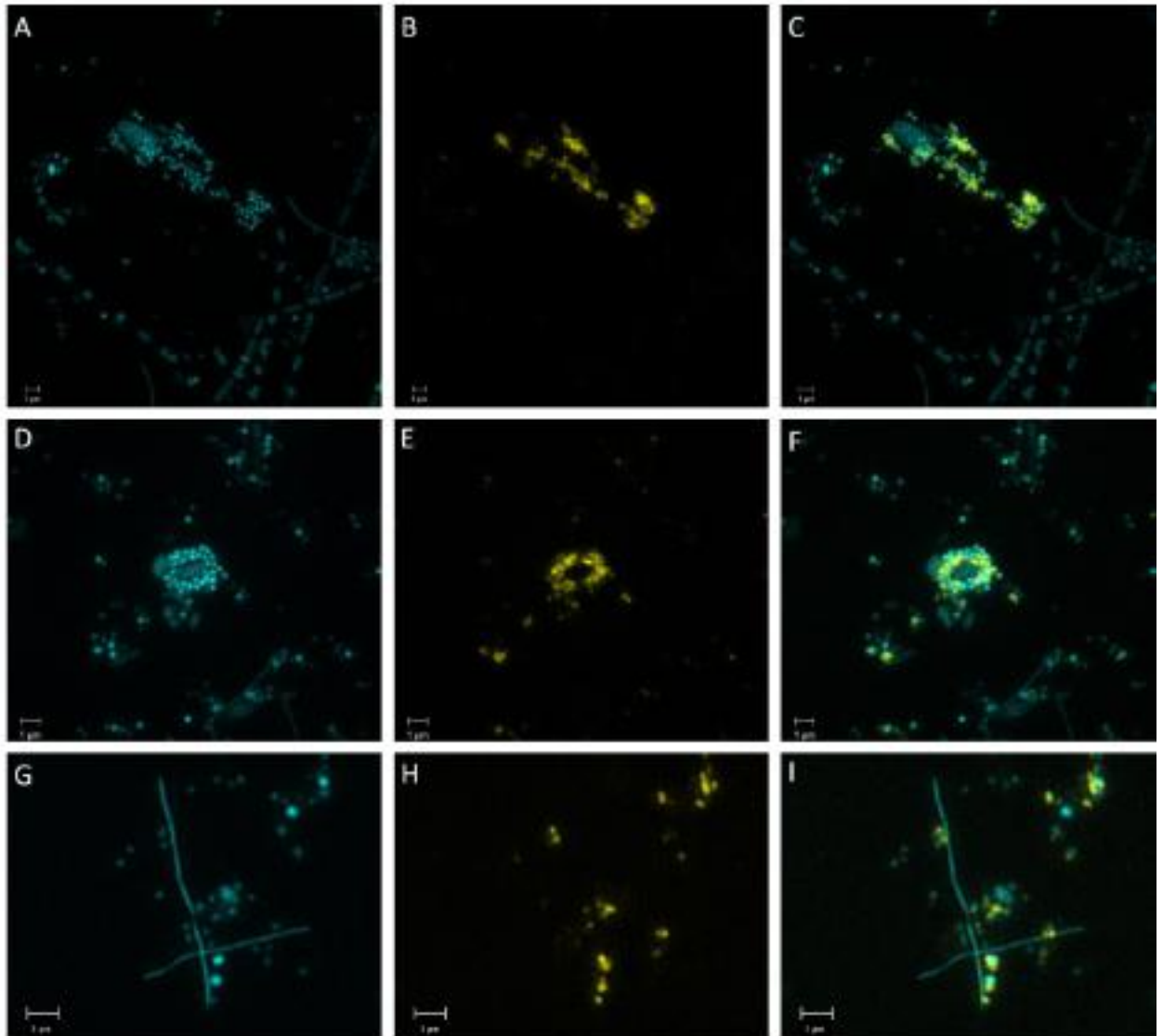
- Kuru, E., Lambert, C., Rittichier, J., Till, R., Ducret, A., Derouaux, A., et al. (2017) Fluorescent D-amino-acids reveal bi-cellular cell wall modifications important for *Bdellovibrio bacteriovorus* predation. *Nature Microbiology* **2**: (12) 1648–1657.
- Lambert, C., Lerner, T.R., Bui, N.K., Somers, H., Aizawa, S.-I., Liddell, S., et al. (2016) Interrupting peptidoglycan deacetylation during *Bdellovibrio* predator-prey interaction prevents ultimate destruction of prey wall, liberating bacterial-ghosts. *Scientific Reports* **6**: (1) 26010.
- Lambowitz, A.M. and Belfort, M. (1993) Introns as mobile genetic elements. *Annual review of biochemistry* **62**: 587–622.
- Langmead, B. and Salzberg, S.L. (2012) Fast gapped-read alignment with Bowtie 2. *Nature methods* **9**: (4) 357–9.
- Lerner, T.R., Lovering, A.L., Bui, N.K., Uchida, K., Aizawa, S., Vollmer, W., and Sockett, R.E. (2012) Specialized peptidoglycan hydrolases sculpt the intra-bacterial niche of predatory *Bdellovibrio* and increase population fitness. *PLoS pathogens* **8**: (2) e1002524.
- Mattick, J.S. (2002) Type IV pili and twitching motility. *Annual review of microbiology* **56**: 289–314.
- McIlroy, S.J., Kirkegaard, R.H., Dueholm, M.S., Fernando, E., Karst, S.M., Albertsen, M., and Nielsen, P.H. (2017) Culture-independent analyses reveal novel *Anaerolineaceae* as abundant primary fermenters in anaerobic digesters treating waste activated sludge. *Frontiers in Microbiology* **8**: 1134.
- Mesnage, S., Dellarole, M., Baxter, N.J., Rouget, J.-B., Dimitrov, J.D., Wang, N., et al. (2014) Molecular basis for bacterial peptidoglycan recognition by LysM domains. *Nature Communications* **5**: (1) 4269.
- Monnappa, A.K., Dwidar, M., Seo, J.K., Hur, J.-H., and Mitchell, R.J. (2014) *Bdellovibrio bacteriovorus* inhibits *Staphylococcus aureus* biofilm formation and invasion into human epithelial cells. *Scientific Reports* **4**: (1) 3811.
- Moreira, D., Zivanovic, Y., López-Archilla, A.I., Iniesto, M., and López-García, P. (2021) Reductive evolution and unique predatory mode in the CPR bacterium *Vamprococcus lugosii*. *Nature Communications* **12**: (1) 2454.
- Nawrocki, E.P., Jones, T.A., and Eddy, S.R. (2018) Group I introns are widespread in archaea. *Nucleic Acids Research* **46**: (15) 7970–7976.
- Nesbø, C.L. and Doolittle, W.F. (2003) Active self-splicing group I introns in 23S rRNA genes of hyperthermophilic bacteria, derived from introns in eukaryotic organelles. *Proceedings of the National Academy of Sciences* **100**: (19) 10806–10811.

- Nielsen, H. and Johansen, S.D. (2009) Group I introns: moving in new directions. *RNA Biology* **6**: (4) 375–383.
- Nudleman, E. and Kaiser, D. (2004) Pulling together with type IV pili. *Journal of Molecular Microbiology and Biotechnology* **7**: (1–2) 52–62.
- Nye, T.M., Fernandez, N.L., and Simmons, L.A. (2020) A positive perspective on DNA methylation: regulatory functions of DNA methylation outside of host defense in Gram-positive bacteria. *Critical Reviews in Biochemistry and Molecular Biology* **55**: (6) 576–591.
- Oren, A. and Garrity, G.M. (2021) Valid publication of the names of forty-two phyla of prokaryotes. *International Journal of Systematic and Evolutionary Microbiology* **71**: (10) 005056.
- Park, J.S., Lee, W.C., Yeo, K.J., Ryu, K.S., Kumarasiri, M., Heseck, D., et al. (2012) Mechanism of anchoring of OmpA protein to the cell wall peptidoglycan of the Gram-negative bacterial outer membrane. *FASEB journal: official publication of the Federation of American Societies for Experimental Biology* **26**: (1) 219–28.
- Paulsson, M., Kragh, K.N., Su, Y.-C., Sandblad, L., Singh, B., Bjarnsholt, T., and Riesbeck, K. (2021) Peptidoglycan-binding anchor is a *Pseudomonas aeruginosa* OmpA family lipoprotein with importance for outer membrane vesicles, biofilms, and the periplasmic shape. *Frontiers in Microbiology* **12**: 639582.
- Pernthaler, A., Pernthaler, J., and Amann, R. (2004) Sensitive multi-color fluorescence *in situ* hybridization for the identification of environmental microorganisms. In *Molecular Microbial Ecology Manual*. G. A. Kowalchuk, F.J. de B., I.M. Head, A.D.L. Akkermans, J.D. van Elsas (ed). Dordrecht: Kluwer Academic Publishers, pp. 711–726.
- Pita, T., Feliciano, J.R., and Leitão, J.H. (2020) Extracellular RNAs in bacterial infections: from emerging key players on host-pathogen interactions to exploitable biomarkers and therapeutic targets. *International Journal of Molecular Sciences* **21**: (24) 9634.
- Pukatzki, S., Ma, A.T., Revel, A.T., Sturtevant, D., and Mekalanos, J.J. (2007) Type VI secretion system translocates a phage tail spike-like protein into target cells where it cross-links actin. *Proceedings of the National Academy of Sciences* **104**: (39) 15508–15513.
- Ramaloko, W.T. and Osei Sekyere, J. (2022) Phylogenomics, epigenomics, virulome and mobilome of Gram-negative bacteria co-resistant to carbapenems and polymyxins: a One Health systematic review and meta-analyses. *Environmental Microbiology* **24**: (3) 1518–1542.

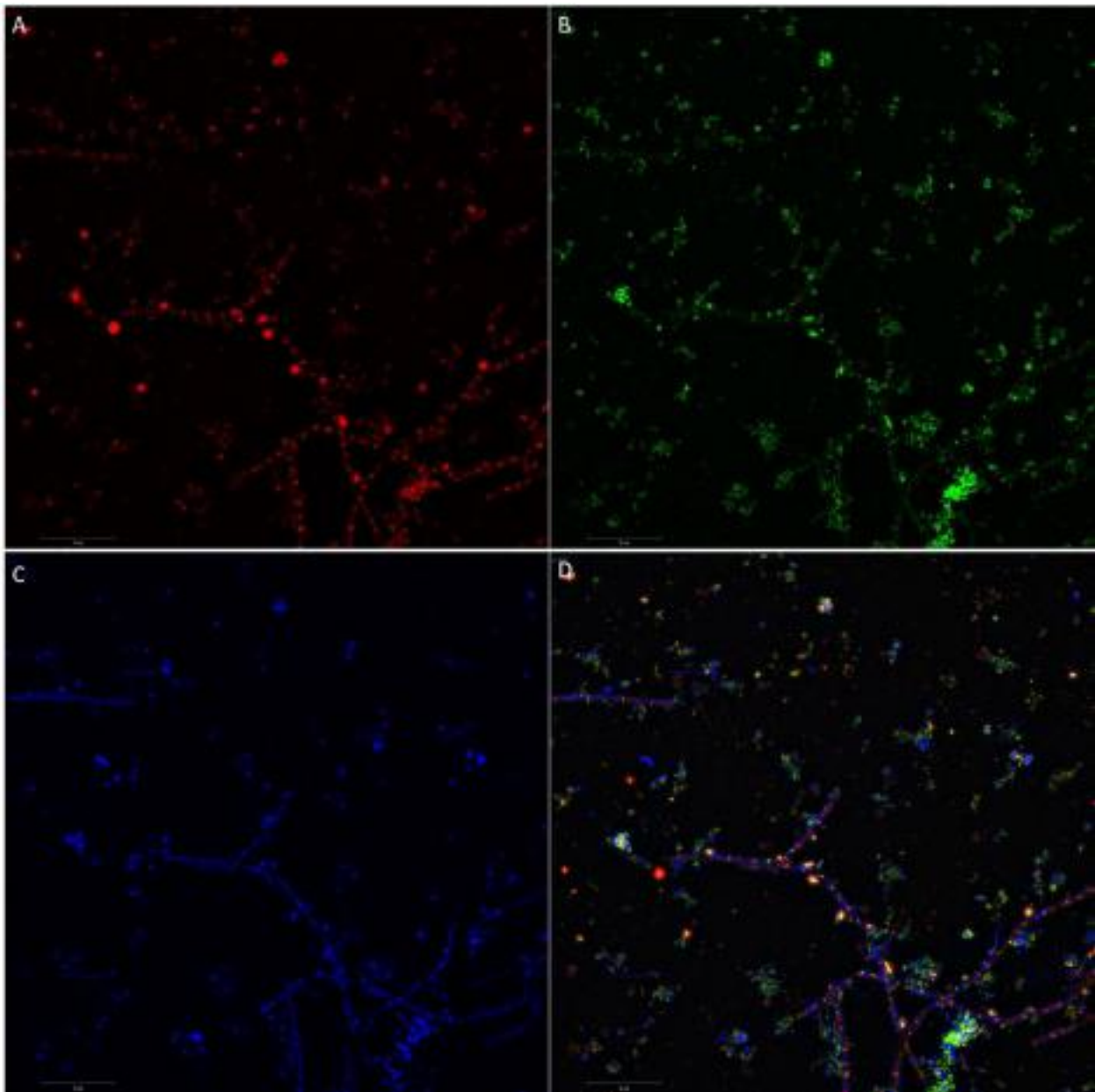
- Raskin, L., Stromley, J.M., Rittmann, B.E., and Stahl, D.A. (1994) Group-specific 16S rRNA hybridization probes to describe natural communities of methanogens. *Applied and Environmental Microbiology* **60**: (4) 1232–1240.
- Reinhold-Hurek, B. and Shub, D.A. (1992) Self-splicing introns in tRNA genes of widely divergent bacteria. *Nature* **357**: (6374) 173–176.
- Rinke, C., Schwientek, P., Sczyrba, A., Ivanova, N.N., Anderson, I.J., Cheng, J.-F., et al. (2013) Insights into the phylogeny and coding potential of microbial dark matter. *Nature* **499**: (7459) 431–437.
- Rombouts, S., Mas, A., Le Gall, A., Fiche, J.-B., Mignot, T., and Nollmann, M. (2023) Multi-scale dynamic imaging reveals that cooperative motility behaviors promote efficient predation in bacteria. *Nature Communications* **14**: (1) 5588.
- Rotaru, A.E., Schauer, R., Probian, C., Mussmann, M., and Harder, J. (2012) Visualization of candidate division OP3 cocci in limonene-degrading methanogenic cultures. *Journal of Microbiology and Biotechnology* **22**: (4) 457–61.
- Sachidanandham, R. and Yew-Hoong Gin, K. (2009) A dormancy state in nonspore-forming bacteria. *Applied Microbiology and Biotechnology* **81**: (5) 927–41.
- Samsudin, F., Ortiz-Suarez, M.L., Piggot, T.J., Bond, P.J., and Khalid, S. (2016) OmpA: A flexible clamp for bacterial cell wall attachment. *Structure* **24**: (12) 2227–2235.
- Schwabe, J., Pérez-Burgos, M., Herfurth, M., Glatter, T., and Søgaard-Andersen, L. (2022) Evidence for a widespread third system for bacterial polysaccharide export across the outer membrane comprising a composite OPX/ β -barrel translocon. *mBio* **13**: (5) e02032-22.
- Schwengers, O., Jelonek, L., Dieckmann, M.A., Beyvers, S., Blom, J., and Goesmann, A. (2021) Bakta: rapid and standardized annotation of bacterial genomes via alignment-free sequence identification. *Microbial Genomics* **7**: (11).
- Seef, S., Herrou, J., de Boissier, P., My, L., Brasseur, G., Robert, D., et al. (2021) A Tad-like apparatus is required for contact-dependent prey killing in predatory social bacteria. *eLife* **10**: e72409.
- Seymour, C.O., Palmer, M., Becraft, E.D., Stepanauskas, R., Friel, A.D., Schulz, F., et al. (2023) Hyperactive nanobacteria with host-dependent traits pervade Omnitrochota. *Nature Microbiology* **8**: (4) 727–744.
- Smith, S.G.J., Mahon, V., Lambert, M.A., and Fagan, R.P. (2007) A molecular Swiss army knife: OmpA structure, function and expression. *FEMS Microbiology Letters* **273**: (1) 1–11.

- Stahl, D.A. and Amann, R. (1991) Development and application of nucleic acid probes in bacterial systematics. In *Sequencing and hybridization techniques in bacterial systematics*. Stackebrandt, E. and G., M. (ed). Chichester, England: John Wiley & Sons, pp. 205–248.
- Stephenson, K. and Hoch, J.A. (2002) Evolution of signalling in the sporulation phosphorelay. *Molecular Microbiology* **46**: (2) 297–304.
- Stoddard, B.L. (2005) Homing endonuclease structure and function. *Quarterly Reviews of Biophysics* **38**: (1) 49–95.
- Takekawa, N., Kawamoto, A., Sakuma, M., Kato, T., Kojima, S., Kinoshita, M., et al. (2021) Two distinct conformations in 34 FlIF subunits generate three different symmetries within the flagellar MS-ring. *mBio* **12**: (2) 10.1128/mbio.03199-20.
- Thiery, S., Turowski, P., Berleman, J.E., and Kaimer, C. (2022) The predatory soil bacterium *Myxococcus xanthus* combines a Tad- and an atypical type 3-like protein secretion system to kill bacterial cells. *Cell Reports* **40**: (11) 111340.
- Tocheva, E.I., Dekas, A.E., McGlynn, S.E., Morris, D., Orphan, V.J., and Jensen, G.J. (2013) Polyphosphate storage during sporulation in the Gram-negative bacterium *Acetoneema longum*. *Journal of Bacteriology* **195**: (17) 3940–3946.
- Tosar, J.P., Witwer, K., and Cayota, A. (2021) Revisiting extracellular RNA release, processing, and function. *Trends in Biochemical Sciences* **46**: (6) 438–445.
- Vermassen, A., Leroy, S., Talon, R., Provot, C., Popowska, M., and Desvaux, M. (2019) Cell wall hydrolases in bacteria: insight on the diversity of cell wall amidases, glycosidases and peptidases toward peptidoglycan. *Frontiers in Microbiology* **10**: 331.
- Vollmer, W. and Tomasz, A. (2002) Peptidoglycan N-acetylglucosamine deacetylase, a putative virulence factor in *Streptococcus pneumoniae*. *Infection and immunity* **70**: (12) 7176–8.
- West-Roberts, J., Valentin-Alvarado, L., Mullen, S., Sachdeva, R., Smith, J., Hug, L.A., et al. (2023) Giant genes are rare but implicated in cell wall degradation by predatory bacteria [Preprint]. *bioRxiv* 2023.11.21.568195.
- Wu, H.-H., Yang, Y.-Y., Hsieh, W.-S., Lee, C.-H., Leu, S.-J.C., and Chen, M.-R. (2009) OmpA is the critical component for *Escherichia coli* invasion-induced astrocyte activation. *Journal of Neuropathology & Experimental Neurology* **68**: (6) 677–690.

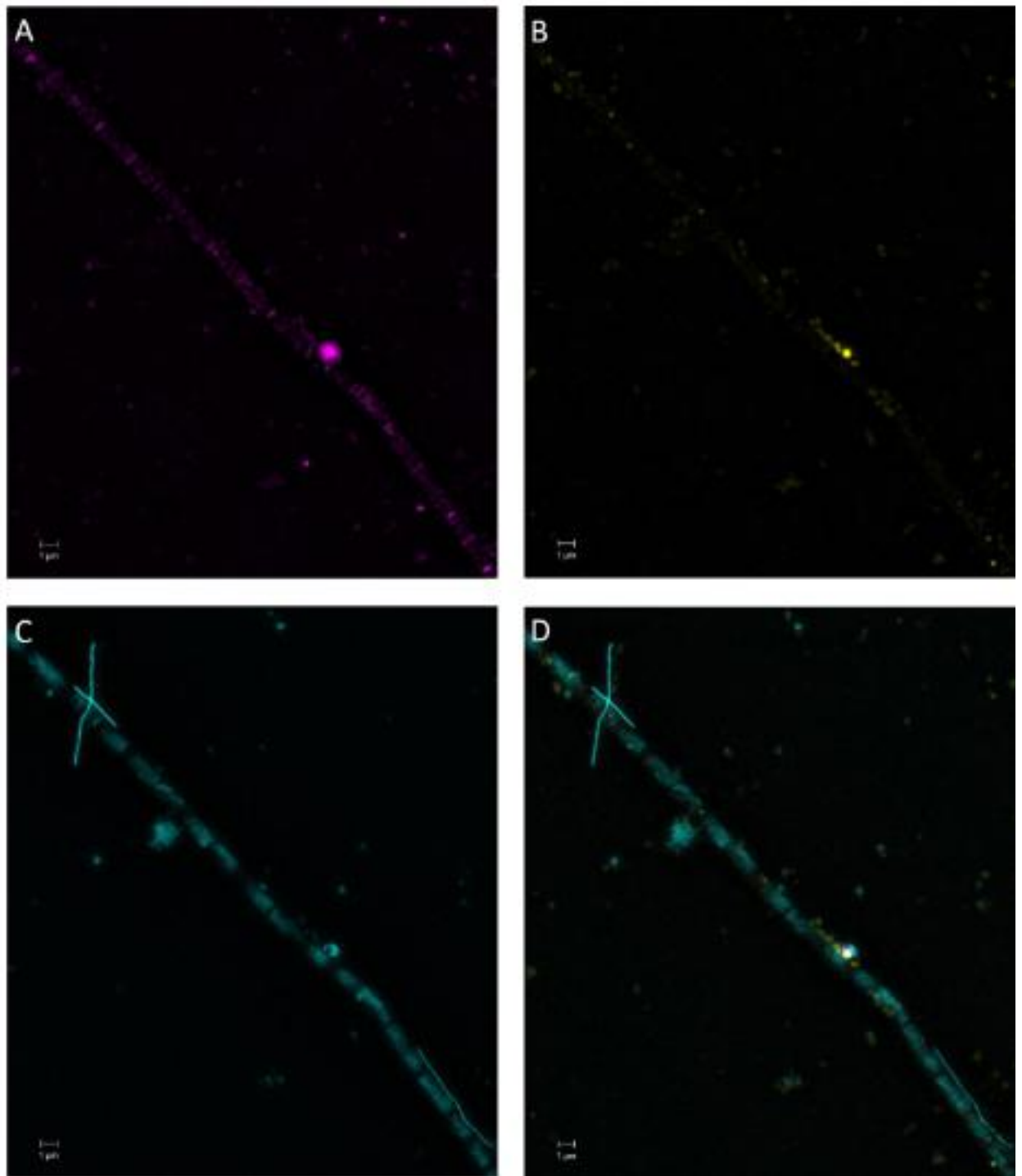
Supplementary material



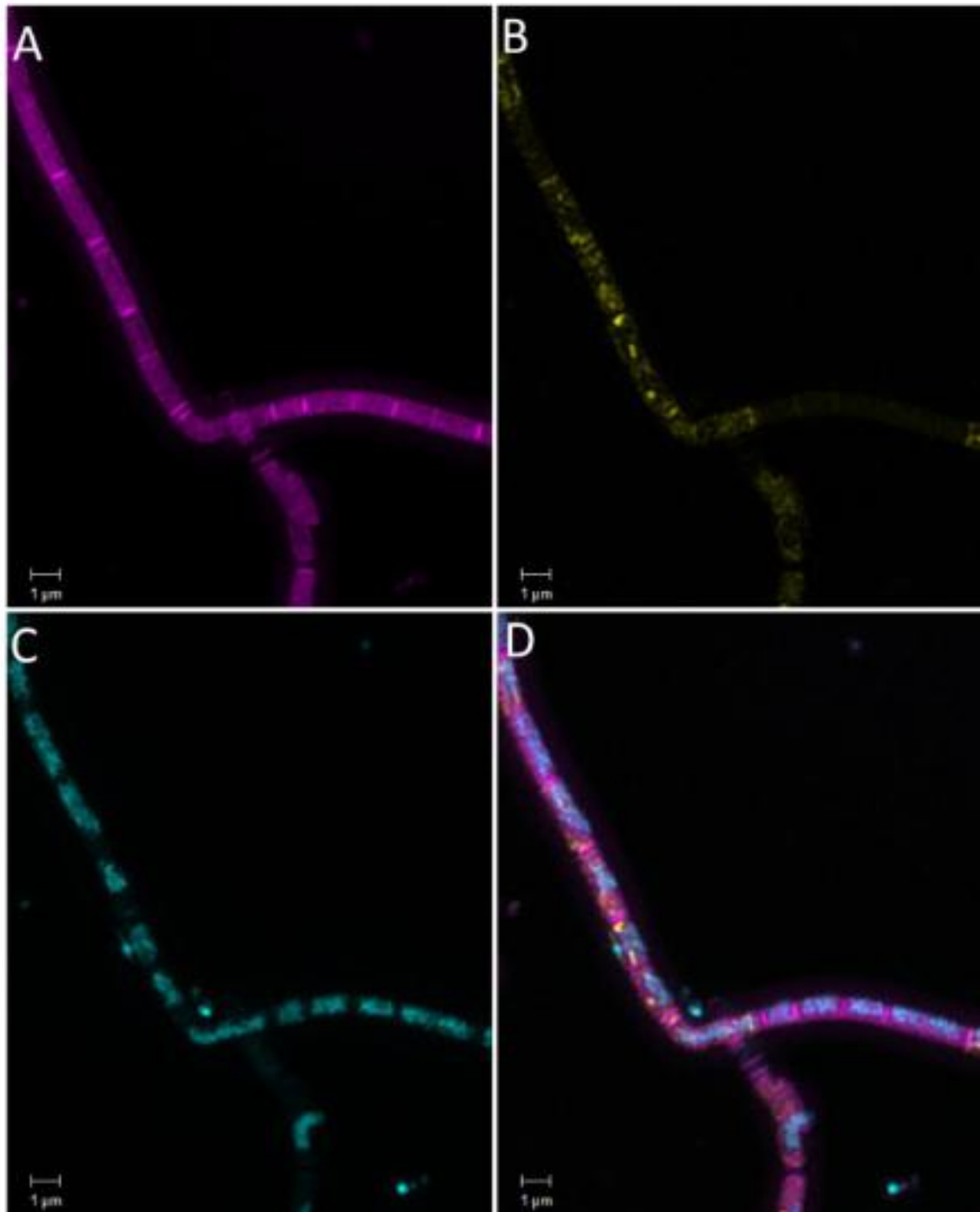
Supplementary figure 1: Structured illumination micrograph of the limonene-degrading methanogenic enrichment culture visualized as a maximum intensity projection. (A, D, and G) DAPI stained cells. (B, E, and H) Signals of OP3-565 CARD-FISH hybridized cells, which identify the cells as *Candidatus Velamenicoccus archaeovor*. (C, F and I) Overlays of the OP3-565 CARD-FISH and DAPI signals. Scale bar is 1 μm .



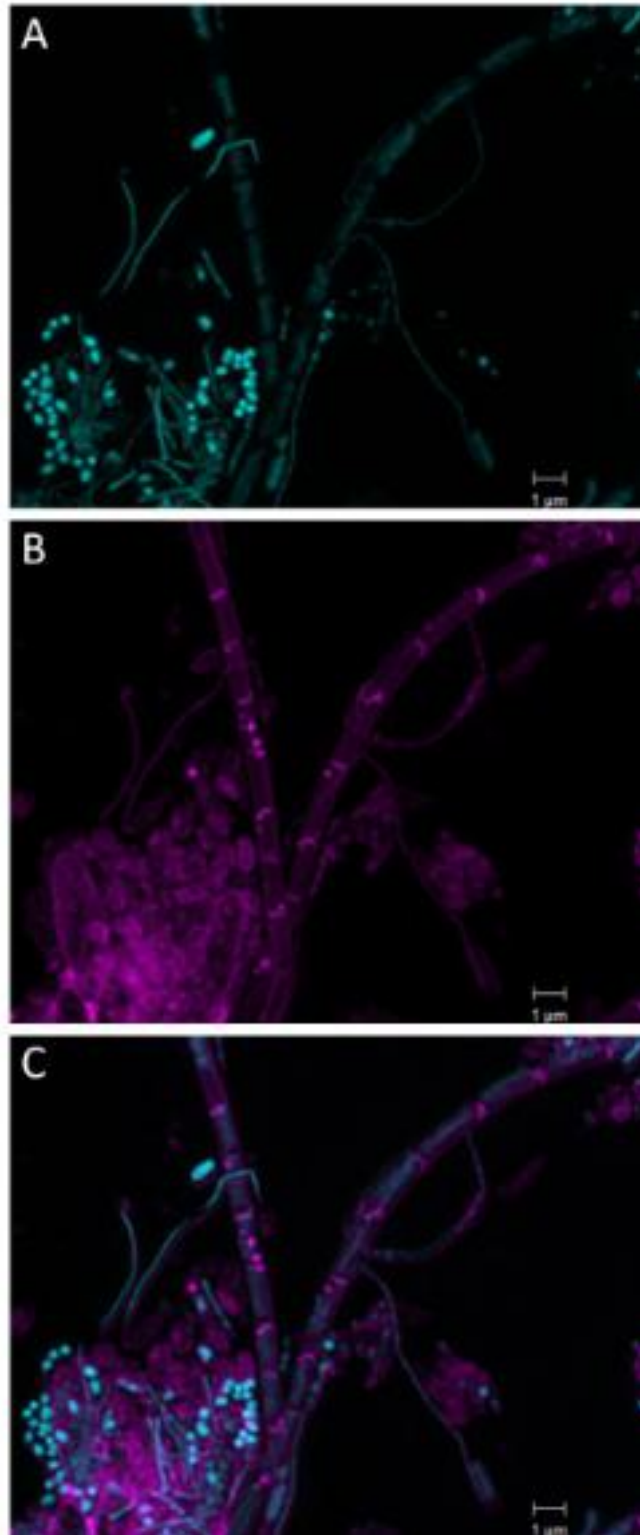
Supplementary figure 2: Structured illumination micrograph of the limonene-degrading methanogenic enrichment culture visualized as a maximum intensity projection. (A) Nile Red stained hydrophobic compounds. (B) Signals of OP3-565 CARD-FISH hybridized cells which identify the cells as *Candidatus Velamenicoccus archaeovor*. (C) DAPI stained cells. (D) Overlay. Scale bar is 10 μm .



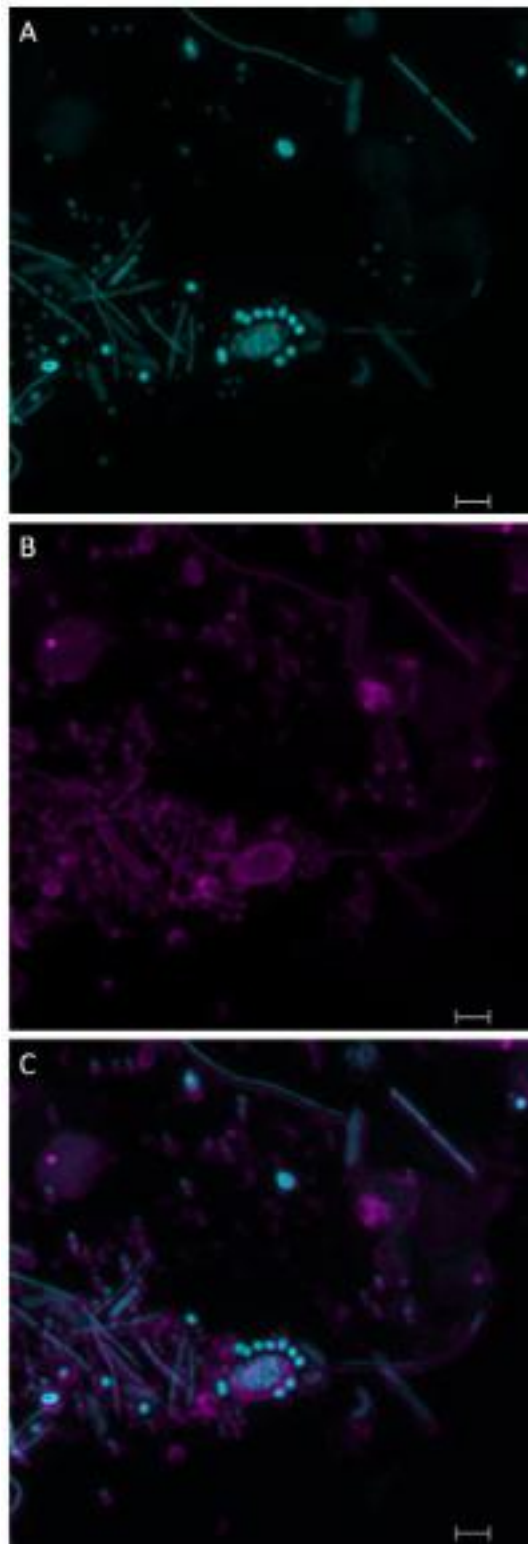
Supplementary figure 3: Structured illumination micrograph of the limonene-degrading methanogenic enrichment culture visualized as a maximum intensity projection. (A) Nile Red stained hydrophobic compounds. (B) Signals of OP3-565 CARD-FISH hybridized cells which identify the cells as *Candidatus Velamenicoccus archaeovorus*. (C) DAPI stained cells. (D) Overlay of DAPI and OP3-565 CARD-FISH. Scale bar is 1 μm .



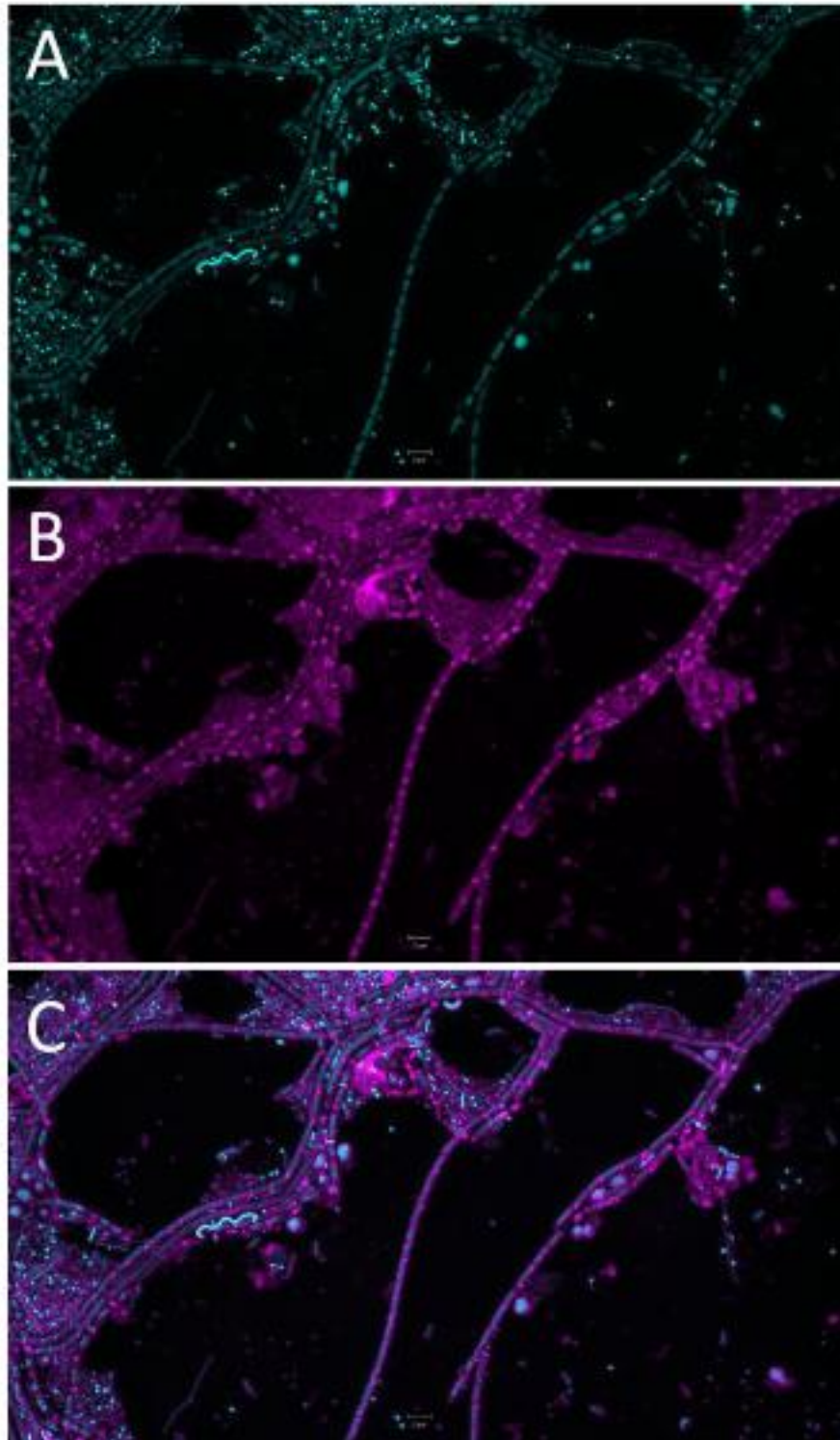
Supplementary figure 4: Structured illumination micrograph of the limonene-degrading methanogenic enrichment culture visualized as a maximum intensity projection. (A) Nile Red stained hydrophobic compounds. (B) Signals of MX825 hybridized cells which identify the cells as *Methanotherix*. (C) DAPI stained cells. (D) Overlay. Scale bar is 1 μm.



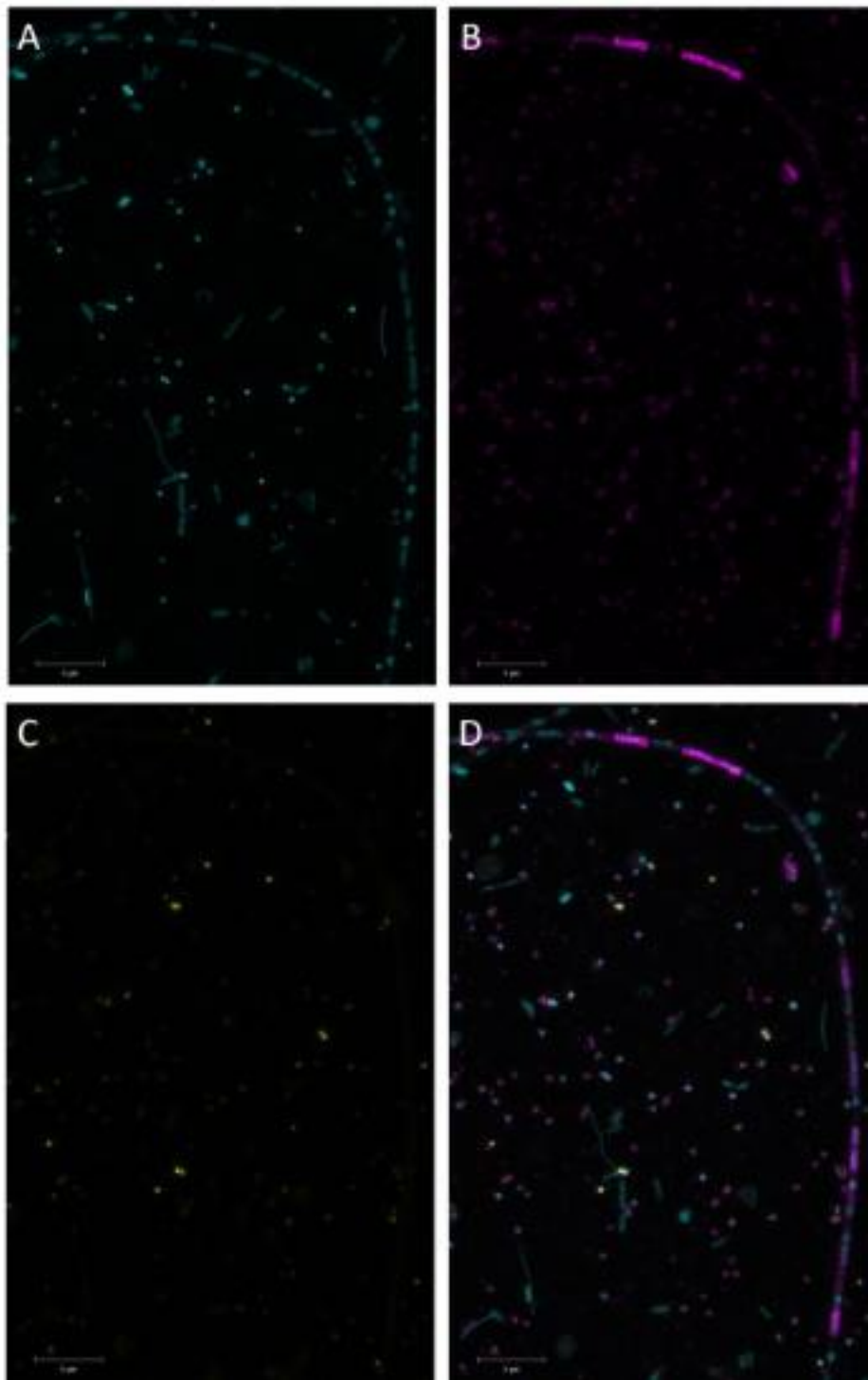
Supplementary figure 5: Structured illumination micrograph of the limonene-degrading methanogenic enrichment culture visualized as a maximum intensity projection. (A) DAPI stained cells. (B) Nile Red stained hydrophobic compounds. (C) Overlay. Scale bar is 1 μm .



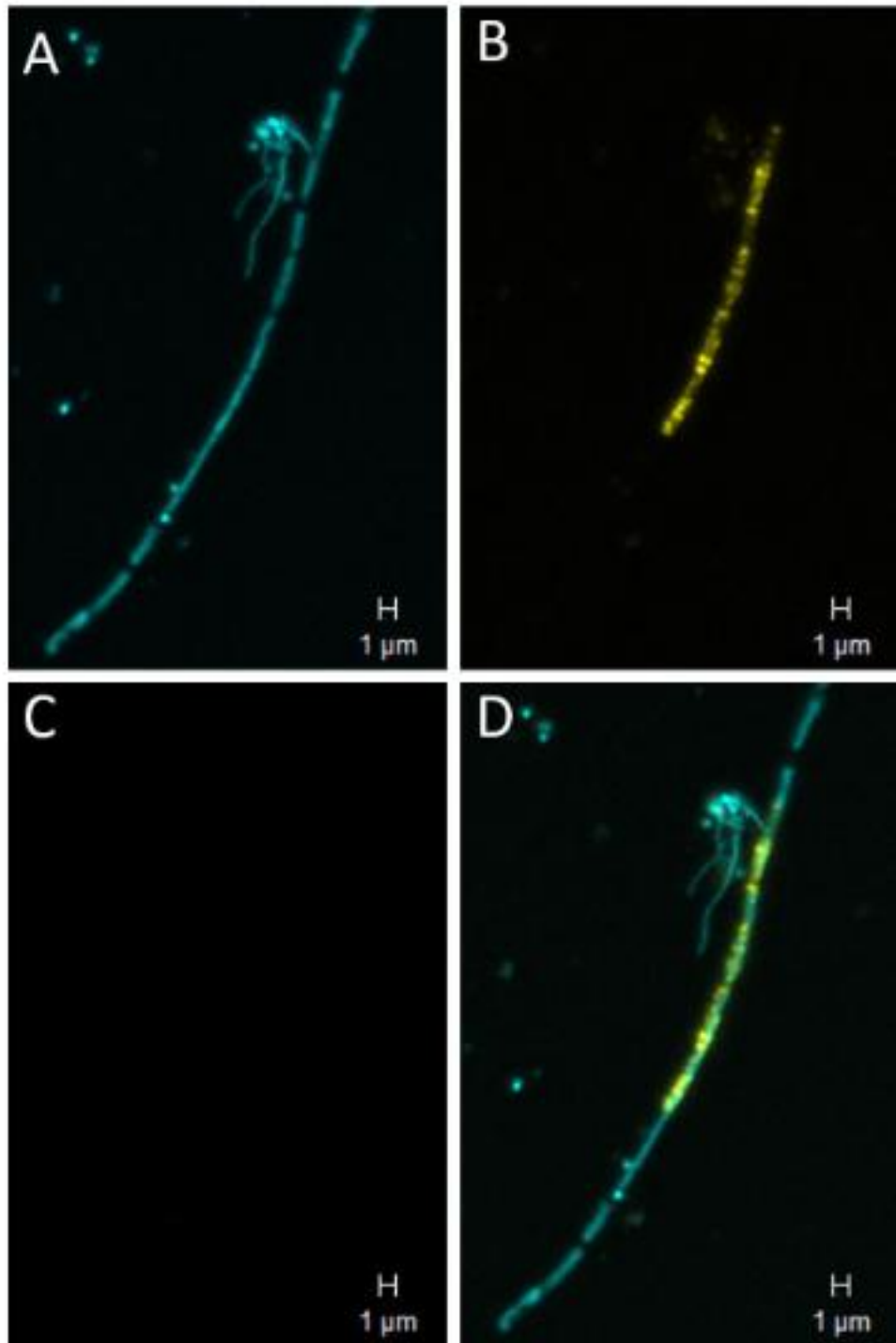
Supplementary figure 6: Structured illumination micrograph of the limonene-degrading methanogenic enrichment culture visualized as a maximum intensity projection. (A) DAPI stained cells. (B) Nile Red stained hydrophobic compounds. (C) Overlay. Scale bar is 1 μm .



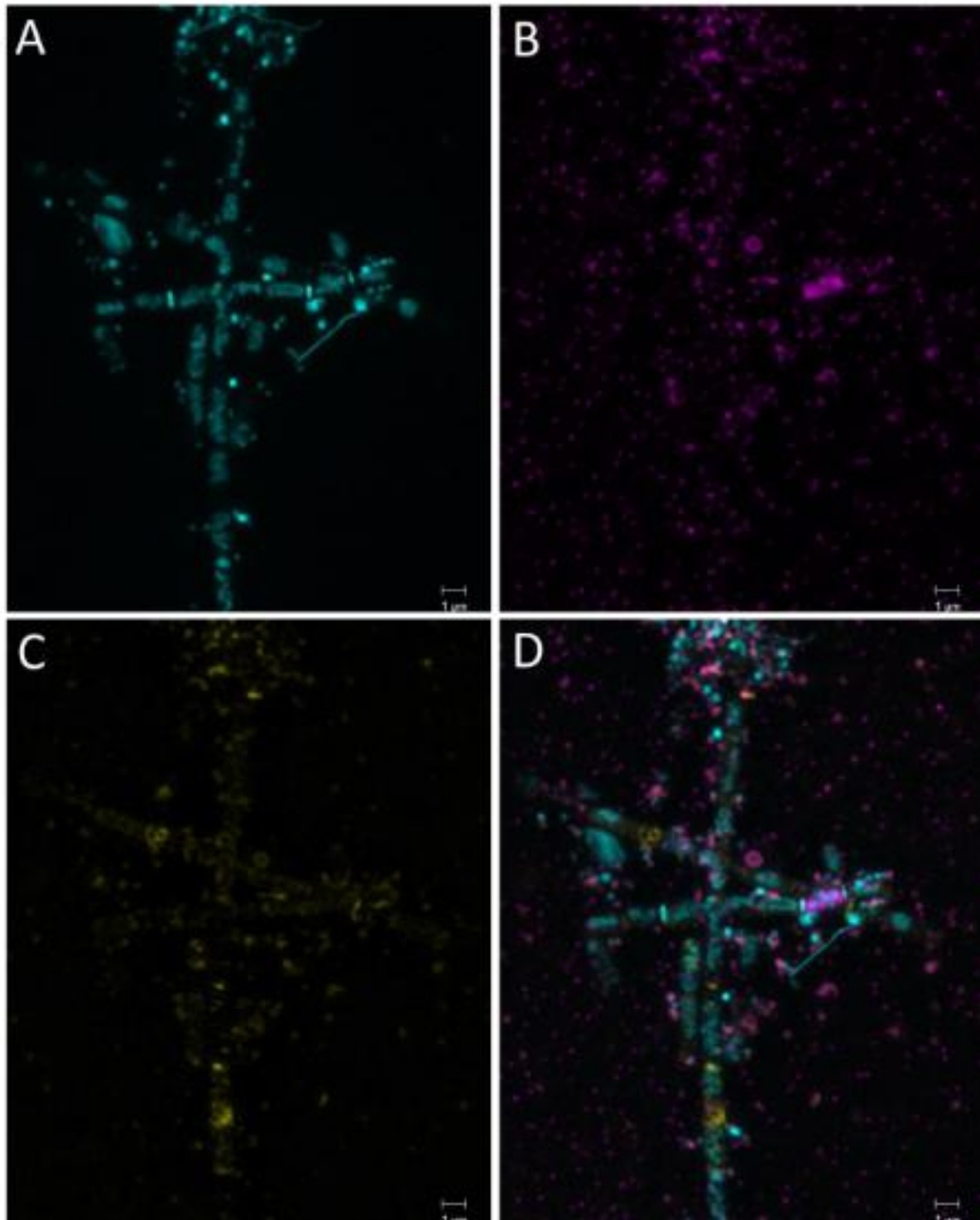
Supplementary figure 7: Structured illumination micrograph of the limonene-degrading methanogenic enrichment culture visualized as a maximum intensity projection. Different morphologies are visible, visualizing the complexity of the enrichment culture. (A) DAPI stained cells. (B) Nile Red stained hydrophobic compounds. (C) Overlay. Scale bar is 1 μm .



Supplementary figure 8: Structured illumination micrograph of the limonene-degrading methanogenic enrichment culture visualized as a maximum intensity projection. (A) DAPI stained cells. (B) Signals of hen2-2309 -CARD-FISH targeting the group I intron found in the 23S rRNA of OP3 LiM. (C) Signals of OP3-565 CARD-FISH hybridized cells, which identify the cells as *Candidatus Velamenicoccus archaeovor* (D) Overlay. Scale bar is 5 μm.



Supplementary figure 9: Structured illumination micrograph of the limonene-degrading methanogenic enrichment culture visualized as a maximum intensity projection. (A) DAPI stained cells. (B) *hen1-2235* and *hen3-2538* CARD-FISH signals, targeting the intron in the 23S rRNA of OP3 LiM. (C) *hen2-rc2309* reverse complement signal. (D) Overlay of the *hen1-2235* and *hen3-2538*, *hen2-rc2309* and DAPI signals. Scale bar is 1 μm .



Supplementary figure 10: Structured illumination micrograph of the limonene-degrading methanogenic enrichment culture visualized as a maximum intensity projection. (A) DAPI stained cells. (B) hen2-2309+4bp signals, probe targeting the intron in the 23S rRNA of OP3 LiM with 4 random basepairs. (C) Signals of EUBI-III and ARCH-915 CARD-FISH hybridized cells which identify cells as *Archaea* or *Bacteria*. (D) Overlay. Scale bar is 1 μm .

Chapter 5: Discussion

In this study, a limonene-degrading methanogenic enrichment culture was characterized, and a metagenome-assembled genome (MAG) assigned to the *Syntrophobacteraceae* family was identified as responsible for the initial step of limonene degradation. Additionally, the predatory bacterium *Candidatus Velamenicoccus archaeovorus*, a known member of this enrichment culture, was visualized. Analyzing the lifestyle of *Candidatus Velamenicoccus archaeovorus* added a new trophic level to this bacterium, shown to prey on *Methanotherix* and on other microorganisms.

5.1 Trophic interactions in the limonene-degrading methanogenic enrichment culture

The microbial community in the limonene-degrading methanogenic enrichment culture demonstrates a complex trophic structure, with key physiologies shown in Figure . The primary degrader, identified as *Syntrophobacteraceae* LiM6, initiates limonene degradation, and eventually, with other syntrophs, produces key intermediates such as acetate, formate, hydrogen, and carbon dioxide. Acetate is utilized by the acetoclastic methanogen *Methanotherix soehngeni* LiM3, while formate and hydrogen are consumed by methanogens such as *Methanoregula* LiM1, *Methanoculleus* LiM2, and *Methanospirillum* LiM14. Biomass generated by these processes is the basis for a new trophic level, the predation by *Candidatus Velamenicoccus archaeovorus*, which preys on the filamentous methanogens like *Methanotherix soehngeni* and other microorganisms, lysing and consuming cell mass. Fermenters like *Lentimicrobium* LiM4, *Anaerolineaceae* LiM8, LiM9, and LiM11, *Aminobacteriaceae* LiM16, and *Mesotoga* LiM7 ferment amino acids and sugars derived from cell material to produce alcohols and fatty acids. These substrates are then utilized by syntrophs (e.g., *Syntrophales* LiM10), contributing to the production of hydrogen, acetate, formate, and carbon dioxide, thus closing the microbial loop within the limonene-degrading methanogenic enrichment culture (Figure 1). The focus of this study was on the initial limonene degradation by *Syntrophobacteraceae* and the predatory behavior of *Candidatus Velamenicoccus archaeovorus*, elucidating their roles within the microbial community. The roles of other community members were described based on metatranscriptomic and metaproteomic data, as

well as their metabolic similarities to relatives in other methanogenic cultures or other environments.

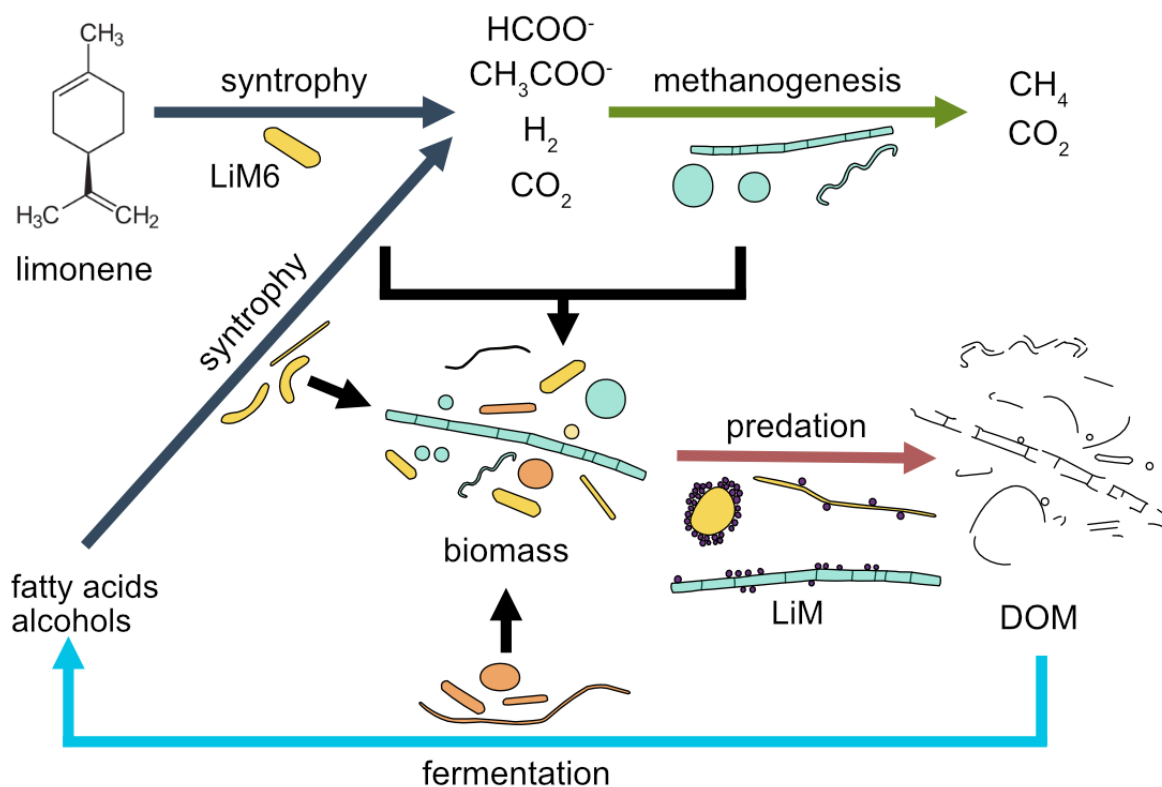


Figure 1: Overview of the main physiologies in the limonene-degrading methanogenic enrichment culture. Limonene is initially degraded by LiM6 (*Syntrophobacteraceae*), producing formate, acetate, hydrogen, and CO₂ utilized by methanogens to produce methane and CO₂. Biomass is lysed and degraded by, e.g., *Candidatus Velamenicoccus archaeovorans*. The resulting dissolved organic matter (DOM) is fermented, and the products (fatty acids and alcohols) are again used by syntrophs to produce acetate, hydrogen, formate, and CO₂. These compounds are consumed by methanogens that convert them into methane and CO₂, closing the microbial loop.

5.2 Necromass in methanogenic communities and its consumers

Necromass, the dead microbial biomass, represents a reservoir of organic matter that can support a variety of microorganisms, especially in energy-limited environments such as methanogenic systems. This necromass consists of proteins, carbohydrates, and other biopolymers released from dead cells and serves as a critical source of nutrients and energy

for other microorganisms (Dong *et al.*, 2018). In the limonene-utilizing enrichment culture, it primarily originates from the predatory behavior of *Candidatus Velamenicoccus* archaeovirus, which attaches to and attacks other cells, consuming them but also releasing necromass into the system (Figure 2).

The degradation of necromass is further carried out by hydrolytic and fermentative bacteria that break down complex macromolecules like proteins, lipids, and polysaccharides into simpler compounds. In the limonene-degrading methanogenic enrichment culture presented here, necromass degradation likely plays a significant role by providing substrates for various community members, introducing another trophic level to the community. This process not only recycles biomass but also creates niches for other bacteria to thrive, increasing the complexity and stability of the microbial ecosystem (Bradley *et al.*, 2018).

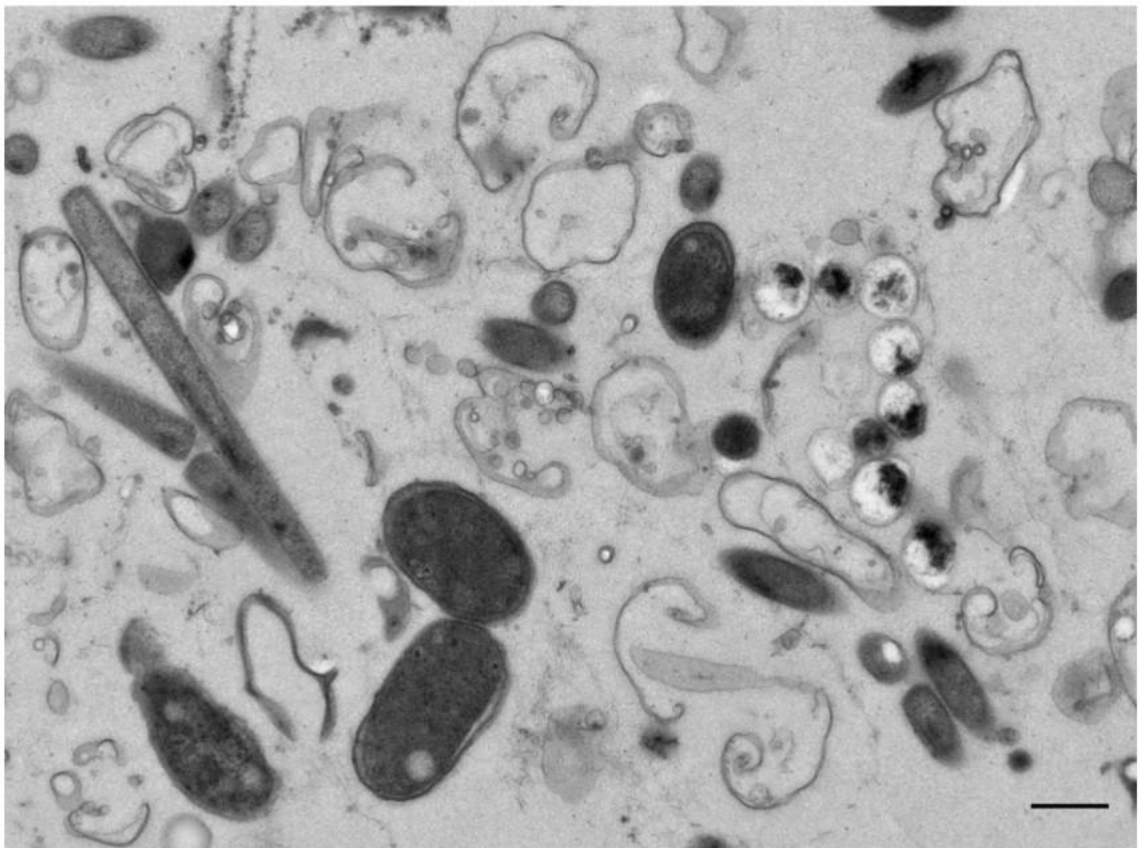


Figure 2: Thin section TEM images of an aged limonene-degrading enrichment culture showing a large amount of cell membrane remains. Small, coccoid-shaped cells were assigned to *Candidatus Velamenicoccus* archaeovirus based on morphology, which were still intact. Scale bar is 500 nm (Kizina *et al.*, 2022).

Members of *Anaerolineaceae* have been described to be involved in the degradation of carbohydrates and cellular components, such as amino acids, producing metabolites like

acetate and hydrogen. These products are crucial for methanogens, supporting syntrophic interactions within the microbial community (Yamada *et al.*, 2005).

In the limonene-degrading methanogenic enrichment culture, members of the *Anaerolineaceae* family (LiM8, LiM9, and LiM11) may perform similar metabolic roles as reported in previous studies. The detection of glycerol kinase in the metaproteome in all three members and dihydroxyacetone kinase subunits in LiM9 and LiM11 suggests the potential use of glycerol, likely derived from the lipid membranes of dead cells, as a substrate (Vizoso Pinto *et al.*, 2004; Gonzalez *et al.*, 2008; Sohlenkamp and Geiger, 2015). The metaproteomic detection of maltodextrin-binding protein in *Anaerolineaceae* LiM8 and LiM11, as well as an ATP-dependent 6-phosphofructokinase detected in *Anaerolineaceae* LiM9, indicates that carbohydrates are also utilized as substrates. Furthermore, the ATP-dependent 6-phosphofructokinase, glycerol kinase as well as the dihydroxyacetone kinase subunits are indicative of fermentative metabolism (Hansen *et al.*, 2002; Clomburg *et al.*, 2022). This suggests that members of *Anaerolineaceae* may play a key role in fermenting available substrates derived from necromass producing metabolites that sustain methanogenic processes. A member of the genus *Lentimicrobium* LiM4, present in the limonene-degrading methanogenic enrichment culture, may be involved in fermenting carbohydrates. A starch-binding protein associating with the outer membrane was detected in the metaproteome, which in other members of the *Bacteroidota* phylum, is involved in the binding of starch (Tuson *et al.*, 2018). The fructose-bisphosphate aldolase detected in *Lentimicrobium* LiM4 is involved in glycolysis and the pyruvate:ferredoxin oxidoreductase is involved in fermentation, indicating that it ferments carbohydrates possibly derived from the necromass in the enrichment culture (Ma *et al.*, 1997; Shams *et al.*, 2014). Other members of the genus *Lentimicrobium* have also been described to ferment glucose, ribose, sucrose, maltose, mannose, pyruvate, and starch, producing metabolites such as acetate, malate, propionate, formate, and hydrogen (Sun *et al.*, 2016).

In *Aminidesulfovibrio aminophilus* LiM27 (formerly known as *Desulfovibrio aminophilus*) multiple proteins like amino acid ABC transporter substrate-binding protein, branched-chain amino acid ABC transporter substrate-binding protein and alanine dehydrogenase indicate that it takes up and ferments amino acids, which are probably derived from the necromass. *Aminidesulfovibrio aminophilus* has been shown to ferment amino acids like serine, glycine, and cysteine in the absence of sulfate, producing acetate and propionate, further supporting the syntrophic interactions within the culture (Baena *et al.*, 1998).

Spirochaetes (LiM30, LiM13 and LiM20) are well-adapted to scavenge and contributing to the degradation of necromass, which is underlined by the metaproteomic detection of D-galactose-binding periplasmic protein (LiM13), D-ribose transporter subunit (LiM13), peptide ABC transporter substrate-binding protein (LiM13) and branched-chain amino acid ABC transporter substrate-binding protein (LiM13). The proteins involved in these processes were detected in the metaproteome in the *Spirochaete* LiM13 indicating that it is more active in necromass degradation than *Spirochaete* LiM30 and LiM20.

Especially in hydrocarbon-contaminated environments, *Spirochaetes* have been shown to ferment proteins and carbohydrates, producing acetate, ethanol, and hydrogen, which serve as key metabolites for other microbial processes, such as sulfate reduction by *Desulfobacterium* species. This metabolic interaction highlights the role of *Spirochaetes* in necromass recycling and maintaining the flow of nutrients and electron donors within anaerobic microbial communities (Dong *et al.*, 2018).

5.3 Comparison with other methanogenic hydrocarbon-degrading communities

Across diverse hydrocarbon-degrading methanogenic environments, the archaeal communities responsible for methane production, for example *Methanotherix*, *Methanoculleus*, and *Methanomicrobiales*, remain relatively consistent regardless of the hydrocarbon substrate.

In contrast, the bacterial populations in these communities show diverse compositions depending on the specific hydrocarbons involved. For instance, *Smithella* species have been identified as the primary degraders of n-alkanes under methanogenic conditions (Ji, Liu, *et al.*, 2020; Ji, Zhou, *et al.*, 2020). A toluene-degrading methanogenic enrichment culture inoculated from anoxic sediment from a gas condensate-contaminated aquifer identified *Desulfosporosinus* as the primary degrader of toluene. *Desulfosporosinus* activates the toluene via fumarate addition. Along with other syntrophic bacteria like *Syntrophaceae* and *Desulfovibrionales*, hydrogen, formate, and acetate were produced and used by methanogens to produce methane (Fowler *et al.*, 2014). A naphthalene-degrading methanogenic enrichment culture originally established from oil reservoir production waters identified an unclassified

species of *Clostridiaceae* as the primary degrader of naphthalene (Toth *et al.*, 2018). The naphthalene is activated via carboxylation, leading to CoA-thioesterification and further degradation. Additionally, bacteria belonging to *Desulfuromonadales* were suggested to play a potential but unclear role in naphthalene degradation (Toth *et al.*, 2018). A methanogenic microbial community enriched from environments contaminated with various polycyclic aromatic hydrocarbons was studied for its ability to degrade specific compounds such as naphthalene, 2-methylnaphthalene, phenanthrene, and anthracene under methanogenic conditions. Distinct microbial profiles were observed depending on the type of hydrocarbon. Naphthalene and 2-methylnaphthalene favored bacteria such as *Alicyclophilus* and *Thauera*, while *Methanomethylovorans* and *Methanobacterium* dominated among the methanogens. The degradation of naphthalene and 2-methylnaphthalene likely occurred through carboxylation, producing intermediates such as 2-naphthoic acid (Ye *et al.*, 2019).

Microbial limonene degradation has been studied in a horizontal-flow anaerobic bioreactor inoculated with anaerobic sludge sourced from a poultry slaughterhouse wastewater treatment system (Pozzi *et al.*, 2024). The bioreactors contained three different support media, coal, polyurethane foam, and gravel, which provided surfaces for microbial communities to attach and grow. As the experimental setup of this experiment might allow traces of oxygen to be present, for example as oxygen was not removed from the limonene by degassing, this system is not directly comparable to the anaerobic enrichment culture of this study. Therefore, it can be viewed as a transition system to oxic systems with minimal traces of oxygen. The microbial community degrading the limonene was dominated by *Synergistetes* (43–57%), *Proteobacteria* (32–42%), *Firmicutes* (7–8%), and *Acidobacteria* (2–3%), while *Actinobacteria*, *Bacteroidetes*, and *Chloroflexi* appeared in lower abundances (1–2%). *Synergistaceae* was the most prevalent family across all these support media, indicating its importance in the microbial consortium, though the specific role of *Synergistaceae* in limonene degradation is less clear. This family is generally involved in syntrophic interactions, assisting in the breakdown of complex organic molecules by producing intermediates like hydrogen and acetate. The primary degradation of limonene was attributed to *Pseudomonas* species (10.6–18.4%) due to their metabolic versatility. *Pseudomonas* species are known for their metabolic versatility and *Pseudomonas putida* as well as *Pseudomonas fluorescens* have been shown to be able to metabolize limonene aerobically (Chatterjee and Bhattacharyya, 2001; Bicas *et al.*, 2008). They likely initiated the breakdown of limonene through pathways that involve enzymes capable of

transforming terpenes, although the study does not explicitly identify the specific enzymes involved in limonene degradation. However, other studies suggest that enzymes such as limonene monooxygenase or cytochrome P450 enzymes are often involved in the initial oxidation of limonene to more degradable intermediates like p-cymene (Harder and Probian, 1995; Eaton, 1997). Although syntrophic interactions between bacteria and methanogenic archaea, such as *Methanotherix* and *Methanospirillum*, were involved in the process, methane production remained low in the bioreactor. This was possibly due to the incomplete anaerobic digestion of limonene, with intermediate compounds like *p-cymene* accumulating in the system and possibly inhibiting full conversion to methane (Trujillo-Reyes *et al.*, 2024). Despite achieving over 95% limonene removal, methane generation was minimal in the horizontal-flow anaerobic bioreactor (Pozzi *et al.*, 2024).

The limonene-degrading methanogenic enrichment culture shows both similarities and unique differences compared to other hydrocarbon-degrading methanogenic communities. Like other systems, it involves syntrophic interactions between bacteria and methanogens, with primary degraders breaking down the hydrocarbon and producing intermediates such as hydrogen, formate, and acetate that methanogens, like *Methanotherix* and *Methanospirillum*, metabolize to methane. However, a key distinguishing feature of this culture is the presence of *Candidatus Velamenicoccus archaeovorus*, a predatory bacterium, which introduces a novel trophic level not observed in other systems. This predatory interaction adds complexity to the ecosystem, creating new ecological niches that can influence the structure and function of the microbial community. While it shares core methanogenic and syntrophic processes with other hydrocarbon-degrading communities, the unique trophic dynamics make this system particularly interesting.

5.4 Novel insights into anaerobic limonene degradation under methanogenic conditions

Within the limonene-degrading methanogenic enrichment culture, a novel enzyme, limonenylsuccinate synthase, and pathway were disclosed, shedding new light on the biochemical processes that catalyze anaerobic terpene degradation. While limonene degradation has previously been observed under denitrifying conditions, for example by enzymes such as limonene dehydrogenase (Puentes-Cala *et al.*, 2018), this study marks the first description of limonene metabolism in methanogenic environments.

The limonenylsuccinate synthase is a fumarate-adding enzyme that activates limonene by catalyzing its addition to fumarate, forming limonenylsuccinate as a key intermediate. This enzyme belongs to the same family as benzylsuccinate synthases, which are key players in the anaerobic degradation of various hydrocarbons, including aromatic compounds like toluene, xylenes, and cresols, as well as aliphatic compounds such as n-alkanes and cycloalkanes (Callaghan, 2013; Heider *et al.*, 2016).

Phylogenetic analysis places limonenylsuccinate synthase on a distinct evolutionary branch within this family of benzylsuccinate synthases, suggesting an adaptation specifically for handling the larger and more complex structure of limonene. This evolutionary divergence is characterized by key modifications in the substrate-binding pocket of the enzyme, which allows it to accommodate limonene's bulky molecular structure, a feature not seen in enzymes specialized for smaller aromatic substrates.

These discoveries further emphasize the adaptability of methanogenic microbial communities to metabolize diverse organic compounds, even under energy-limited conditions where electron acceptors like nitrate or sulfate are absent. The presence of such specialized enzymes suggests that methanogenic systems may play an important role in the anaerobic degradation of plant-derived hydrocarbons, particularly in environments rich in natural terpenes such as wetlands, peatlands, and engineered anaerobic digesters. This, in turn, contributes to carbon cycling in both natural and engineered anaerobic environments.

5.5 Omnitrophota: origins, classification, and characteristics

One particularly unique member of the limonene-degrading methanogenic enrichment culture is the predatory *Candidatus Velamenicoccus archaeovorus*, which belongs to the candidate bacterial phylum Omnitrophota within the Planctomycetota-Verrucomicrobiota-Chlamydiota (PVC) superphylum. Omnitrophota represents a diverse group of ultra-small (<0.3 μ m) bacteria found in various habitats, including water, sediments, soils, and anoxic environments (Seymour *et al.*, 2023). Despite their broad environmental distribution, Omnitrophota is not completely understood due to challenges in cultivation. Its members have mostly been studied through metagenomics and single-cell genomics. This phylum is characterized by reduced metabolic capacities, small genome size, and a host-associated or symbiotic lifestyle (Seymour *et al.*, 2023). Within Omnitrophota, *Candidatus Velamenicoccus archaeovorus* is a significant representative, particularly due to its predatory lifestyle and its interaction with its prey, which consists of archaea and bacteria in the limonene-degrading methanogenic enrichment culture.

Phylogenetically, Omnitrophota is divided into six major classes - Velamenicoccia, Omnitrophia, Gorgyraia, Aquivivencia, and two unnamed clades, 2-02-FULL-51-18 and 4484-213 (Figure 3) - each showing genomic reduction that ranges from 0.86 to 3.20 Mb (Seymour *et al.*, 2023) (Seymour *et al.*, 2023). This reduction is characteristic of bacteria with highly specialized host-associated or symbiotic lifestyles (Toft and Andersson, 2010). The Velamenicoccia class is notable for host-dependent and parasitic traits, including species like *Candidatus Velamenicoccus archaeovorus*, which shows metabolic versatility by relying on acetogenesis via the Wood-Ljungdahl pathway and likely utilizing carbon monoxide as an alternative substrate (Seymour *et al.*, 2023). This reflects the metabolic flexibility seen across Omnitrophota members, with some capable of acetogenesis while others are facultative fermenters or mixotrophs. Omnitrophia, on the other hand, demonstrates more complex metabolic capabilities, including aerobic respiration and partial denitrification pathways. Gorgyraia species are largely anaerobic, relying on acetogenesis, while Aquivivencia species appear more adapted to oxic environments, capable of aerobic respiration and diverse metabolic functions. The two unnamed clades, 2-02-FULL-51-18 and 4484-213, are less understood but exhibit phylogenetic divergence from the other classes, indicating unique evolutionary pathways.

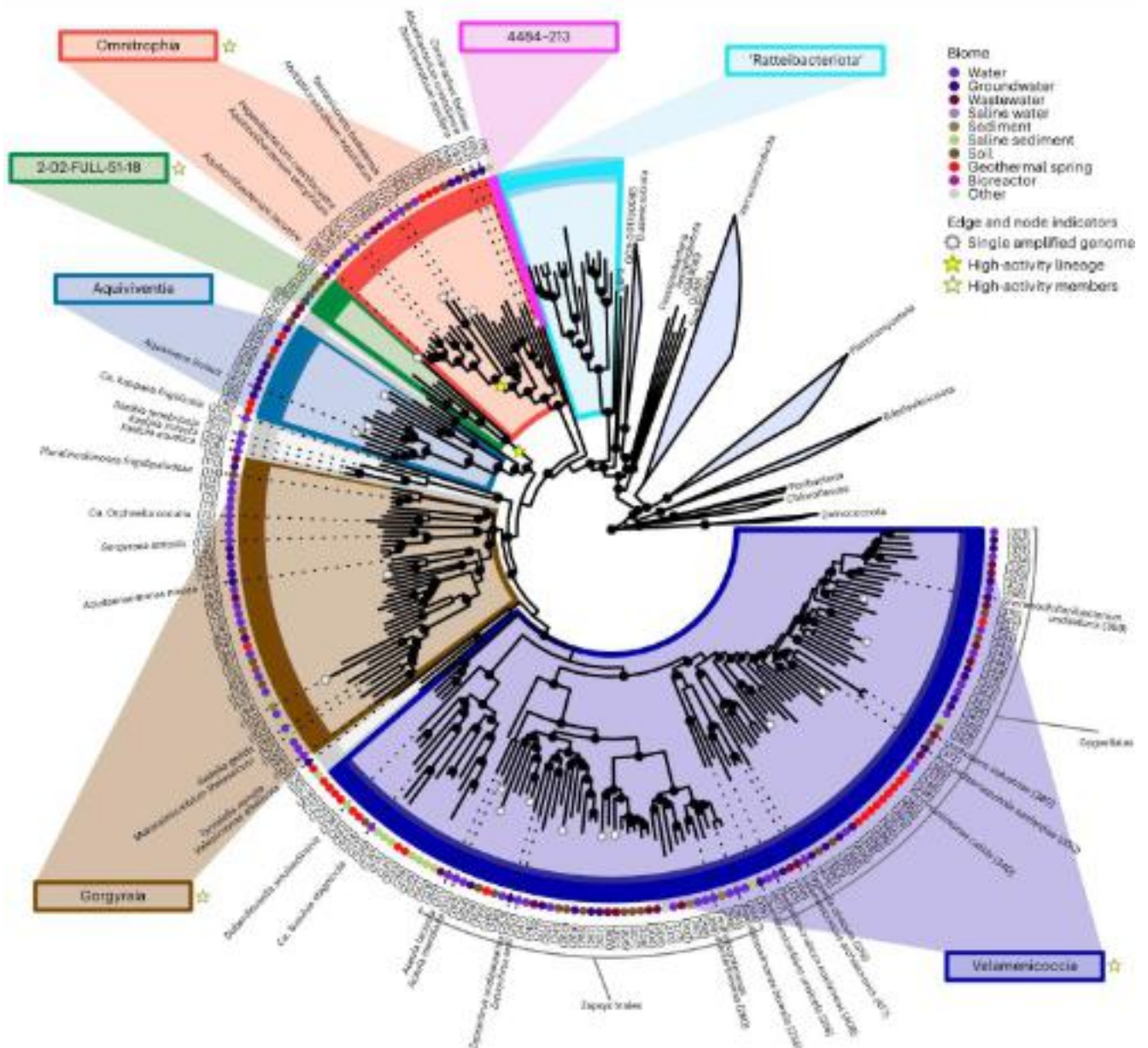


Figure 3: Maximum-likelihood phylogenetic tree of 204 Omnitrophota species representatives. The tree was constructed using the Bac120 marker gene set and confirms that Omnitrophota is part of the *Planctomycetota-Verrucomicrobiota-Chlamydiota* (PVC) superphylum. It reveals six distinct classes within Omnitrophota originating from different environments (Seymour et al., 2023).

Recent studies have expanded the understanding of Omnitrophota's metabolic potential, revealing its capacity for both aerobic and anaerobic respiration, as well as fermentative lifestyles (Perez-Molphe-Montoya *et al.*, 2022). The discovery of new clades, including several

novel genomes from groundwater ecosystems, emphasizes the metabolic diversity within the phylum. Some members of Omnitrophota, exhibit the Wood-Ljungdahl pathway and are key players in nitrogen and sulfur cycling, suggesting their involvement in carbon fixation and organic matter degradation in oligotrophic environments. This metabolic versatility, observed across different environmental gradients, highlights the ecological resilience of Omnitrophota, particularly in groundwater and aquatic systems (Perez-Molphe-Montoya *et al.*, 2022).

5.6 A group I Intron in the limonene-degrading enrichment culture

The limonene-degrading methanogenic enrichment culture not only contains the predatory bacterium *Candidatus Velamenicoccus archaeovorans* but also contains a genetic element - a Group I intron located in the 23S rRNA gene of this predator. This intron was visualized using Fluorescence *in situ* Hybridization (FISH) which revealed its horizontal transfer to its prey, *Methanothrix soehngenii*, a filamentous acetoclastic methanogen. The horizontal transfer of this Group I intron into *Methanothrix* is a novel observation that sheds light on the interplay between predation, genetic exchange, and microbial evolution within complex methanogenic ecosystems. Group I introns are ribozymes, which means they have catalytic properties, enabling them to self-splice from RNA transcripts, leaving the host gene functional. These introns are typically found in rRNA, tRNA, and mRNA genes across a variety of organisms, including bacteria, archaea, fungi, and even in some viral genomes (Cech, 1988). In prokaryotic systems, these elements are less common but have been increasingly recognized for their role in genomic plasticity, horizontal gene transfer, and metabolic adaptation in diverse environments (Hausner *et al.*, 2014).

Group I introns are believed to have ancient origins, dating back to the early stages of RNA evolution (Cech, 1988). They likely played a role in early molecular evolution, influencing the development of RNA-based genetic systems. The catalytic properties of Group I introns, particularly their self-splicing mechanism, suggest that they may have contributed to the transition from RNA-based life forms to the more complex DNA-RNA-protein systems observed today (Woodson and Chauhan, 2009). Over time, these introns evolved from selfish elements to become integral parts of some microbial genomes. Group II introns, a different type of RNA enzymes, have been described to have functional roles, such as regulating gene expression or contributing to genome stability by assisting in the repair of DNA (Lambowitz and Zimmerly, 2011).

Group I introns likely originated as selfish genetic elements, relying on horizontal gene transfer to propagate themselves across species (Goddard and Burt, 1999). Their persistence in microbial genomes suggests that they can provide selective advantages, such as enabling genome adaptation under stress conditions or enhancing the efficiency of gene expression through precise RNA splicing mechanisms (Edgell *et al.*, 2000). In archaea and bacteria, Group I introns often include genes for homing endonucleases, enzymes that further promote intron mobility and insertion into new genomic sites. Introns can contribute to the evolutionary adaptability of microbial communities, allowing organisms to respond to environmental stressors (Edgell *et al.*, 2000).

In this limonene-degrading culture, the presence of a Group I intron within the genome of *Candidatus Velamenicoccus archaeovorus* suggests potential roles in shaping the dynamics of the microbial community. Specifically, Group I introns have been implicated in promoting genome rearrangements (Hausner *et al.*, 2014). These properties can lead to more genetic diversity and may provide an evolutionary advantage, especially in competitive or nutrient-limited environments, such as those found in methanogenic ecosystems. The ability of Group I introns to mediate gene mobility and exchange could foster metabolic flexibility in *Candidatus Velamenicoccus archaeovorus*, potentially contributing to its predatory lifestyle and interactions with other community members.

5.7 Outlook and future directions

The limonene-degrading methanogenic enrichment culture provides many aspects for further exploration.

The limonenylsuccinate synthase identified in this culture could be analyzed in regard to its structure. However, heterologous expression of fumarate-adding enzymes in common systems like *E. coli* can be problematic due to issues such as improper folding and post-translational modifications. Alternative expression systems such as *Azoarcus* or *Thauera* species, which are more closely related to anaerobic hydrocarbon-degrading bacteria, could be explored (Achong *et al.*, 2001; Heider *et al.*, 2016). Expression in these systems may allow for the biochemical characterization of limonenylsuccinate synthase, providing insights into its kinetics, substrate

specificity, and structural features. Additionally, the structure could be revealed through crystallization or predictive models like AlphaFold giving insights into the enzyme's active site architecture and its evolutionary adaptations to limonene's unique structure.

To enhance our understanding of the interactions within the limonene-degrading microbial community, creating a FISH probe specifically targeting *Syntrophobacteraceae* would allow for more precise visualization of this primary limonene degrader *in situ*. Probes targeting other key members in the community could be considered, to visualize each community member and understand its role. It would complement the metagenomic data and help identify potential physical associations between community members.

An in-depth analysis of the existing metagenomic, metatranscriptomic, and metaproteomic datasets could enhance the understanding of the community's interactions. By examining gene expression and protein presence, how these organisms contribute to the breakdown of organic matter and support syntrophic interactions within the community can be elucidated.

The discovery of a Group I intron in *Candidatus Velamenicoccus archaeovorans* opens up a new perspective on horizontal gene transfer and genetic exchange in methanogenic ecosystems. Future studies could explore the prevalence and role of such introns in other methanogenic or hydrocarbon-degrading systems. Understanding how these genetic elements influence microbial evolution, genomic stability, or gene regulation could provide insights into the adaptive strategies of microorganisms in energy-limited environments.

Co-cultivation experiments involving *Candidatus Velamenicoccus archaeovorans* and its prey, such as *Methanotheroxiphosphaerium soehngenii*, could help clarify the ecological role of this predator. Controlled studies in co-culture systems may allow for a detailed analysis of predator-prey dynamics, including how predation impacts biomass turnover and overall community structure. Such studies could also investigate how necromass recycling affects methanogenesis and other syntrophic interactions within the community.

Given the identification of limonene degradation pathways in this system, it would be valuable to investigate the presence of limonenylsuccinate synthase and similar enzymes in other environments, such as citrus waste or terpene-rich ecosystems, where natural plant-derived hydrocarbons are abundant. Bioinformatics searches through environmental metagenomic databases could reveal whether similar metabolic strategies exist in different contexts, potentially leading to novel discoveries of hydrocarbon-degrading microbial communities in industrial or natural settings.

While metagenomic analyses have provided a wealth of information, cultivating and characterizing this culture in the laboratory remains essential to confirm hypotheses generated from bioinformatics. By isolating individual members of the community and studying their metabolism under controlled conditions more experimentally validated insights can be revealed.

References

- Achong, G.R., Rodriguez, A.M., and Spormann, A.M. (2001) Benzylsuccinate Synthase of *Azoarcus* sp. strain T: cloning, sequencing, transcriptional organization, and its role in anaerobic toluene and *m*-xylene mineralization. *J Bacteriol* **183**: (23) 6763–6770.
- Baena, S., Fardeau, M.L., Labat, M., Ollivier, B., Garcia, J.L., and Patel, B.K.C. (1998) *Desulfovibrio aminophilus* sp. nov., a novel amino acid degrading and sulfate reducing bacterium from an anaerobic dairy wastewater lagoon. *Systematic and Applied Microbiology* **21**: (4) 498–504.
- Bicas, J.L., Fontanille, P., Pastore, G.M., and Larroche, C. (2008) Characterization of monoterpene biotransformation in two pseudomonads. *Journal of Applied Microbiology* **105**: (6) 1991–2001.
- Bradley, J.A., Amend, J.P., and LaRowe, D.E. (2018) Necromass as a limited source of energy for microorganisms in marine sediments. *Journal of Geophysical Research: Biogeosciences* **123**: (2) 577–590.
- Callaghan, A.V. (2013) Enzymes involved in the anaerobic oxidation of n-alkanes: from methane to long-chain paraffins. *Frontiers in Microbiology* **4**: 89.
- Cech, T.R. (1988) Conserved sequences and structures of group I introns: building an active site for RNA catalysis — a review. *Gene* **73**: (2) 259–271.
- Chatterjee, T. and Bhattacharyya, D. (2001) Biotransformation of limonene by *Pseudomonas putida*. *Applied Microbiology and Biotechnology* **55**: (5) 541–546.
- Clomburg, J.M., Cintolesi, A., and Gonzalez, R. (2022) *In silico* and *in vivo* analyses reveal key metabolic pathways enabling the fermentative utilization of glycerol in *Escherichia coli*. *Microbial biotechnology* **15**: (1) 289–304.
- Dong, X., Greening, C., Bröls, T., Conrad, R., Guo, K., Blaskowski, S., et al. (2018) Fermentative Spirochaetes mediate necromass recycling in anoxic hydrocarbon-contaminated habitats. *The ISME journal* **12**: (8) 2039–2050.
- Eaton, R.W. (1997) p-Cymene catabolic pathway in *Pseudomonas putida* F1: cloning and characterization of DNA encoding conversion of p-cymene to p-cumate. *Journal of Bacteriology* **179**: (10) 3171–3180.
- Edgell, D., R., Belfort, M., and Shub, D., A. (2000) Barriers to intron promiscuity in bacteria. *Journal of Bacteriology* **182**: (19) 5281–5289.

- Fowler, S.J., Gutierrez-Zamora, M.-L., Manefield, M., and Gieg, L.M. (2014) Identification of toluene degraders in a methanogenic enrichment culture. *FEMS Microbiology Ecology* **89**: (3) 625–636.
- Goddard, M.R. and Burt, A. (1999) Recurrent invasion and extinction of a selfish gene. *Proceedings of the National Academy of Sciences* **96**: (24) 13880–13885.
- Gonzalez, R., Murarka, A., Dharmadi, Y., and Yazdani, S.S. (2008) A new model for the anaerobic fermentation of glycerol in enteric bacteria: trunk and auxiliary pathways in *Escherichia coli*. *Metabolic Engineering* **10**: (5) 234–245.
- Hansen, T., Musfeldt, M., and Schönheit, P. (2002) ATP-dependent 6-phosphofructokinase from the hyperthermophilic bacterium *Thermotoga maritima*: characterization of an extremely thermophilic, allosterically regulated enzyme. *Archives of Microbiology* **177**: (5) 401–409.
- Harder, J. and Probian, C. (1995) Microbial degradation of monoterpenes in the absence of molecular oxygen. *Applied and Environmental Microbiology* **61**: (11) 3804–3808.
- Hausner, G., Hafez, M., and Edgell, D.R. (2014) Bacterial group I introns: mobile RNA catalysts. *Mobile DNA* **5**: (1) 8.
- Heider, J., Szaleniec, M., Martins, B.M., Seyhan, D., Buckel, W., and Golding, B.T. (2016) Structure and function of benzylsuccinate synthase and related fumarate-adding glycol radical enzymes. *Microbial Physiology* **26**: (1–3) 29–44.
- Ji, J.-H., Liu, Y.-F., Zhou, L., Irfan, M., Mbadinga, S.M., Pan, P., et al. (2020) Methanogenic biodegradation of C13 and C14 n-alkanes activated by addition to fumarate. *International Biodeterioration & Biodegradation* **153**: 104994.
- Ji, J.-H., Zhou, L., Mbadinga, S.M., Irfan, M., Liu, Y.-F., Pan, P., et al. (2020) Methanogenic biodegradation of C9 to C12n-alkanes initiated by *Smithella* via fumarate addition mechanism. *AMB Express* **10**: (1) 23.
- Kizina, J., Jordan Sebastian, F.A., Martens Gerrit, A., Lonsing, A., Probian, C., Kolovou, A., et al. (2022) *Methanosaeta* and “*Candidatus Velamenicoccus archaeovor*”. *Applied and Environmental Microbiology* **88**: (7) e02407-21.
- Lambowitz, A.M. and Zimmerly, S. (2011) Group II introns: mobile ribozymes that invade DNA. *Cold Spring Harbor perspectives in biology* **3**: (8) a003616.
- Ma, K., Hutchins, A., Sung, S.-J.S., and Adams, M.W.W. (1997) Pyruvate ferredoxin oxidoreductase from the hyperthermophilic archaeon, *Pyrococcus furiosus*, functions as a CoA-dependent pyruvate decarboxylase. *Proceedings of the National Academy of Sciences* **94**: (18) 9608–9613.

- Perez-Molphe-Montoya, E., Küsel, K., and Overholt, W.A. (2022) Redefining the phylogenetic and metabolic diversity of phylum Omnitrophota. *Environmental Microbiology* **24**: (11) 5437–5449.
- Pozzi, E., Motteran, F., de Mello, B.S., Rodrigues, B.C.G., and Sarti, A. (2024) Biomass profiling in a horizontal-flow anaerobic bioreactor used for limonene degradation. *Current Microbiology* **81**: (10) 323.
- Puentes-Cala, E., Liebeke, M., Markert, S., and Harder, J. (2018) Limonene dehydrogenase hydroxylates the allylic methyl group of cyclic monoterpenes in the anaerobic terpene degradation by *Castellaniella defragrans*. *Journal of Biological Chemistry* **293**: (24) 9520–9529.
- Seymour, C.O., Palmer, M., Becraft, E.D., Stepanauskas, R., Friel, A.D., Schulz, F., et al. (2023) Hyperactive nanobacteria with host-dependent traits pervade Omnitrophota. *Nature Microbiology* **8**: (4) 727–744.
- Shams, F., Oldfield, N.J., Wooldridge, K.G., and Turner, D.P.J. (2014) Fructose-1,6-bisphosphate aldolase (FBA)—a conserved glycolytic enzyme with virulence functions in bacteria: ‘ill met by moonlight.’ *Biochemical Society Transactions* **42**: (6) 1792–1795.
- Sohlenkamp, C. and Geiger, O. (2015) Bacterial membrane lipids: diversity in structures and pathways. *FEMS Microbiology Reviews* **40**: (1) 133–159.
- Sun, L., Toyonaga, M., Ohashi, A., Tourlousse, D.M., Matsuura, N., Meng, X.Y., et al. (2016) *Lentimicrobium saccharophilum* gen. nov., sp. nov., a strictly anaerobic bacterium representing a new family in the phylum *Bacteroidetes*, and proposal of *Lentimicrobiaceae* fam. nov. *International Journal of Systematic and Evolutionary Microbiology* **66**: (7) 2635–2642.
- Toft, C. and Andersson, S.G.E. (2010) Evolutionary microbial genomics: insights into bacterial host adaptation. *Nature Reviews Genetics* **11**: (7) 465–475.
- Toth, C.R.A., Berdugo-Clavijo, C., O’Farrell, C.M., Jones, G.M., Sheremet, A., Dunfield, P.F., and Gieg, L.M. (2018) Stable isotope and metagenomic profiling of a methanogenic naphthalene-degrading enrichment culture. *Microorganisms* **6**: (3) 65.
- Trujillo-Reyes, Á., Pérez, A.G., Cuéllar, S.G., Serrano, A., Cubero-Cardoso, J., Jeison, D., and Feroso, F.G. (2024) Evaluation of toxic effect of monoterpene compounds on anaerobic digestion. *Journal of Environmental Chemical Engineering* **12**: (2) 112035.
- Tuson, H.H., Foley, M.H., Koropatkin, N.M., and Biteen, J.S. (2018) The starch utilization system assembles around stationary starch-binding proteins. *Biophysical Journal* **115**: (2) 242–250.

- Vizoso Pinto, M.G., Pasteris, S.E., and Strasser de Saad, A.M. (2004) Glycerol catabolism by *Pediococcus pentosaceus* isolated from beer. *Food Microbiology* **21**: (1) 111–118.
- Woodson, S.A. and Chauhan, S. (2009) Group I ribozymes as a paradigm for RNA folding and evolution. In *Non-Protein Coding RNAs*. Springer Series in Biophysics. Walter, N.G., Woodson, S.A., and Batey, R.T. (eds). Berlin, Heidelberg: Springer Berlin Heidelberg, pp. 145–166.
- Yamada, T., Sekiguchi, Y., Imachi, H., Kamagata, Y., Ohashi, A., and Harada, H. (2005) Diversity, localization, and physiological properties of filamentous microbes belonging to Chloroflexi subphylum I in mesophilic and thermophilic methanogenic sludge granules. *Applied and Environmental Microbiology* **71**: (11) 7493–7503.
- Ye, Q., Liang, C., Chen, X., Fang, T., Wang, Y., and Wang, H. (2019) Molecular characterization of methanogenic microbial communities for degrading various types of polycyclic aromatic hydrocarbon. *Journal of Environmental Sciences* **86**: 97–106.

Acknowledgments

First and foremost, I am very thankful to my supervisor, **Prof. Dr. Jens Harder**, for giving me the opportunity to work on this project, trusting in my abilities, and allowing me to explore this fascinating enrichment culture. Jens, thank you for your constant guidance, support, and mentorship. Thank you for always being available for questions, providing valuable feedback whenever needed, and detailed data quality control. Your supportive supervision and in-depth knowledge of microbiology and biochemistry made you a great supervisor.

I thank **Prof. Dr. Rudolf Amann** for welcoming me into the molecular ecology group and for the insightful feedback on my project. Thank you, Rudi, for always being present and for many great seminar discussions. Your expertise in microbiology, including many other fields, and deep understanding of methods has been truly inspiring and helpful throughout these years.

Thank you, **Dr. Gunter Wegener**, for being part of my thesis committee and reviewing my thesis. Thank you for being available, your feedback, and the new perspectives you offered.

My thanks go to **Prof. Dr. Michael Friedrich** for being on my thesis committee and for offering insightful perspectives and feedback. Thank you for being the chair of my examination board.

Thank you, **Dr. Katrin Knittel**, for being part of my thesis committee and always providing valuable input. Your expertise, especially in microscopy, has been very helpful and has contributed to beautiful microscopic images.

I thank **Louison** and **Leonard** for being on the examination board.

Thank you, **Sabine**, for teaching me cultivation methods and for always being very supportive and helpful.

Thank you to **Fengqing, Chandni, Anke, Sebastian, and Coto** for your advice and support in bioinformatics. Your input has helped me navigate the computational aspects of this work. I would also like to thank the IT for technical support.

A big thank you goes to **Andreas Ellrott** for technical support and for the great introduction to the microscopes. Thank you, for explaining the complex physical principles which underlie high-resolution microscopy in simple and clear ways.

Acknowledgments

I thank **Kathi, Mirja, and Jörg** for giving me valuable and practical advice and support in the lab.

To my office mates, **Saskia, Anni, and Greta R.** - thank you for many memorable moments, your advice, and your support, as well as the many fun activities, be it board games or triathlons. Thank you to **Clara, Sarah, Dolma, Aman, and Greta G.** for encouragement, support, adventures, and friendship.

Thank you to my family, especially my mom, and to my friends for their constant support and for always being there for me.

I thank the entire Molecular Ecology group. Thank you for the wonderful breakfasts, insightful seminars, and retreats. Thank you all for your contributions, inspiration, support, and for having me as a PhD student.

Universität Bremen
Fachbereich 2
Prüfungsamt Chemie
z. Hd. Frau Frauke Ernst
Leobener Straße

28359 Bremen
Deutschland

Prüfungsamt
Chemie

Frauke Ernst
Geschäftsstelle
Fachbereich 02
Leobener Str. / NW2
D-28359 Bremen

Verwaltungspavillon 06
Tel. 0421 218-62802
Fax 0421 218-9862802
frauke.ernst@uni-bremen.de
www.fb2.uni-bremen.de

Versicherung an Eides Statt

Name, Vorname	Lonsing, Almud
Matrikel-Nr.	
Straße	Celsiusstraße 1
Ort, PLZ	Bremen, 28359

Ich, Almud Lonsing,

versichere an Eides Statt durch meine Unterschrift, dass ich die vorstehende Arbeit selbständig und ohne fremde Hilfe angefertigt und alle Stellen, die ich wörtlich dem Sinne nach aus Veröffentlichungen entnommen habe, als solche kenntlich gemacht habe, mich auch keiner anderen als der angegebenen Literatur oder sonstiger Hilfsmittel bedient habe.

Ich versichere an Eides Statt, dass ich die vorgenannten Angaben nach bestem Wissen und Gewissen gemacht habe und dass die Angaben der Wahrheit entsprechen und ich nichts verschwiegen habe.

Die Strafbarkeit einer falschen eidesstattlichen Versicherung ist mir bekannt, namentlich die Strafandrohung gemäß § 156 StGB bis zu drei Jahren Freiheitsstrafe oder Geldstrafe bei vorsätzlicher Begehung der Tat bzw. gemäß § 161 Abs. 1 StGB bis zu einem Jahr Freiheitsstrafe oder Geldstrafe bei fahrlässiger Begehung.

Ort, Datum / Unterschrift