TECHNISCHE UNIVERSITÄT MÜNCHEN

Lehrstuhl für Grundwasserökologie

BSS and beyond – The structure and function of anaerobic hydrocarbon degrader communities in the environment as traced via gene markers for catabolic key reactions

Frederick Conrad Reinhard von Netzer

Vollständiger Abdruck der von der Fakultät Wissenschaftszentrum Weihenstephan für Ernährung, Landnutzung und Umwelt der Technischen Universität München zur Erlangung des akademischen Grades eines

Doktors der Naturwissenschaften

genehmigten Dissertation.

Vorsitzende(r): Univ.-Prof. Dr. S. Scherer

Prüfer der Dissertation:

1. Univ.-Prof. Dr. R. U. Meckenstock
2. Univ.-Prof. Dr. W. Liebl

Die Dissertation wurde am 30. August 2012 bei der Technischen Universität München eingereicht und durch die Fakultät Wissenschaftszentrum Weihenstephan für Ernährung, Landnutzung und Umwelt am 17. Juli 2013 angenommen.

« La science est faite suivant les données fournies par un coin de l'étendue. Peut-être ne convient-elle pas à tout le reste qu'on ignore, qui est beaucoup plus grand et qu'on ne peut découvrir. »

Aus „Bouvard et Pécuchet“ von Gustave Flaubert, 1881

Abbreviation index

16S rRNA	ribosomal RNA, small subunit
ASS	alkylsuccinate synthase
assA	gene for alkylsuccinate synthase, alpha subunit
ARB	Software for calculating phylogenetic trees (Ludwig *et al.* 2004), not an abbreviation (derived from lat. *arbor*, tree)
ATP	Adenosine triphosphate
bamB	gene for benzoyl-CoA reductase, beta subunit, in obligate anaerobes
bamA	gene for ring cleaving hydrolase, alpha subunit
bbs	operon involved in the degradation of benzylsuccinate
bcr	gene for benzoyl-CoA reductase in facultative anaerobes
BSA	bovine serum albumin
BSS	benzylsuccinate synthase
bssA	gene for benzylsuccinate synthase, alpha subunit
bp	base pairs
BTEX	benzene, toluene, ethylbenzene, xylenes
bzd	gene for benzoyl reductase
cAMP	cyclic adenosine monophosphate
CoA	coenzyme A
DFG SPP 1319	Priority program of the German research funding organisation "Biologische Umsetzungen von Kohlenwasserstoffen in Abwesenheit von Sauerstoff: Von molekularer zu globaler Ebene"
DNA	deoxyribonucleic acid
	DNA-SIP stable isotope probing based on labelled DNA
dNTP	deoxyribonucleotide
et al.	*et alii*
FAE	fumarate-adding enzyme
FAM	5′-6-carboxyfluorescein (dye for labelling primers for T-RFLP)
*Hae*III	restriction enzyme of *Haemophilus aegypticus*
kb	kilobase, 1000 base pairs
MAS	methylalkylsuccinate synthase
masD	gene for methylalkylsuccinate synthase, delta subunit
*Msp*I	restriction enzyme of *Moraxella sp.* ATCC 49670
NMS	naphthylmethylsuccinate synthase
nmsA	gene for naphthylmethylsuccinate synthase, alpha subunit
OTU	organizational taxonomic unit
PAH	polycyclic aromatic hydrocarbon
PCR	polymerase chain reaction
PEG	polyethylene glycol
qPCR	quantitative (=real time) polymerase chain reaction
RNA	ribonucleic acid
rRNA	ribosomal RNA
SAM	S-adenosylmethionine
SIP	stable isotope probing
s. l.	*sensu lato*
s. str.	*sensu stricto*
spp.	species (plural)
*Taq*I	restriction enzyme of *Thermus aquaticus*
T-RFLP	terminal restriction fragment length polymorphism
T-RF	terminal restriction fragment
U	unit of enzyme activity, 1 µmol/min

Zusammenfassung

Kohlenwasserstoffe sind gefährliche Substanzen für die meisten Organismen. Aufgrund ihrer chemischen Eigenschaften sind sie schwer abbaubar und können, besonders unter Abwesenheit von Sauerstoff, in der Umwelt sehr beständig sein. Dennoch sind Bakterien in der Lage, Kohlenwasserstoffe unter anoxischen Bedingungen abzubauen, was durch die Existenz verschiedener anaerober Abbauwege für Kohlenwasserstoffe belegt ist. Unter ihnen ist die Aktivierung durch Fumarat-Addition am besten untersucht. Dieser besondere Aktivierungsmechanismus – die radikalische Addition von Fumarat an eine Methyl- oder Methylengruppe durch Fumarat-addierende Enzyme (FAE) – ist besonders geeignet für die Entwicklung eines Marker-Gen-Ansatzes für den Nachweis sowie die Charakterisierung der Struktur und Funktion anaerober Populationen von Kohlenwasserstoff-Abbauern in der Umwelt. FAE können eine große Auswahl an Kohlenwasserstoffen aktivieren: aromatische Verbindungen über Benzylsuccinat- (BSS), Alkane und Alkene über Alkylsuccinat- und 2-Methylnaphthalin über Naphthylmethylsuccinat-Synthasen (NMS). Die Gene der Untereinheiten von BSS und ASS sind bereits als spezifische, funktionelle Marker-Gene (*bssA*, *assA*) für anaerobe Kohlenwasserstoff-Abbauer etabliert. Zusätzlich wurden jedoch in mehreren aktuellen, auf Reinkulturen, Anreicherungen und Umweltproben beruhenden Studien neue und tief abzweigende FAE-Gen-Linien entdeckt: Clostridien-*bssA* und ihre Homologe, die ‚F-' und ‚T-Cluster' *bssA*-Linien sowie *nmsA*. In dieser Doktorarbeit war es möglich, die bisher nicht identifizierten ‚F1-' und ‚F2-Cluster'-*bssA* den *Desulfobulbaceae* bzw. *Peptococcaceae* zuzuordnen. Somit konnten zwei in der Umwelt relevante, neuartige *bssA*-Linien phylogenetisch positioniert werden.

Jedoch ist die Präsenz solcher neuartiger FAE-Linien an kontaminierten Umweltstandorten noch nicht eindeutig gezeigt, da die verfügbaren Methoden für den Nachweis von FAE-Genen nicht auf die neuen Gen-Linien ausgerichtet sind. Deswegen wurde hier eine Auswahl neuer, synchroner Primer-Paare entwickelt, um die Breite der bekannten FAE so vollständig wie möglich abzudecken, *bssA* von Clostridien, *nmsA* und *assA* mit eingeschlossen. Es war jedoch nicht möglich, einen einheitlichen PCR-Nachweis zu entwickeln, der die gesamte Diversität der FAE-Gene abdeckt. Die hier neu entwickelten, sich gegenseitig ergänzenden Ansätze wurden mit DNA von verschiedenen Kohlenwasserstoff-abbauenden Reinkulturen, Anreicherungen und Umweltproben marinen und terrestrischen Ursprungs getestet. Dadurch konnten, teilweise sogar unerwartete FAE-Genlinien, die bisher noch nicht an den untersuchten Standorten beschrieben wurden, nachgewiesen werden: Neben verschiedenen

bssA-Linien der Deltaproteobakterien sowie der Clostridien wurden auch zum ersten Mal *nmsA*-homologe Umweltsequenzen gefunden. Diese Ergebnisse werden durch eine sequenzierungsunabhängige, auf der terminalen Restriktions-Fragment Längen-Polymorphismus (T-RFLP)-Analytik basierenden Fingerprinting-Methode unterstützt. Durch diese T-RFLP-Methode ist nun mit geringem zeitlichem und finanziellem Aufwand die Diagnostik von FAE-Gen-Populationen in umfangreichen Probensätzen möglich.

Um eine möglichst umfassende Charakterisierung der Populationsstruktur von anaeroben Kohlenwasserstoff-Abbauern in Umweltproben weiter voranzutreiben, wurde hier erstmals ein Verfahren zur Amplikon-Pyrosequenzierung von FAE-Genen entwickelt. Es wird in der vorliegenden Arbeit gezeigt, dass durch eine solche „Pyrotag-Sequenzierung“ von FAE-Gen-Fragmenten verschiedene systematische Fehler der klassischen Markergen-Analytik umgangen werden. Dadurch ist nun auch für schwierige Proben, die durch methodische Beschränkungen einer Charakterisierung intrinsischer FAE-Gene über T-RFLP oder Klonierung und Sequenzierung bislang nicht zugänglich waren, erstmals eine gesättigte Diversitätsanalyse dieser Populationen möglich. Zusammenfassend bietet diese Dissertation verschiedene wichtige analytische Werkzeuge und neuartige Einblicke in die Biodiversität sowie Ökologie anaerober Kohlenwasserstoff-Abbauer in ihren natürlichen Habitaten. Diese Erkenntnisse erlauben erstmals die Etablierung einer umfassenden ökologischen Perspektive sowie neuer Konzepte und Lösungen in der Charakterisierung und dem Management kontaminierter Standorte.

Abstract

Hydrocarbons are a class of compounds generally harmful for organisms. Due to their chemical characteristics, hydrocarbons can be very persistent in the environment, especially under oxygen-limited conditions. Yet, bacteria are able to utilise hydrocarbons under anoxic conditions: several anaerobic hydrocarbon degradation pathways were recently discovered. Amongst them, fumarate addition is the most intensively investigated one. Its unique and conserved activation mechanism – the radical addition of fumarate to a methyl or methylene group – is ideally suited to develop a marker gene concept for the involved fumarate-adding enzymes (FAE). The application of this marker gene will help to detect as well as to study the structure and function of anaerobic hydrocarbon degrader populations in the environment. FAEs are known to activate a variety of hydrocarbons: aromatic compounds via benzylsuccinate synthase (BSS), alkanes and alkenes via alkylsuccinate synthase (ASS) and 2-methylnaphthalene via naphthylmethylsuccinate synthase (NMS). The genes coding subunits of BSS and ASS have already been established as specific functional marker genes (*bssA*, *assA*) for anaerobic hydrocarbon degraders. However, several recent studies based on pure cultures, laboratory enrichments and environmental samples have shown the existence of new and deeply branching FAE gene lineages, such as clostridial *bssA* and homologues, the 'F-' and 'T-cluster' *bssA* lines, as well as *nmsA*. In this thesis, it was possible to affiliate the as-yet unidentified 'F1-' and 'F2-cluster' *bssA* to members of the *Desulfobulbaceae* and *Peptococcaceae*, respectively. Thus, two major novel lineages of *bssA* in the environment are now phylogenetically placed.

However, as available FAE gene detection assays were not designed to target such novel lineages, their presence in contaminated environments remains obscure. Therefore, a suite of new synchronous primer sets was developed for detecting the comprehensive range of FAE gene markers known to date, including clostridial *bssA*, as well as *nmsA* and *assA*. Still, it was not possible to develop one unified assay recovering the complete diversity of FAE genes. The new complementary assays were tested with DNA from a range of hydrocarbon-degrading pure cultures, enrichments and environmental samples of marine and terrestrial origin. They revealed the presence of several, partially unexpected FAE gene lineages not described for these environments before: Apart from diverse deltaproteobacterial as well as clostridial *bssA* also, for the first time, environmental *nmsA* homologues. These findings are backed up by sequencing-independent dual-digest T-RFLP diagnostics capable of rapidly monitoring and identifying FAE gene populations in large sample sets.

For facilitating future research and gaining new in-depth insights into anaerobic degrader community structure, pyrotag amplicon sequencing and a pioneering analysis pipeline for FAE gene surveys was also developed. I show that pyrotag sequencing of FAE gene fragments circumvents several classical biases of marker gene analyses and allows for a saturated FAE diversity analysis also in samples not amenable by FAE gene T-RFLP or cloning and sequencing before. In summary, this thesis provides several important analytical tools and novel insights into the biodiversity and ecology of anaerobic hydrocarbon degraders in the environment, which fosters the incorporation of an ecological perspective as well as new concepts and solutions in contaminated site monitoring and management.

1. Introduction

1.1 A general overview of hydrocarbon contamination in the environment

Hydrocarbons are a class of diverse compounds that are exclusively composed of carbon and hydrogen. Despite their seemingly simple composition of only two elements, hydrocarbons display widely varying structures and chemical properties. Hydrocarbons exist in different states of matter at room temperature: gaseous, liquid or solid. Hydrocarbons can be classified either according to their general structural arrangement or the saturation level of the covalent bonds between carbon atoms (Wilkes and Schwarzbauer 2010). On structural level, it is possible to distinguish between (i) aliphatic hydrocarbons (single, branched alka(e)nes or closed chains, the cycloalkanes), (ii) aromatic hydrocarbons, sometimes substituted with an aliphatic chain (i.e. classical BTEX compounds like benzene, toluene, ethylbenzene and xylenes) and (iii) polycyclic aromatic hydrocarbons (PAHs, e.g. naphthalene, benzopyrene etc.). Polycyclic aromatic hydrocarbons can also carry additional aliphatic side chain substitutions (e.g. methylnaphthalenes). The saturation level of the covalent bond between carbon atoms is, along with the steric structure of a hydrocarbon, determining the reactivity and stability of the compound due to mesomeric effects. Saturated hydrocarbons have single electron bonds between carbon atoms (e.g. alkanes in general), unsaturated hydrocarbons have two (e.g. alkenes) or even three bonds, such as alkyne compounds (Wilkes and Schwarzbauer 2010).

Although some hydrocarbons can be produced naturally from organisms as secondary metabolites, they are usually regarded as important and dangerous contaminants. Hydrocarbons are on account of their properties generally harmful for organisms and also very persistent in the environment, especially with increasing complexity of the molecular structure (Wilkes and Schwarzbauer 2010).There are different physico-chemical properties of hydrocarbons that are the reason for their comparable resistance to biodegradation: Hydrocarbons with low molecular weight are generally volatile at room temperature and often excellent solvents. In contrast, hydrocarbons with high molecular weight or long chain length are hydrophobic, solid waxes with low mobility. Aromatic hydrocarbons are stabilised by high resonance energy levels due to electron delocalization originating from the characteristically alternating saturated and unsaturated bonds within the compound's ring structure. Per definition, hydrocarbons lack functional groups such as e.g. carboxyl, hydroxyl

and amine groups or other substitutions which could serve as a point of attack for biochemical reactions (Wilkes and Schwarzbauer 2010).

Notwithstanding, many bacteria are in fact able to utilize hydrocarbons as a carbon source and as an electron donor, despite the toxicity of many compounds. The toxicity of hydrocarbons to bacteria is due to their high solubility in the lipid bilayer of cell membranes. This alters the biochemistry of the membrane in such a manner, that e.g. important ion gradients of the cell are no longer maintainable. Hydrocarbon-degrading bacteria often alter the lipid composition of their membranes in order to maintain its functionality during hydrocarbon exposure (Sikkema *et al.* 1995, Duldhardt *et al.* 2010). Bacteria have also evolved ways to cope with the low solubility of the more complex, hydrophobic hydrocarbons by expressing biosurfactants for enhancing their bioavailability. Another strategy is the observed growth of bacteria along the phase interface of hydrophilic growth medium and hydrophobic hydrocarbon substrate (Wackett and Hershberger 2001, Rojo 2010).

Methane and halogenated hydrocarbons are only to be mentioned briefly in this overview as special members of the diverse group of hydrocarbon compounds. Methane is the smallest hydrocarbon, but plays a significant role in the global carbon cycle. In microbiology, methane is of interest as a product of methanogenic archaea (Formolo 2010) as well as a substrate for methanotrophic bacteria and archaea (Shima and Thauer 2005, Knittel and Boetius 2010, Murrell 2010). Methane is under oxic conditions easily degraded by methanotrophic bacteria via methane monooxygenases. Under anoxic conditions, methane is degraded by syntrophic consortia consisting of bacterial and archeal partners. The key enzyme for anaerobic methanothrophy is the methyl-coenzyme M reductase. This enzyme was previously considered to be strictly linked to the last step in methanogenesis by methanogenic archaea (Knittel and Boetius 2010).

Halogenated hydrocarbons are known as important pollutants in the environment mostly due to anthropogenic causes, although some of them can also be produced naturally by different organisms (Wackett and Hershberger 2001, Murphy 2010). The crucial step for the degradation of halogenated hydrocarbons is the dehalogenation triggered by various dehalogenases. The subsequent degradation of the dehalogenated hydrocarbon metabolite may follow common catabolic pathways for hydrocarbon degradation. But, depending on the reaction mechanisms for dehalogenation and the resulting metabolites, this is not necessarily the case (Fetzner 2010, Pieper *et al.* 2010, Zinder 2010). However, methane and halogenated

compounds are not within the scope of this thesis. In general, different biochemical pathways are involved in their initial activation and degradation than the activation pathway on which my thesis is actually focusing.

Hydrocarbons occur naturally in large reservoirs like coal beds and oil deposits, formed over geological timescales by diagenesis of buried biomass. Such buried reservoirs can resurface naturally as e.g. hydrocarbon seeps. Marine hydrocarbon seeps, especially deep sea hydrothermal vents releasing natural gas, represent unique ecosystems mostly independent from phototrophic primary production (except for the oxygen provided from surface systems as electron acceptor). Easily accessible terrestrial hydrocarbon seeps, e.g. asphalt pits in Western Asia, were exploited in small scale for early human industrial use since the dawn of civilisation. The large scale exploitation and usage of hydrocarbon deposits has begun since the industrial revolution. Today, coal and oil are used in huge amounts as energy source as well as important raw material for the chemical industry. During the processing of coal and oil for further use, hydrocarbon-containing waste products are formed. Many occasions of accidental (and, unfortunately, also deliberate) release into pristine ecosystems have been reported. Such contaminations are serious threats to ecosystems, food webs and water resources due to the physico-chemical characteristics and the resulting toxicity of hydrocarbons discussed above. Contaminations manifest themselves e.g. in leachates from hydrocarbon deposits like tar-oil waste from former gas works sites, accidental leakage of hydrocarbon storages and spills occurring at production sites or during transportation of hydrocarbon resources, oil tanker accidents being a prominent example (e.g. Gertler *et al.* 2010, Meckenstock *et al.* 2010).

1.2 Hydrocarbon degradation in the subsurface

Aquifer ecosystems, an important resource for human drinking water production, are especially threatened by hydrocarbon contamination. While the contamination is at first restricted to the point of its origin, the more hydrophilic compounds spread with groundwater flow, forming a contamination plume. In tar-oil contamination scenarios, especially the comparably hydrophilic BTEX-compounds such as benzene, toluene, ethylbenzene and xylenes are key substances in plumes. This plume can thus affect a larger space in the subsurface than the original contamination source. Moreover, subsurface conditions are usually unfavourable for a rapid contamination removal. Electron acceptors, especially

oxygen, are quickly depleted but only slowly replenished by infiltration or groundwater flow (Griebler and Lueders 2008).

The removal of such contaminations can either happen by engineering measures or natural attenuation. Engineering approaches for hydrocarbon removal include methods like e.g. excavation, spatial containment of the contamination source as well as bioremediation methods. Bioremediation methods imply contamination removal driven by biological processes, e.g. amendment of electron acceptors via aeration or bioaugmentation by adding microbes with biodegradative potential to the contaminated site (Wackett and Hershberger 2001). In contrast, natural attenuation implies, according to its definition, passive contaminant removal processes by physico-chemical, abiotic reactions (e.g. dispersion, dilution or adsorption) as well as degradation via the microbial community naturally occurring at the contaminated site (Röling and van Verseveld 2002). Therefore, microbial degradation is the key process for effectively reducing contaminant loads in natural attenuation.

Compared to the catabolism of other compounds, the crucial part for the degradation of hydrocarbons is the initial activation. As mentioned above, this is due to missing functional groups serving as potential points of attack for reactions leading further into central catabolic pathways. Anaerobic and aerobic pathways for degradation of aromatic hydrocarbons share common principles (see Figure 1). As evolutionary perspectives prevent the development of completely distinct catabolic pathways for every single compound, funnelling pathways exist where a compound is initially activated and converted by subsequent reactions to a central metabolite. This central metabolite is then further degraded to compounds usable for the common central metabolism of the bacterial cell, such as acetyl-CoA, or completely oxidised to CO_2 (Fuchs *et al.* 2011).

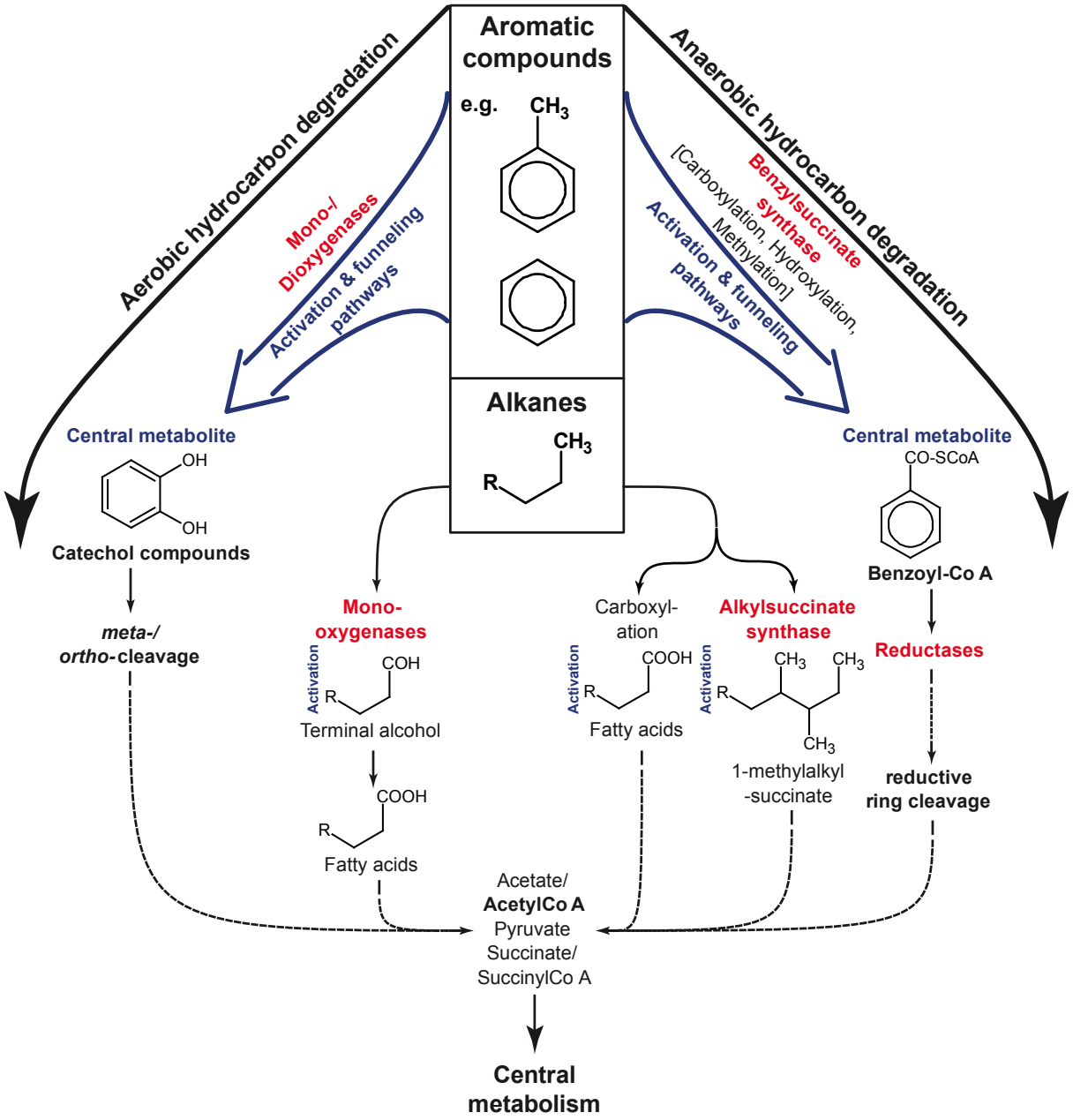

Figure 1. A general overview on fundamental hydrocarbon degradation pathways. Key enzymes whose genes are utilised as functional gene markers are set in red, general reaction types are given otherwise. Anaerobic activation of benzene is still under investigation, with carboxylation being the most likely candidate at the moment (Meckenstock and Mouttaki 2011).

As long as oxygen is available, hydrocarbons can be catabolized rather quickly by bacteria via aerobic degradation pathways. The aerobic degradation of hydrocarbons is well investigated, in contrast to anaerobic degradation. Hydrocarbons are activated by the addition of hydroxyl groups by various mono- or dioxygenases. Oxygen is strictly necessary as a co-substrate for the activation reaction. Monooxygenases introduce only one hydroxyl group from a single oxygen atom, the second is oxidized to water. These enzymes are involved in the breakdown of either aliphatic or substituted aromatic compounds, including benzene. Aliphatic compounds are thus converted into the corresponding alcohol via terminal oxidation by a monooxygenase. This alcohol is then turned into its fatty acid equivalent, ready for subsequent β-oxidation (Rojo 2010). Dioxygenases incorporate both oxygen atoms as two separate hydroxyl groups. Aromatic compounds are metabolised by various upper funnelling pathways into a few different catechol-compounds as central metabolites. The ring structure of the central metabolite is opened by either a *meta*- or *ortho*-cleavage. The metabolites from the ring cleavage are further degraded by subsequent reactions before entering into the central metabolism. Aerobic degradation is generally more effective than anaerobic degradation fuelled by the reduction of alternative electron acceptors. Oxygen is a superior electron acceptor from a thermodynamic (high redox potential) as well as from a kinetic (rapid diffusion through the lipid bilayer) point of view (Fuchs *et al.* 2011).

As hydrocarbon pollution often exceeds the availability of oxygen, particularly in aquatic and subsurface environments (i.e. aquifers), the ecological role of hydrocarbon degradation under anoxic conditions is at least equally important. It was believed for a long time that hydrocarbons, specifically the persistent aromatic compounds, are only degradable in the presence of oxygen (Fuchs *et al.* 2011). This is due to the fact that bacteria involved in anaerobic degradation and their biochemistry were elusive to the hitherto available methods for a long time. Also, degradation rates depending on less optimal electron acceptors than oxygen are generally much lower (Wackett and Hershberger 2001). Only in the recent decades, it was possible to isolate bacteria capable of using hydrocarbons as electron donor under anoxic conditions. The examination of these strains led to the discovery of novel metabolic pathways and activation reactions. Consequently, the investigation of the relevance of these pathways and the corresponding organisms for degradation *in situ* is now important (Widdel *et al.* 2010), which is in line with the objectives of this thesis.

Since oxygen is not available under anoxic conditions for activating reactions, other strategies are necessary. This involves special reactions, often with initially high activation energies.

Currently, there are three general strategies known to be involved in anaerobic hydrocarbon activation (Heider 2007): (i) Anaerobic addition of fumarate via a glycyl radical enzyme to a methyl or methylene group (see Figure 2, explained in depth in chapter *1.3*). (ii) Oxygen-independent hydroxylation, known to be involved in the degradation of ethylbenzene and related compounds (Johnson *et al.* 2001, Kniemeyer and Heider 2001). The ethylbenzene dehydrogenase catalyses the addition of water to the methylene group of the ethyl-chain. This reaction results in a hydroxyl group which serves as a further point of attack for the degradation towards the central metabolite of anaerobic ring cleavage. (iii) Carboxylation, proposed for alkanes (Callaghan *et al.* 2006), methylnaphthalenes (Sullivan *et al.* 2001), naphthalene and benzene (Meckenstock and Mouttaki 2011) as well as for phenantrene (Davidova *et al.* 2007). Carboxylation is still hard to elucidate as a direct activation reaction, since carboxylated hydrocarbons are also general central intermediates of other activation pathways (Meckenstock and Mouttaki 2011). A methylation reaction as an alternative activation (with fumarate addition as a following step) is neither shown nor excluded for naphthalene (Musat *et al.* 2009). Additionally, there was the recent discovery of oxygenases using oxygen released from nitrate reduction for methane and hexadecane in anoxic systems. However, the underlying mechanisms are not yet fully understood (Zedelius *et al.* 2011, Ettwig *et al.* 2012).

For anaerobic degradation of aromatic hydrocarbons, there is a well-characterized central metabolite, analogous to the catechol compounds in aerobic aromatic hydrocarbon degradation (see Figure 1). All aforementioned anaerobic activation pathways funnel aromatic (and maybe also polyaromatic) compounds to benzoyl-CoA. The aromatic ring is desaturated in several steps by benzoyl-CoA reductases. There are initially two systems known for the initial dearomatization step: Either ATP-dependent in facultative anaerobes (benzoyl-CoA reductase BcrABCD in *Thauera aromatica* and *Azoarcus* spp.) or ATP-independent in strict anaerobes (benzoyl-CoA reductase BamBCDEFGHI in *Geobacter metallireducens*). Subsequently, 6-oxocyclohex-1-en-1-carbonyl-CoA is formed, which is in turn cleaved by the ring cleaving hydrolase (BamA). The phototroph *Rhodopseudomonas palustris* differs here from the ATP-dependent reductase-driven pathway as ring cleavage happens as a four electron ring reduction to cyclohex-1-ene-carboxy-CoA. After subsequent β-oxidative-like reactions, carbon dioxide and three molecules of acetyl-CoA are formed, which are in turn funnelled into the central metabolism (Fuchs *et al.* 2011).

1.3 Hydrocarbon activation via fumarate adding enzymes

Fumarate-addition was first reported for the activation of toluene by the benzylsuccinate synthase (BSS) in *Thauera aromatica* strain K172 (Biegert *et al.* 1996). Fumarate adding enzymes (FAE) like BSS belong to the family of pyruvate formate lyases, which are all glycyl radical enzymes (Lehtiö and Goldman 2004). Glycyl radical enzymes need to be activated by radical SAM (S-adenosylmethionine) enzymes. Radical SAM enzymes create a radical glycyl group within the FAE. In the case of BSS, this glycyl radical creates an enzyme-bound benzyl radical from toluene. The fumarate is then added at its double bond to the enzyme-bound benzyl radical. The thus formed benzylsuccinate is subsequently degraded via further activation to CoA-thioesters and reactions similar to β-oxidation to benzoyl-CoA, the aforementioned central metabolite of anaerobic aromatic hydrocarbon degradation. The fumarate is cleaved from the intermediates and recycled for further activation as shown in Figure 2. BSS contains four subunits (α, α', β, γ), encoded by the *bssABC* genes. The involved activating radical SAM enzyme is encoded by the *bssD* gene. (Leuthner *et al.* 1998). The enzymes of the subsequent degradation of benzylsuccinate to benzoyl-CoA are encoded by the *bbs* genes. The *bss* and *bbs* operons were found to be present throughout all anaerobic toluene-degrading pure cultures isolated to date (Carmona *et al.* 2009).

The substrate range of FAE is not only limited to the activation of aromatic compounds such as toluene, xylenes and also ethylbenzene (Heider 2007). The same activation reaction is furthermore used in alkylsuccinate synthases (ASS, also called methylalkylsuccinate synthase MAS) for long (Kropp *et al.* 2000) as well as short (Kniemeyer *et al.* 2007) chain alka(e)nes and in naphthylmethylsuccinate synthases (NMS) for 2-methylnaphthalene activation (Annweiler *et al.* 2000). The analogous activation reactions for BSS, NMS and ASS are illustrated in Figure 2. Cyclohexane may also be activated by fumarate addition (Musat *et al.* 2010). Furthermore, cresols (Heider 2007) and also linear alkylbenzenesulfonate detergents (Lara-Martín *et al.* 2010) have been reported to be activated via fumarate addition. Alternative initial degradation reactions are known for 2-methylnaphthalene (Sullivan *et al.* 2001) and alka(e)nes via the afore mentioned carboxylation (Heider 2007). The structure of the operons encoding ASS/MAS and NMS follow the same pattern as explained above for BSS. Compared to the current knowledge on compounds activated via FAE, oxygen-independent hydroxylation and carboxylation pathways are only well understood for a small number of compounds. Therefore, fumarate-addition can be considered as a key reaction for anaerobic hydrocarbon degradation (von Netzer *et al.* 2013).

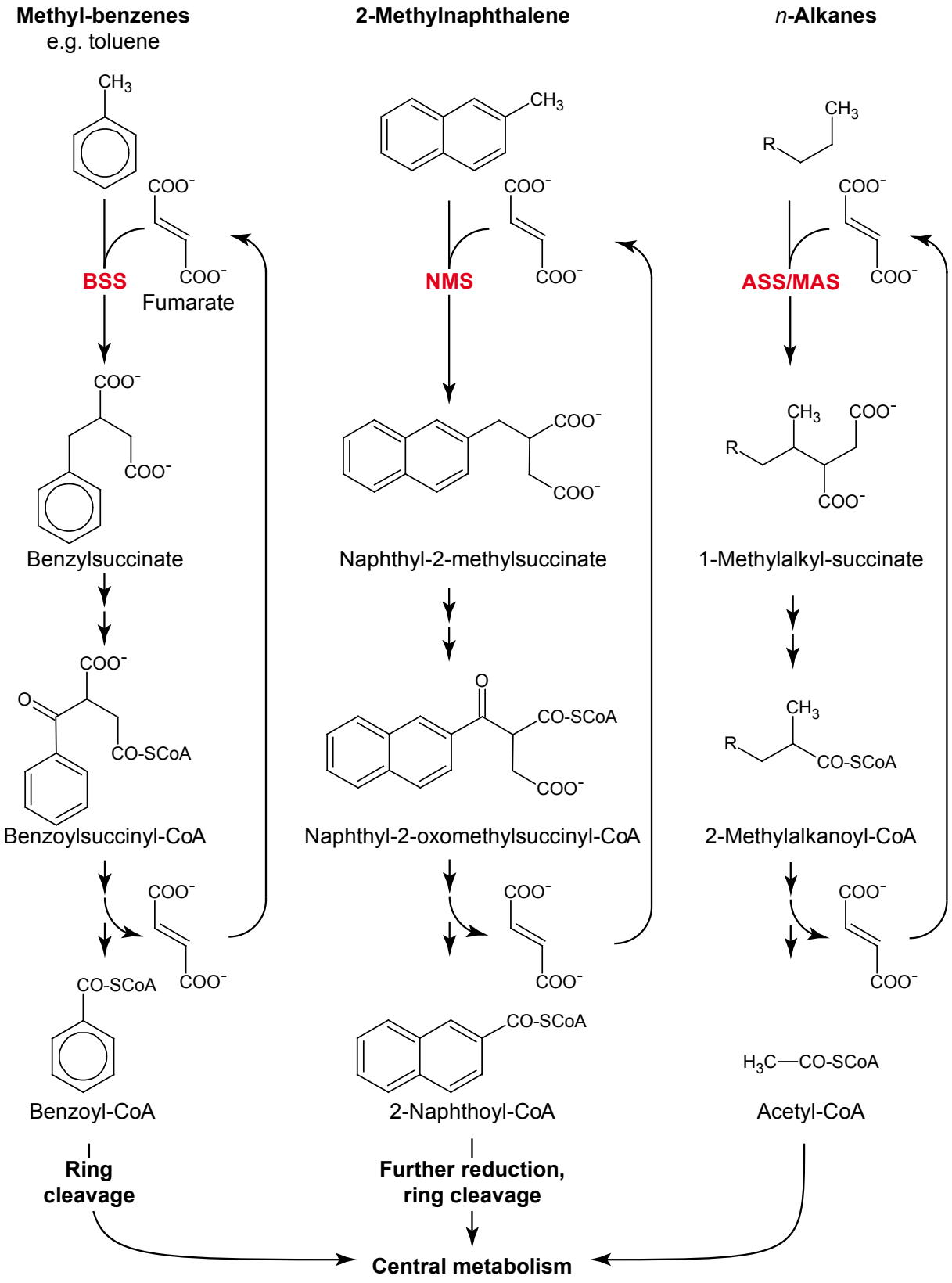

Figure 2. Initial activation of toluene, 2-methylnaphthalene and *n*-alkanes by the FAEs benzylsuccinate synthase (BSS), naphthylmethylsuccinate synthase (NMS) and alkylsuccinate synthase (ASS)/methylalkylsuccinate synthase (MAS). Subsequent degradation steps are simplified, also the recycling of the fumarate.

A wide diversity of bacterial cultures and enrichments is known to use fumarate-addition for activating hydrocarbons and related substances while respiring different electron acceptors (see Figure 3 for 16S rRNA gene phylogeny as well as Figure 4 and Figure 15 for FAE gene phylogeny, see also Widdel *et al.* 2010). Nitrate-reducing, toluene-degrading *Rhodocyclaceae* (*Betaproteobacteria*) such as *Thauera aromatica* K172 (Biegert *et al.* 1996), *Aromatoleum aromatica* EbN1 (Rabus and Widdel 1995, Kube *et al.* 2004) and *Azoarcus* sp. T (Beller and Spormann 1999) were among the first strains known for fumarate addition. Other nitrate-reducing *Azoarcus* spp. enrichments were found to degrade alkanes via fumarate addition (Ehrenreich *et al.* 2000, Rabus *et al.* 2001, Zedelius *et al.* 2011). Recently, the iron- and nitrate-reducing toluene degrader *Georgfuchsia toluolica*, also a member of the family *Rhodocyclaceae*, was isolated from a landfill-leachate contaminated aquifer (Weelink *et al.* 2009). Nitrate-reducing *Rhodocyclaceae* are often facultative anaerobes, for *Thauera* sp. DNT-1 it is reported that it degrades toluene under anoxic (via BSS) and oxic (via toluene dioxygenase TOD) conditions (Shinoda *et al.* 2004). Within the *Alphaproteobacteria*, so far only a nitrate-reducing *Magnetospirillum* sp. TS-6 strain was found to degrade toluene via fumarate-addition (Shinoda *et al.* 2005). The involved BSS is closely related to BSS found in *Rhodocyclaceae* (see Figure 4 or Figure 15 for more detail).

A considerable diversity of strains using FAEs for hydrocarbon activation is reported within the *Deltaproteobacteria*. Different *Geobacter* spp. are known, either from metabolite- or genome-based information, to degrade toluene via BSS (Kane *et al.* 2002, Winderl *et al.* 2007). Members of the *Geobacteraceae* are generally known for their ability of using different electron acceptors. However, BSS-mediated toluene degradation was until now only shown under iron-reducing conditions. It can thus be hypothesized that hydrocarbon degradation by *Geobacter* spp. may also be possible under e.g. uranium- (U(VI)) or manganese- (Mn(IV)) reducing conditions. Among the sulphate-reducing *Deltaproteo-bacteria*, several strains belonging to *Desulfobacteraceae* and *Syntrophobacteraceae* were found to utilize FAE for degrading hydrocarbons. Among the *Desulfobacteraceae*, there are e.g. the toluene-degrading *Desulfobacula toluolica* (Rabus *et al.* 1993), alka(e)ne-degrading *Desulfatibacillum alkenivorans* (So and Young 1999) and a few 2-methylnaphtalene-degrading enrichments (Safinowski and Meckenstock 2004, Musat *et al.* 2009). For the *Syntrophobacteraceae*, the toluene-degrading strain PRTOL (Beller *et al.* 1996) and the alkane-degrading *Desulfoglaeba alkanexedens* (Davidova *et al.* 2006) were found.

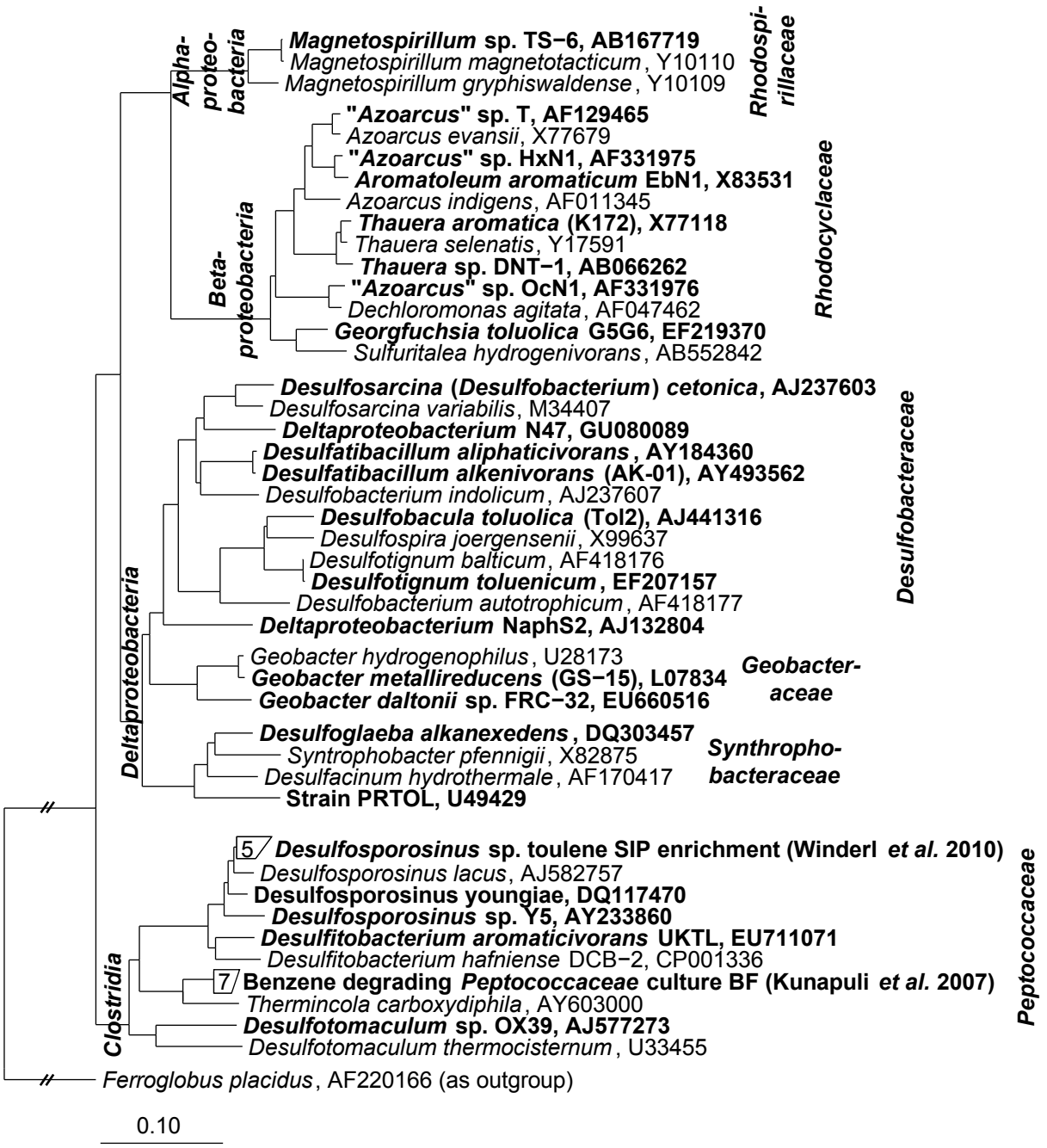

Figure 3. 16S rRNA based phylogeny of cultures and enrichments known to use FAE (similar selection as in FAE based phylogenetic tree in Figure 4 and Figure 15, hydrocarbon degraders with bold letters) and closely related non-hydrocarbon degrading bacteria. Hydrocarbon-degrading *Azoarcus* spp. are set in quotation marks as it is suggested that they belong to a different genus due to their metabolic differences from endophytic *Azoarcus* spp. (Wöhlbrand *et al.* 2007). The tree was calculated based on the 16S rRNA Living Tree Project ARB database version 108 (Munoz et al. 2011, database accessible at http://www.arb-silva.de/projects/living-tree/) in ARB (Ludwig *et al.* 2004) using Phylip distance matrix with the Fitch algorithm in similar manner as Winderl *et al.* (2007).

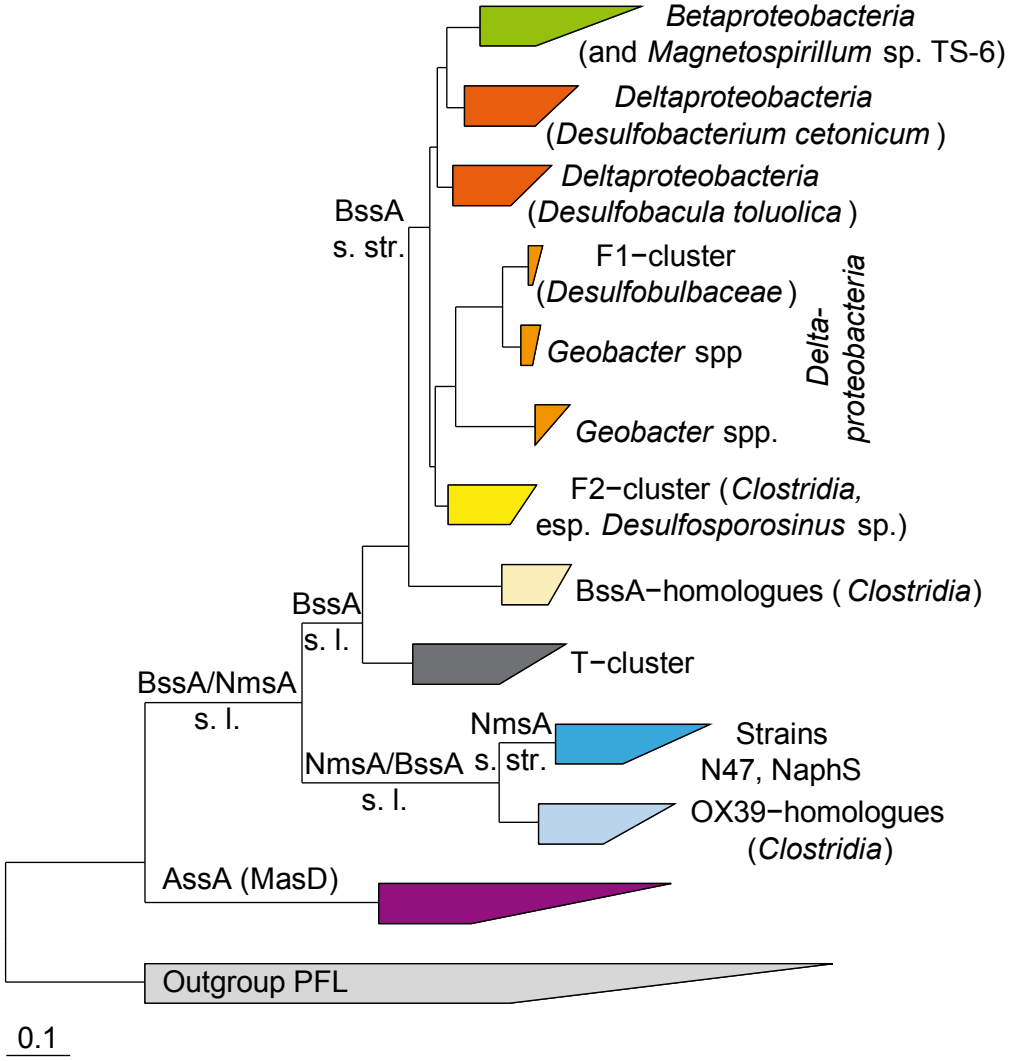

Figure 4. General overview of the amino acid sequence-based phylogeny of known pure culture and environmental FAE-sequences. Based on the phylogenetic tree as displayed in Figure 15, calculated as described in chapter *2.3.5*.

For a considerable time, it seemed that FAEs were limited to *Proteobacteria*. Nonetheless, some Gram-positive toluene degraders were isolated recently. They fall all within the family *Peptococcaceae*, belonging to the class *Clostridia*. Those strains and enrichments are either sulphate-reducers such as toluene-degrading *Desulfotomaculum* spp. (Morasch *et al.* 2004) and *Desulfosporosinus* spp., (Liu *et al.* 2004, Lee *et al.* 2009), or iron reducers like the toluene-degrading *Desulfitobacterium aromaticivoran*s UKTL (Kunapuli *et al.* 2010) and benzene-degrading strain BF (Kunapuli *et al.* 2007). A *bss*-homologue was found in BF, but its role and substrate is still unknown, as benzene degradation was shown to be initiated by carboxylation in this strain (Abu Laban *et al.* 2009). The *Desulfosporosinus* spp. enrichment strain Y5 was found to utilize toluene also under arsenate- (As(V)) reducing conditions (Liu

et al. 2004). Despite these important advances, the knowledge on clostridial BTEX-degraders and their relevance in the environment is still moot.

There is also the possibility of FAE-mediated hydrocarbon degradation under fermentative or syntrophic conditions. This was first demonstrated by R. U. Meckenstock (1999) for co-cultures consisting of *Wolinella succinogenes* and either with the deltaproteobacterial, sulphate-reducing toluene degrader TRM1 (with nitrate or fumarate as electron acceptor) or with *Geobacter metallireducens* (with fumarate as electron acceptor). Both degraders cannot use fumarate or nitrate as electron acceptors themselves, but *W. succinogenes* served as an electron sink, thus enabling fermentative degradation of toluene. The involvement of FAE in hydrocarbon degradation under methanogenic conditions was later also shown for toluene (Washer and Edwards 2007, Fowler *et al.* 2012), alkanes (Mbadinga *et al.* 2011) and PAHs (Berdugo-Clavijo *et al.* 2012). It is remarkable that the presumed toluene- and PAH-degrading key players in the methanogenic enrichments were also affiliated to *Clostridia* (Washer and Edwards 2007, Fowler *et al.* 2012).

1.4 Methods for detecting FAE-mediated hydrocarbon degradation in the environment

As previously stated, contaminant removal via anaerobic degradation is a very important process in subsurface environments. Among the different hydrocarbon activation modes, FAE provide the largest substrate versatility. Therefore, FAE genes are currently quite suitable for tracing microbial natural attenuation of hydrocarbons. There are several general and specific strategies for tracing FAE-mediated hydrocarbon degradation and the involved degraders in the environment: Compound specific stable isotope analysis, determination of characteristic metabolites (metabolic markers) and the detection of involved genes (genetic markers).

Compound specific stable isotope analysis for determining stable isotope fractionation is used to generally distinguish physico-chemical attenuation processes from microbial contamination degradation. This method relies on the effect of different reaction types on the natural stable isotope signature of a compound. While physico-chemical processes such as transport, dilution and sorption do not affect isotope signatures, microbial degradation favours molecules containing lighter isotopes. Thus, molecules with heavy isotopes will enrich during degradation. This enrichment is not observable in the isotope signature if a measured concentration decrease at a contaminated site originates from sorption or dilution. If stable isotope enrichment factors are available from laboratory experiments, it may be also possible

to identify the initial activation mechanism active at a contaminated field site. Thus, due to the distinct isotopic enrichment factors of different degradation pathways, it is possible to distinguish e.g. between aerobic and anaerobic BTEX degradation. However, while this method helps determining if the decrease of a compound is related to degradation, it cannot identify the involved organisms (reviewed by Meckenstock *et al.* 2004).

Metabolic markers are used as a pathway specific detection system for tracing anaerobic hydrocarbon degradation. A metabolic marker is a characteristic intermediate specific for a certain degradation pathway. Depending on its stability, a metabolite may thus provide a snapshot of an active degradation process. This method is established for various characteristic metabolites of aerobic and anaerobic hydrocarbon degradation pathways (Callaghan 2012). For the detection of FAE-mediated hydrocarbon degradation, benzylsuccinate, 2-methylnaphthalene succinate and alkylsuccinate can be used as metabolic markers. (e.g. Elshahed *et al.* 2001, Young and Phelps 2005, Oka *et al.* 2011). However, tracing metabolites as markers for degradation processes requires knowledge of the involved pathways and elaborate analytical equipment such as mass spectrometers. An identification of the involved bacteria is hardly possible in this manner (Callaghan 2012).

A genetic marker targets either a gene for a key enzyme of a metabolic or respiratory pathway (functional marker) or a gene for a highly conserved cellular structure – i.e. the 16S subunit of the rRNA (phylogenetic marker). In contrast to the methods mentioned above, the usage of genetic markers is simpler, more versatile and has a more diagnostic yield of information. This approach relies on various established molecular biology techniques such as gene detection via amplification by specific primers in PCR (Polymerase chain reaction), microscopic detection via FISH (fluorescence in situ hybridization) and detection of comprehensive sets of genetic markers at once via microarrays. Phylogenetic markers targeting the highly conserved 16S rRNA genes resolve the composition of the total microbial community. Functional markers are able to determine the presence or absence of certain catabolic pathways in a microbial community. Depending on the source material, functional gene markers can be used to detect the catabolic potential (DNA, e.g. Taylor *et al.* 2002, Song and Ward 2005) or actual gene expression (mRNA, e.g. Wilson *et al.* 1999, Yagi *et al.* 2010) in hydrocarbon degrader communities. The knowledge of the identity of the hydrocarbon-degrading bacteria present at an impacted site and their population structure is crucial for assessing natural attenuation, e.g. electron acceptor and donor, substrate range, growth conditions etc. (reviews by Galvão *et al.* 2005, Weiss and Cozzarelli 2008). In this thesis, I

will focus on PCR-dependent detection of genetic markers, especially functional markers for FAE.

The existing detection systems are not yet optimized for a comprehensive, ecological investigation of degradation processes: Either the detection systems mostly target specific single, model hydrocarbon compounds as e.g. BTEX or they are solely based on existing pure cultures. Thus, it may not be possible to detect or to identify novel degraders based solely on stable isotope fractionation ratios or metabolites. Also, metabolic markers can, depending on protocol and detection system, only target a limited class of metabolites. Optimally, a comprehensive detection system for a wide range of hydrocarbons should be employed for monitoring in order to get an overview of *in situ* degradation processes. This is due to the fact that hydrocarbons mostly do not occur as single compounds at impacted sites (e.g. Kleindienst *et al.* 2012, Winderl *et al.* 2007).

In contrast, genetic markers offer more versatility for gaining better insights into degrader ecology. Tracing the general bacterial community and also known degraders on 16S rRNA gene level alone is not sufficient as it is not possible to allocate specific catabolic potentials to unknown FAE-containing degraders. This is prevented by the polyphyletic distribution of the ability to degrade hydrocarbons via fumarate-addition, as illustrated in Figure 3: Many close relatives of known hydrocarbon degraders cannot use hydrocarbons and clear degrader identification is thus only based on the available isolates. Yet with comprehensive, but specific functional gene markers targeting FAE, it is possible to test the relevance of known hydrocarbon-degrading isolates *in situ*, and also to find novel, unknown degraders in the environment (Winderl *et al.* 2007).

The detection of hydrocarbon degraders via molecular targeting of FAE genes is already established for some subsets of FAE (e.g. Beller *et al.* 2002, see chapter *1.5* and Table 1 for more detail). But apart from FAE genes, several other functional markers are also employed successfully for detecting this functional guild: The benzoyl-CoA reductases *bcr*/*bzd* (Hosoda *et al.* 2005, Song and Ward 2005) or *bamB* (Löffler *et al.* 2011) as well as the ring cleaving hydrolase *bamA* (Kuntze *et al.* 2008, Staats *et al.* 2011) of the central metabolism of aromatic hydrocarbon degradation. All have been used in combined assays to detect degraders in enrichments and environmental samples (Kuntze *et al.* 2011).

Other examples for long-standing functional gene markers can be found for analysing e.g. aromatic ring-hydroxylating dioxygenases (Taylor *et al.* 2002), methanotrophic and

methanogenic communities. The genes for methane monooxygenases are employed on different platforms, e.g. primers for PCR detection (McDonald and Murrell 1997), microarrays (Bodrossy *et al.* 2003) and pyrosequencing assays (Lüke and Frenzel 2011). Also, genes for methyl-coenzyme M reductases are used for detection of microbial communities responsible for methanogenesis as well as anaerobic methanotrophy (Friedrich 2005).

Although FAE display wide substrate ranges, due to their unique biochemistry – addition of a fumarate radical – they harbour specific protein motifs whose sequences are ideal targets for creating functional gene PCR assays. Thus, FAE genes are ideally suited as a functional gene marker for analysing anaerobic hydrocarbon degrader communities.

1.5 Gene detection assays for FAE genes

Several primers targeting the genes for the α-subunit of BSS and ASS are already in use as markers for the detection of certain hydrocarbon degrader lineages (see Table 1 for an extensive overview) in the environment. The first primers for normal and quantitative PCR targeting *bssA* genes of nitrate-reducing *Betaproteobacteria* were introduced by Beller *et al.* (2002). This assay was later updated for sulphate-reducing bacteria (Beller *et al.* 2008). The primers of Washer and Edwards (2007) were specifically designed for a toluene-degrading, methanogenic microcosm. The assay of Winderl *et al.* (2007) extended the range of detectable hydrocarbon-degrading microbes to iron- and sulphate-reducing *Deltaproteobacteria* and revealed, as mentioned above, novel, site-specific degrader populations at different tar-oil impacted aquifers in Germany. Staats *et al.* (2011) applied altered primers first developed by Botton *et al.* (2007) which target *bssA* of iron-reducing degraders at an aquifer contaminated by landfill leachate. The retrieved *bssA* sequences were related to the betaproteobacterial *bssA* sequence of *Georgfuchsia toluolica* (Weelink *et al.* 2009), rather than the *Geobacter* populations expected from *in situ* 16S rRNA gene studies. Recently, Callaghan *et al.* (2010) introduced assays also for ASS genes, evolved from existing *bssA* primers, on the basis of the small number of pure culture *assA* sequences available. These optimized primer sets were applied to DNA extracted from propane- and paraffin-degrading enrichments as well as several aquifer, freshwater and estuarine habitats contaminated with alkanes, revealing for the first time a similar diversity of *assA* genes in the environment as already known for *bssA* genes. Two additional recent primer sets of Kolukirik *et al.* (2011), used for detecting *bssA*

and *assA* in hydrocarbon-degrading microcosm with sediment from the Marmara Sea, provide only a short amplicon. Additionally, the regions amplified by these primers are far apart so that there is no phylogenetic comparison possible between *bssA* and *assA* sequences.

However, these established *bssA* and *assA* gene detection assays were not designed to detect FAE in a comprehensive as well as comparable manner. Additionally, most of them were not designed to target the novel, deeply-branching FAE sequences that have emerged recently, but rather a small selection of strains long known from pure culture studies. Consequently, the presence and detectability of novel, deeply branching FAE gene pools in different environments remains obscure. Already Winderl *et al.* (2007) retrieved several unassigned, deeply-branching *bssA* lineages in their pioneering study, the so-called T- and F-clusters. Using the same primers on a tar-oil contaminated aquifer, Yagi *et al.* (2010) found new sequences related to the F2-cluster. Similarly, Herrmann *et al.* (2009) reported on xylene-degrading enrichments sequences related to T-cluster *bssA* homologues and *Desulfotomaculum* sp. OX39, branching from *nmsA*-sequence clusters. These studies corroborated the existence of new and deeply-branching FAE-lineages, in addition to the known BSS, NMS and ASS lineages.

Furthermore, several novel anaerobic hydrocarbon degraders belonging to *Clostridia* were recently discovered: *Desulfitobacterium aromaticivorans* UKTL, using fumarate-addition for toluene activation (Kunapuli *et al.* 2010) and strain BF, possessing a *bss*-homologous operon (Abu Laban *et al.* 2010) despite utilizing benzene and not toluene. Their FAE genes are not detectable with established *bssA* primers. Moreover, the NMS genes recently described in naphthalene degrading marine strains NaphS2, NaphS3, NaphS6 (Musat *et al.* 2009) and aquifer sediment enrichment strain N47 (Selesi *et al.* 2010) are not targeted by available primers at all. Consequently, the recovery of FAE genes of putatively novel degraders is not yet fully warranted by the existing assays. Thus, there is a strong need for the development and application of more comprehensive primers for FAE genes in general. These primers should be less biased than the established primers for known proteobacterial FAE sequence types. The design of such new primers was one important aim of this thesis.

Table 1. Currently published primers (Beller *et al.* 2002, Washer and Edwards 2007, Winderl *et al.* 2007, Botton *et al.* 2007, Callaghan *et al.* 2008, Beller *et al.* 2008, Callaghan *et al.* 2010, Staats *et al.* 2011, Kolukirik *et al.* 2011) for *bssA* and *assA* employed for environmental samples. Primers are positioned on the *bss* operon of *Thauera aromatica* K172 (Leuthner *et al.* 1998), see also Figure 5.

	forward primer					reverse primer					
	sequence (5'-3') Protein (K127)	***bss* operon** (12448 bp, K127)		***bssA*** (6560-9145 bp, K127 bss operon)		**sequence** (5'-3') reverse complement Protein (K127)	***bss* operon** (12448 bp, K127)		***bssA*** (6560-9145 bp, K127 bss operon)		**Amplicon (bp)**
Beller *et al.* (2002)	AC GAC GGY GGC ATT TCT C D D G G I S P	8744	8762	2184	2202	G CAT GAT SGG YAC CGA CA TG TCG GTR CCS ATC GTA C L S V P I M R	8859	8876	2299	2286	132
Washer &Edwards (2007)	BssA327f C GAA TTC ATC NTC GGC TAC C D E F I V G Y H	6886	6905	326	345	BssA2004r GTC GTC RTT GCC CCA YTT NGG CCN AAR TGG GGC AAY GAC GAC P K W G H D D	8543	8563	1983	1973	1677
	MBssA1516f AG ACC CAG AAG ACC AGG TC K A Q K T R S	8055	8073	1495	1513	MBssA2446r ATG CTT TTC AGG CTC CCT CT AG AGG GAG CCT GAA AAG CAT Q R E P E K H	8988	9007	2428	2417	952
	BssA1985f CN AAR TGG GGC AAY GAC GA P K W G N D D	8544	8562	1984	2002	BssA2524r AT GAT SGT GTT YTG SCC RTA GGT ACC TAY GGS CAR AAC ACS ATC AT T Y G Q N T I I	9062	9084	2524	2494	540
	bssA1-416f CA GAA CAC AAA GTA TGC CC L K S K Y S P	6972	6989	412	429	bssA1-749r TGG AGT TCC AGA GGT CGA TT AA TCG ACC TCT GGA ACT CCA K I D N W K A	7284	7304	744	714	332
	bssA2-413f TAC CTG CAG AGC AAG TAC GC Y L K S K Y S	6968	6987	408	427	bssA2-713r A CTT CAG CTT ATC GGC GTT C G AAC GCC GAT AAG CTG AAG T W H A P S G L E	7252	7271	711	681	303
Winderl *et al.* (2007)	7772f GAC ATG ACC GAC GCS ATY CT D M T D A I L	7772	7791	1212	1231	8546r TC GTC GTC RTT GCC CCA YTT AAR TGG GGC AAY GAC GAC GA K W G H D D D	8546	8565	1986	2005	793
Botton et al. (2007)	TC GAY GAY GGC TGC ATG GA L D D G C M E	8322	8340	1762	1780	TT CTG GTT YTT CTG CAC GTG CAG AAR AAC CAG AA V Q K N Q K	8819	8835	2259	2275	513
Callaghan *et al.* (2008)	*bssA* 1230F GAC ATG ACC GAY GCC ATY CT D M T D A I L	7772	7791	1212	1231	*bssA* 2000R TC GTC GTC RTT GCC CCA YTT AAR TGG GGC AAY GAC GAC GA K W G H D D D	8546	8565	1986	2005	793
Beller *et al.* (2008)	SRBf GTS CCC ATG ATG CGC AGC V P I M R S	8865	8882	2305	2322	SRBr C GAC ATT GAA CTG CAC GTG RTC G C GAY CAC GTG CAG TTC AAT GTC G I D H V Q F N V V	8939	8961	2379	2401	96
Callaghan *et al.* (2010)	Primer Set 1 TTT GAG TGC ATC CGC CAY GGI CT F E C I R H G L	7879	7902	1319	1342	Primer Set 1 TC GTC RTT GCC CCA TTT IGG IGC GCY CCY AAA TGG GGC AAY GAC GA A P K W G H D D	8540	8562	1980	2002	683
	Primer Set 5 TTY GAG TGY ATN CGC CAS GGC F E C I R H G	7879	7902	1319	1342	Primer Set 5 TC RTC ATT NCC CCA YTT NGG CCN AAR TGG GGN AAT GAY GA P K W G H D D	8543	8562	1983	2002	683
	Primer Set 9 CC NAC CAC NAA GCA YGG G L A G R R	8037	8053	1477	1493	Primer Set 9 TC GTC RTT GCC CCA TTT IGG IGC GCY CCY AAA TGG GGC AAY GAC GA A P K W G H D D	8540	8562	1980	8518	525
Staats *et al.* (2011)	bssA3f TC GAY GAY GGS TGC ATG GA L D D G C M E	8322	8340	1762	1780	bssAr TT CTG GTT YTT CTG CAC GTG CAG AAR AAC CAG AA V Q K N Q K	8819	8835	2259	2275	513
Kolukirik *et al.* (2011)	bssA_715f WGG ATC GMC AAG ATC GAY RA W I D K I D N	7274	7293	714	733	bssA_1107r GNC TWA ARG TYT CMG ARC AR RT GYT CKG ARA CYT TSA GNC R L K V S E H	7647	7666	1087	1106	392
	assA_1578f K GAY TTT GAG SAS CTT TTC S A T F D Q L W E	8194	8213	1634	1653	assA_1967r GA CGA CGA YTA YGT GGA CGA TCG TCC ACR TAR TCG TCG TC D D D D A D V	8558	8577	1998	2017	383

1.6 Affiliation of novel benzylsuccinate synthase gene clusters detected in the environment

In a pioneering study on different tar-oil contaminated aquifer sediments in Germany, several novel *bssA* sequence clusters were discovered (Winderl *et al.* 2007): The F1- and F2-clusters were found in Flingern aquifer sediment, the T-cluster at the Testfeld Süd aquifer. Both clusters showed no affiliation to sequences from at that time known degrader cultures (see Figure 4). The Flingern aquifer is a well-studied former gas work site in Düsseldorf characterised by tar-oil contamination. Sulphate reduction was found as the dominating redox process for the degradation of typical BTEX and PAH contaminants. Toluene is the dominating contaminant (10 µg l^{-1} – 60 mg l^{-1}). The contamination forms a characteristic plume with the groundwater flow with an upper plume fringe, a plume core and a lower plume fringe (Anneser *et al.* 2008). Here, two *bssA* sequence clusters unidentified at that time were found: the F1-cluster, branching from *Geobacteraceae bssA* sequences, and the F2-cluster, remote from known proteobacterial *bssA* sequences (Winderl *et al.* 2007). On 16S rRNA gene level, the lower plume fringe was found to be dominated by *Deltaproteobacteria* (mainly *Geobacteraceae* and *Desulfobulbaceae*) and sulphate-reducing *Clostridia* (Winderl *et al.* 2008).

Testfeld Süd is a tar-oil contaminated aquifer in the vicinity of Stuttgart, also located at a former gas work site with mostly gravel in the saturated zone. The seepage velocity is 0.5 – 3.9 m per day. The dominant redox process is sulphate reduction, the important contaminants are BTEX, PAH, phenol and cyanide. Among those, naphthalene was found to be dominant (14 µg l^{-1} – 86 mg l^{-1}) within the contamination plume (Herfort *et al.* 1998, Bockelmann *et al.* 2001, Zamfirescu and Grathwohl 2001, Griebler *et al.* 2004, Winderl *et al.* 2007). Here, the deeply branching T-cluster *bssA* sequences were detected, which have no close relatives amongst genes from hydrocarbon-degrader pure cultures (Winderl *et al.* 2007).

For F1-cluster *bssA,* phylogenetic analysis suggested an affiliation to *Geobacter* spp. (Winderl *et al.* 2007). For the F2- and T-cluster *bssA* sequences, there was no phylogenetic affiliation possible with the then available *bssA* reference sequences. In order to unravel the affiliation of those clusters in this thesis, I analysed samples from stable isotope probing (SIP) experiments previously performed in the context of the theses of Christian Winderl (2007, for T-cluster *bssA*) and Giovanni Pilloni (2011, for F1-cluster *bssA*). This resulted in two collaborative papers in the context of this thesis (Pilloni *et al.* 2011, Winderl *et al.* 2010), where my specific contributions were the degrader community analyses on *bssA* level.

By using ^{13}C-labeled substrates (in both cases toluene), heavy isotopes are incorporated by active degraders into their DNA. The resulting heavier DNA of the degraders able to use the amended substrate is then separated from the lighter DNA of other bacteria not utilizing the labelled substrate by isopycnic centrifugation. The thus isolated degrader genomes can subsequently be analysed by various molecular techniques (Lueders 2010), including functional marker gene screening. The advantage of SIP is the possibility of functionally dissecting a natural community under controlled conditions, without a significant loss of diversity as it would be the case in a normal enrichment.

This gives the means to relate the labelled gene markers for *bssA* to the key players identified via 16S rRNA gene analysis. While assays based on 16S rRNA gene can give information of overall community composition and thus general microbiome characteristics (e.g. general metabolic or respiratory modes etc. known for a taxonomic unit), targeting functional genes like *bssA* can show the presence of not only known, but also unknown degraders not related to pure culture isolates. In the thesis here at hand and the resulting, aforementioned collaborative studies, it is demonstrated how novel hydrocarbon-degraders can be identified with stable isotope probing (SIP) in combination with general (16S rRNA) as well as functional (*bssA*) markers.

1.7 Ecology of hydrocarbon-degrading communities: Dynamics and stability of community composition

As the knowledge on *in situ* degrader community composition advances, there is a need for a more elaborate theoretical framework for understanding the general principles of community assembly and functions. Such a theoretical framework will help to understand the drivers of microbial natural attenuation. Degrader populations can indeed react to disturbances, therefore a better understanding of how degradation performance is maintained is needed. Such concepts describing population dynamics and composition already exist in classical ecology. But microbial ecology is traditionally not well linked with concepts and theories of 'macro'-ecology, because the scientific development of microbiology and ecology was largely disconnected. There are several recent literature reviews demanding a more stringent application of classical ecological concepts in microbial ecology (Horner-Devine *et al.* 2004, McMahon *et al.* 2007, Prosser *et al.* 2007).

Ecological concepts that could be important for a better understanding of microbial natural attenuation are the insurance and intermediate disturbance hypotheses. The insurance hypothesis states that a diverse, functionally redundant community is more resistant and resilient towards environmental change than a specialised, low-diversity community. This may result in stable community functions despite of external disturbance (Yachi and Loreau 1999). The concept is evidently very interesting for situations where stable contaminant degradation is desired. Although subsurface groundwater systems were perceived as stable environments, mostly undisturbed by e.g. dynamics of electron donor/acceptor, recent studies showed that aquifers can be more dynamic than expected due to groundwater table fluctuations. Hence, unexpected reactions of microbial communities towards subsurface ecosystem dynamics were observed in contaminated and oligotrophic aquifers (Pilloni 2011, Lin *et al.* 2012).

The intermediate disturbance hypothesis (Connell 1978) explains the maintenance of a high diversity where ecological niches are occupied by several species. The maximum species diversity is reached with an intermediate frequency and intensity of disturbances, especially if generation times are in sync with the disturbance frequency. In this manner, the effect of interspecific competition for resources is reduced, resulting in diverse and seemingly functional redundant communities. In turn, the intermediate disturbance hypothesis helps to understand community assembly and is relevant to unravel whether aquifer dynamics may enhance or even oppose microbial natural attenuation.

For testing such hypotheses for anaerobic degraders, it is consequently necessary to re-evaluate and enhance the detection systems at hand for this functional guild. This means to detect all members of a degrader community as precisely as possible. Otherwise, important key players and potential seeding populations may be missed. Culture-independent, molecular methods are best suited at the moment for a comprehensive census of microbial populations.

In this thesis, anaerobic degrader communities are established as an ideal model for the evaluation of general ecological concepts for microbes: Community structure can easily be monitored with 16S rRNA and functional gene markers readily define functions. With the enhanced detection systems for FAE genes and high throughput methods introduced in this thesis, it thus becomes possible to dissect a degrader population to finest detail. Greater knowledge on degrader population ecology may then provide in future novel solutions and concepts for contaminated site management.

1.8 High-throughput screening of *bssA* gene diversity

With the continuous expansion of the known FAE gene diversity and the consequential growth of sequence databases, the development of a high throughput screening approach for environmental samples becomes desirable. Such a screening approach will enable the standardised, repetitive observation of anaerobic hydrocarbon-degrading microbial communities in spatial and temporal resolution. Additional, it should be possible to monitor seeding populations of degraders due to an enhanced detection of rare community members. Both are not possible with the standard cloning and sequencing procedure, due to its limited sample size.

Terminal restriction fragment length polymorphism (T-RFLP) is one possibility for the rapid screening of amplicon pools and the community composition in spatial as well as temporal resolution. Amplicons generated with a fluorescent-labelled primer are first digested with a restriction enzyme. The terminal-labelled fragments are then separated and visualised via capillary gel electrophoresis; each fragment has a characteristic length, optimally representing one taxon within a community. The identity of the taxon may be elucidated by cross-referencing observed T-RF lengths with *in silico* digested sequences. T-RFLP is well established using 16S rRNA gene markers (e.g. Lueders and Friedrich 2003) and is not at all laborious or cost-intensive compared to microarrays or next-generation sequencing. In this thesis, I developed a novel, dual-digest T-RFLP fingerprinting method for sequencing-independent diagnostics of major FAE gene lineages in environmental samples.

Microarrays are also well-established tools in microbial diversity studies on either functional gene and/or 16S rRNA gene levels. For example, the PhyloChip targets 16S rRNA genes of bacteria as well as archaea for specifically identifying the members of complex microbial communities (Brodie *et al.* 2006, Andersen *et al.* 2010). The GeoChip designed by He *et al.* (2007) detects a huge, comprehensive set of functional genes simultaneously. In its latest iteration, the GeoChip 4.0 (Hazen *et al.* 2010) covers genes involved in carbon, nitrogen, phosphorous and sulphur cycling, energy metabolism, metal reduction, resistance to antibiotics and metals, organic contaminant degradation, stress response including genes from bacterial phages and important human pathogens. However, this array contains neither the most comprehensive nor the most recent sequence sets of FAE genes necessary for characterising anaerobic degrader communities in detail. There are also microarrays for analysing specific subsets of functional genes, such as an array for exclusively detecting and quantifying methanotrophs by targeting *pmoA* (Bodrossy *et al.* 2003). Another recent example

for a specific functional gene microarray is the hydrocarbon degradation array by Vilchez-Vargas *et al.* (2013) for surveying the catabolic potential of microbial communities for especially aromatic and aliphatic environmental pollutants. This array includes mostly probes for aerobic degradation, but also a limited number of anaerobic degradation genes. However, novel FAE gene lineages are not included here. Therefore, this array is also not yet applicable for screening FAE gene communities in depth.

The development of a high throughput platform for FAE gene screening based on a microarray was suggested in the initial proposal for this thesis. However, the recent developments in next-generation sequencing have made a different, more promising, option feasible. Although a microarray can detect rare community members, it is always dependent on a well-curated sequence and probe database. Thus, it can only detect sequences for which specifically developed probes are at hand (Wagner *et al.* 2007). In contrast, a sequencing assay allows the detection also of novel sequence types with degenerate primers. With next-generation sequencing, it is also possible, due to the high yield of sequence reads, to overcome the problem of short PCR artefacts lowering good sequence output of classical cloning approaches with limited library sizes. There are currently three generic approaches in environmental next-generation sequencing: the analysis of metagenomes on DNA level, of transcriptomes on mRNA level and amplicon-based sequencing of marker genes. Metagenomic approaches are used for gaining comprehensive, non-targeted insights on the genomic potentials in a given sample (on both functional and 16S rRNA gene level), covering optimally all genomes (or all transcribed genes) of the dominating organisms present. Amplicon-based sequencing of marker genes is a much more specific approach for targeting the (deep) diversity of a given microbial gene pool. This allows for a more targeted as well as detailed examination of important markers like 16S rRNA or functional gene markers present in a given sample than a metagenomic approach.

Among the several next-generation, high throughput sequencing platforms in use (Metzker 2010), the Illumina (e.g. Degnan and Ochman 2012), 454-Pyrosequencing (Novais and Thorstenson 2011) and Ion Torrent (Rothberg *et al.* 2011) systems are the most prevalent in microbiology (Glenn 2011, Loman *et al.* 2012). They all use sequencing by synthesis in contrast to chain-terminating sequencing in the established Sanger sequencing (Sanger *et al.* 1977). Sequencing by synthesis allows a massively parallel, automated high throughput work flow. The differences between the next-generation sequencing methods relate to the detection principles used for determining DNA sequences. Illumina uses reversible dye terminators,

pyrosequencing the detection of pyrophosphate released during DNA synthesis and Ion Torrent the detection of conductive changes during synthesis. Among these still rapidly developing methods, pyrosequencing offers the longest read lengths at the moment, with the Roche 454 GS FLX+ system currently reaching up to ~ 1 kb (Loman *et al.* 2012). The 454 GS FLX Titanium system used in this thesis can provide read lengths of up to ~ 500 bp. This is still more than the currently reached read length of Illumina (~ 150 bp) and IonTorrent (~ 100 bp) sequencing (Glenn 2011, Degnan and Ochman 2012, Loman *et al.* 2012).

The read length offered by pyrosequencing approaches is ideally suited for gene-specific amplicon-sequencing. Amplicon pyrosequencing is rather similar to the sequencing of classic amplicon clone libraries, but without the caveats of ligation, transformation and small library size. While amplicon based pyrosequencing of 16S rRNA genes is already well-established and in use as a standard method (Pilloni *et al.* 2012), functional gene amplicon pyrosequencing is currently still emerging. This is due to the fact that there are well-established, ready-to-use bioinformatic resources publicly available for automated 16S rRNA sequence classification (e.g. RDP by Cole *et al.* 2009 and Greengenes by DeSantis *et al.* 2006) but not yet for functional genes in general. This is on account of the fact that every gene marker requires its own validated database and analysis pipeline setup. Thus, developing a first workflow for FAE gene pyrosequencing and classification was one further aim of this thesis. This will enable future high throughput characterisation of anaerobic hydrocarbon degrader communities.

1.9 Aims and scope of this thesis

The aim of this thesis is to elaborate a better understanding of the structure and function of anaerobic hydrocarbon degrader communities in the environment. Gene markers for their key catabolic reactions – the activation of hydrocarbons by fumarate addition are used as a tool to achieve this aim. At the moment, FAE display the highest substrate versatility. But current FAE gene based detection systems do not exploit this versatility and important degrader community members may be missed in the environment. Here, the application of FAE genes as molecular markers offers unique means to understand the ecology of anaerobic hydrocarbon degrader communities in a comprehensive manner. This is important for advancing the general understanding of contaminant breakdown in aquifers, an important ecosystems services provided by microbes

Three main objectives are defined:

I. Several novel *bssA*-lineages tentatively involved in the degradation of toluene are to be identified and affiliated to respective key players on 16S rRNA gene level. This is achieved with DNA samples from SIP experiments with tar-oil contaminated aquifer sediments conducted within the theses of Christian Winderl (2007) and Giovanni Pilloni (2011):

    a. In unfractionated DNA from the Testfeld Süd sediment SIP experiment, microbial populations are traced over time on 16S rRNA gene and *bssA* level. This is done to unambiguously identify clostridial *bssA* and to verify that these degrader populations were initially present in the SIP inoculum and not just enriched during incubation (my contribution to Winderl *et al.* 2010).

    b. *bssA* populations are identified in iron- and sulphate-reducing SIP incubations of aquifer sediment from the Flingern site. Thus, the affiliation of the unidentified F1-cluster *bssA* previously found *in situ* to either *Geobacteraceae* or *Desulfobulbaceae* is to be unravelled (my contribution to Pilloni *et al.* 2011).

II. A comprehensive assessment of FAE gene pools in different terrestrial and marine environments impacted by hydrocarbons is to be facilitated via the development of more universal detection assays. It is hypothesised that the existing FAE gene detection systems do not yet optimally recover the full diversity of catabolic lineages, especially for clostridial *bssA* and *nmsA*-homologues (von Netzer *et al.* 2013):

a. Can one universal primer pair be developed for a wide-ranging recovery of FAE lineages in different pure cultures, enrichments and samples from hydrocarbon-impacted environments? Or are parallel primer sets targeting overlapping regions within the FAE gene alignment necessary to achieve this aim? This question is tackled with own aquifer DNA samples and with samples provided by cooperation partners within the DFG priority programme 1319.

b. For the routine and high-throughput detection of FAE genes in environmental samples, a rapid, sequencing-independent fingerprinting method is introduced. This dual-digest T-RFLP approach is tested and verified for tracing hydrocarbon degraders in space and time in two exemplary systems:

   i. Spatial distribution of hydrocarbon degraders over different depths of the sulfidogenic, lower fringe of the toluene plume in the tar-oil contaminated Flingern aquifer.

   ii. Temporal dynamics of anaerobic hydrocarbon degraders in a butane-amended SIP enrichment provided by DFG SPP 1319 cooperation Sara Kleindienst (2012). This sediment originates from the marine Guaymas hydrocarbon seep.

III. A pipeline for a next-generation sequencing approach of FAE gene pools is to be established. With the growing knowledge on environmental FAE gene diversity, it becomes increasingly important to overcome the classically low sequence yields of outdated cloning and sequencing approaches. Functional gene pyrosequencing may overcome those problems, which is to be demonstrated here, for the first time, for anaerobic hydrocarbon-degrader communities. In this manner, comprehensive diversity coverage and deep sequence screening of anaerobic hydrocarbon degrader gene pools now becomes possible.

2. Material and methods

2.1 Affiliation of novel FAE gene clusters to *Clostridia* via stable isotope probing in Testfeld Süd sediment

Sampling, incubation and biomass collection for DNA extraction was performed by Christian Winderl. I conducted DNA extractions, subsequent 16S rRNA gene and *bssA* T-RFLP for samples of time points 1, 8, 29, 64 and 86 days after incubation in order to trace the emergence of clostridial toluene degraders during incubation.

2.1.1 Sampling site and sample acquisition

Sampling was carried out at a former gas works site (Testfeld Süd) which has been intensively studied and described elsewhere (Herfort *et al.* 1998, Bockelmann *et al.* 2001, Zamfirescu and Grathwohl 2001, Griebler *et al.* 2004; see also chapter *1.1*). Aquifer sediment from the bottom of monitoring well B49 (at ~ 7.50 m depth, groundwater table at ~ 3.25 m) was retrieved in December 2006 with an aqua-sampler (Bürkle, Bad Bellingen, Germany). The sampling was identical to the previous sampling campaign for intrinsic *bssA* gene diversity screening conducted by Winderl *et al.* (2007). Sampled sediments and groundwater were immediately filled into sterile glass bottles without gaseous headspace to minimize oxygen exposure and transported to the laboratory under cooling (Winderl *et al.* 2010).

2.1.2 Incubation of sediments

Replicates of ~8 g (wet weight) of freshly sampled sediment material were anoxically incubated in sterile 120 ml serum bottles containing 50 ml of low salt artificial groundwater medium with a headspace of N_2/CO_2 (80/20). The medium contained all components of the freshwater mineral medium described by Widdel and Bak (1992) in a 1:10 dilution, except for the bicarbonate buffer, which was added to 30 mM final concentration. The artificial groundwater medium was adjusted to pH 7 (Winderl *et al.* 2010).

Table 2. Components for freshwater medium according to Widdel and Bak (1992) in 1:10 dilution.

Compound	Concentration	
	g/l	mM
NaCl	0.1	1.7
$MgCl_2$	0.04	0.42
KH_2PO_4	0.02	0.15
NH_4Cl	0.025	0.47
KCl	0.05	0.67
$CaCl_2$ x 2 H_2O	0.015	0.12

In addition, 1 mM Na_2S was used as reducing agent, 10 mM Na_2SO_4 was added as electron acceptor and 5 mM cAMP was supplemented in order to stimulate microbial activity (Bruns *et al.* 2002). To warrant constantly low *in situ* concentrations of toluene during SIP incubation, 0.3 g of Amberlite XAD7 adsorber resin (Sigma-Aldrich, Munich, Germany) was added to each bottle (Morasch *et al.* 2001). An amount of 5 ml of either non-labelled ($^{12}C_7H_8$) or fully labelled ($^{13}C_7H_8$) toluene (Sigma-Aldrich, Munich, Germany) was injected through butyl rubber stoppers with a gastight syringe. Before incubating the bottles with sediment, the toluene was allowed to adsorb to XAD7 for 2 days. This resulted in a reservoir of 0.96 mM toluene in each bottle. However, actual concentrations were maximally 0.1 times as high due to sorption to XAD7. After sediment addition, the bottles were gassed with N_2/CO_2 (80:20 v/v), sealed anoxically with butylstoppers and incubated statically for over 133 days at 16 °C in the dark. A total of 15 replicate bottles were prepared for each series (^{12}C and ^{13}C) for successive time-dependent termination (Winderl *et al.* 2010).

2.1.3 Nucleic acid extraction

The sediment and biomass was collected as described by Winderl *et al.* (2010) via centrifugation at 4000 rpm (3345 g) and 4 °C for 10 min with a Megafuge 1.0 R (Heraeus Instruments, Hanau, Germany). Pellets were frozen immediately at -20 °C until nucleic acid extraction. For the analyses presented in this thesis contributing to the Testfeld Süd study (chapter *3.1.1*), I extracted total DNA from microcosm sediment with labelled (^{13}C) and unlabelled (^{12}C) toluene for the inoculum and selected successive time points (1, 8, 29, 64 and 86 days after incubation). Approximately 2 ml of incubated sediment was used for DNA extraction following the protocol of Lueders *et al.* (2004), with adaptations as described by Winderl *et al.* (2007):

First, I incubated ~ 0.8-1 g aliquots of the sediment in 2 ml cryo cups (Biozym, Hessisch Oldendorf, Germany) with 750 µl phosphate buffer (120 mM Na_2HPO_4, 125 mM Tris, 25 mM NaCl, pH 8, Winderl *et al.* 2010), 40 µl lysozyme (50 mg/ml in TE buffer, Sigma-Aldrich, St. Louis, MO) and 10 µl proteinase K (10 mg/ml Sigma-Aldrich, St. Louis, MO) at 30 °C for 15 min. This was followed by supplementing 100 µl sodium dodecyl sulphate (20 % w/v, Sigma-Aldrich, St. Louis, MO) and further incubation for 15 min at 65 °C.

These incubation steps were followed by two rounds of bead beating (45 s, 6 m/s and 20 s, 6.5 m/s) with zirconia-silica beads (Carl Roth, Karlsruhe, Germany). I purified the thus released DNA in the supernatant with one volume of phenol-chloroform-isoamylalcohol (25:24:1, pH

8.0, Sigma-Aldrich, St. Louis, MO). This step was followed by another extraction of the supernatant in one volume of chloroform-isoamylalcohol (24:1, Sigma-Aldrich, St. Louis, MO). The organic and water phase was separated by centrifugation (5 min, 4 °C, 20817 g) in Phase Lock Gel Heavy tubes (Eppendorf, Hamburg, Germany).

I precipitated the DNA dissolved in the supernatant with 2 volumes of PEG precipitation solution (30% polyethylene glycol 6000 and 1,6 M NaCl, Sigma-Aldrich, St. Louis, MO) at 4 °C for one hour, followed by centrifugation for 30 min at 20817 g. After carefully discarding the PEG precipitation solution supernatant, the DNA pellet was washed in 70 % (v/v) ice-cold ethanol. The ethanol was removed by pipetting (without disturbing the pellet) and by briefly drying the DNA pellet at room temperature (< 5min) prior to eluting with EB buffer (Qiagen, Hilden, Germany). The volume for elution depended on the expected yield, which resulted in DNA concentrations of ~ 10-20 ng/µl. Extracted DNA was stored at -20 °C or for long term storage at -80 °C.

2.1.4 PCR and T-RFLP

I conducted the 16S rRNA gene PCR with the aforementioned unfractionated DNA as described by Winderl *et al.* (2008) with the primers Ba27f-FAM (Weisburg *et al.* 1991, labelled with 5′-6-carboxyfluorescein) and Ba907r (Muyzer *et al.* 1995). A PCR reaction with a total volume of 50 µl contained nuclease-free water (Promega, Fitchburg, WI), 1x PCR buffer, 1.5 mM $MgCl_2$ (Fermentas, Thermo Fisher Scientific, Waltham, MA), 10 µg BSA (Roche, Basel, CH) 0.1 mM dNTPs (Fermentas, Thermo Fisher Scientific, Waltham, MA), 0.3 µM primer (biomers.net, Ulm, Germany) 1.25 U Taq polymerase (Fermentas, Thermo Fisher Scientific, Waltham, MA) and DNA in the range of ~ 10 - 100 ng. The cycling conditions were set to the following conditions: 5 min initial degradation at 94 °C, 25 cycles of amplification (30 s at 94 °C, 30 s at 52 °C, 60 s at 72 °C) and 5 min at 72 °C for terminal extension.

For *bssA* gene PCR, I used the original 7772f/8546r primer pair (Winderl *et al.* (2007), see also or Table 4). The reverse primer 8546r was FAM-labelled for T-RFLP. The PCR reaction contained the same reagents as described above for 16S rRNA gene PCR, the cycling conditions were set to 3 min initial degradation, 28 cycles of amplification (30 s at 94 °C, 30 s at 52 °C instead at 58 °C, 60 s at 72 °C) and 5 min at 72 °C of terminal extension (Winderl *et al.* 2010).

For T-RFLP, I processed the amplicons as described by Lueders *et al.* (2006). Amplicons were at first purified with MinElute columns (Qiagen, Hilden, Germany) according to the manufacturer's protocol. The purified 16S rRNA and *bssA* amplicons were digested with *Msp*I at 37 °C respectively *Taq*I at 65 °C for two hours (Winderl *et al.* 2010). The digested amplicons were desalted with DyeEx columns (Qiagen, Hilden, Germany) following the manufacturer's instructions. 1 µl of desalted digest was mixed with 13 µl Hi-Di formamide (Applied Biosystems, Darmstadt, Germany) which contained in 1:400 dilution a 6-carboxy-X-rhodamine-labeled MapMarker 1000 ladder (BioVentures, Murfeesboro, TN). The thus prepared amplicons were denatured at 95 °C for 5 min, cooled down to 4°C and separated by capillary electrophoresis on an ABI 3730 DNA analyser (Applied Biosystems, Darmstadt, Germany). Electrophoresis was performed with POP-7 polymer in a 50 cm capillary array under the following conditions: 10 s injection time, 2 kV injection voltage, 7 kV run voltage, 66 °C run temperature and 63 min analysis time. The electropherograms were evaluated with the Gene Mapper 5.1 software (Applied Biosystems, Darmstadt, Germany).

I identified T-RF peaks of *bssA* amplicons by comparing *in silico* T-RF length of available *bssA* sequences with measured peak data. Further T-RFLP data analysis was conducted for 16S rRNA and *bssA* data with the T-REX online software (http://trex.biohpc.org, Culman *et al.* 2009) using the default workflow. Noise filtering was done based on peak height with standard deviation multiplier set to 1. T-RFs were defined by aligning peaks with a clustering threshold of 2 bp for 16S rRNA and 1 bp for *bssA*-data. Due to lower peak diversity, it is possible to set a more stringent clustering threshold for *bssA* peaks.

I constructed the phylogenetic trees displayed in Figure 8 with 16S rRNA gene and *bssA* clones of fractionated DNA originally sequenced in the thesis of Christian Winderl (2007) and published in Winderl *et al.* (2010). Both trees were calculated in ARB (Ludwig *et al.* 2004) via the internal ARB neighbour joining algorithm. The 16S rRNA gene phylogenetic tree in Figure 8A is based on the alignment of the 16S rRNA gene Living Tree Project ARB database version 108 (Munoz *et al.* 2011, database accessible at www.arb-silva.de). The Felsenstein correction algorithm provided within ARB was used for tree calculation. For calculating the BssA dendrogram in Figure 8B, original *bssA* gene sequences were translated within ARB into amino acid sequences and aligned according the FAE alignment established by Winderl *et al.* 2007.The amino acid tree was calculated using neighbour-joining with the Kimura correction algorithm implemented in ARB (Ludwig *et al.* 2004).

2.2 Stable isotope probing for identifying toluene degraders in Flingern aquifer sediments

Sampling, incubation, DNA extraction, ultracentrifugation and subsequent DNA fractionation were done by Giovanni Pilloni. I performed additional *bssA* T-RFLP on selected density-resolved fractions as indicated below for identifying the key toluene degraders on functional gene level.

2.2.1 Sampling site and sample acquisition

Sampling was performed within a well-studied (Anneser *et al.* 2008, Anneser *et al.* 2010, Winderl *et al.* 2008, see also chapter *1.1*) tar-oil contaminated aquifer in Düsseldorf-Flingern, Germany, in September 2008. Sediments ranging between 6.9 and 7.75 m (the lower plume fringe) below the ground surface were sampled from intact direct-push liners and immediately dispensed into 1 l glass bottles full of anoxically autoclaved (under N_2/CO_2) deionized water. Bottles were directly closed without gaseous headspace to minimize oxygen exposure of sediments and transported to the laboratory under cooling at 4 °C (Pilloni *et al.* 2011).

2.2.2 Incubation of sediments

Replicates (~ 8 g wet weight) of homogeneously mixed sediment material were transferred (in an anoxic tent) into sterile 120 ml serum bottles containing 50 ml of artificial anoxic groundwater medium as described previously (Winderl *et al.* 2010, see also chapter *2.1.2*). To ensure constantly low *in situ*-like concentrations of toluene during SIP incubation, 0.3 g of Amberlite XAD7 absorber resin (Sigma-Aldrich, Munich, Germany) in each microcosm was loaded with a total of 1 mM of either non-labelled ($^{12}C_7$) or fully labelled ($^{13}C_7$) toluene (Sigma-Aldrich, Munich, Germany), as previously described (Winderl *et al.* 2010, see also chapter *2.1.2* above). Parallel incubation series were amended with either 10 mM of sodium sulphate or 40 mM of amorphous Fe(III) oxyhydroxide prepared by titration of a solution of $FeCl_3$ to a pH of 7 with NaOH (Lovley and Phillips 1986). Twelve replicate microcosms for each treatment were set up. Abiotic control bottles (autoclaved three times), amended with each electron acceptor and unlabelled toluene, were also prepared to exclude the occurrence of abiotic redox reactions. The bottles were incubated statically for over 4 months in the dark at 16 °C, which was close to *in situ* aquifer temperatures of 14 – 16 °C (Pilloni *et al.* 2011).

2.2.3 Nucleic acid extraction and ultracentrifugation

At selected time points, a pair of bottles (^{12}C and ^{13}C incubation) for each electron acceptor series (sulphate, ferric iron) was sacrificed for DNA-SIP analyses. Sediment, XAD7 and

biomass were collected by centrifugation at 3350 g at 4 °C for 10 min using a Megafuge 1.0 R (Heraeus Instruments, Hanau, Germany). Pellets were frozen immediately at -20 °C and DNA extracted as described previously (Winderl *et al.* 2008, see also chapter *2.1.3* above), with minor modifications (~12 h of DNA precipitation at 4 °C) For each single extract, replicate extractions were pooled in maximum 100 ml of elution buffer (Qiagen, Hilden, Germany) and stored frozen (-20 °C) for downstream analyses. Approximately 1 µg of PicoGreen-quantified (Invitrogen, Darmstadt, Germany) DNA extract was loaded onto a gradient medium of CsCl (average density 1.71 g mL^{-1}, Calbiochem, Merck, Darmstadt, Germany) in gradient buffer (0.1 M Tris-HCl at pH 8, 0.1 M KCl, 1 mM EDTA) and centrifuged (180,000 g at 65 h) as described elsewhere in detail (Lueders 2010, Winderl *et al.* 2010). Thirteen fractions from each gradient were collected from 'heavy' to 'light' using a Perfusor V syringe pump (Braun, Melsungen, Germany). Refractometric measurement of fraction buoyant density and the recovery of DNA from gradient fractions were also performed as previously published (Lueders 2010, Pilloni *et al.* 2011).

I conducted the *bssA* T-RFLP analysis with gradient fractions (Table 3) selected according to 16S rRNA gene copy number per fraction as determined in the thesis of Giovanni Pilloni (2011). I conducted the PCR for detecting *bssA* as described above (chapter *2.1.4*), with annealing temperature set at 58 °C as initially described by Winderl *et al.* (2007). I performed T-RFLP and all following downstream analyses as described in chapter *2.1.4*.

Table 3. DNA fractions of SIP incubations of Giovanni Pilloni (2011) amended with ferric iron and sulphate used for *bssA* T-RFLP analysis. The results are displayed in Figure 12.

	Ferric iron		**Sulphate**	
	CsCl [g/ml]	Gradient fraction	CsCl [g/ml]	Gradient fraction
$^{12}C_7$-Toluene	1.6950	NH09	1.6906	NF09
$^{13}C_7$-Toulene	1.6994	NI09	1.6961	NG09

2.3 Enhanced gene detection assays for fumarate-adding enzymes

New primer sets for FAE genes and a dual-digest T-RFLP for rapidly tracing degrader populations were developed in this thesis as described in detail below. For this, I cultivated and extracted DNA of *Desulfotignum toluenicum* DSMZ 18732, *Desulfatibacillum alkenivorans* DSMZ 16219, *Desulfatibacillum aliphaticivorans* DSMZ 15576, *Desulfitobacterium aromaticivorans* UKTL and *Desulfosporosinus* sp. enrichments. DNA

from other strains (*Azoarcus* sp. T, *Geobacter metallireducens* DSMZ 7210, strains BF and N47) was already present in the Institute of Groundwater Ecology sample collection. DNA extracts of Guaymas and Gulf of Mexico hydrocarbon seeps as well as from the Nyegga methane seeps were provided by DFG SPP 1319 collaborators Sara Kleindienst and Katrin Knittel. DNA extracts from sediment enrichments of Paclele Mici mud volcano and Gölzau aquifer sediment were provided by DFG SPP 1319 collaborators Friederike Gründger and Martin Krüger. Flingern aquifer sediment and DNA extracts were provided by Giovanni Pilloni. Cloning, sequencing and subsequent analyses of FAE amplicons were performed by me. The SIP experiment with Guaymas sediment and subsequent DNA extraction was performed by Sara Kleindienst, I conducted FAE gene-based T-RFLP analysis.

2.3.1 Primer design

I designed novel primers using an updated and curated version of the *bssA* ARB-database originally set up by Winderl *et al.* (2007). Several new *bssA*, *nmsA* and *assA* (also known as *masD*) sequences were added (see Figure 15). The primer sets 'FAE-B', 'FAE-N' and 'FAE-KM', presented in Table 4, were designed to match the same target region as the original primers of Winderl *et al.* (2007), which were also applied in this study (see Figure 5). The binding sites for the new primers were found by searching the FAE-alignment for conserved, distinctive protein motifs common for mayor gene clusters and using the DNA sequences of those protein motifs for primer design (von Netzer *et al.* 2013).

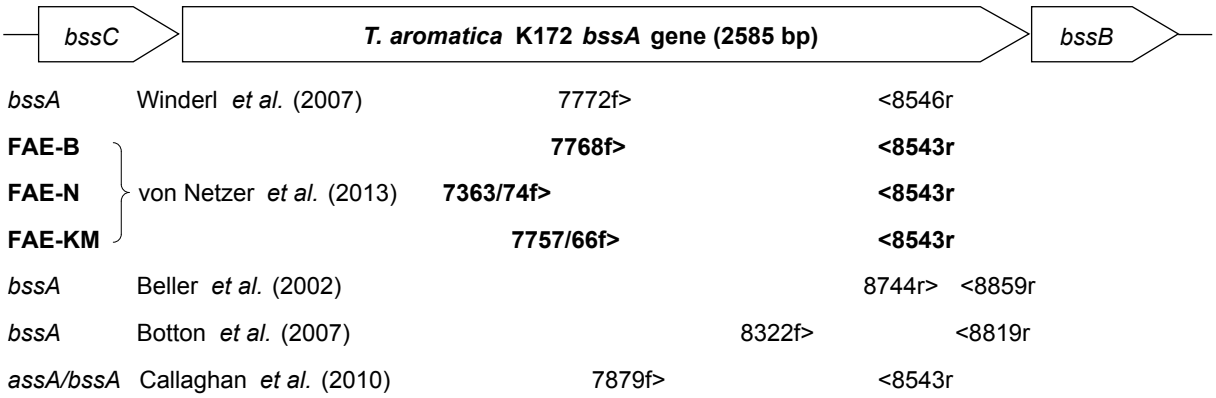

Figure 5. Graphical overview of primer localization within the reference *bssA* gene of *Thauera aromatica* K172 (Leuthner *et al.* 1998).

Table 4. Primer sets designed and employed in this thesis for targeting FAE genes.

Primer set	Annealing temperature (°C)	Primers, position[1]	Sequence (5'-3')	targeted FAE lineages[2]
***bssA* primers[3]**	58/52	7772f	GAC ATG ACC GAC GCS ATY CT	*bssA* s. l.
		8546r	TC GTC GTC RTT GCC CCA YTT	
FAE-B	58	7768f	C AAY GAT TTA ACC RAC GCC AT	clostridial *bssA*, *bssA* s. l., *nmsA*
		8543r	TC GTC RTT GCC CCA YTT NGG	
FAE-N	58	7363f	TC GCC GAG AAT TTC GAY TTG	*nmsA* s. str.
		7374f	TTC GAY TTG AGC GAC AGC GT	
		8543r	TC GTC RTT GCC CCA YTT NGG	
FAE-KM	58	FAE-Kf 7757f-1	TCG GAC GCG TGC AAC GAT CTG A	*assA*
		7757f-2	TCG GAC GCG TGC AAC GCC CTG A	
		FAE-Mf 7766f	TGT AAC GGC ATG ACC ATT GCG CT	
		8543r	TC GTC RTT GCC CCA YTT NGG	

[1] *T. aromatica* 172 *bss* operon position (Leuthner *et al.* 1998)

[2] targeting as intended *in silico*, see lineages in Figure 3

[3] Primer set by Winderl *et al.* (2007, 2010)

2.3.2 Samples and DNA extraction

Desulfotignum toluenicum DSMZ 18732, *Desulfatibacillum alkenivorans* DSMZ 16219, *Desulfatibacillum aliphaticivorans* DSMZ 15576 and *Geobacter metallireducens* DSMZ 7210 were obtained from the Deutsche Sammlung von Mikroorganismen und Zellkulturen (DSMZ, Braunschweig, Germany). The strains were grown anaerobically in 120 ml serum bottles with butyl stoppers and the media specified by the DSMZ. A toluene-degrading *Desulfosporosinus* sp. enrichment (Winderl *et al.* 2010, see also chapters *2.1.1* - *2.1.2*), *Desulfitobacterium aromaticivorans* UKTL, *Azoarcus* sp. T, strain BF and N47 originated from the strain collection of the Institute of Groundwater Ecology (Helmholtz Zentrum München, Germany). I cultivated the *Desulfosporosinus* sp. enrichment and *D. aromaticivorans* UKTL under anoxic conditions as described by Winderl *et al.* (2010) respectively Kunapuli *et al.* (2010). DNA from collected biomass was extracted by the bead-beating and phenol-chloroform extraction protocol described above (chapter *2.1.3*). *Azoarcus* sp. T was cultivated and extracted as described in Winderl *et al.* (2007). Strain BF was cultivated as described in Kunapuli *et al.* (2007) and the DNA was extracted as previously described (Abu Laban *et al.* 2010). Strain N47 was cultivated and extracted as described in Selesi *et al.* (2010).

Flingern tar-oil contaminated aquifer sediment was sampled in June 2009 and extracted as described in Pilloni *et al.* (2012) with a total of four replicate sediment DNA extracts per chosen depth. Three depths (6.85 m, 7.15 m, 7.25 m below surface) spanning the lower plume fringe were chosen, as those depths were previously identified as distinct degradation hot spot (Anneser *et al.* 2010, Jobelius *et al.* 2010).

The DFG SIP 1319 cooperation partners Martin Krüger and Friederike Gründger provided DNA originating from different terrestrial hydrocarbon-degrading enrichments (Paclele Mici mud volcano and Gölzau refinery contaminated aquifer). DNA was extracted with the Fast DNA Spin Kit for soil (MP biomedicals, Irvine, CA). Further information on the sampling sites and enrichments can be found in Alain *et al.* (2006) and Feisthauer *et al.* (2010).

DNA from marine samples was provided by DFG SPP 1319 collaborators Sara Kleindienst and Katrin Knittel. Gulf of Mexico hydrocarbon seep sediment was sampled from Station 161 as described in Orcutt *et al.* (2010). DNA was extracted by the collaborators according to the protocol of Zhou *et al.* (1996) and clean-up was performed with the Wizard DNA Clean-up system (Promega, Fitchburg, WI). Marine samples from the Nyegga area methane seeps were taken at pockmarks located on the edge of the Norwegian continental slope, the sampling is described in Van Gaever *et al.* (2010). The DNA was extracted with the Power Soil RNA extraction kit and DNA elution accessory kit (MO BIO Laboratories, Carlsbad, CA). Sediment samples from the Amon mud volcano (Mastalerz *et al.* 2009) in the eastern Nile deep-sea fan (station 929, dive 240) and Guaymas Basin hydrocarbon seeps (Didyk and Simoneit 1989) in the Gulf of California (dive 5473) were collected from below a microbial mat. The DNA was extracted as described by Sara Kleindienst (2012) from 2-20 cm sediment depth (Amon mud volcano) and from 0-10 cm sediment depth (Guaymas, fine grained sediments) according to the protocol described by Zhou *et al.* (1996) and based on mechanical, chemical and enzymatic lyses.

2.3.3 PCR amplification

I amplified FAE gene fragments via PCR using available and newly developed primer sets (see Table 4) in a mastercycler ep gradient (Eppendorf, Hamburg, Germany) with the following cycling conditions: 3 min initial degradation, 30 – 40 cycles of amplification (30 s at 94 °C, 30 s at 52 °C or 58 °C, 60 s at 72 °C) and 5 min at 72 °C of terminal extension. The annealing temperature was 58 °C for all primers, as initially published (Winderl *et al.* 2007), but also lowered comparatively to 52 °C (Winderl *et al.* 2010, see also chapter *2.1.4*) for the

bssA primer set. The 50 µl PCR reaction contained nuclease-free water (Promega, Fitchburg, WI), 1x PCR buffer, 1.5 mM $MgCl_2$ (Fermentas, Thermo Fisher Scientific, Waltham, MA), 10 µg BSA (Roche, Basel, CH) 0.1 mM dNTPs, 0.3 µM primer (biomers.net, Ulm, Germany) 1.25 U Taq polymerase (Fermentas, Thermo Fisher Scientific, Waltham, MA) and DNA in the range of ~ 0.2 - 20 ng. Amplicon quality was checked with gel electrophoresis in a 1.5 % agarose gel. Clean-up of the amplicons was done with the PCRextract Kit (5Prime, Hamburg, Germany) according to the manufacturer's protocol before further processing (von Netzer *et al.* 2013).

2.3.4 Dual-digest T-RFLP

The reverse primers 8546r and 8543r were labelled for T-RFLP with 5′-6-carboxyfluorescein (FAM). After amplicon purification, these were digested in two separate preparations with *Taq*I (Fermentas, Thermo Fisher Scientific, Waltham, MA) and *Hae*III (New England Biolabs, Ipswich, MA) for two hours in a thermocycler at 65 and 37 °C, respectively. The digests were purified with DyeEx columns (Qiagen, Hilden, Germany). The T-RFLP run was performed on an ABI 3730 DNA Analyzer (Applied Biosystems, Foster City, CA) as described by Lueders *et al.* (2006) and chapter *2.1.4* above. The electropherograms were processed with the programs PeakScanner 1.0 and GeneMapper 4.0 (Applied Biosystems, Foster City, CA). Further analysis was performed with the T-REX online software (Culman *et al.* 2009) using the default settings. Noise filtering was done based on peak height with standard deviation multiplier set to 1. Terminal restriction fragments (T-RFs) were defined by aligning peaks with a clustering threshold of 1 bp. T-RF affiliation in dual fingerprints was assigned via searching *Taq*I digests first for database entries with a corresponding *in silico* T-RF. If multiple entries with potentially matching T-RFs were present, the corresponding data from the *Hae*III digests was used for unambiguous T-RF affiliation to a defined FAE gene lineage (von Netzer *et al.* 2013)

2.3.5 Cloning and sequencing

I cloned selected purified, non-labelled FAE gene amplicons with the TOPO-XL Kit (Invitrogen, Carlsbad, CA) following the manufacturer's protocol. The sequencing of the Flingern samples (*bssA* primers, 52 °C: all three depths; primer sets FAE-B and FAE-N; only the 6.85 m depth; 30 clones were picked per five samples) was performed on an ABI 3730 DNA Analyzer (Applied Biosystems, Foster City, CA) with Big Dye terminator v3.1 (Applied Biosystems) chemistry according to the manufacturer's instructions. Samples of

Amon mud volcano (12 clones picked generated with *bssA* primers, 58 °C), Gulf of Mexico (*bssA* and FAE-N primers, 58 °C, for each primer 12 clones picked) and Guaymas hydrocarbon seeps (with primer sets FAE-N, FAE-KM for each primer 12 clones picked) were sequenced by GATC (Konstanz, Germany). I assembled the resulting forward and reverse sequence reads for each sample individually with the SeqMan II software module of the Lasergene 5 suite (DNASTAR, Madison, WI). Remaining vector sequences were cropped manually from the sequences by recognition of primer binding sites. The contigs were screened using BLASTN (http://blast.ncbi.nlm.nih.gov/Blast.cgi) for similarities to published FAE genes. Subsequently, I inserted the FAE related sequences into the ARB database, using the existing alignment, the original trace files and T-RFLP restriction data for correcting ambiguous codes. FAE gene tree reconstruction (Figure 4 and Figure 15) was done in ARB (Ludwig *et al.* 2004) based on an amino acids alignment via a Phylip distance matrix with the Fitch algorithm as described in Winderl *et al.* (2007). The sequences were deposited with GenBank under the accession numbers JX219271 - JX219368 (von Netzer *et al.* 2013).

2.3.6 Guaymas hydrocarbon seep sediment incubation with butane for stable isotope probing

Different SIP incubations were performed in thesis of Sara Kleindienst (2012) in order to examine alkane-degrading marine benthic microorganisms. DNA from Guaymas hydrocarbon seep sediment incubated with butane was available for FAE gene detection and monitoring of degrader community dynamics in this thesis. Anoxic sediments samples from the Guaymas Basin hydrocarbon seeps in the Gulf of California were collected as described above. (chapter *2.3.2*). Prior to the main experiment, the sediments were stored for 4 months at 4 °C. Subsets of sediment were screened in pre-incubations for highest alkane-degrading microbial activities, determined by sulphide and $^{13}CO_2$-production (for measurements methods and further results, please refer to the thesis of Sara Kleindienst (2012)). 15 ml sediment slurries with artificial anoxic seawater medium according to Widdel and Bak (1992), fully labelled ^{13}C-butane (Campro Scientific, Berlin, Germany) and unlabelled butane for controls were used for the SIP incubations. The chosen concentration of butane (2.1 mM) was higher than the natural concentrations in order to prevent substrate-diffusion limitations of the biodegradation. The sediments for SIP were incubated at 28 °C until harvesting for molecular analyses. Sediments were harvested by centrifugation (10 min, 4000 g), the pellets were stored at -80 before DNA extraction, performed as described above (chapter *2.3.2*) according to the protocol of Zhou *et al.* (1996). I conducted FAE-PCR with samples extracted after 57 d

and 113 d of incubation with primer sets *bssA*, FAE-B, FAE-N and FAE-KM as described above (chapter *2.3.3*). I performed the subsequent FAE gene T-RFLP and data analysis as previously described (chapter *2.3.4*).

2.4 Establishment of a pyrosequencing analysis pipeline for *bssA* amplicons

Flingern aquifer sediment DNA was extracted and provided by Giovanni Pilloni. Gölzau aquifer sediment enrichment DNA was extracted and provided by DFG SPP 1319 collaborators Friederike Gründger and Martin Krüger. The sediments were extracted as described in chapter *2.3.2*. I prepared the amplicons, the pyrosequencing mix was prepared by the molecular ecology group's technical assistant Katrin Hörmann and 454-pyrosequencing was provided by Helmholtz Zentrum München, research unit environmental genomics. Subsequent data analysis and pipeline establishment was done by me.

2.4.1 Pyrotag library preparation and sequencing

I chose two sample categories for a first FAE gene pyrosequencing trial: (i) DNA from the Flingern aquifer sediment samples (same as used for cloning described above in chapter *2.3.2*) retrieved from depths 6.85 m, 7.15m, 7.25m as amplified by barcoded primers 7772f (Winderl *et al.* 2007) and 8543r (von Netzer *et al.* 2013) as well as DNA extracted from depth 6.85 m as amplified by barcoded primers 7768 and 8543r (FAE-B). These samples were chosen as reference data from T-RFLP and classical sequencing was already available from previous experiments. (ii) DNA from Gölzau aquifer sediment enrichment (see chapter *2.3.2*) with benzene amplified by barcoded primers 7772f and 8543r (see also Table 4 for primer sequences without barcodes) under two different annealing conditions. This sample was chosen as a "worst case" example, since it exhibited extensive difficulties in T-RFLP and classical sequencing approaches, which are further explained below.

Table 5. Samples used for a first FAE gene pyrosequencing trial.

Sample name	Site	Primer sets
F3A	Flingern 6.85 m	7772f 8543r
F4A	Flingern 7.15 m	7772f 8543r
F5A	Flingern 7.25 m	7772f 8543r
F3B	Flingern 6.85 m	7768f 8543r
G3-58	Gölzau enrichment	7772f 8543r
G3-52	Gölzau enrichment	7772f 8543r

The primers were barcoded and extended with the respective A/B adapters, key sequences and Multiplex identifiers as recommended by 454/Roche (Basel, CH), the full-length primer and adapter sequences can be found in Table 6. The amplicons were generated with the same PCR cycling conditions as described above (chapter *2.3.3*) with 40 cycles and either 58 or 52 °C annealing temperature as indicated in Table 4. Amplicon quality was checked by electrophoresis on a 1.5 % agarose gel. Amplicons were not excised from a gel, but the complete PCR reaction was used for further analysis. Purification and pooling of amplicons was done according to the 454/Roche amplicon pyrosequencing protocol. Emulsion PCR, emulsion breaking and sequencing on a 454 GS FLX pyrosequencer were performed as recommended by the GS FLX Titanium protocols (all by Roche). The 6 amplicons were sequenced together in a pool of 30 mixed amplicons (16S rRNA) on a $1/4^{th}$ of an FLX picotitre plate.

Table 6. Primers used for creating pyrotag libraries.

Sample	Annealing (°C)	MID #	Adaptor	Primer	Sequences		
					Adaptor	MID	Primer
F3A	52	33	A	7772f	CGT ATC GCC TCC CTC GCG CCA TCA G	ATA GAG TAC T	GAC ATG ACC GAC GCS ATY CT
		33	B	8543r	CTA TGC GCC TTG CCA GCC CGC TCA G	ATA GAG TAC T	TCG TCR TTG CCC CAY TTN GG
F4A	52	34	A	7772f	CGT ATC GCC TCC CTC GCG CCA TCA G	CAC GCT ACG T	GAC ATG ACC GAC GCS ATY CT
		34	B	8543r	CTA TGC GCC TTG CCA GCC CGC TCA G	CAC GCT ACG T	TCG TCR TTG CCC CAY TTN GG
F5A	52	35	A	7772f	CGT ATC GCC TCC CTC GCG CCA TCA G	CAG TAG ACG T	GAC ATG ACC GAC GCS ATY CT
		35	B	8543r	CTA TGC GCC TTG CCA GCC CGC TCA G	CAG TAG ACG T	TCG TCR TTG CCC CAY TTN GG
F3B	58	41	A	7768f	CGT ATC GCC TCC CTC GCG CCA TCA G	TAG TGT AGA T	AAY GAT TTA ACC RAC GCC AT
		37	B	8543r	CTA TGC GCC TTG CCA GCC CGC TCA G	TAC ACA CAC T	TCG TCR TTG CCC CAY TTN GG
G3-58	58	40	A	7772f	CGT ATC GCC TCC CTC GCG CCA TCA G	TAC GCT GTC T	GAC ATG ACC GAC GCS ATY CT
		40	B	8543r	CTA TGC GCC TTG CCA GCC CGC TCA G	TAC GCT GTC T	TCG TCR TTG CCC CAY TTN GG
G3-52	52	38	A	7772f	CGT ATC GCC TCC CTC GCG CCA TCA G	TAC ACG TGA T	GAC ATG ACC GAC GCS ATY CT
		38	B	8543r	CTA TGC GCC TTG CCA GCC CGC TCA G	TAC ACG TGA T	TCG TCR TTG CCC CAY TTN GG

2.4.2 Sequence processing and analysis

The general overview of the sequence analysis pipeline developed here is shown in Figure 6. The raw data was first filtered and trimmed according to its quality based on 454 GS FLX protocols and read-length. Then I classified the reads using a self-made local BLASTN-database consisting of FAE gene sequences. In order to obtain the final classification, the BLASTN results were again filtered normalizing positive matches to read length (Identities/Query value in Figure 6).

The first quality filtering of the pyrosequencing reads was performed via the automatic amplicon pipeline of the GS Run Processor (Roche) with a modified valley filter (vfScanAllFlows false instead of TiOnly) for sequence extraction. I trimmed the initial raw reads using the 'trim'-function of Greengenes (DeSantis *et al.* 2006) with the following settings: good-quality score 20, window size 40 bp and window threshold 90%. Reads were

batched (forward and reverse) per samples based on MID identifiers with BioEdit 7.0.5.3 (Hall 1999) and reads below 250 bp after trimming were excluded (Pilloni *et al.* 2011, 2012).

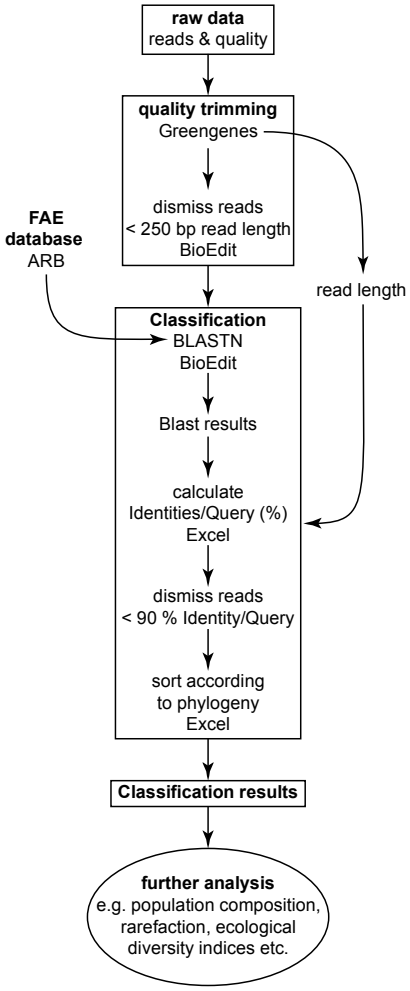

Figure 6. Schematic representation of FAE pyrosequencing analysis pipeline

For total community composition analysis, I constructed a local BLAST database using BLAST implemented within BioEdit. All sequences of the FAE gene ARB database (created by Winderl *et al.* 2007), updated to contain 497 different FAE and 5 pyruvate formate lyase sequences) were exported unaligned as a fasta file. Due to BLAST search file definition restrictions, the ARB sequences names needed to be replaced with simpler identifiers for sequence (a three digit number) and an additional value string for organisational taxonomic unit (OTU). The OTUs were based on the tree structure displayed in Figure 4 and Figure 15; the identifiers were set as shown in Table 7. The reads were thus classified with BLASTN with the parameters as displayed in Figure 7.

Table 7. OTU set-up for FAE BLAST database. OTU levels are based on the phylogeny as shown in Figure 4 and Figure 15.

Identifier	OTU levels				
|1-1-1-1	*bssA/nmsA* s. l.	*bssA* s. l.	*bssA* s. str.	*Betaproteobacteria*	
|1-1-1-2				*Deltaproteobacteria* 1 (*Desulfobacterium* spp.)	
|1-1-1-3				*Deltaproteobacteria* 2 (strain TRM1, PRTOL)	
|1-1-1-4-1				*Deltaproteobacteria* 3	F1-cluster
|1-1-1-4				(F1, *Geobacter* spp., *Desulfobacula* spp.)	
|1-1-1-5-2	*bssA/nmsA* s.l.	*bssA* s. l.	*bssA* s. str.	*Clostridia* 1	F2-cluster
|1-1-1-5				(*Desulfosporosinus* spp.)	
|1-1-2	*bssA/nmsA* s.l.	*bssA* s. l.	*Clostridia* 2 (strain BF)		
|1-1-3-6	*bssA/nmsA s.l.*	*bssA* s. l.	Testfeld Süd-cluster	T1-cluster	
|1-1-3-7				T2-cluster	
|1-2-4	*bssA/nmsA* s.l.	*nmsA/bssA* s. l.	*nmsA* s. str.		
|1-2					
|1-2-5			*bssA* OX39		
|2-3	*assA* (*masD*)	*assA* (*masD*) s. str.			
|2					
|2-4		hexadecane degrading methanogenic consortium			
|2-5		Brand *assA* homologues			
|3	Outgroup PFL				

Figure 7. Screenshot of BLASTN parameters used for classification via BioEdit. -G cost to open a gap, -E cost to extend a gap, -W word size, - q penalty for a nucleotide mismatch, r reward for a nucleotide match

The tabular classification output was imported into Excel, where possible duplicates would be removed and read length data from Greengenes trimming added (for a detailed description of the workflow within Excel, please refer to this thesis's appendix). The results were sorted according to score and the Identity/Query (%) value was calculated with the BLASTN sequence identity value (number of bp found to be sequentially identical) and the Greengenes read length data. The Identity/Query value was essential to sort out false-positive FAE gene identifications due to short sequence motifs being found within the reads and returning high BLASTN score values. The trend of the Identity/Query values suggested a cut-off below 90 %, as illustrated later by Figure 17. The remaining results were then sorted according to the

OTU name string of the database and the phylogenetic read composition thus analysed and compared with existing clone and T-RFLP data (presented also in chapter *3.2.3*). Rarefaction was calculated in the statistical program environment R (version 2.15.1, Ihaka and Gentleman 1996) with the package "vegan" (version 2.0.3, Oksanen *et al.* 2012) via the function "rarecurve". For calculation of rarefaction, the highest OTUs levels F1, F2, T1 and T2 were binned in their respective directly lower ranking OTUs.

3. Results

3.1 Affiliation of novel, unidentified FAE gene clusters via stable isotope probing

Unaffiliated *bssA* lineages were previously found at the tar-oil contaminated aquifers at Testfeld Süd and Flingern (Winderl *et al.* 2007), the deeply-branching Testfeld Süd (T) sequence cluster respective the F1- and F2-cluster (see Figure 4). Sediments from both sites were analysed via stable isotope probing (SIP) in order to identify key degraders and to link the labelled catabolic and 16S rRNA genes (Winderl *et al.* 2010, Pilloni *et al.* 2011). My contribution to the Testfeld Süd SIP study was the analysis of unfractionated DNA over different time points of the experiment via 16S rRNA gene as well as *bssA*-T-RFLP. The aim was to observe if the identified clostridial toluene degrading key players were detectable from the beginning of the incubation on both total population and functional gene level. For the Flingern aquifer study, I performed *bssA* gene T-RFLP with fractionated DNA from sulphate and ferric iron reducing SIP incubations in order to unravel the affiliation of F1-cluster *bssA* sequences to either *Desulfobulbaceae* or *Geobacteraceae*.

3.1.1 Tracing the affiliation of unidentified FAE gene clusters to *Clostridia* with stable isotope probing in Testfeld Süd sediment.

As shown in Winderl *et al.* (2010), clones related to *Desulfosporosinus* spp. (*Peptococcaceae*) were found to dominate in labelled DNA at the end (100 days) of the SIP experiment on both 16S rRNA gene and *bssA* level. The Testfeld Süd SIP *bssA* sequences clustered together with the previously unidentified F2-cluster *bssA* sequences, making it likely that this *bssA* sequence cluster is of clostridial origin. The phylogenetic trees in Figure 8 are a synopsis of these results, which are displayed in more detail in the aforementioned publication (Winderl *et al.* 2010). It is noteworthy that the F2-cluster *bssA* in Figure 8 is rather remote from the only other at the time known clostridial hydrocarbon degrader *Desulfotomaculum* sp. OX39, which is discussed at a later point (chapter *4.4*). In the following part, I focus in further detail on the results elaborated this thesis and their original contribution to the paper of Winderl *et al.* (2010).

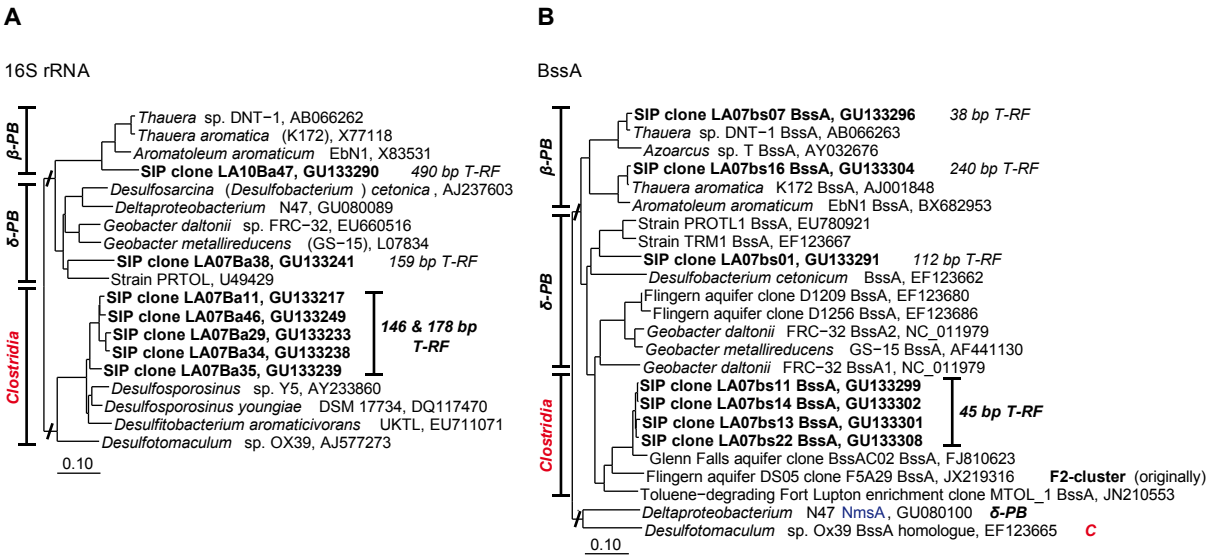

Figure 8. Phylogenetic systematic (both neighbour joining) of clones (bold) from SIP incubations at Testfeld Süd on community (16S rRNA gene, part A) and functional level (BssA, part B) published by Winderl *et al.* (2010). *β-PB Betaproteobacteria*; *δ-PB Deltaproteobacteria*; *C Clostridia*. Representative T-RFs measured for the respective gene lineages are indicated.

In *bssA* gene T-RFLP which I performed (results shown for the inoculum and 86 days after incubation in Figure 9) with PCR conditions set with an annealing temperature at 58 °C, two peaks were found to dominate: 240 bp and 80 bp. The peak at 240 bp was attributed to SIP microcosm "heavy" DNA *bssA* clone LA07bs16 (*in silico* T-RF 242 bp) and clusters next to *Thauera aromatica* T1 *bssA* (see also Figure 8B). It was not possible to assign the 80 bp peak to the *bssA* clones from this experiment. Because *T. aromatica* T1 was not identified as the key degrader on 16S rRNA gene level, this illustrates the high selectivity of the *bssA* 7772f-8546r primer pair for *bssA* genes of *Betaproteobacteria*.

With lowered PCR selectivity achieved by using 52 °C as annealing temperature, it was finally possible to observe two additional major, identifiable peaks (see Figure 9): 364 bp, assigned to Testfeld Süd cluster *bssA* sequences (*in silico* T-RF 365 bp) and the peak at 45 bp below the normal quality cut-off (50 bp). The cut-off is usually necessary as it is often not possible to distinguish short fragments from primer-dimers artefacts. But here, it was possible to assign the T-RF with 45 bp to F2-cluster related *bssA* clones (*in silico* T-RF 50 bp, for phylogeny, see Figure 8B). The peak formerly observed at 80 bp with the annealing temperature set at 58 °C might be shifted under a lowered annealing temperature to 89 bp. It was not possible to attribute identities to other minor peaks appearing at 53 bp and 112 bp.

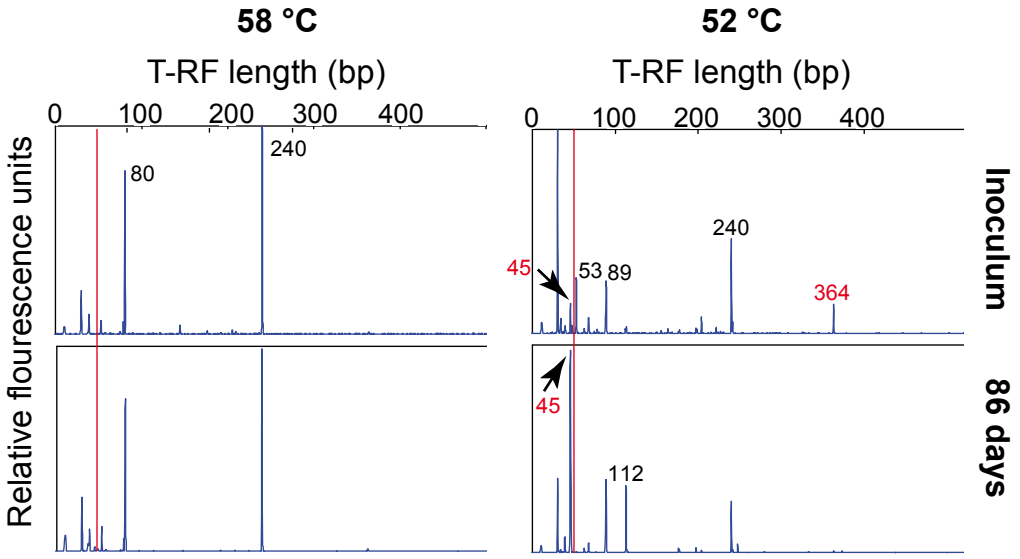

Figure 9. Comparison of *bssA* gene T-RFLP pattern with different annealing temperatures, here shown for the inoculum and 86 days after incubation. Peaks are labelled with the corresponding T-RF size in bp. The T-RFs of *Desulfosporosinus* sp. *bssA* (45 bp) and T-cluster *bssA* homologues (364 bp) are labelled with red letters; the red line indicates the normal cut-off at 50 bp.

The peak (364 bp) representing the Testfeld Süd *bssA* sequence cluster clearly declined over time in ^{12}C and ^{13}C toluene treatments, as presented in Figure 10. It was not possible to tentatively correlate a clear decrease for any population on 16S rRNA gene level as shown in Figure 11. Therefore, the phylogenetic affiliation of this *bssA* homologous cluster (Figure 4) remains an enigma.

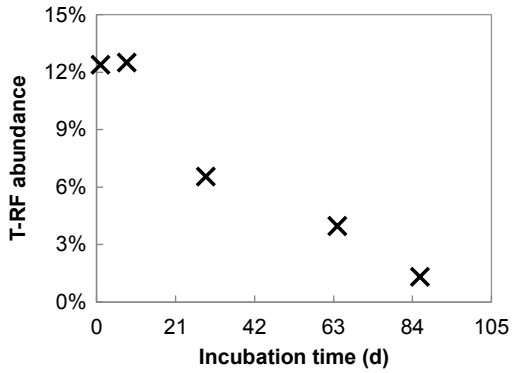

Figure 10. Decline of 364 bp *bssA* gene T-RF identified as Testfeld Süd cluster sequences.

The sizes of T-RFs calculated *in silico* (50 bp) and measured (45 bp) representing Testfeld Süd SIP *bssA* sequences related to the F2-cluster are below to the normal T-RFLP analysis cut-off (<50 bp) as illustrated in Figure 9. Therefore, the quantitative development of the putatively clostridial SIP *bssA* clones was only traceable in raw electropherogram data. It was possible to observe the increase of the 45 bp peak over time (Figure 9), which was assigned to *Desulfosporosinus* spp. *bssA*. The size shift of 5 bp between observed and calculated T-RF may be due to the Map Marker sizing curve no longer being linear in this size region. Also, it was possible to observe on 16S rRNA gene level a clear increase of *Desulfosporosinus* spp. T-RFs (146 bp, 178 bp) during incubation (Figure 11). Thus, a clear identification of the F2-lineage was not only possible on clone level but also by T-RFLP on both functional (*bssA*) and 16S rRNA gene levels. These results demonstrate that the putative *Desulfosporosinus* spp. degraders were already naturally present in the aquifer sediment from the beginning of the SIP incubation.

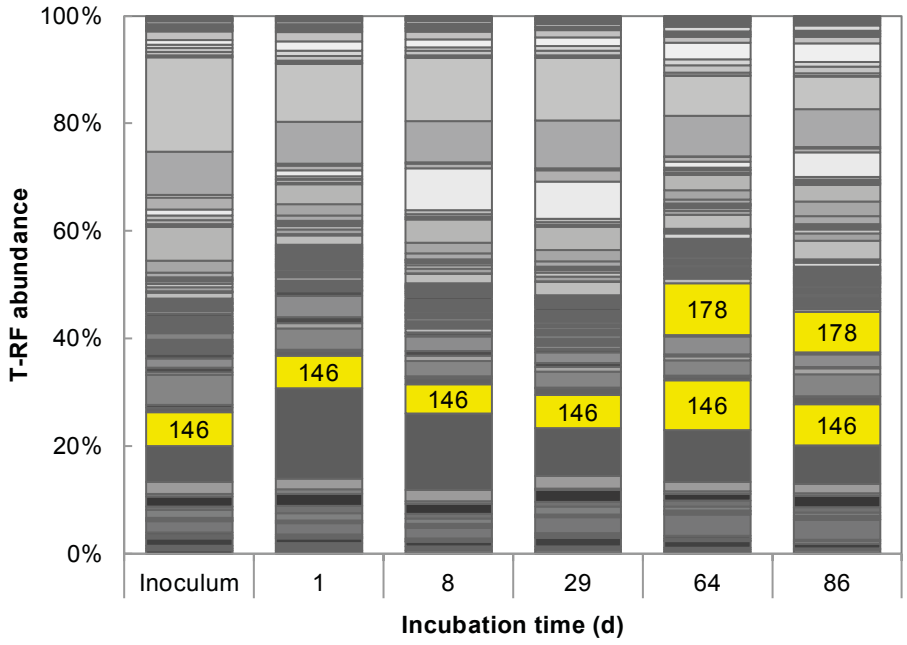

Figure 11. Increasing abundance of T-RF typical for *Desulfosporosinus* spp. (146 bp, 178 bp) over time in total, ^{13}C-labelled DNA as traced on 16S rRNA gene level (fragment length in the diagram shown as bp). The full dataset is displayed in the supplementary, Table 14.

3.1.2 Stable isotope probing for identifying key anaerobic, sulfidogenic toluene degraders in Flingern aquifer as members of *Desulfobulbaceae*

The previous studies on Flingern sediment were not able to clarify if the key degraders found on *bssA* level – the F1-cluster *bssA* – belong to either *Geobacteraceae* or *Desulfobulbaceae* (Winderl *et al.* 2007, 2008). On 16S rRNA gene level, Pilloni *et al.* (2011) showed via amplicon pyrosequencing of total ^{13}C-labelled DNA and additional T-RFLP of fractionated DNA that *Desulfobulbaceae* are responsible for toluene degradation under sulphate reducing conditions in Flingern sediment. In the treatments with iron, *Geobacteraceae* were clearly not involved in toluene degradation, but *Betaproteobacteria* related to *Georgfuchsia* spp. and *Clostridia* related to the genus *Thermincola*.

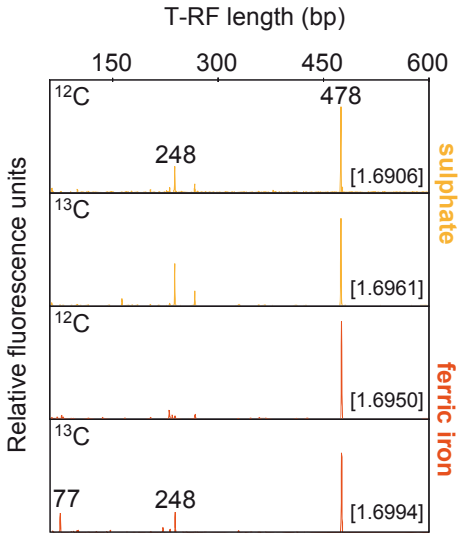

Figure 12. Comparison of *bssA* T-RFLP results from heavy DNA fractions (buoyant density in square brackets) from SIP incubations of Flingern sediment with ^{12}C or ^{13}C-tolulene as electron donor and either sulphate (two upper electropherograms) or ferric iron (lower two electropherograms) as electron acceptor. Redrawn according to Pilloni *et al.* (2011). 478 bp: F1-cluster *bssA*; 278 bp and 77 bp: unassigned.

It was not possible to find any peaks associated to the aforementioned *Georgfuchsia* spp. and *Thermincola* spp. populations found on 16S rRNA gene level with the *bssA*-T-RFLP analysis I performed with samples as indicated in Table 3 and Figure 12. As presented in Figure 12, a *bssA* gene T-RF of 478 bp was found to dominate the heavy fractions of the density-resolved DNA from both sulphate and iron reducing SIP incubations. This peak was assigned to hither to unidentified F1-cluster *bssA* sequences in the FAE gene database. Since the DNA from sulphate-reducing SIP incubations was on 16S rRNA gene level clearly dominated by *Desulfobulbaceae*, it was clear that that both markers must be linked. Moreover, a peak at 248

bp was found in fractions from sulphate treatments and a peak at 77 bp in fractions from iron treatments. Both peaks were not assignable to *in silico* T-RF of known *bssA* sequences (Pilloni *et al.* 2011). Still, these results confirm now the affiliation of F2-cluster *bssA* to members of *Desulfobulbaceae*.

3.2 Enhanced gene detection assay for fumarate-adding enzymes uncover anaerobic hydrocarbon degraders in terrestrial and marine systems

The results I presented in the previous chapters point already towards the general problem that the existing detection systems for FAE genes are not yet optimized. This is especially the case for deeply-branching *bssA* homologues of putative clostridial origin. The T-RFLP analysis based on restriction with solely one enzyme (*Taq*I) failed to produce directly interpretable data as the T-RF of clostridial F2-cluster *bssA* sequences was in the same size range as putative primer artefacts. Additionally, there are other FAE genes as *nmsA* and *assA*/*masD,* which are not detectable with the 7772f-8546r *bssA* primer initially published by Winderl *et al.* (2007). Here, I present the development and verification of an enhanced assay capable to detect such FAE genes and also a dual-digest T-RFLP for a more rapid and sequencing-independent FAE gene screening. This was tested in samples from different environments, some of them kindly provided by cooperation partners within the DFG SPP 1319. My results reveal the hitherto underestimated catabolic potential of different FAE lineages at the investigated sites.

3.2.1 Establishment of optimised primer sets

No universally conserved protein motif that would allow the development of a single, universal primer pair was found across the different FAE gene lineages (shown in Figure 4 and Figure 15). Therefore, I designed *in silico* several primer sets called FAE-B, FAE-N and FAE-KM (as shown in Table 4) to optimally target a common FAE gene region for robust phylogenetic analysis. All newly developed forward primers bind to sequence motifs not targeted before by available assays prior to this study. The new FAE-B forward primer was predicted to target betaproteobacterial *bssA*, the recently discovered clostridial *bssA* homologues (F2-cluster, strain BF), as well as *nmsA* homologues (strain N47, NaphS-strains). The two new FAE-N forward primers specifically target a sequence motif only found in *nmsA* of strain N47 and NaphS-strains, 400 bp upstream of the other forward primer sites. This primer set was aimed to distinguish *nmsA sensu stricto* (s. str.) amplicons from the homologous *bssA/nmsA sensu lato* (s. l.) amplicons retrieved by set FAE-B (see lineages

either Figure 4 or Figure 15). The new FAE-KM set contained two different forward primer combinations: forward primer combination FAE-K to target sequences related to *assA* genes of *Desulfatibacillum alkenivorans* and forward primer FAE-M to target relatives of strain HxN1 and OcN1 *masD*. FAE-K forward primers consisted of two different primers with two varying positions, instead of one with two ambiguity codes, to avoid four possible homopolymers and subsequent loss of specificity. The primer sets FAE-B, FAE-N and FAE-KM shared the same modified reverse primer 8543r, as the targeted sequence motif is conserved in all known FAE genes. This is documented for other, similar primers targeting this region (Washer and Edwards 2007, Winderl *et al.* 2007, Callaghan *et al.* 2010) as presented in Table 1 (von Netzer *et al.* 2013).

3.2.2 Primer performance in pure cultures and environmental samples

The FAE-gene lineage-specific performance of the different primer sets predicted *in silico* was verified, wherever possible, using pure culture DNA as shown in Table 8. The *bssA* primer set published in Winderl *et al.* (2007) did not amplify FAE genes of *Desulfitobacterium aromaticivorans* UKTL, clostridial strain BF, strain N47, *Desulfatibacillum alkenivorans* PF2803 or *D. aliphaticivorans* CV2803. Apart from the *bssA* of *Azoarcus* sp. T, the new primer set FAE-B did indeed amplify FAE genes of these selected clostridial strains and also *nmsA* of strain N47. As intended, set FAE-N only amplified *nmsA* of strain N47. Primer set FAE-KM amplified only the *assA* of *D. alkenivorans* PF2803 and *D. aliphaticivorans* CV2803. A combination of all possible forward primers with one reverse primer in a single PCR reaction failed to produce amplicons (von Netzer *et al.* 2013).

Subsequently, the performance of the primer sets was tested with samples originating from different hydrocarbon-degrading enrichments or impacted environmental systems. The samples in Table 9 were chosen as they were exposed to and actively degrading different aliphatic or aromatic hydrocarbons: single compounds (enrichments), confined mixtures (e.g. natural gas at marine seeps) or complex mixtures of hydrocarbons (oil, asphalt). According to the known hydrocarbon substrates of those samples and the potential FAEs supposedly involved in their degradation by previous studies, FAE lineages expected to be recovered by the new primer sets for FAE were predicted: (i) *bssA/nmsA* s. l. as amplified by primers *bssA* and FAE-B, (ii) *nmsA* s. str. as amplified by primer set FAE-N and (iii) *assA* as amplified by primer set FAE-KM (see Figure 4 or Figure 15). The results are presented in Table 9 (von Netzer *et al.* 2013).

Table 8. Experimentally tested performance of different FAE gene-targeted primers in selected strains and enrichments degrading hydrocarbons via fumarate addition (von Netzer *et al.* 2013).

FAE lineage		Culture	Substrate	Reference	Amplicon production with primer set[1]			
					bssA primers	FAE-B	FAE-N	FAE-KM
bssA s. l.	*Proteobacteria*	*Azoarcus* sp. T	Toluene	Achong *et al.* (2001)	+	+	-	-
		Geobacter metallireducens GS-15	Toluene	Lovley *et al.* (1993)	+	-	-	-
		Desulfotignum toluenicum H3	Toluene	Ommedal and Torsvik (2007)	+	-	-	-
	Clostridia	*Desulfosporosinus* sp. enrichment	Toluene	Winderl *et al.* (2010)	+	+	-	-
		Desulfitobacterium aromaticivorans UKTL	Toluene	Kunapuli *et al.* (2010)	-	+	-	-
		Strain BF	Benzene	Abu Laban *et al.* (2010)	-	+	-	-
nmsA		Strain N47	2-Methyl-naphthalene	Selesi *et al.* (2010)	-	+	+	-
assA	*Proteobacteria*	*Desulfatibacillum alkenivorans* PF2803	Alkanes	Cravo-Laureau *et al.* (2004b)	-	-	-	+
		Desulfatibacillum aliphaticivorans CV2803	Alkanes	Cravo-Laureau *et al.* (2004a)	-	-	-	+

[1] (+) A clear amplicon with the correct size was observed, (-) no amplicon was obtained.

For samples exposed to confined mixture of hydrocarbons, like those from the mud volcanoes and the Nyegga methane seeps, amplicons were obtained only for single FAE gene lineages. In contrast, several major FAE gene lineages were consistently obtained for samples exposed to a broader mixture of hydrocarbons. For example, all expected FAE lineages were detected for the respective substrates used for the Gölzau aquifer sediment enrichments, but also additional FAE lineages were present. Multiple FAE gene lineages were also detected at the Flingern aquifer and the Gulf of Mexico hydrocarbon seep. In the Guaymas Basin hydrocarbon seep sediments, only *nmsA* s. str. and *assA* could be detected (von Netzer *et al.* 2013).

Table 9. Verification of FAE assay performance in DNA from selected hydrocarbon-impacted enrichments and environmental samples (von Netzer *et al.* 2013)

Sample	Region	Reference	Environment, sample type	Substrate	Expected FAE lineages	Results[1]			
						bssA/nmsA s. l.		*nmsA*s. str.	*assA*
						bssA primer	FAE-B	FAE-N	FAE-KM
Paclele Mici mud volcano	Romania	(Alain *et al.* 2006)	terrestial, enrichment	Hexadecane	*assA*	-	-	-	+
Amon mud volcano	Mediterranean Sea, Nile deep-sea fan	(Kleindienst 2012, Kleindienst *et al.* 2012)	marine, sediment	Methane, ethane, propane, butane	*assA*-like?	+[3]	-	-	-
Nyegga methane seeps	Norwegian continental ridge	(Van Gaever *et al.* 2010)	marine, sediment	Methane, natural gas	*assA*-like?	-	-	-	+
Chapopote asphalt volcano station GeoB106-17-6	Gulf of Mexico	(Bohrmann *et al.* 2008, Kleindienst *et al.* 2012)	marine, sediment	Asphalt, alkanes, aromatics, methane	FAE	+	+	-	+
Hydrocarbon seep station 161	Gulf of Mexico	(Orcutt *et al.* 2010, Kleindienst *et al.* 2012)	marine, sediment	Oil, asphalt, alkanes, aromatics	FAE	+[3]	+[3]	+[3]	+
Guaymas Basin hydrocarbon seep	Gulf of California	(Kleindienst 2012, Kleindienst *et al.* 2012)	marine, sediment	Methane, alkanes, aromatics	FAE, esp. *assA*	-	-	+[3]	+[3]
Gölzau aquifer former refinery	Germany	(Feisthauer *et al.* 2010)	terrestial, enrichment	Toluene	*bssA*	+	-	+	-
				Butane	*assA*	+	+	+	+
				Methylnaphthalene	*nmsA*	+	-	+	-
Flingern aquifer former gas work site	Germany	(Pilloni *et al.* 2012)	terrestial, sediment	BTEX, PAH	*bssA/ nmsA* s. l.	+[3]	+[3]	+[3]	+[2]

[1] (+) a clear amplicon with the expected size was observed, (-) no amplicon.

[2] 40 PCR cycles were necessary to obtain amplicon.

[3] amplicons were actually cloned and sequenced as displayed in Figure 1.

In order to substantiate these qualitative amplicon-based results, selected samples were chosen for sequencing in small clone libraries: Amon mud volcano (6 FAE gene sequences out of 25 clones with *bssA* primers, 58 °C), Gulf of Mexico hydrocarbon seep station 161 (9 FAE gene sequences out of 25 clones with *bssA* primers, 58 °C, 1 FAE gene sequence out of 25 clones with FAE-B primers and 6 FAE gene sequences out of 25 clones with FAE-N primers) and Guaymas Basin hydrocarbon seep (8 FAE gene sequences out of 25 clones with FAE-N primers and 2 FAE gene sequences out of 25 clones with FAE-KM primers). The clones (Figure 12) were clearly affiliated to the expected FAE lineages, the rest of the libraries were sequences that resulted from unspecific amplification. Amon mud volcano clones were mostly related to *Desulfobacula toluolica bssA*, with one assigned to more deeply branching *nmsA*-homologues. Guaymas Basin hydrocarbon seep clones generated with set FAE-KM were related to betaproteobacterial *masD* and *Desulfoglaeba alkanexedens assA* sequences. From all three sites, clones related to *nmsA* s. str. of NaphS-strains (all from set FAE-N) or *bssA*-homologues were retrieved with primer sets FAE-N, FAE-B and *bssA* at 52 °C (von Netzer *et al.* 2013). In general, these results demonstrate that also FAE genes other than *bssA* are widely distributed in hydrocarbon-impacted environments.

3.2.3 Hidden FAE lineages revealed by dual-digest T-RFLP analysis in the tar-oil contaminated Flingern aquifer in spatial resolution

A more extensive assay performance test and the introduction of a sequencing-independent T-RFLP-based FAE gene lineage identification was conducted with sediment from the well-investigated tar-oil contaminated aquifer in Flingern (Winderl *et al.* 2007, 2008; see also chapter *1.5*). This aquifer is known to host two specific *bssA*-populations at the lower plume fringe, the hot-spot of sulfidogenic BTEX degradation (Anneser *et al.* 2010, Jobelius *et al.* 2010): dominating *Desulfobulbaceae* F1-cluster *bssA* (Pilloni *et al.* 2011, Winderl *et al.* 2007) and the less abundant clostridial F2-cluster *bssA* (Winderl *et al.* 2007, 2010). As the sediment studied here is from a later sampling period (2009), I was not entirely convinced to expect the same *bssA*-populations. I already introduced the use of a *bssA*-based T-RFLP fingerprinting system allowing sequencing independent lineage identification at reasonable confidence from samples of which sequences are known from prior clone library analysis (Winderl *et al.* 2010, see also chapter *3.1*). However, several different *bssA* lineages were observed *in silico* to display identical T-RFs with *Taq*I restriction (see Figure 15; e.g. 50 bp T-RF for certain *Clostridia*, *Beta-* and *Deltaproteobacteria*). In order to obtain better diagnostic confidence, I

now introduce a second, parallel digest with *Hae*III in addition to the primary *Taq*I-digests to unambiguously identify T-RF (von Netzer *et al.* 2013).

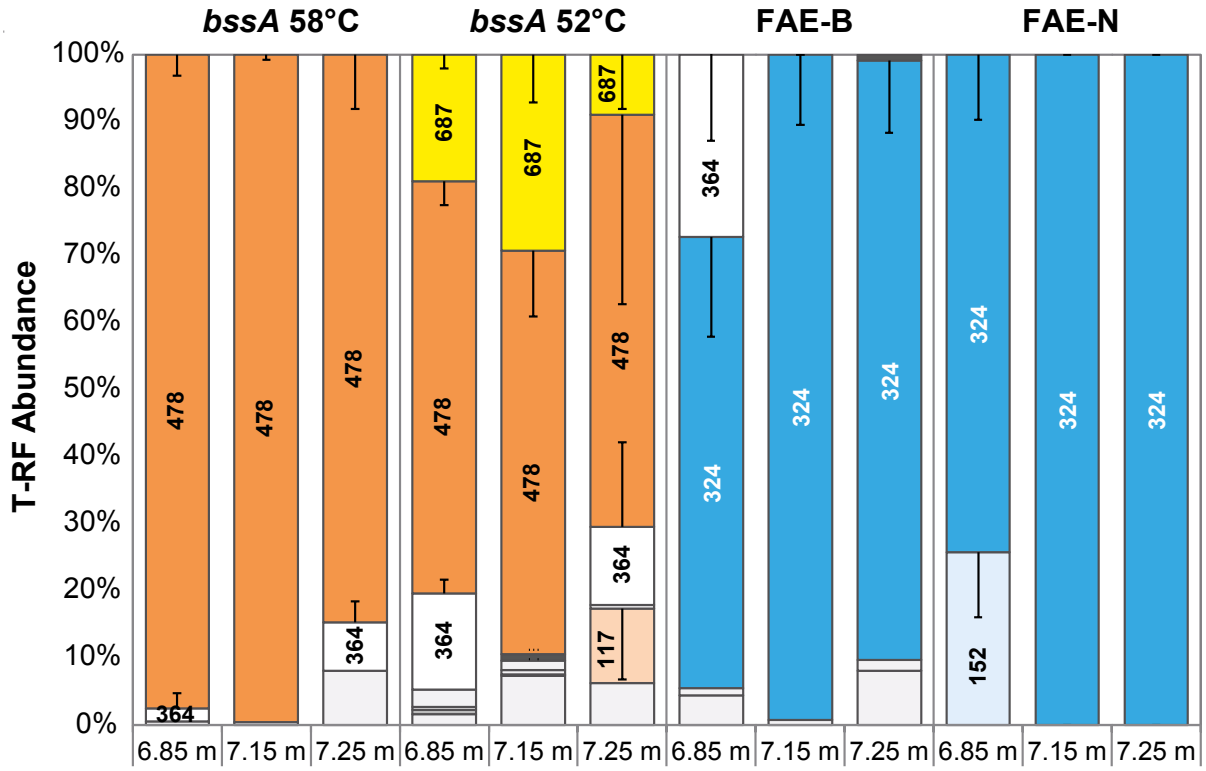

Figure 13. Comparative FAE gene T-RF retrieval with different assays in distinct depths of the tar-oil contaminated Flingern aquifer. T-RF abundance is averaged over results from four independent DNA extractions; error bars represent the standard deviation. The identity of represented FAE lineages is given in Table 10 (von Netzer *et al.* 2013).

The application of the different primer sets and *Taq*I digested T-RFLP for the three chosen depths from the Flingern aquifer revealed the presence of a number of tentative *bssA* T-RFs as shown in Figure 13. The identity of those *bssA* lineages was assigned via T-RF sizes predicted for sequences in the FAE ARB database and further confirmed by parallel *Hae*III digest as well as, for proof-of-principle, by cloning and sequencing (see Table 10 and Figure 15, *Hae*III digest raw data not shown). With the standard *bssA* primers and settings, the *Desulfobulbaceae* F1-cluster *bssA* T-RF (478 bp, *Taq*I) clearly dominated over the three chosen depths, as it was expected from previous work (Winderl *et al.* 2007, 2010). But, already at a lowered annealing temperature (for reduced PCR selectivity, as previously established by Winderl *et al.* (2010) also mentioned in chapter *3.1.1*), a T-RF representing the

clostridial F2-cluster *bssA* (687 bp, *Taq*I) and the unidentified Testfeld Süd T1-cluster (Winderl *et al.* 2007) *bssA* homologues (364 bp, *Taq*I) became detectable. By applying the new FAE-B primer set, strain N47 *nmsA*-like sequences (324 bp *Taq*I) became detectable in all three depths as dominating populations, while F1- and F2-cluster disappeared. With the FAE-N primers, it was furthermore possible to detect a T-RF (152 bp *Taq*I) at 6.85 m depth tentatively assigned to the *nmsA*-associated *bssA* homologue previously detected for *Desulfotomaculum* sp. strain OX39 (von Netzer *et al.* 2013).

Table 10. Sequencing-independent identification of T-RF lineages from Flingern sediment via dual-digest T-RFLP[1] (see also Figure 10; von Netzer *et al.* 2013).

Observed fragment size (bp)		*In silico* fragment size (bp)		Candidate lineage	No. of FAE clones	Primer sets
***Taq*I**	***Hae*III**	***Taq*I**	***Hae*III**			
364	398	362	397	unidentified *bssA* homologues unaffiliated (Testfeld Süd T1)	8	*bssA* 52°C, *bssA* 58°C, FAE-B
478	51	474	52	*Deltaproteobacteria bssA Desulfobulbaceae* (F1-cluster)	14	*bssA* 52°C, *bssA* 58°C
678	304	670	300	Clostridial *bssA Desulfosporosinus-related* (F2-cluster)	22	*bssA* 52°C
324	505	329	506	*Deltaproteobacteria nmsA Desulfobacteraceae* (N47-related)	18	FAE-B, FAE-N
152	130	155	132	Clostridial *bssA* homologues *Desulfotomaculum*-related	1	FAE-N
-	132	35	132	*Betaproteobacteria bssA Thauera* sp. DNT-1	2	FAE-B
117	76	113	75	*Deltaproteobacteria bssA Desulfobacterium cetonicum bssA*	1	*bssA* 52°C

[1] Observed T-RFs from the *Taq*I digest are also shown in Figure 13, data for *Hae*III digest is not shown.

[2] *In silico* T-RFs are given for the labelled 8543r primer, hence a difference of 3 bp must be considered to T-RFs delivered by the 'classical' 8546r-labeled primer pair as shown in Figure 13.

Based on previous FAE gene sequence data from the site (Winderl *et al.* 2007), one minor T-RF (117/76 bp; *Taq*I/*Hae*III), found with FAE-B primers at 7.25 m, was not assignable *in silico* and was thus identified by cloning and sequencing to represent a new relative of *Desulfobacterium cetonicum bssA* (67 % amino acid sequence identity, clone F5A05, Figure 15). Another FAE-population became detectable at 6.85 m with primer set FAE-B (35 bp T-RF), which was revealed by *Hae*III-digest and cloning to be related to betaproteobacterial *bssA* (see Table 4, clone F3B22 in Figure 15.). This population was not at all detectable in standard *Taq*I digested fingerprints, as such short fragments (<50 bp) are usually not considered since they cannot be effectively differentiated from primer-dimers and other 'short' artefacts usually observed in T-RFLP raw data. All expected *Hae*III T-RFs

(summarised in Table 10), corresponding to the primary *Taq*I T-RFs in Figure 15, and as predicted via clones from the same samples (Figure 15), were readily detectable (*Hae*III data not shown). However, recovery of comparative T-RF abundance in both assays was more qualitative, not rigidly semi-quantitative, as illustrated exemplary for one Flingern sample in Figure 14 (von Netzer *et al.* 2013).

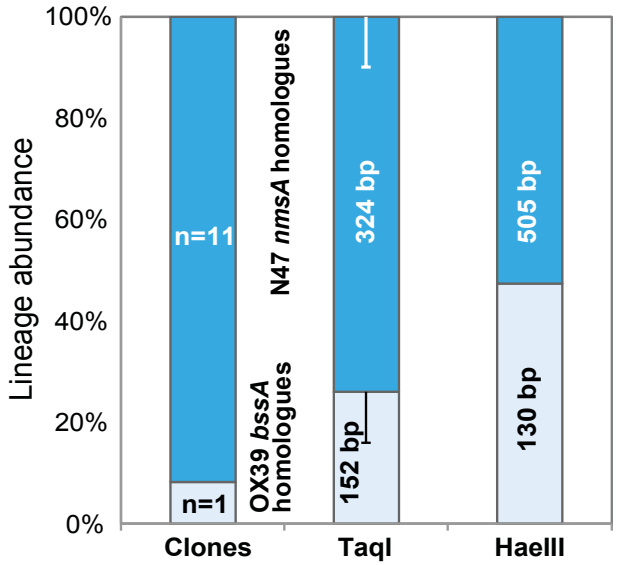

Figure 14. Exemplary comparison of semi-quantitative FAE gene-lineage recovery in Flingern sediments (6.85 m) as retrieved via cloning and sequencing (a total of 12 clones) and dual-digest T-RFLP analyses with the FAE-N primer set (von Netzer *et al.* 2013).

The results show spatial distinctions in degrader community composition which would have been missed with just one set of primers. In all analysed sediment depths, *bssA* of *Desulfobulbaceae* (F1-cluster) is dominating. However, there is also a considerable amount of clostridial F2-cluster *bssA* present in the lower plume fringe (7.15 m). *nmsA* sequences related to N47 are dominant throughout the three different depths. In the proximity to the plume core (6.85 m), there are also *bssA* sequences related to the clostridial *Desulfotomaculum* sp. OX39 and T-cluster *bssA*. Sequences related to T-cluster *bssA* are also present below the plume fringe (7.25 m), but not directly in the plume fringe (7.15 m).

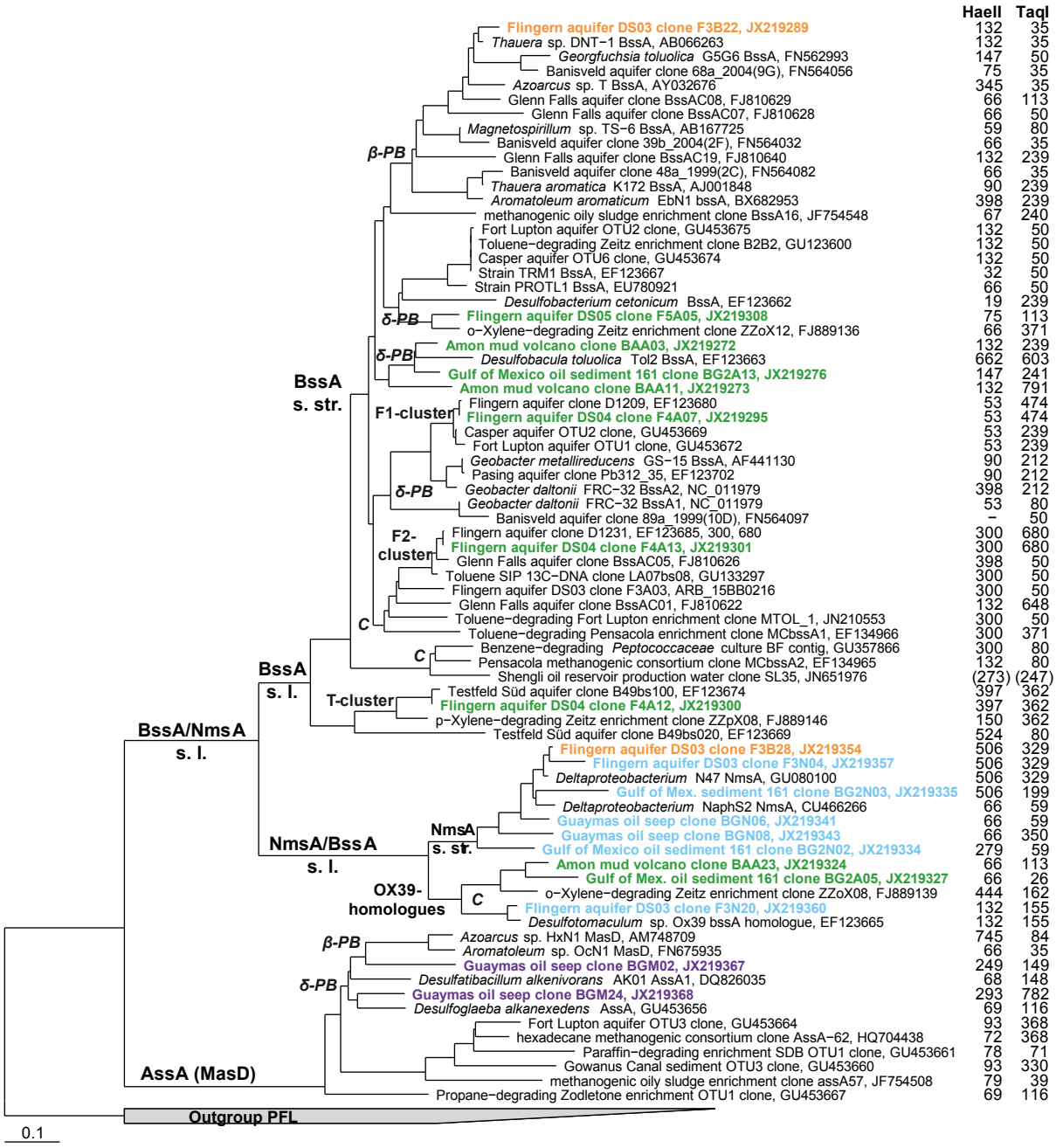

Figure 15. Overview of current FAE gene phylogeny. The numbers behind the sequence entries give the T-RFs predicted *in silico* for amplicons generated with the 8543r primer and a digest with *Hae*III and *Taq*I, respectively. The T-RF sizes in brackets are tentatively extrapolated, as some sequence entries ends before the primer binding site. Abbreviations of FAE lineages (as identified via enrichments or pure cultures): C *Clostridia*, β-P *Betaproteobacteria*, δ-P *Deltaproteobacteria*. Sequences retrieved in this study are coloured according to primer sets: set FAE-B orange, bssA 52°C green, set FAE-N blue, set FAE-KM purple (von Netzer *et al.* 2013).

3.2.4 Tracing hydrocarbon-degrader population dynamics and stability in Guaymas SIP incubations

My next aim was to demonstrate that my new FAE primer sets and T-RFLP approach can actually be applied to resolve temporal changes of degrader community structure and diversity. I analysed DNA from different time points of a SIP experiment done in the context of the thesis of Sara Kleindienst (2012). The SIP experiment was performed with sediment from the marine hydrocarbon seep in Guaymas Basin in the Gulf of California (Didyk and Simoneit 1989). At the Guaymas Basin hydrocarbon seep, there are mainly methane, alkanes and aromatic hydrocarbons present (Kleindienst *et al.* 2012). The SIP experiment analysed here was amended with butane (Kleindienst 2012) which was presumed to be a major driver of intrinsic degrader community dynamics. DNA of three time points (inoculum = 0 d, 57 d and 113 d) was available for FAE-analysis.

As presented in Table 11, I was able to detect FAE genes in inoculum DNA with the primer sets FAE-N and FAE-KM, but neither with the primer set FAE-B nor classical *bssA* primers. Therefore, bacteria containing *assA* and *nmsA* genes seemed to be present, but not degraders with genes falling into the *bssA/nmsA* s. l. sequence branch. After 57 days of incubation with butane, the community composition changed dramatically: *nmsA* genes were no longer detectable by the FAE-N primer set. However, I could still obtain amplicons with primer sets *bssA*, FAE-B and also still with FAE-KM. The same amplification pattern was again obtained after 113 days of incubation, as illustrated in Table 11.

Table 11. Development of different major FAE gene populations during the SIP incubation with butane on amplicon detection level.

	***bssA/nmsA* s. l.**		***nmsA/bssA s. l.***	***assA***
	***bssA* primers**	**FAE-B**	**FAE-N**	**FAE-KM**
Inoculum	-	-	+	+
57 d	+	+	-	+
113 d	+	+	-	+

The community composition as resolved by *Taq*I T-RFLP with primer sets FAE-N and FAE-KM is shown in Figure 16. For *nmsA* genes, a population characterized by a T-RF at 348 bp dominated in the inoculum (Figure 16B). This T-RF could be assigned to Guaymas oil seep clone BGN08 which in turn is related to *Deltaproteobacterium* strain NaphS2 (Figure 15). For *assA* genes, two populations were found to dominate (Figure 16A): a T-RF with 78 bp and 147 bp, among other minor populations with T-RF abundance < 10 %. The peak with 147 bp

was assigned to *assA*-related Guaymas oil seep clone BGM02 (*in silico* T-RF 149 bp, Figure 15); the peak with 78 bp was not identifiable. A minor population with a T-RF of 788 bp (2.3 % abundance) was identified as *assA*-related Guaymas oil seep clone BGM24 (*in silico* T-RF 782 bp, Figure 15); all other minor populations remained unassigned. During the course of the incubation with butane, an unidentified population with 438 bp T-RF increased dramatically its abundance. It was additionally possible to observe an unidentified population with a T-RF of 78 bp increasing until 57 days and a subsequently decreasing in its abundance. This was also observable on a lower abundance level for an unassigned T-RF with 54 bp. The previously dominating T-RF population with 147 bp clearly decreased over time. The T-RFLP of *bssA* primer and FAE-B amplicons (data not shown) was inconclusive, also after an additional *Hae*III T-RFLP.

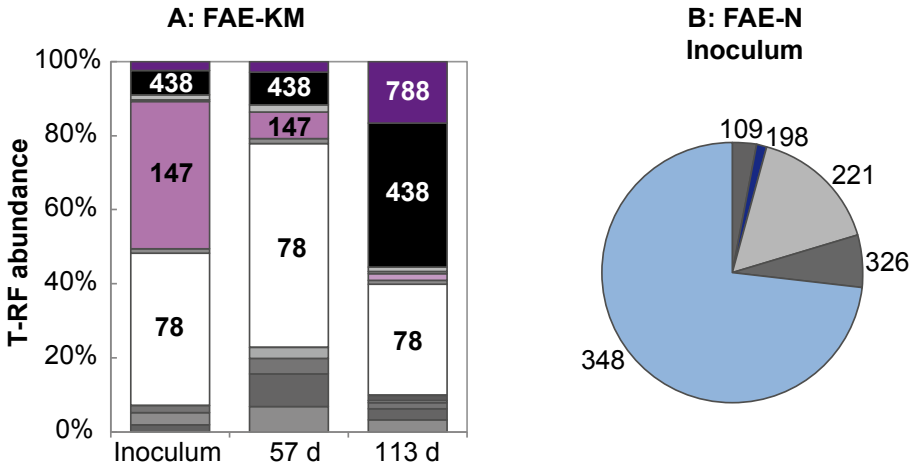

Figure 16. Community composition of *assA* (A) and *nmsA/bssA* s. l. (B) genes retrieved from Guaymas SIP experiment, T-RF are given in bp. Identifiable T-RF are in Figure 16A at 147 bp Guaymas oil seep clone BGM02, 788 bp Guaymas oil seep clone BGM24 and in Figure 16B at 198 bp Gulf of Mexico sediment 161 clone BG2N03 and at 348 bp Guaymas oil seep clone BGN08.

Thus, I was able to trace the dynamics of different FAE gene populations during incubation with butane. The population changed its composition but diversity was largely maintained on *assA* gene level.

3.3 Establishment of a pyrosequencing pipeline for *bssA* amplicons

One challenge in the monitoring of functional genes via degenerate primers as shown in Table 4 is balancing between comprehensively detecting a broad range of targeted genes and meaningless artefacts. Additionally, primers may have a high yield in one environmental system but a low performance in other systems, especially if degeneracies need to be as

frequent as for the new FAE primers I developed in this thesis (Table 4). This hinders further amplicon screening by T-RFLP or classical cloning and sequencing, which subsequently results in a relatively low output of usable sequences, as encountered in chapter *3.2.2*.

Potentially, biases of PCR and low insert yields of clone libraries can be overcome by the massive numbers of reads initially delivered with next-generation sequencing. Here, I was able to demonstrate this advantage by applying for the first time pyrotag sequencing on FAE gene markers. Several libraries of FAE amplicons were generated from Flingern sediment DNA and also from a methanogenic, benzene-degrading enrichment with Gölzau sediment provided by Martin Krüger and Friederike Gründger. These samples (see Table 12) were chosen as abundant FAE community data gained by cloning as well as T-RFLP existed for the Flingern samples (see chapter *3.2*). Furthermore, a strong formation of PCR artefacts for the Gölzau enrichment was observed (see Figure 20), which therefore remained up to now inaccessible to FAE community analyses via classical cloning and sequencing or T-RFLP.

Table 12. Samples used for pyrosequencing libraries and respective statistics concerning number of reads, read length and FAE classification yield.

	Libraries		F3A	F4A	F5A	F3B	G3-58	G3-52
			Flingern aquifer sediment 2009				Gölzau sediment enrichment (benzene)	
			6.85 m	7.15 m	7.25 m	6.85 m		
	Primer set		7772f 8543r	7772f 8543r	7773f 8543r	7768f 8543r (FAE-B)	7772f 8543r	7772f 8543r
	Annealing (°C)		52	52	52	58	58	52
reads	raw data		3527	3639	5974	7503	4205	6962
	after trimming		3438	3588	5565	6480	3919	6811
	≥ 250 bp		1755	2155	2221	2010	1563	4009
	forward		503	607	664	648	555	1530
	reverse		1252	1548	1557	1362	1008	2479
length	raw	average	375	407	301	478	282	340
		maximum	782	739	859	1639	933	761
	trimmed	average	236	258	207	316	199	258
		maximum	505	549	560	1038	514	567
FAE	total		1574	1902	1386	1084	1060	347
	% of raw data reads		45	52	23	14	25	5

This first demonstration was indeed successful and unprecedented total numbers of FAE gene reads were obtained from all samples. By applying different quality trimming strategies (quality trimming in Greengenes and discarding sequences shorter than 250 bp after trimming), the number of reads as well as the average and maximum read length dropped significantly, as illustrated in Table 8. Still, after classification, the number of total FAE reads was much higher than in standard clone libraries, despite yields in the range from 5 – 52 %. This is also illustrated by the rarefaction curves in Figure 17.

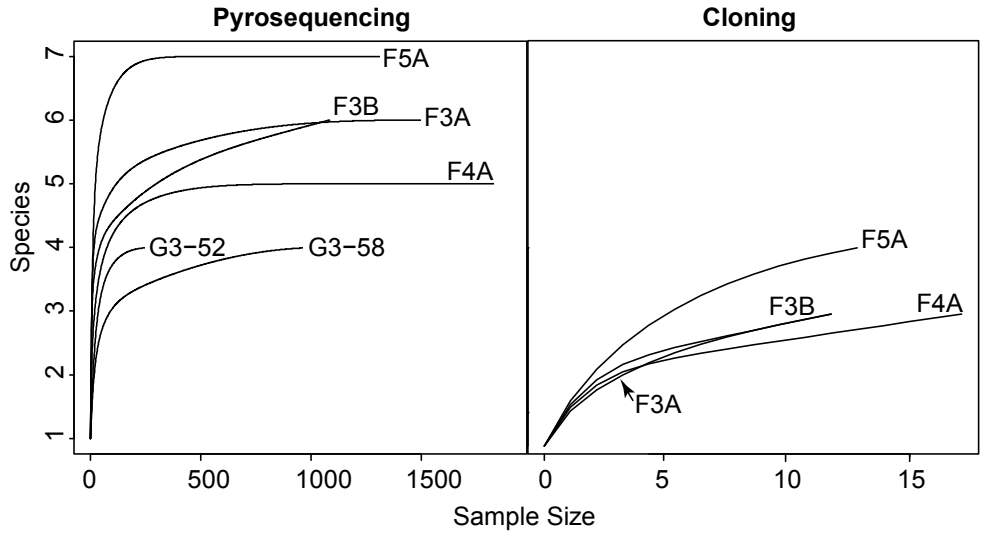

Figure 17. Rarefaction curves of pyrosequencing and cloning libraries.

In order to find a cut-off value for FAE-classification, the Identity/Query values were plotted versus the number of reads as shown in Figure 18. The Identity/Query value is the quotient of the similarity (in bp) to a FAE gene match and the read length, giving the total similarity of a read to a FAE gene. All the curves display a steep descent from ~ 90% to ~15% Identity/Query, therefore, an appropriate FAE identification cut-off was chosen at 90% Identity/Query. Only for the F3B library, there is a small plateau below 60 % Identity/Query which may be due to PCR artefacts.

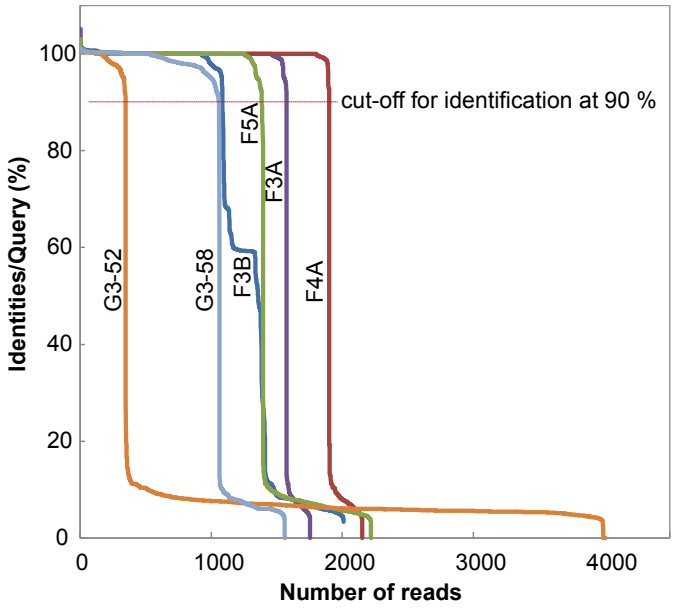

Figure 18. Plot of Identities/Query versus number of reads for determination of classification cut-off, indicated by the dotted red line.

The community composition of the libraries from Flingern aquifer sediment samples (F3A, F4A, F5A and F3B) was in general highly comparable with the data from prior T-RFLP and cloning, as shown in Figure 19. Thus, in the pryotag libraries created with *bssA*-primers (F3A, F4A, F4A), reads related to the F1-cluster dominated. *bssA* sequences related to the clostridial F2-cluster became also detectable. With the FAE-B primer (library F3B), it was moreover possible to detect the presence of *nmsA* s. str. sequences. Additionally, pyrosequencing revealed the presence of betaproteobacterial *bssA* sequences which were not detectable by T-RFLP and only marginally present in some clone libraries (see also Table 9). The abundance of the single OTUs was different, especially when compared with cloning. Except for betaproteobacterial *bssA*-homologous sequences, which were present with higher abundances in pyrosequencing libraries than in T-RFLP data, all other lineages were present in the same relations in pyrosequencing libraries and T-RFLP. The rarefaction curves in Figure 17 show saturation for all pyrosequencing samples, except for F3B, where a positive slope was still observable beyond ~ 1000 reads. In comparison, the cloning libraries never reached saturation within any attainable sample size.

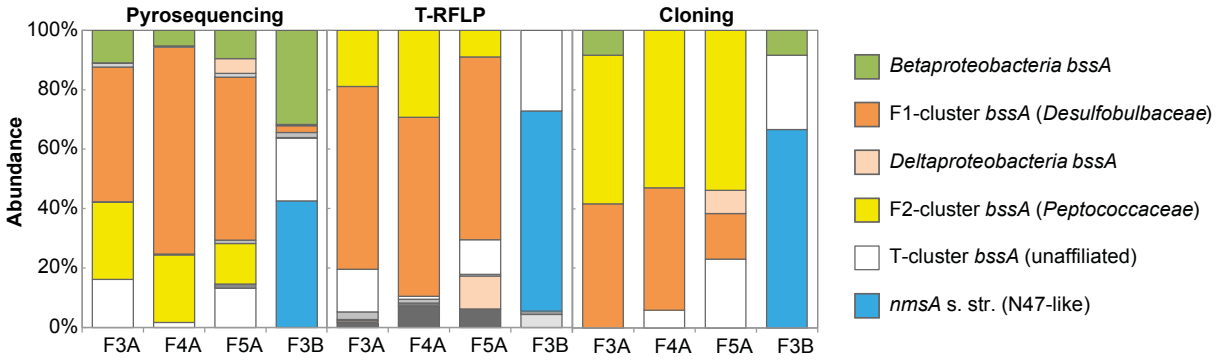

Figure 19. Comparison of community composition of pyrosequencing libraries, T-RFLP and cloning libraries. The libraries F3A, F4A and F5A were obtained with primer set 7772f 8543r, F3B with primer set 7768 8543r (FAE-B).

Both Gölzau aquifer sediment enrichment libraries (G3-52 and G3-58) reached saturation regardless of read abundance as indicated in Table 12. As shown in Figure 20A, the amplicon quality for the two libraries exemplarily generated at different annealing temperatures differed drastically, but the community composition of FAE-specific reads still remained largely similar (Figure 20B). The community was mainly dominated by *bssA*-like sequences related to the OTU of strain BF *bssA*-homologues. This deeply-branching cluster of clostridial origin is especially interesting, because the function of the strain BF *bssA*-homologue in hydrocarbon degradation is yet entirely unknown (Abu Laban *et al.* 2009). As the community composition did not differ despite the obvious amplicon quality differences pointed out in Figure 20A, this is a nice demonstration for the abilities of the application of next-generation sequencing for hydrocarbon degrader community analysis via FAE genes.

Thus, I demonstrate here for the first time a new pipeline for generating and analysing FAE gene pyrotags. With this method, it was possible to detect additional degrader community components not traceable by alternative T-RFLP or clone library approaches. Pyrotag screening reached saturation in diversity coverage, which was never possible with classical cloning and sequencing to date.

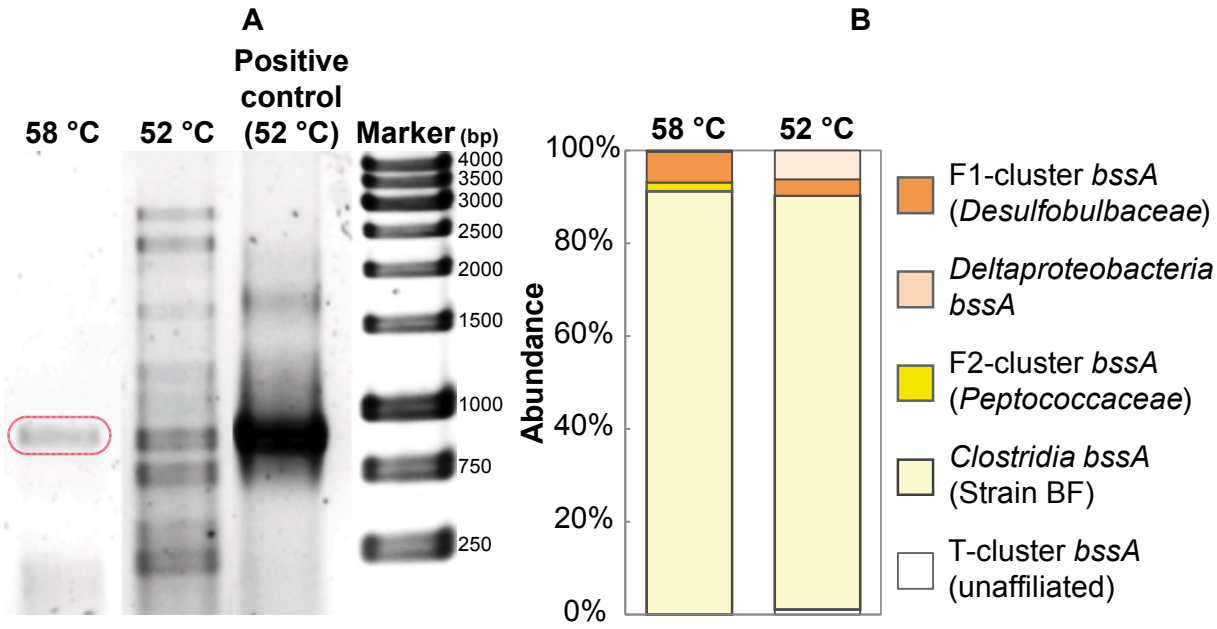

Figure 20. Test of pyrosequencing robustness towards artefacts. A: Agarose gel with G3 *bssA* primer amplicon. Note the band weakness at 58 °C and the artefacts at 52 °C annealing temperature. B: Community composition of the G3 pyrosequencing libraries.

4. Discussion

This thesis shows that FAE gene pools are more diverse and widespread in hydrocarbon-impacted environments than previously known. The detectability of FAE genes was indeed independent of the sample origin: marine or aquifer sediment, natural occurrence of hydrocarbons or anthropogenic contamination. This emphasises the general importance of anaerobic hydrocarbon degradation in the environment, as the genetic potential for degradation seems to be widespread. Amongst all described anaerobic hydrocarbon degradation pathways, activation by fumarate addition is the most intensively investigated to date, and has been demonstrated for a great number of substances. Other pathways such as oxygen-independent hydroxylation or carboxylation are either known for only a small selection of substances (Heider 2007) or researchers are just beginning to investigate those pathways in more detail (Meckenstock and Mouttaki 2011, Mouttaki *et al.* 2012). Thus, my results substantiate that FAE genes can be regarded as comprehensive genetic marker for investigating the ecology of anaerobic hydrocarbon degraders and degradation in the environment. This thesis shows also that hydrocarbon-degrading microbial communities are much more complex than currently perceived already on FAE gene level alone. With increasing knowledge about the biochemistry and substrate versatility of other key reactions in anaerobic hydrocarbon degradation, it should be possible to gain in the near future an even more comprehensive picture of degrader populations in the environment.

The present study focuses on functional gene analyses (amplicon fingerprinting and sequencing) for degrader detection and identification, in contrast to other strategies using combined approaches such as compound stable isotope analysis and detection of metabolites (Griebler *et al.* 2004), 16S rRNA and functional gene approaches (e.g. Winderl *et al.* 2008) or metabolite and functional gene detection (e.g. Oka *et al.* 2011) for tracing biodegradation and degraders at contaminated sites. Of course, a comprehensive detection of degradation on all levels distinguishing abiotic and biotic natural attenuation via compound stable isotopes, active pathways as traced via metabolites and community structure as revealed by phylogenetic and functional gene markers should be the favoured approach in general. But such a comprehensive screening is not always feasible in many cases due to project limitations. However, the new primer sets here introduced are able to survey directly and rapidly (in fingerprinting) the structure of natural FAE gene pools in hydrocarbon-degrading terrestrial and marine systems. This offers novel possibilities for application such as contaminated site management, which is discussed further below (chapter *4.10*).

4.1 Affiliation of *bssA* in SIP experiments with Testfeld Süd and Flingern aquifer sediment

First, I aimed to clarify the phylogenetic affiliation of the novel, deeply-branching *bssA* clusters F1 and F2 from the Flingern aquifer, as well as sequence cluster T from Testfeld Süd. This was achieved as additional contribution to the SIP experiments performed by Christian Winderl (2007) and Giovanni Pilloni (2011). My own contributions to these projects were T-RFLP analyses on functional gene level with *bssA* primers and, for Testfeld Süd SIP, also on 16S rRNA gene level. While the T-cluster *bssA* homologues remain yet unaffiliated, it was nevertheless possible to affiliate the F2-cluster *bssA* to *Desulfosporosinus* spp. (members of *Peptococcaceae*, belonging to *Clostridia*) in the Testfeld Süd SIP experiment and the F1-cluster *bssA* to *Desulfobulbaceae* in the Flingern SIP experiment. The results of my analyses demonstrate a need for extending the current detection systems towards more deeply-branching FAE genes, as discussed further below.

4.1.1 Testfeld Süd sediment SIP: Identification of clostridial *bssA* genes

During the incubation of Testfeld Süd sediment with toluene, my results revealed a clear decrease of the T-RF population associated with T-cluster *bssA* sequences. However, it was not possible to tentatively match any 16S rRNA gene T-RF to this declining *bssA* T-RF, so that the deeply-branching T-cluster initially detected at the site remains unidentified. Herrmann *et al.* (2009) discovered sequence types of which some cluster within the T-cluster in sulphate reducing enrichments with xylenes from aquifer sediment of the former industrial site in Zeitz (see Figure 15). Thus, it can be assumed that this deeply-branching FAE gene cluster may be solely involved in the degradation of xylenes and not toluene. *p*- and *o*-xylene are indeed present at the Testfeld Süd aquifer in the same concentration (~ 3-4 µg/l) as toluene (Bockelmann *et al.* 2001, Zamfirescu and Grathwohl 2001)

Instead, dominating *bssA* populations from the Testfeld Süd SIP incubations were related to F2-cluster *bssA* sequences and affiliated to *Desulfosporosinus* spp., (*Peptococcaceae*, *Clostridia*). My results from 16S rRNA as well as *bssA* gene T-RFLP prove that this particular population was present from the beginning of the experiment as a member of the natural aquifer community. However, it is not possible to define whether the two closely related T-RFs (146 and 178 bp) found on 16S rRNA gene level actually represent two closely

related, functionally identical subpopulations of *Desulfosporosinus* spp. or whether this was one population with two slightly varying 16S rRNA operons (Pei *et al.* 2010).

These results alleviate a very central critique of SIP being not much more meaningful than a 'classical' enrichment with a loss of original population diversity and results hardly linked to the initial inoculum. The suggested affiliation of the *bssA* F2-cluster to *Clostridia* was recently corroborated by findings of Fowler *et al.* (2012). They found *bssA* sequences related to F2-cluster *bssA* sequences in a methanogenic enrichment with toluene and sediment from a gas condensate contaminated aquifer (see 'Toluene-degrading Fort Lupton enrichment clone MTOL_1' in Figure 15). DNA of this methanogenic, toluene-degrading SIP incubation was also dominated by *Clostridia* sequence types on bacterial 16S rRNA gene level. On the whole, these results further support the previously unrecognized importance of *Peptococcaceae* in BTEX degradation and subsurface environments; which is discussed in more detail below (chapter *4.8*).

With the annealing temperature set at 58 °C, it was not possible to trace the T-RFs of T- and F2-cluster *bssA* sequences as well as with 52 °C. This implies that primer performance is not always optimal and further development is necessary. Otherwise, important members of a degrader community may be missed due to PCR bias. The *bssA* primer (7772f/8546r) seems also to be selective for betaproteobacterial *bssA*, which can play only a minor role in this sulphate reducing, nitrate-free setting (Winderl *et al.* 2010). This finding motivated the design of more advanced primers as subsequently done in this thesis. Furthermore, *Taq*I alone seems to be not an optimal restriction enzyme for tracking the clostridial *bssA* sequences identified here. This was due to the necessary analysis cut-off for peaks with fragment lengths below 50 bp, as shorter fragments cannot always be distinguished from primer artefacts. However, here it was possible to clearly distinguish artefact-derived T-RF peaks from the 45 bp peak in question. In order to solve that problem, a dual-digest T-RFLP analysis of FAE gene pools with *Taq*I and *Hae*III was proposed in thesis.

4.1.2 Flingern sediment SIP: Affiliation of F1 *bssA* gene cluster to *Desulfobulbaceae*

The results from the Flingern SIP study reveal the central role of *Desulfobulbaceae* for *in situ* toluene degradation driven by sulphate reduction. It is the first example for identification of a novel *bssA*-carrying lineage dominant under *in situ* conditions. The dominance of the F1 *bssA* gene cluster and *Desulfobulbaceae* at the Flingern aquifer was already shown by Winderl *et al.* (2008). However, it was at that time not possible to prove that the F1 *bssA* gene cluster

actually belonged to *Desulfobulbaceae*. Additionally, these results suggest that the F1-cluster *bssA* sequences might be acquired by lateral gene transfer, since they fall into a branch with otherwise geobacterial *bssA* sequences (Figure 15). This demonstrates the polyphyletic nature of *bssA* phylogeny (Figure 4 and Figure 15). Another example for lateral gene transfer is the *bssA* gene of toluene degrading strain *Magnetospirillum* sp. TS-6, an *Alphaproteobacterium*. This *bssA* gene is placed amongst the sequences of nitrate-reducing *Betaproteobacteria* which may have served as a donor. For the F1 *bssA* gene cluster, it was already speculated that *Geobacteraceae* might be the original donor species, due to their close phylogenetic relationship on functional gene level (Figure 15, Winderl *et al.* 2008). Therefore, *bssA* or FAE genes in general display a low diagnostic rigour as a phylogenetic marker. The polyphyletic distribution of FAE genes hints also towards their genetic mobility. This may point to a possibility of the emergence of novel degraders in long-term contamination degradation scenarios due to lateral gene transfer (van der Meer 2006).

For the Flingern aquifer sediment SIP incubations, an ideal scenario would show the absence of F1-cluster *bssA* sequence T-RF in iron incubations instead of the observed presence in both treatments. This could be explained by a memory effect from detecting DNA of dead or inactive *Desulfobulbaceae* cells, which is confirmed on 16S rRNA gene level by Giovanni Pilloni (2011). Such a problem could be solved by using mRNA, but gaining the necessary yields from sediment extractions is problematic. At the same time, the *bssA*-T-RFLP data shows that the employed *bssA* primers are not ideal for tracing *Georgfuchsia* spp. and *Thermincola* spp., which were identified as key players in the incubations with ferric iron. The annealing temperature of the classic *bssA* primer sets was not lowered to 52 °C as in the previous Testfeld Süd SIP experiment, in order to enhance the detectability of F1-cluster *bssA* genes. It can be argued that with either a lowered annealing temperature or by also lowering the T-RFLP quality cut-off, the clostridial and betaproteobacterial populations found on 16S rRNA gene level in these iron-reducing SIP-incubations may also have been detectable on functional gene level.

This demonstrates the need for more general, optimized FAE primers with lower PCR selection for certain populations, which is elaborated in chapter *3.2.3*. Another way to unravel the presence of *Desulfobulbaceae bssA* genes in both incubations would be the introduction of a qPCR for different *bssA* gene lineages, as T-RFLP provides only semi-quantitative information. Thus, it would be possible to normalize the T-RF peak results and then to compare the copy numbers to sequence abundance on 16S rRNA gene level, expecting a far

lower copy number of *Desulfobulbaceae bssA* genes in the incubations with ferric iron than in those with sulphate.

4.2 Performance of enhanced FAE gene assays in pure cultures and environmental samples

The tests for FAE lineage detectability on amplicon level (Table 8 and Table 9) clearly show that major FAE gene populations are not detectable with the established *bssA* primer set alone. Thus, not all catabolic potentials for fumarate-addition may be detected with this assay, as I demonstrate here for terrestrial and marine environmental samples. Additionally, the clones retrieved in this study prove that the synchronous application of my newly developed primer sets is able to recover a comprehensive range of proteobacterial and clostridial FAE gene lineages. The primer sets are also individually specific for the targeted lineages (i.e. FAE-B for *bssA*/*nmsA* s. l., FAE-N for *nmsA* s. str. and FAE-KM for *assA*). Currently, it is not possible to propose a single forward primer for detecting all FAE genes. Nevertheless, the 8543r reverse primer used in all new primer combinations effectively amplified FAE gene fragments in all different scenarios. Furthermore, the new FAE-KM primer set for *assA* detection was introduced, because the primary *assA* primers published by Callaghan *et al.* (2010) generate FAE amplicons not fully overlapping with the established *bssA*-7772f/8546r amplicons (Table 1, Figure 5), so that a robust phylogenetic comparison with sequence data from the new FAE primer sets would be impaired.

In part, the presented novel primer sets are based on a combination of different forward primers, which may potentially complicate the interpretation of amplification results. Yet, mixing of different primers is commonly used in multiplex PCRs (Settanni and Corsetti 2007). The application of such methods is not only common to the analysis of nutritional or clinical samples, but also established for the analysis of complex environmental samples (e.g. Volkmann *et al.* 2007, Xiong *et al.* 2012). For example, Knaebel and Crawford (1995) employed multiplex primers for the detection of aerobic contaminant degradation genes. There, amplification was shown to be very sensitive, but not linear for all template ratios. Still, the influence of sample origin, targeted gene and inhibitors from the sampled material was shown to introduce more considerable bias to PCR results (e.g. on template ratios and detectability) than the use of multiple primers. Because of this, the introduction of multiple forward primers in some primer sets was considered to be connected to an acceptable bias,

especially since catabolic gene analyses and fingerprinting are already known to be much less robust in a semi-quantitative manner than that of ribosomal markers (Lueders and Friedrich 2003).

Another assay performance parameter of concern was the low frequency of actual FAE gene clones observed especially for the libraries from the marine sites. Non-intended inserts were mostly shorter primer concatemers or similar PCR artefacts. With Flingern aquifer DNA, yield of 'good' clones was always substantially higher (between 40 – 60%). This is why the verification of the new T-RFLP diagnostics was confined to the Flingern samples, as PCR artefacts certainly complicate the interpretation of T-RFLP fingerprints. However, also other studies have reported low FAE sequence yields and cloning efficiency for certain sample types (Callaghan *et al.* 2010, Winderl *et al.* 2007), which may indicate general PCR problems in applying these degenerate primers to complex and low-quality environmental DNA. Here, a possible remedy would be the introduction of cloning-independent, high-throughput next-generation pyrotag sequencing of FAE amplicons as presented in this thesis. This method can be expected to produce sufficient amounts of 'good' reads, even if short reads from biases in pyrotag amplification are abundant.

The detection of community assembly on a qualitative level is a prerequisite for more quantitative approaches. Only with the knowledge of the full spectrum of the lineages present at a site it is possible to design in turn quantitative, specific assays as demonstrated by Winderl *et al.* (2008). With such quantitative data it is then possible to offer estimates on the on-going degradation processes. Such estimation of toluene degradation rates in microcosms based on chemical and functional marker analyses (*bssA* of nitrate-reducing *Betaproteobacteria*) is demonstrated by Kazy *et al.* (2010).

4.3 Insights into degrader diversity and on site hydrocarbon degradation processes

As the present study was entirely based on gene detection in extracted DNA, assumptions on actually on-going hydrocarbon-degrading processes in the examined samples are not straightforward. For example, several distinct FAE gene lineages were qualitatively detected in the Gölzau enrichments; despite they had been cultivated with mono-substrates for over a year. This either suggests a surprising diversity of catabolic potentials involved in mono-substrate biodegradation in these enrichments, or indicates a memory effect from initially

diverse *in situ* degrader populations for various hydrocarbons in Gölzau. At any rate, a more detailed investigation of catabolic potentials and genes in the Gölzau samples is still under way.

Assumptions on on-going hydrocarbon-degrading processes are also not straightforward for the sites with natural gas seepage. The occurrence of hydrocarbons (> C_4) other than natural gas is known for the Paclele Mici mud volcano (Alain *et al.* 2006), but for the Amon mud volcano and Nyegga methane seeps, there was to date no assessment for the presence of hydrocarbons higher than C_4 (Mastalerz *et al.* 2009, Van Gaever *et al.* 2010, Vanreusel *et al.* 2009). While the detection of putative *assA*-homologues is expected for natural gas seepages due to the potential degradation of short chain alkanes like propane via fumarate addition (Callaghan *et al.* 2010, Kniemeyer *et al.* 2007), it was unexpected to exclusively discover *bssA*-fragments related to the gene of *Desulfobacula toluolica* in Amon mud volcano sediments. This could imply two different scenarios: (i) the microbes carrying the *bssA*-sequence types found in Amon mud volcano sediments may be actually involved in the degradation of short-chain alkanes like butane or propane, despite the phylogenetic placement of these markers with toluene-degrading *bssA* s. str. (ii) The detection of these FAE gene types was attributed to a putative presence of aromatic hydrocarbon carry-over from nearby oil deposits via natural gas ebullition as known from Amon mud volcano's neighbour Isis mud volcano (Mastalerz *et al.* 2007). The latter would imply that the true alkane catabolizing gene markers at the site remained undetected, despite our comprehensive synchronous PCR screening, or could potentially even proceed via mechanisms other than fumarate addition. Considering that FAE substrate specificity for methylated aromatics and non-gaseous alkanes do usually not overlap amongst degrader isolates (Rabus *et al.* 2011), this still may be possible for short-chain alkanes like butane or propane.

Although the primer set FAE-B does detect a selection of several clostridial *bssA* genes and homologues *in silico* as well in pure culture screenings, it was only possible to find the clostridial F2-cluster *bssA* sequences in Flingern sediments using the 'classical' *bssA* primer pair at lowered annealing temperature. It seems that the FAE-B primer set has, despite its proven ability to amplify clostridial *bssA* in pure culture DNA, a bias for *nmsA* in environmental samples as observed in T-RFLP data retrieved from Flingern (see Figure 13). A lowered annealing temperature, as employed successfully for extending the range of the *bssA* primer set, may help to overcome this limitation, and still needs to be tested.

4.4 Novel *nmsA* sequences found in different environments

With the new primer sets FAE-B and FAE-N, the first evidence is provided that *nmsA* and homologous sequences can be detected in the environment, regardless of sample origin. This raises a question about the relevance of the respective catabolic pathway – addition of fumarate to methylnaphthalene. Recent work by Oka *et al.* (2011) showed metabolites of 2-methylnaphthalene carboxylation to be present at a contaminated aquifer under a manufactured gas plant. Additionally, methylation and subsequent fumarate-addition was first presumed to be the degradation pathway for naphthalene (Safinowski and Meckenstock 2006) but recent results suggest carboxylation (Mouttaki *et al.* 2012) as principal degradation mechanism in *nmsA* containing strain N47 (Selesi *et al.* 2010). This discovery abates the potential of *nmsA* as a marker for naphthalene degradation. DiDonato *et al.* (2010) found that fumarate-addition to 2-methylnaphthalene and toluene is possible with NMS alone for strain NaphS2. Therefore, the *nmsA* sequences retrieved in this study could potentially not only play a role in 2-methylnaphthalene catabolism (a common compound of complex hydrocarbon mixtures), but also in toluene degradation. This concept may also help to better explain the phylogenetic placement of the *bssA*-homologue previously retrieved from the toluene-degrading clostridial strain OX39 (Winderl *et al.* 2007), which does not degrade 2-methylnaphthalene (Morasch *et al.* 2004). Yet, its *bssA*-homologous sequence is untypically placed in the *nmsA* branch with respect to the now widely documented 'typical' clostridial *bssA* sequence type (F2-cluster, see Figure 15). Thus, the new sequences related to OX39 (Amon mud volcano clone BAA23, Gulf of Mexico sediment clone BG2A05 and Flingern aquifer clone F3N20) may be non-PAH degraders of putative clostridial affiliation utilizing aromatics.

4.5 Hidden FAE gene lineages revealed in spatial resolution by sequencing-independent T-RFLP screening in Flingern sediments

The results for the Flingern aquifer depth transect demonstrate, in contrast to previous findings (Winderl *et al.* 2007), that there are indeed different intrinsic FAE gene populations to be found in this contaminated aquifer. The apparent dominance of the *Desulfobulbaceae* F1-cluster *bssA* reported previously may thus be an effect of PCR selection. It is shown that the clostridial F2-cluster *bssA*-sequences are not consistently detected over the three selected depths by the standard *bssA* primer set (58 °C), while putative toluene degraders related to

Desulfosporosinus spp. have been identified in sediments from the same site sampled in 2006 (Pilloni *et al.* 2011), and are abundant in 16S rRNA pyrotag libraries of sediments taken in 2009 (Pilloni *et al.* 2012). The spatially resolved T-RFLP data presented here are largely consistent with classical *bssA* amplicon screening. Still, with lowered PCR annealing temperature and the new primer sets developed here, the presence of subpopulations other than the *Desulfobulbaceae* F1-cluster *bssA* becomes more evident in the different sediment transects. In this manner, it is now possible to observe distinct FAE gene pools for each depth transect.

The identification of degrader lineages based on T-RFLP results was straightforward at the Flingern site, as there is already ample *bssA* sequence data available from there (Winderl *et al.* 2007, 2008), facilitating the matching of T-RFs observed in dual-digest fingerprints to defined lineages. It is not possible to claim that the new approach presented here has absolute diagnostic capability, sequencing will still be necessary to identify FAE lineage if there are no sequences with matching *in silico* T-RFs in the database. Nevertheless, the double digest can add diagnostic security in case of overlapping T-RF sizes for unrelated FAE lineages (e.g. a *Hae*III T-RF of 132 bp or a *Taq*I T-RF of 50 bp could both represent distinct beta- and deltaproteobacterial as well as clostridial *bssA* lineages, as illustrated in Figure 15) but the dual-digest combination would mutually exclude options. Furthermore, the dual-digest approach helps to prevent the 'missing' of OTUs, as shown here for the 35 bp *Betaproteobacteria bssA* fragment with *Taq*I. The employment of the advanced T-RFLP approach reduced the cloning and sequencing effort for the analysis of the Flingern samples as they displayed the same T-RF lineages over the three analysed depths in a highly reproducible manner for the *Taq*I-digests. Thus, a dual-digest T-RFLP can be used as a tool for tracking and analysing FAE-population of a site with minimized sequencing effort.

4.6 Degrader dynamics as traced in Guaymas Basin hydrocarbon seep sediment SIP incubations

DNA samples from several time points of a SIP experiment from Sara Kleindienst (2012) with butane and sediment from the Guaymas Basin hydrocarbon seep were analysed with the newly introduced assays for FAE genes. The change from the complex hydrocarbon mixture present at the hydrocarbon seep (Kleindienst *et al.* 2012) to a single substrate can be interpreted as a large disturbance to the original degrader community. Indeed, it was possible

to observe the change of FAE gene populations during the incubation: the initial general community composition, where only *nmsA/bssA* s. l. and *assA* was detectable, was not sustained with butane (as indicated in Table 11 and Figure 16). While the initial degrader community containing *nmsA/bssA* s. l. genes, mainly dominated by the 348 bp T-RF population, was seemingly not able to degrade butane, the initially dominating *assA* population was replaced. Lacking FAE gene qPCR data, it is however not possible to quantify and compare the degrader populations in order to investigate if changes of general FAE gene abundance were also observable. Additionally, *bssA/nmsA* s. l. became clearly detectable as amplicons over time, but not yet resolvable on community level. As the data is based on DNA, it can be assumed that the observed changes are indeed dynamics of individual degrader abundance and not only differential gene transcription in a single degrader with multiple FAE genes.

A degrader population with the capability to use several different hydrocarbon compounds would result in a community more resilient to changes on electron donor level and therefore, no change of community composition would be detectable. An example for such a degrader would be the NaphS2 strain, which was revealed to degrade 2-methylnaphthalene and toluene even with the same NMS (DiDonato *et al.* 2010). And as the changes are different on the main FAE gene levels, an effect of a single bacterium possessing several FAE-genes can also be excluded. As a hypothesis for future work, it may be postulated that the higher diversity of *assA* genes (two major lineages instead of one as for *nmsA*-population, not including minor populations for both populations) may be beneficial for the subsequent existence of the *assA* community in the sense of functional redundancy in the terms of the insurance hypothesis (Yachi and Loreau 1999).

The appearance of the *bssA* gene populations during incubation may be due to an as yet uncharacterized detection limit of the used primer sets for low functional gene abundance. Yet, these results prove the general utility of the introduced primer sets and T-RFLP method for monitoring degrader dynamics, which will foster in future new insights into the ecology of hydrocarbon degrader communities.

4.7 FAE-amplicon pyrosequencing

Pyrosequencing of FAE gene amplicons was performed with samples from Flingern sediment (the same as in chapter *3.2.3*) in order to obtain deeper insights into degrader diversity and to

compare pyrotag results to previous results. The pyrotag data was in general consistently comparable with earlier T-RFLP or clone library data. Additionally, the method was tested on a sample (Gölzau aquifer sediment enrichment with benzene) with high abundance of PCR artefacts as observed on gel electrophoresis (see Figure 20A, chapter *3.3*). All samples exhibited abundant FAE sequence reads, even with low yields as indicated in Table 12.

In addition, the saturation of the rarefaction curves and the different sequencing yields (from 5 – 52 %) showcases the robustness of FAE amplicon pyrosequencing, even in the case of weak bands and severe PCR artefacts as for the Gölzau enrichments. This gives an optimistic outlook for further implementation of this method for the degenerate primer sets FAE-N and FAE-KM as well as for searching and investigating further 'rare' FAE gene populations (e.g. the Testfeld Süd related sequences found in Flingern sediment) in the future. Although an assembly to complete amplicons is at the moment not possible due to short overlaps, this should be possible with 'third generation sequencing' with longer read lengths (up to 1000 bp). Thus, in future a reliable comparison between T-RFLP and pyrosequencing reads may be possible, as already described for 16S amplicon pyrosequencing (Pilloni *et al.* 2012). Even if pyrotag amplicon sequencing becomes more and more routine, T-RFLP may be still the method of choice for rapidly and comparably cheaply tracing hydrocarbon-degrading communities in the future, especially if T-RF peaks in all samples were once reliably identified by a single pyrosequencing run.

Pyrotag amplicon sequencing also revealed an important flaw of the single enzyme T-RFLP as pointed out earlier: Populations below the T-RF size exclusion threshold (usually < 50 bp) may be missed, even if they represent a major part of the population as shown in Figure 18 for sample F3B. The same is true for the clone libraries where rare populations like the betaproteobacterial *bssA* were missed. The rarefaction curves provide primary evidence on the coverage needed for a saturated recovery of FAE gene pools. In general, overlaps between pyrosequencing and T-RFLP results were consistent. The advantage of pyrosequencing here is that it is possible to distinguish clearly between primer artefacts and small T-RFs, and thus to analyse FAE gene pools in a manner not affected this kind of artefact.

Moreover, the presence of *bssA*-homologous sequences in the methanogenic benzene degrading Gölzau enrichment related to the iron-reducing, benzene degrading strain BF is of particular interest as most recent research suggests carboxylation (Abu Laban *et al.* 2009, Meckenstock and Mouttaki 2011) and not methylation with subsequent fumarate addition as

the activating step for benzene degradation. Thus, the role of the *bssA*-homologues found in strain BF and the Gölzau enrichment remains an enigma.

The analysis pipeline set up here is as a first proof of principle and does not require complex bioinformatic tools such as mothur (Schloss *et al.* 2009), which is already used for analysing *pmoA* gene amplicon pyrosequencing results (Lüke and Frenzel 2011). A protein blast was not used here, as it would have been more sensitive to errors in reading frame. Homopolymers are a known difficulty for pyrosequencing, and may thus explain the small yield of actual FAE sequences achieved here compared to established 16S rRNA gene based gene pyrosequencing. Wrongly sequenced homopolymers can thus produce reading frame shifts, which in turn prevent correct identification. The BLASTN algorithm was chosen as it is able to find similar sequences as well for being able to identify sequences.

4.8 *Clostridia* as key players in subsurface anaerobic hydrocarbon degradation

Only recently, it was possible to show the relevance of clostridial lineages, especially those related to *Desulfosporosinus* spp., in hydrocarbon-impacted subsurface environments (Winderl *et al.* 2010, Fowler *et al.* 2012) The importance of *Clostridia* seems to have been generally underestimated in subsurface processes. This may be due to shortcomings of culture dependent isolation techniques which in general favour fast growing bacteria tolerant to relatively high substrate concentrations. In the SIP experiment with Testfeld Süd sediment conducted by Christian Winderl (2007), high substrate concentrations were prevented by the use of the adsorber resin XAD7. This adsorber resin was initially introduced in the successful isolation of the primary clostridial toluene degrader *Desulfotomaculum* sp. OX39 (Morasch *et al.* 2001, Morasch *et al.* 2004) and also key for isolating the other clostridial strains in the laboratory of R. U. Meckenstock; such as the toluene degrader *Desulfitobacterium aromaticivorans* (Kunapuli *et al.* 2010) as well as the benzene degrading strain BF (Abu Laban *et al.* 2009). Therefore, it can be speculated that low substrate concentrations and cultivation under close *in situ* conditions may be generally necessary for enriching novel clostridial hydrocarbon degraders.

The recent advent of culture-independent techniques showed the presence of clostridial species in several subsurface ecosystems (Detmers *et al.* 2004), even as a dominant taxon in the deep subsurface (Chivian *et al.* 2008). Clostridia display some unique features which may influence the ecology of subsurface degradation processes. Many *Clostridia* are able to

perform incomplete oxidation of a substrate under electron acceptor limitation. The effect of such an incomplete degradation of a contaminant may be such that a toxic compound is degraded to a less toxic compound (e.g. alcohols or organic acids, most prominently acetate) which is in turn a potential substrate for methanogenesis. This is especially important as aquifer ecosystems are generally electron acceptor limited (Griebler and Lueders 2008). Another feature of *Clostridia* that is not as common in *Proteobacteria* is the ability to survive unfavourable conditions in persistent stages, such as endospores. Therefore, *Clostridia* could be perceived as important seeding populations after disturbance events.

4.9 General insights into the diversity and ecology of anaerobic hydrocarbon degraders

Hydrocarbon pollutions are often complex mixtures of different compounds. These diverse compounds may thus provide ecological niches for different degraders. For FAE mediated degradation, proteobacterial toluene degraders and the degraders of aliphatic hydrocarbons have been shown to be substrate specific (Rabus *et al.* 2011). Thus, those strains could coexist in a complex hydrocarbon plume. However, some organisms, such as the aforementioned NaphS2-strain (DiDonato *et al.* 2010), are able to degrade both methylnaphthalene and toluene. And there are some strains with several FAE genes in their genome (*Geobacter daltonii* FRC-32*, Desulfatibacillum alkenivorans* AK-01). Therefore, a certain flexibility of electron donor usage by such degrader subpopulations towards different hydrocarbon substrates may be expected. This is important for disturbance scenarios, when one functional redundant degrader population may be replaced by another existing population. However, this would reduce degrader diversity and may result in a reduced resistance against disturbance. It should not be forgotten, that hydrocarbon degraders are also able to thrive on other substrates. Suspected key players may thus not be involved directly in hydrocarbon degradation, as demonstrated for the Banisveld aquifer (Staats *et al.* 2011) mentioned earlier. There, *Geobacteraceae* were found not to be involved in initial hydrocarbon breakdown, even though they were abundant on 16S rRNA gene level. It can be hypothesised that such a population could also replace the key degrader after a disturbance event. In brief, for obtaining a qualified insight on on-going processes and potential degradation capacities after disturbances, it is necessary to apply tools for a comprehensive and thorough screening of the community present at contaminated sites. Such tools were developed and applied in this thesis.

4.10 Possibilities of an application in site remediation

In this thesis, several molecular methods based on PCR are introduced for detecting anaerobic hydrocarbon degraders using FAE genes as a functional marker. As hydrocarbon pollution is an on-going threat to pristine groundwater environments, a critical evaluation of the application possibilities of such methods in site remediation strategies is presented here.

The application potential of the here introduced primer sets, T-RFLP method and amplicon pyrosequencing assay are manifold. The use of different primers sets gives certain flexibility, if a question is focusing on the detection specific FAE gene lineages. The different primer sets can be also used for excluding the presence of certain lineages, giving the simple amplicon-based detection a more diagnostic edge. The T-RFLP method provides a simple and rapid method for following degrader community structure in space and time. A good standardisation with *in silico* T-RFs is however still essential for unleashing the full diagnostic potential of the dual-digest T-RF as shown in this thesis. The T-RFLP results of this thesis show also already a rich untapped diversity of uncultured FAE genes which remain to be determined/discovered with either novel single-cell, next-generation sequencing techniques or classical culture based methods.

However, for a simple monitoring of natural attenuation mediated degradation correlating with geochemical data by a governmental authority or private companies, it would not even be necessary to know the identities of FAE gene T-RF peaks, as the relations of FAE T-RF peak intensity should correlate in a robust way with degradation data of a hydrocarbon compound. This has already been shown for qPCR estimation of biodegradation by Kazy *et al.* (2010). The usage of more advanced techniques such as pyrosequencing, apart from providing detailed sequence information, provides also reference sequences for a T-RFLP-based monitoring. Such detailed sequencing information delivers not only the presence of major key players, but also their identity. With additional literature data it may be also possible to provide preliminary information about putative substrate usage patterns or electron acceptor requirements. For example, Acosta-González *et al.* (2013) showed for *bssA*-like genes in different hydrocarbon-polluted sediments a substrate-dependent clustering.

This showcases also the versatility of functional gene markers compared to other methods such as metabolites or isotope fractionation. With the FAE gene markers used in this thesis, it is possible to provide clues on a catabolic potential in different environments (terrestrial aquifers, marine samples) without the need of expensive analytical equipment. However, if

the means are available, an optimized monitoring approach should ideally be based on a parallel approach: Compound stable isotope analysis for distinguishing abiotic and biotic processes, detection of metabolites for determining active pathways and molecular ecology approaches for identifying key players. In this manner, an ecological perspective can be incorporated into contaminated site monitoring during natural attenuation or bioremediation approaches such as bioaugmentation.

5. Conclusions

The aim of this thesis was to advance the detection systems at hand for anaerobic hydrocarbon-degrading bacteria based on fumarate-adding enzyme (FAE) gene markers. Thus, important unidentified FAE gene lineages detected in the environment were to be assigned to defined degrader populations, and novel, comprehensive insights into the biodiversity and ecology of the respective anaerobic degraders in contaminated systems were to be provided. Most importantly, the following advances were elaborated:

(i) The SIP experiment with Testfeld Süd aquifer sediments revealed the existence of a novel clostridial *bssA* sequence cluster, previously described as 'F2-cluster' *bssA*. The results from the Flingern SIP study substantiated the role of *Desulfobulbaceae* as key players for toluene degradation *in situ* and demonstrate a possible lateral gene transfer of their F1-cluster *bss* between iron- and sulphate-reducing *Deltaproteobacteria*. Problematic *bssA* detection encountered in both SIP experiments demonstrated the need for enhanced *bssA* primer sets and T-RFLP detection assays.

(ii) Comprehensive FAE primer sets were developed and applied; they were necessary to recover full diversities of FAE-gene lineages in different hydrocarbon-impacted terrestrial and marine systems. Especially, novel primer systems were presented and evaluated here for detecting clostridial *bssA* homologues, *nmsA* and *assA*. It was not possible to develop one single primer pair spanning the complete diversity of FAE genes. Via amplicon screening as well as sequencing-independent T-RFLP diagnostics it was revealed that certain FAE gene-lineages are missed with available *bssA* primer sets in different cultures, enrichments and environments, especially clostridial and *nmsA*-homologues. The assessment of degrader populations and FAE gene pools in different hydrocarbon-impacted environments reveals the presence of several FAE gene lineages not detected in these environments before: proteobacterial *nmsA* homologues and, to a lesser extent, clostridial FAE. This gives the opportunity to extend the knowledge of hydrocarbon biodegradation and the ecology of involved bacterial populations in future studies addressing bioremediation and natural attenuation as shown for the exemplary monitoring of FAE population in a SIP experiment with butane over time.

(iii) For the first time, FAE gene pyrotag sequencing was successfully applied for unravelling degrader community structure in environmental samples. Overall recovery of FAE gene pools was surprisingly robust between samples with varying amplicon quality and classification yields. As a method, pyrosequencing of FAE gene amplicons is thus capable of overcoming crucial PCR artefacts encountered in classical cloning and sequencing. This will allow, together with T-RFLP screening, a high-throughput, cloning-independent workflow for tracing hydrocarbon degrader populations over time. With this method, high-quality data can be generated to unravel the ecological controls of hydrocarbon-degrading microbial communities, such as functional redundancy and the role of rare community members.

6. References

Abu Laban N., Selesi D., Jobelius C. J., Meckenstock R. U. (2009). Anaerobic benzene degradation by Gram-positive sulfate-reducing bacteria. *FEMS Microbiology Ecology* **68:** 300-311.

Abu Laban N., Selesi D., Rattei T., Tischler P., Meckenstock R. U. (2010). Identification of enzymes involved in anaerobic benzene degradation by a strictly anaerobic iron-reducing enrichment culture. *Environmental Microbiology* **12:** 2783-2796.

Achong G. R., Rodriguez A. M., 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. *Journal of Bacteriology* **183:** 6763-6770.

Acosta-González A., Rosselló-Móra R., Marqués S. (2013). Diversity of benzylsuccinate synthase-like (*bssA*) genes in hydrocarbon-polluted marine sediments suggests substrate-dependent clustering. *Applied and Environmental Microbiology* **79:** 3667-3676.

Alain K., Holler T., Musat F., Elvert M., Treude T., Krüger M. (2006). Microbiological investigation of methane - and hydrocarbon-discharging mud volcanoes in the Carpathian Mountains, Romania. *Environmental Microbiology* **8:** 574-590.

Andersen G. L., He Z., DeSantis T. Z., Brodie E. L., Zhou J. (2010). The use of microarrays in microbial ecology. In: Liu W.-T., Jansson J. K. (eds). *Environmental Molecular Microbiology*. Horizon Scientific Press: Norfolk, UK. pp 87-109.

Anneser B., Einsiedl F., Meckenstock R. U., Richters L., Wisotzky F., Griebler C. (2008). High-resolution monitoring of biogeochemical gradients in a tar oil-contaminated aquifer. *Applied Geochemistry* **23:** 1715-1730.

Anneser B., Pilloni G., Bayer A., Lueders T., Griebler C., Einsiedl F., Richters L. (2010). High resolution analysis of contaminated aquifer sediments and groundwater—what can be learned in terms of natural attenuation? *Geomicrobiology Journal* **27:** 130 - 142.

Annweiler E., Materna A., Safinowski M., Kappler A., Richnow H. H., Michaelis W., Meckenstock R. U. (2000). Anaerobic degradation of 2-methylnaphthalene by a sulfate-reducing enrichment culture. *Applied and Environmental Microbiology* **66:** 5329-5333.

Beller H., Spormann A., Sharma P., Cole J., Reinhard M. (1996). Isolation and characterization of a novel toluene-degrading, sulfate- reducing bacterium. *Applied and Environmental Microbiology* **62:** 1188-1196.

Beller H. R., Spormann A. M. (1999). Substrate range of benzylsuccinate synthase from *Azoarcus* sp. strain T. *FEMS Microbiology Letters* **178:** 147-153.

Beller H. R., Kane S. R., Legler T. C., Alvarez P. J. (2002). A real-time polymerase chain reaction method for monitoring anaerobic, hydrocarbon-degrading bacteria based on a catabolic gene. *Environmental Science & Technology* **36:** 3977-3984.

Beller H. R., Kane S. R., Legler T. C., McKelvie J. R., Sherwood Lollar B., Pearson F., Balser L., Mackay D. M. (2008). Comparative assessments of benzene, toluene, and xylene natural attenuation by quantitative polymerase chain reaction analysis of a catabolic gene, signature metabolites, and compound-specific isotope analysis. *Environmental Science & Technology* **42:** 6065-6072.

Berdugo-Clavijo C., Dong X., Soh J., Sensen C. W., Gieg L. M. (2012). Methanogenic biodegradation of two-ringed polycyclic aromatic hydrocarbons. *FEMS Microbiology Ecology* **81:** 124-133.

Biegert T., Fuchs G., Heider J. (1996). Evidence that anaerobic oxidation of toluene in the denitrifying bacterium *Thauera aromatica* is initiated by formation of benzylsuccinate from toluene and fumarate. *European Journal of Biochemistry* **238:** 661-668.

Bockelmann A., Ptak T., Teutsch G. (2001). An analytical quantification of mass fluxes and natural attenuation rate constants at a former gasworks site. *Journal of Contaminant Hydrology* **53:** 429-453.

Bodrossy L., Stralis-Pavese N., Murrell J. C., Radajewski S., Weilharter A., Sessitsch A. (2003). Development and validation of a diagnostic microbial microarray for methanotrophs. *Environmental Microbiology* **5:** 566-582.

Bohrmann G., Spiess V., Böckel B., Boetius A., Boles M., Brüning M., Buhmann S., Cruz Melo C., Dalthorp-Moorhouse M., Dehning K., Ding F., Escobar-Briones E., Enneking K., Felden J., Fekete N., Freidank T., Gassner A., Gaytan A., Geersen J., Hinrichs K.-U., Hohnberg J., Kasten S., Keil H., Klar S., Klaucke I., Kuhlmann J., MacDonald I., Meinecke G., Mortera C., Naehr T., Nowald N., Ott C., Pacheco Muñoz J., Pelaez J. R., Ratmeyer V., Renken J., Reuter M., Sackmann V., Sahling H., Schubotz F., Schewe F., Stephan S., Thal J., Trampe A., Truscheit T., Viehweger M., T. Wilhelm T., Wegner G., Wenzhöfer F., Zabel M. (2008). Report and preliminary results of R/V Meteor cruise M67/2a and 2b, Balboa–Tampico–Bridgetown, 15 March–24 April 2006. Fluid seepage in the Gulf of Mexico. *Berichte aus dem Fachbereich Geowissenschaften, Universität Bremen* **263**.

Botton S., van Harmelen M., Braster M., Parsons J. R., Röling W. F. M. (2007). Dominance of *Geobacteraceae* in BTX-degrading enrichments from an iron-reducing aquifer. *FEMS Microbiology Ecology* **62:** 118-130.

Brodie E. L., DeSantis T. Z., Joyner D. C., Baek S. M., Larsen J. T., Andersen G. L., Hazen T. C., Richardson P. M., Herman D. J., Tokunaga T. K., Wan J. M., Firestone M. K. (2006). Application of a high-density oligonucleotide microarray approach to study bacterial population dynamics during uranium reduction and reoxidation. *Applied and Environmental Microbiology* **72:** 6288-6298.

Bruns A., Cypionka H., Overmann J. (2002). Cyclic AMP and acyl homoserine lactones increase the cultivation efficiency of heterotrophic bacteria from the central Baltic Sea. *Applied and Environmental Microbiology* **68:** 3978-3987.

Callaghan A. V., Gieg L. M., Kropp K. G., Suflita J. M., Young L. Y. (2006). Comparison of mechanisms of alkane metabolism under sulfate-reducing conditions among two bacterial isolates and a bacterial consortium. *Applied and Environmental Microbiology* **72:** 4274-4282.

Callaghan A. V., Wawrik B., Ni Chadhain S. M., Young L. Y., Zylstra G. J. (2008). Anaerobic alkane-degrading strain AK-01 contains two alkylsuccinate synthase genes. *Biochemical and Biophysical Research Communications* **366:** 142-148.

Callaghan A. V., Davidova I. A., Savage-Ashlock K., Parisi V. A., Gieg L. M., Suflita J. M., Kukor J. J., Wawrik B. (2010). Diversity of benzyl- and alkylsuccinate synthase genes in hydrocarbon-impacted environments and enrichment cultures. *Environmental Science & Technology* **44:** 7287-7294.

Callaghan A. V. (2012). Metabolomic investigations of anaerobic hydrocarbon-impacted environments. *Current Opinion in Biotechnology* in press, 10.1016/j.copbio.2012.08.012.

Carmona M., Zamarro M. T., Blazquez B., Durante-Rodriguez G., Juarez J. F., Valderrama J. A., Barragan M. J., Garcia J. L., Diaz E. (2009). Anaerobic catabolism of aromatic compounds: A genetic and genomic view. *Microbiology and Molecular Biology Reviews* **73:** 71-133.

Chivian D., Brodie E. L., Alm E. J., Culley D. E., Dehal P. S., DeSantis T. Z., Gihring T. M., Lapidus A., Lin L.-H., Lowry S. R., Moser D. P., Richardson P. M., Southam G., Wanger G., Pratt L. M., Andersen G. L., Hazen T. C., Brockman F. J., Arkin A. P., Onstott T. C. (2008). Environmental genomics reveals a single-species ecosystem deep within earth. *Science* **322:** 275-278.

Cole J. R., Wang Q., Cardenas E., Fish J., Chai B., Farris R. J., Kulam-Syed-Mohideen A. S., McGarrell D. M., Marsh T., Garrity G. M., Tiedje J. M. (2009). The Ribosomal Database Project: Improved alignments and new tools for rRNA analysis. *Nucleic Acids Research* **37:** D141-D145.

Connell J. H. (1978). Diversity in tropical rain forests and coral reefs. *Science* **199:** 1302-1310.

Cravo-Laureau C., Matheron R., Cayol J.-L., Joulian C., Hirschler-Réa A. (2004a). *Desulfatibacillum aliphaticivorans* gen. nov., sp. nov., an *n*-alkane- and *n*-alkene-degrading, sulfate-reducing bacterium. *International Journal of Systematic and Evolutionary Microbiology* **54:** 77-83.

Cravo-Laureau C., Matheron R., Joulian C., Cayol J.-L., Hirschler-Rea A. (2004b). *Desulfatibacillum alkenivorans* sp. nov., a novel *n*-alkene-degrading, sulfate-reducing bacterium, and emended description of the genus *Desulfatibacillum*. *International Journal of Systematic and Evolutionary Microbiology* **54:** 1639-1642.

Culman S., Bukowski R., Gauch H., Cadillo-Quiroz H., Buckley D. (2009). T-REX: software for the processing and analysis of T-RFLP data. *BMC Bioinformatics* **10:** 171.

Davidova I. A., Duncan K. E., Choi O. K., Suflita J. M. (2006). *Desulfoglaeba alkanexedens* gen. nov., sp. nov., an *n*-alkane-degrading, sulfate-reducing bacterium. *International Journal of Systematic and Evolutionary Microbiology* **56:** 2737-2742.

Davidova I. A., Gieg L. M., Duncan K. E., Suflita J. M. (2007). Anaerobic phenanthrene mineralization by a carboxylating sulfate-reducing bacterial enrichment. *The ISME Journal* **1:** 436-442.

Degnan P. H., Ochman H. (2012). Illumina-based analysis of microbial community diversity. *The ISME Journal* **6:** 183-194.

DeSantis T. Z., Hugenholtz P., Larsen N., Rojas M., Brodie E. L., Keller K., Huber T., Dalevi D., Hu P., Andersen G. L. (2006). Greengenes, a chimera-checked 16S rRNA gene database and workbench compatible with ARB. *Applied and Environmental Microbiology* **72:** 5069-5072.

Detmers J., Strauss H., Schulte U., Bergmann A., Knittel K., Kuever J. (2004). FISH shows that *Desulfotomaculum* spp. are the dominating sulfate-reducing bacteria in a pristine aquifer. *Microbial Ecology* **47:** 236-242.

DiDonato R. J., Jr., Young N. D., Butler J. E., Chin K.-J., Hixson K. K., Mouser P., Lipton M. S., DeBoy R., Methé B. A. (2010). Genome sequence of the deltaproteobacterial strain NaphS2 and analysis of differential gene expression during anaerobic growth on naphthalene. *PLoS ONE* **5:** e14072.

Didyk B. M., Simoneit B. R. T. (1989). Hydrothermal oil of Guaymas Basin and implications for petroleum formation mechanisms. *Nature* **342:** 65-69.

Duldhardt I., Gaebel J., Chrzanowski L., Nijenhuis I., Härtig C., Schauer F., Heipieper H. J. (2010). Adaptation of anaerobically grown *Thauera aromatica*, *Geobacter sulfurreducens* and *Desulfococcus multivorans* to organic solvents on the level of membrane fatty acid composition. *Microbial Biotechnology* **3:** 201-209.

Ehrenreich P., Behrends A., Harder J., Widdel F. (2000). Anaerobic oxidation of alkanes by newly isolated denitrifying bacteria. *Archives of Microbiology* **173:** 58-64.

Elshahed M. S., Gieg L. M., McInerney M. J., Suflita J. M. (2001). Signature metabolites attesting to the *in situ* attenuation of alkylbenzenes in anaerobic environments. *Environmental Science & Technology* **35:** 682-689.

Ettwig K. F., Speth D. R., Reimann J., Wu M. L., Jetten M. S. M., Keltjens J. T. (2012). Bacterial oxygen production in the dark. *Frontiers in Microbiology* **3**.

Feisthauer S., Siegert M., Seidel M., Richnow H. H., Zengler K., Gründger F., Krüger M. (2010). Isotopic fingerprinting of methane and CO_2 formation from aliphatic and aromatic hydrocarbons. *Organic Geochemistry* **41:** 482-490.

Fetzner S. (2010). Aerobic degradation of halogenated aliphatics. In: Timmis K. N. (ed). *Handbook of Hydrocarbon and Lipid Microbiology*. Springer: Berlin Heidelberg. pp 865-885.

Formolo M. (2010). The microbial production of methane and other volatile hydrocarbons. In: Timmis K. N. (ed). *Handbook of Hydrocarbon and Lipid Microbiology*. Springer: Berlin Heidelberg. pp 113-126.

Fowler S. J., Dong X., Sensen C. W., Suflita J. M., Gieg L. M. (2012). Methanogenic toluene metabolism: community structure and intermediates. *Environmental Microbiology* **14:** 754-764.

Friedrich M. W. (2005). Methyl-Coenzyme M Reductase Genes: Unique Functional Markers for Methanogenic and Anaerobic Methane-Oxidizing Archaea. In: Jared R. L. (ed). *Methods in Enzymology*. Academic Press. pp 428-442.

Fuchs G., Boll M., Heider J. (2011). Microbial degradation of aromatic compounds — from one strategy to four. *Nature Reviews Microbiology* **9:** 803-816.

Galvão T. C., Mohn W. W., de Lorenzo V. (2005). Exploring the microbial biodegradation and biotransformation gene pool. *Trends in Biotechnology* **23:** 497-506.

Gertler C., Yakimov M. M., Malpass M. C., Golyshin P. N. (2010). Shipping-related accidental and deliberate release into the environment. In: Timmis K. N. (ed). *Handbook of Hydrocarbon and Lipid Microbiology*. Springer: Berlin Heidelberg. pp 243-256.

Glenn T. C. (2011). Field guide to next-generation DNA sequencers. *Molecular Ecology Resources* **11:** 759-769.

Griebler C., Safinowski M., Vieth A., Richnow H. H., Meckenstock R. U. (2004). Combined application of stable carbon isotope analysis and specific metabolites determination for assessing *in situ* degradation of aromatic hydrocarbons in a tar oil-contaminated aquifer. *Environmental Science & Technology* **38:** 617-631.

Griebler C., Lueders T. (2008). Microbial biodiversity in groundwater ecosystems. *Freshwater Biology* **54:** 649-677.

Hall T. A. (1999). BioEdit: a user-friendly biological sequence alignment editor and analysis program for Windows 95/98/NT. *Nucleic Acids Symposium Series* **41:** 95-98.

Hazen T. C., Dubinsky E. A., DeSantis T. Z., Andersen G. L., Piceno Y. M., Singh N., Jansson J. K., Probst A., Borglin S. E., Fortney J. L., Stringfellow W. T., Bill M., Conrad M. E., Tom L. M., Chavarria K. L., Alusi T. R., Lamendella R., Joyner D. C., Spier C., Baelum J., Auer M., Zemla M. L., Chakraborty R., Sonnenthal E. L., D'haeseleer P., Holman H.-Y. N., Osman S., Lu Z., Van Nostrand J. D., Deng Y., Zhou J., Mason O. U. (2010). Deep-sea oil plume enriches indigenous oil-degrading bacteria. *Science* **330:** 204-208.

He Z., Gentry T. J., Schadt C. W., Wu L., Liebich J., Chong S. C., Huang Z., Wu W., Gu B., Jardine P., Criddle C., Zhou J. (2007). GeoChip: A comprehensive microarray for investigating biogeochemical, ecological and environmental processes. *The ISME Journal* **1:** 67-77.

Heider J. (2007). Adding handles to unhandy substrates: Anaerobic hydrocarbon activation mechanisms. *Current Opinion in Chemical Biology* **11:** 188-194.

Herfort M., Ptak T., Hümmer O., Teutsch G., Dahmke A. (1998). Testfeld Süd: Einrichtung der Testfeldinfrastruktur und Erkundung hydraulisch-hydrogeochemischer Parameter des Grundwasserleiters. *Grundwasser* **3:** 159-166.

Herrmann S., Vogt C., Fischer A., Kuppardt A., Richnow H.-H. (2009). Characterization of anaerobic xylene biodegradation by two-dimensional isotope fractionation analysis. *Environmental Microbiology Reports* **1:** 535-544.

Horner-Devine M. C., Carney K. M., Bohannan B. J. (2004). An ecological perspective on bacterial biodiversity. *Proceedings of the Royal Society B: Biological Sciences* **271:** 113-122.

Hosoda A., Kasai Y., Hamamura N., Takahata Y., Watanabe K. (2005). Development of a PCR method for the detection and quantification of benzoyl-CoA reductase genes and its application to monitored natural attenuation. *Biodegradation* **16:** 591-601.

Ihaka R., Gentleman R. (1996). R: a language for data analysis and graphics. *Journal of Computational and Graphical Statistics* **5:** 299-314.

Jobelius C., Ruth B., Griebler C., Meckenstock R. U., Hollender J., Reineke A., Frimmel F. H., Zwiener C. (2010). Metabolites indicate hot spots of biodegradation and biogeochemical gradients in a high-resolution monitoring well. *Environmental Science & Technology* **45:** 474-481.

Johnson H. A., Pelletier D. A., Spormann A. M. (2001). Isolation and characterization of anaerobic ethylbenzene dehydrogenase, a novel Mo-Fe-S enzyme. *Journal of Bacteriology* **183:** 4536-4542.

Kane S. R., Beller H. R., Legler T. C., Anderson R. T. (2002). Biochemical and genetic evidence of benzylsuccinate synthase in toluene-degrading, ferric iron-reducing *Geobacter metallireducens*. *Biodegradation* **13:** 149-154.

Kazy S., Monier A., Alvarez P. (2010). Assessing the correlation between anaerobic toluene degradation activity and *bssA* concentrations in hydrocarbon-contaminated aquifer material. *Biodegradation* **21:** 793-800.

Kleindienst S. (2012). Hydrocarbon-degrading sulfate-reducing bacteria in marine hydrocarbon seep sediments. PhD thesis, Universität Bremen, Bremen.

Kleindienst S., Ramette A., Amann R., Knittel K. (2012). Distribution and *in situ* abundance of sulfate-reducing bacteria in diverse marine hydrocarbon seep sediments. *Environmental Microbiology* **14:** 2689-2710.

Knaebel D. B., Crawford R. L. (1995). Extraction and purification of microbial DNA from petroleum-contaminated soils and detection of low numbers of toluene, octane and pesticide degraders by multiplex polymerase chain reaction and Southern analysis. *Molecular Ecology* **4:** 579-592.

Kniemeyer O., Heider J. (2001). Ethylbenzene dehydrogenase, a novel hydrocarbon-oxidizing molybdenum/iron-sulfur/heme enzyme. *Journal of Biological Chemistry* **276:** 21381-21386.

Kniemeyer O., Musat F., Sievert S. M., Knittel K., Wilkes H., Blumenberg M., Michaelis W., Classen A., Bolm C., Joye S. B., Widdel F. (2007). Anaerobic oxidation of short-chain hydrocarbons by marine sulphate-reducing bacteria. *Nature* **449:** 898-901.

Knittel K., Boetius A. (2010). Anaerobic methane oxidizers. In: Timmis K. N. (ed). *Handbook of Hydrocarbon and Lipid Microbiology*. Springer: Berlin Heidelberg. pp 2023-2032.

Kolukirik M., Ince O., Ince B. (2011). Increment in anaerobic hydrocarbon degradation activity of Halic Bay sediments via nutrient amendment. *Microbial Ecology* **61:** 871-884.

Kropp K. G., Davidova I. A., Suflita J. M. (2000). Anaerobic oxidation of *n*-dodecane by an addition reaction in a sulfate-reducing bacterial enrichment culture. *Applied and Environmental Microbiology* **66:** 5393-5398.

Kube M., Heider J., Amann J., Hufnagel P., Kuhner S., Beck A., Reinhardt R., Rabus R. (2004). Genes involved in the anaerobic degradation of toluene in a denitrifying bacterium, strain EbN1. *Archives of Microbiology* **181:** 182-194.

Kunapuli U., Lueders T., Meckenstock R. U. (2007). The use of stable isotope probing to identify key iron-reducing microorganisms involved in anaerobic benzene degradation. *The ISME Journal* **1:** 643-653.

Kunapuli U., Jahn M. K., Lueders T., Geyer R., Heipieper H. J., Meckenstock R. U. (2010). *Desulfitobacterium aromaticivorans* sp. nov. and *Geobacter toluenoxydans* sp. nov., iron-reducing bacteria capable of anaerobic degradation of monoaromatic hydrocarbons. *International Journal of Systematic and Evolutionary Microbiology* **60:** 686-695.

Kuntze K., Shinoda Y., Moutakki H., McInerney M. J., Vogt C., Richnow H.-H., Boll M. (2008). 6-Oxocyclohex-1-ene-1-carbonyl-coenzyme A hydrolases from obligately anaerobic bacteria: characterization and identification of its gene as a functional marker for aromatic compounds degrading anaerobes. *Environmental Microbiology* **10:** 1547-1556.

Kuntze K., Vogt C., Richnow H.-H., Boll M. (2011). Combined application of PCR-based functional assays for the detection of aromatic-compound-degrading anaerobes. *Applied and Environmental Microbiology* **77:** 5056-5061.

Lara-Martín P. A., Gómez-Parra A., Sanz J. L., Gonzáez-Mazo E. (2010). Anaerobic degradation pathway of linear alkylbenzene sulfonates (LAS) in sulfate-reducing marine sediments. *Environmental Science & Technology* **44:** 1670-1676.

Lee Y.-J., Romanek C. S., Wiegel J. (2009). *Desulfosporosinus youngiae* sp. nov., a spore-forming, sulfate-reducing bacterium isolated from a constructed wetland treating acid mine drainage. *International Journal of Systematic and Evolutionary Microbiology* **59:** 2743-2746.

Lehtiö L., Goldman A. (2004). The pyruvate formate lyase family: Sequences, structures and activation. *Protein Engineering, Design & Selection* **17:** 545-552.

Leuthner B., Leutwein C., Schulz H., Horth P., Haehnel W., Schiltz E., Schagger H., Heider J. (1998). Biochemical and genetic characterization of benzylsuccinate synthase from *Thauera aromatica*: A new glycyl radical enzyme catalysing the first step in anaerobic toluene metabolism. *Molecular Microbiology* **28:** 615-628.

Lin X., McKinley J., Resch C. T., Kaluzny R., Lauber C. L., Fredrickson J., Knight R., Konopka A. (2012). Spatial and temporal dynamics of the microbial community in the Hanford unconfined aquifer. *The ISME Journal* **6:** 1665-1676.

Liu A., Garcia-Dominguez E., Rhine E. D., Young L. Y. (2004). A novel arsenate respiring isolate that can utilize aromatic substrates. *FEMS Microbiology Ecology* **48:** 323-332.

Löffler C., Kuntze K., Vazquez J. R., Rugor A., Kung J. W., Böttcher A., Boll M. (2011). Occurrence, genes and expression of the W/Se-containing class II benzoyl-coenzyme A reductases in anaerobic bacteria. *Environmental Microbiology* **13:** 696-709.

Loman N. J., Misra R. V., Dallman T. J., Constantinidou C., Gharbia S. E., Wain J., Pallen M. J. (2012). Performance comparison of benchtop high-throughput sequencing platforms. *Nature Biotechnology* **30:** 434-439.

Lovley D. R., Phillips E. J. P. (1986). Organic matter mineralization with reduction of ferric iron in anaerobic sediments. *Applied and Environmental Microbiology* **51:** 683-689.

Lovley D. R., Giovannoni S. J., White D. C., Champine J. E., Phillips E. J. P., Gorby Y. A., Goodwin S. (1993). *Geobacter metallireducens* gen. nov. sp. nov., a microorganism capable of coupling the complete oxidation of organic compounds to the reduction of iron and other metals. *Archives of Microbiology* **159:** 336-344.

Ludwig W., Strunk O., Westram R., Richter L., Meier H., Yadhukumar, Buchner A., Lai T., Steppi S., Jobb G., Forster W., Brettske I., Gerber S., Ginhart A. W., Gross O., Grumann S., Hermann S., Jost R., Konig A., Liss T., Lussmann R., May M., Nonhoff B., Reichel B., Strehlow R., Stamatakis A., Stuckmann N., Vilbig A., Lenke M., Ludwig T., Bode A., Schleifer K. H. (2004). ARB: A software environment for sequence data. *Nucleic Acids Research* **32:** 1363-1371.

Lueders T., Friedrich M. W. (2003). Evaluation of PCR amplification bias by terminal restriction fragment length polymorphism analysis of small-subunit rRNA and *mcrA* genes by using defined template mixtures of methanogenic pure cultures and soil DNA extracts. *Applied and Environmental Microbiology* **69:** 320-326.

Lueders T., Manefield M., Friedrich M. W. (2004). Enhanced sensitivity of DNA- and rRNA-based stable isotope probing by fractionation and quantitative analysis of isopycnic centrifugation gradients. *Environmental Microbiology* **6:** 73-78.

Lueders T., Kindler R., Miltner A., Friedrich M. W., Kaestner M. (2006). Identification of bacterial micropredators distinctively active in a soil microbial food web. *Applied and Environmental Microbiology* **72:** 5342-5348.

Lueders T. (2010). Stable isotope probing of hydrocarbon-degraders. In: Timmis K. N. (ed). *Handbook of Hydrocarbon and Lipid Microbiology*. Springer: Berlin Heidelberg. pp 4011-4026.

Lüke C., Frenzel P. (2011). Potential of *pmoA* amplicon pyrosequencing for methanotroph diversity studies. *Applied and Environmental Microbiology* **77:** 6305-6309.

Mastalerz V., de Lange G. J., Dählmann A., Feseker T. (2007). Active venting at the Isis mud volcano, offshore Egypt: Origin and migration of hydrocarbons. *Chemical Geology* **246:** 87-106.

Mastalerz V., de Lange G. J., Dählmann A. (2009). Differential aerobic and anaerobic oxidation of hydrocarbon gases discharged at mud volcanoes in the Nile deep-sea fan. *Geochimica et Cosmochimica Acta* **73:** 3849-3863.

Mbadinga S. M., Wang L.-Y., Zhou L., Liu J.-F., Gu J.-D., Mu B.-Z. (2011). Microbial communities involved in anaerobic degradation of alkanes. *International Biodeterioration and Biodegradation* **65:** 1-13.

McDonald I. R., Murrell J. C. (1997). The particulate methane monooxygenase gene *pmoA* and its use as a functional gene probe for methanotrophs. *FEMS Microbiology Letters* **156:** 205-210.

McMahon K. D., Martin H. G., Hugenholtz P. (2007). Integrating ecology into biotechnology. *Current Opinion in Biotechnology* **18:** 287-292.

Meckenstock R. U. (1999). Fermentative toluene degradation in anaerobic defined syntrophic cocultures. *FEMS Microbiology Letters* **177:** 67-73.

Meckenstock R. U., Morasch B., Griebler C., Richnow H. H. (2004). Stable isotope fractionation analysis as a tool to monitor biodegradation in contaminated acquifers. *Journal of Contaminant Hydrology* **75:** 215-255.

Meckenstock R. U., Lueders T., Griebler C., Selesi D. (2010). Microbial hydrocarbon degradation at coal gasification plants. In: Timmis K. N. (ed). *Handbook of Hydrocarbon and Lipid Microbiology*. Springer: Berlin Heidelberg. pp 2293-2312.

Meckenstock R. U., Mouttaki H. (2011). Anaerobic degradation of non-substituted aromatic hydrocarbons. *Current Opinion in Biotechnology* **22:** 406-414.

Metzker M. L. (2010). Sequencing technologies - the next generation. *Nature Reviews Genetics* **11:** 31-46.

Morasch B., Annweiler E., Warthmann R. J., Meckenstock R. U. (2001). The use of a solid adsorber resin for enrichment of bacteria with toxic substrates and to identify metabolites: degradation of naphthalene, o-, and m-xylene by sulfate-reducing bacteria. *Journal of Microbiological Methods* **44:** 183-191.

Morasch B., Schink B., Tebbe C. C., Meckenstock R. U. (2004). Degradation of *o*-xylene and *m*-xylene by a novel sulfate-reducer belonging to the genus *Desulfotomaculum*. *Archives of Microbiology* **181:** 407-417.

Mouttaki H., Johannes J., Meckenstock R. U. (2012). Identification of naphthalene carboxylase as a prototype for the anaerobic activation of non-substituted aromatic hydrocarbons. *Environmental Microbiology* **14:** 2770-2774.

Munoz R., Yarza P., Ludwig W., Euzéby J., Amann R., Schleifer K.-H., Oliver Glöckner F., Rosselló-Móra R. (2011). Release LTPs104 of the all-species living tree. *Systematic and Applied Microbiology* **34:** 169-170.

Murphy C. D. (2010). Halogenated organic compounds – Carbon-halogen bond formation. In: Timmis K. N. (ed). *Handbook of Hydrocarbon and Lipid Microbiology*. Springer: Berlin Heidelberg. pp 375-383.

Murrell J. C. (2010). The aerobic methane oxidizing bacteria (methanotrophs). In: Timmis K. N. (ed). *Handbook of Hydrocarbon and Lipid Microbiology*. Springer: Berlin Heidelberg. pp 1953-1966.

Musat F., Galushko A., Jacob J., Widdel F., Kube M., Reinhardt R., Wilkes H., Schink B., Rabus R. (2009). Anaerobic degradation of naphthalene and 2-methylnaphthalene by strains of marine sulfate-reducing bacteria. *Environmental Microbiology* **11:** 209-219.

Musat F., Wilkes H., Behrends A., Woebken D., Widdel F. (2010). Microbial nitrate-dependent cyclohexane degradation coupled with anaerobic ammonium oxidation. *The ISME Journal* **4:** 1290-1301.

Muyzer G., Teske A., Wirsen C., Jannasch H. (1995). Phylogenetic relationships of *Thiomicrospira* species and their identification in deep-sea hydrothermal vent samples by denaturing gradient gel electrophoresis of 16S rDNA fragments. *Archives of Microbiology* **164:** 165-172.

Novais R. C., Thorstenson Y. R. (2011). The evolution of Pyrosequencing® for microbiology: From genes to genomes. *Journal of Microbiological Methods* **86:** 1-7.

Oka A. R., Phelps C. D., Zhu X., Saber D. L., Young L. Y. (2011). Dual biomarkers of anaerobic hydrocarbon degradation in historically contaminated groundwater. *Environmental Science & Technology* **45:** 3407-3414.

Oksanen J., Blanchet F. G., Kindt R., Legendre P., Minchin P. R., O'Hara R. B., Simpson G. L., Solymos P., Stevens M. H. H., Wagner H. (2012). vegan: Community Ecology Package. R package version 2.0-3. http://CRAN.R-project.org/package=vegan

Ommedal H., Torsvik T. (2007). *Desulfotignum toluenicum* sp. nov., a novel toluene-degrading, sulphate-reducing bacterium isolated from an oil-reservoir model column. *International Journal of Systematic and Evolutionary Microbiology* **57:** 2865-2869.

Orcutt B. N., Joye S. B., Kleindienst S., Knittel K., Ramette A., Reitz A., Samarkin V., Treude T., Boetius A. (2010). Impact of natural oil and higher hydrocarbons on microbial diversity, distribution, and activity in Gulf of Mexico cold-seep sediments. *Deep Sea Research Part II: Topical Studies in Oceanography* **57:** 2008-2021.

Pei A. Y., Oberdorf W. E., Nossa C. W., Agarwal A., Chokshi P., Gerz E. A., Jin Z., Lee P., Yang L., Poles M., Brown S. M., Sotero S., DeSantis T., Brodie E., Nelson K., Pei Z. (2010). Diversity of 16S rRNA genes within individual prokaryotic genomes. *Applied and Environmental Microbiology* **76:** 3886-3897.

Pieper D. H., González B., Cámara B., Pérez-Pantoja D., Reineke W. (2010). Aerobic degradation of chloroaromatics. In: Timmis K. N. (ed). *Handbook of Hydrocarbon and Lipid Microbiology*. Springer: Berlin Heidelberg. pp 839-864.

Pilloni G. (2011). Distribution and dynamics of contaminant degraders and microbial communities in stationary and nonstationary contaminant plumes. PhD thesis, Technische Universität München.

Pilloni G., von Netzer F., Engel M., Lueders T. (2011). Electron acceptor-dependent identification of key anaerobic toluene degraders at a tar-oil-contaminated aquifer by Pyro-SIP. *FEMS Microbiology Ecology* **78:** 165-175.

Pilloni G., Granitsiotis M. S., Engel M., Lueders T. (2012). Testing the limits of 454 pyrotag sequencing: Reproducibility, quantitative assessment and comparison to T-RFLP fingerprinting of aquifer microbes. *PLoS ONE* **7:** e40467.

Prosser J. I., Bohannan B. J., Curtis T. P., Ellis R. J., Firestone M. K., Freckleton R. P., Green J. L., Green L. E., Killham K., Lennon J. J., Osborn A. M., Solan M., van der Gast C. J., Young J. P. (2007). The role of ecological theory in microbial ecology. *Nature Reviews Microbiology* **5:** 384-392.

Rabus R., Nordhaus R., Ludwig W., Widdel F. (1993). Complete oxidation of toluene under strictly anoxic conditions by a new sulfate-reducing bacterium. *Applied and Environmental Microbiology* **59:** 1444-1451.

Rabus R., Widdel F. (1995). Anaerobic degradation of ethylbenzene and other aromatic hydrocarbons by new denitrifying bacteria. *Archives of Microbiology* **163:** 96-103.

Rabus R., Wilkes H., Behrends A., Armstroff A., Fischer T., Pierik A. J., Widdel F. (2001). Anaerobic initial reaction of *n*-alkanes in a denitrifying bacterium: Evidence for (1-methylpentyl)succinate as initial product and for involvement of an organic radical in *n*-hexane metabolism. *Journal of Bacteriology* **183:** 1707-1715.

Rabus R., Jarling R., Lahme S., Kühner S., Heider J., Widdel F., Wilkes H. (2011). Co-metabolic conversion of toluene in anaerobic *n*-alkane-degrading bacteria. *Environmental Microbiology* **13:** 2576-2586.

Rojo F. (2010). Enzymes for aerobic degradation of alkanes. In: Timmis K. N. (ed). *Handbook of Hydrocarbon and Lipid Microbiology*. Springer: Berlin Heidelberg. pp 781-797.

Röling W. F., van Verseveld H. W. (2002). Natural attenuation: What does the subsurface have in store? *Biodegradation* **13:** 53-64.

Rothberg J. M., Hinz W., Rearick T. M., Schultz J., Mileski W., Davey M., Leamon J. H., Johnson K., Milgrew M. J., Edwards M., Hoon J., Simons J. F., Marran D., Myers J. W., Davidson J. F., Branting A., Nobile J. R., Puc B. P., Light D., Clark T. A., Huber M., Branciforte J. T., Stoner I. B., Cawley S. E., Lyons M., Fu Y., Homer N., Sedova M., Miao X., Reed B., Sabina J., Feierstein E., Schorn M., Alanjary M., Dimalanta E., Dressman D., Kasinskas R., Sokolsky

T., Fidanza J. A., Namsaraev E., McKernan K. J., Williams A., Roth G. T., Bustillo J. (2011). An integrated semiconductor device enabling non-optical genome sequencing. *Nature* **475:** 348-352.

Safinowski M., Meckenstock R. U. (2004). Enzymatic reactions in anaerobic 2-methylnaphthalene degradation by the sulphate-reducing enrichment culture N 47. *FEMS Microbiology Letters* **240:** 99-104.

Safinowski M., Meckenstock R. U. (2006). Methylation is the initial reaction in anaerobic naphthalene degradation by a sulfate-reducing enrichment culture. *Environmental Microbiology* **8:** 347-352.

Sanger F., Nicklen S., Coulson A. R. (1977). DNA sequencing with chain-terminating inhibitors. *Proceedings of the National Academy of Sciences* **74:** 5463-5467.

Schloss P. D., Westcott S. L., Ryabin T., Hall J. R., Hartmann M., Hollister E. B., Lesniewski R. A., Oakley B. B., Parks D. H., Robinson C. J., Sahl J. W., Stres B., Thallinger G. G., Van Horn D. J., Weber C. F. (2009). Introducing mothur: Open-source, platform-independent, community-supported software for describing and comparing microbial communities. *Applied and Environmental Microbiology* **75:** 7537-7541.

Selesi D., Jehmlich N., von Bergen M., Schmidt F., Rattei T., Tischler P., Lueders T., Meckenstock R. U. (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:** 295-306.

Settanni L., Corsetti A. (2007). The use of multiplex PCR to detect and differentiate food- and beverage-associated microorganisms: A review. *Journal of Microbiological Methods* **69:** 1-22.

Shima S., Thauer R. K. (2005). Methyl-coenzyme M reductase and the anaerobic oxidation of methane in methanotrophic *Archaea*. *Current Opinion in Microbiology* **8:** 643-648.

Shinoda Y., Sakai Y., Uenishi H., Uchihashi Y., Hiraishi A., Yukawa H., Yurimoto H., Kato N. (2004). Aerobic and anaerobic toluene degradation by a newly isolated denitrifying bacterium, *Thauera* sp. strain DNT-1. *Applied and Environmental Microbiology* **70:** 1385-1392.

Shinoda Y., Akagi J., Uchihashi Y., Hiraishi A., Yukawa H., Yurimoto H., Sakai Y., Kato N. (2005). Anaerobic degradation of aromatic compounds by *Magnetospirillum* strains: Isolation and degradation genes. *Bioscience, Biotechnology, and Biochemistry* **69:** 1483-1491.

Sikkema J., de Bont J. A., Poolman B. (1995). Mechanisms of membrane toxicity of hydrocarbons. *Microbiolological Reviews* **59:** 201-222.

So C. M., Young L. Y. (1999). Isolation and characterization of a sulfate-reducing bacterium that anaerobically degrades alkanes. *Applied and Environmental Microbiology* **65:** 2969-2976.

Song B., Ward B. B. (2005). Genetic diversity of benzoyl coenzyme A reductase genes detected in denitrifying isolates and estuarine sediment communities. *Applied and Environmental Microbiology* **71:** 2036-2045.

Staats M., Braster M., Röling W. F. M. (2011). Molecular diversity and distribution of aromatic hydrocarbon-degrading anaerobes across a landfill leachate plume. *Environmental Microbiology* **13:** 1216-1227.

Sullivan E. R., Zhang X., Phelps C., Young L. Y. (2001). Anaerobic mineralization of stable-isotope-labeled 2-methylnaphthalene. *Applied and Environmental Microbiology* **67:** 4353-4357.

Taylor P. M., Medd J. M., Schoenborn L., Hodgson B., Janssen P. H. (2002). Detection of known and novel genes encoding aromatic ring-hydroxylating dioxygenases in soils and in aromatic hydrocarbon-degrading bacteria. *FEMS Microbiology Letters* **216:** 61-66.

van der Meer J. R. (2006). Environmental pollution promotes selection of microbial degradation pathways. *Frontiers in Ecology and the Environment* **4:** 35-42.

Van Gaever S., Raes M., Pasotti F., Vanreusel A. (2010). Spatial scale and habitat-dependent diversity patterns in nematode communities in three seepage related sites along the Norwegian Sea margin. *Marine Ecology* **31:** 66-77.

Vanreusel A., Andersen A. C., Boetius A., Connelly D., Cunha M. R., Decker C., Hilario A., Kormas K. A., L. Maignien L., Olu K., Pachiadaki M., Ritt B., Rodrigues C., Sarrazin J., Tyler T., Van Gaever S., Vanneste H. (2009). Biodiversity of cold seep ecosystems along the European margins. *Oceanography* **22:** 110-127.

Vilchez-Vargas R., Geffers R., Suárez-Diez M., Conte I., Waliczek A., Kaser V. S., Kralova M., Junca H., Pieper D. H. (2013). Analysis of the microbial gene landscape and transcriptome for aromatic pollutants and alkane degradation using a novel internally calibrated microarray system. *Environmental Microbiology* **15:** 1016-1039.

Volkmann H., Schwartz T., Kirchen S., Stofer C., Obst U. (2007). Evaluation of inhibition and cross-reaction effects on real-time PCR applied to the total DNA of wastewater samples for the quantification of bacterial antibiotic resistance genes and taxon-specific targets. *Molecular and Cellular Probes* **21:** 125-133.

von Netzer F., Pilloni G., Kleindienst S., Krüger M., Knittel K., Gründger F., Lueders T. (2013). Enhanced gene detection assays for fumarate-adding enzymes allow uncovering anaerobic hydrocarbon degraders in terrestrial and marine systems. *Applied and Environmental Microbiology* **79:** 543-552.

Wackett L. P., Hershberger C. D. (2001). *Biocatalysis and biodegradation - Microbial transformation of organic compounds*. ASM Press, Washington.

Wagner M., Smidt H., Loy A., Zhou J. (2007). Unravelling microbial communities with DNA-microarrays: Challenges and future directions. *Microbial Ecology* **53:** 498-506.

Washer C. E., Edwards E. A. (2007). Identification and expression of benzylsuccinate synthase genes in a toluene-degrading methanogenic consortium. *Applied and Environmental Microbiology* **73:** 1367-1369.

Weelink S. A. B., Doesburg W. v., Saia F. T., Rijpstra W. I. C., Röling W. F. M., Smidt H., Stams A. J. M. (2009). A strictly anaerobic betaproteobacterium *Georgfuchsia toluolica* gen. nov., sp. nov. degrades aromatic compounds with Fe(III), Mn(IV) or nitrate as an electron acceptor. *FEMS Microbiology Ecology* **70:** 575-585.

Weisburg W. G., Barns S. M., Pelletier D. A., Lane D. J. (1991). 16S ribosomal DNA amplification for phylogenetic study. *Journal of Bacteriology* **173:** 697-703.

Weiss J. V., Cozzarelli I. M. (2008). Biodegradation in contaminated aquifers: Incorporating microbial/molecular methods. *Ground Water* **46:** 305-322.

Widdel F., Bak F. (1992). Gram-negative mesophilic sulfate-reducing bacteria. In: Balows A., Trüper H. G., Dworkin M., Harder W., Schleifer K. H. (eds). *The Prokaryotes*, 2. edn. Springer Verlag: Berlin, Heidelberg, New York. pp 3352-3378.

Widdel F., Knittel K., Galushko A. (2010). Anaerobic hydrocarbon-degrading microorganisms: An overview. In: Timmis K. N. (ed). *Handbook of Hydrocarbon and Lipid Microbiology*. Springer: Berlin Heidelberg. pp 1997-2021.

Wilkes H., Schwarzbauer J. (2010). Hydrocarbons: An introduction to structure, physico-chemical properties and natural occurrence. In: Timmis K. N. (ed). *Handbook of Hydrocarbon and Lipid Microbiology*. Springer: Berlin Heidelberg. pp 1-48.

Wilson M. S., Bakermans C., Madsen E. L. (1999). *In situ*, real-time catabolic gene expression: extraction and characterization of naphthalene dioxygenase mRNA transcripts from groundwater. *Applied and Environmental Microbiology* **65:** 80-87.

Winderl C. (2007). Functional and phylogenetic diversity of anaerobic BTEX-degrading microorganisms in contaminated aquifers. PhD thesis, Eberhard-Karls-Universität, Tübingen.

Winderl C., Schaefer S., Lueders T. (2007). Detection of anaerobic toluene and hydrocarbon degraders in contaminated aquifers using benzylsuccinate synthase (*bssA*) genes as a functional marker. *Environmental Microbiology* **9:** 1035-1046.

Winderl C., Anneser B., Griebler C., Meckenstock R. U., Lueders T. (2008). Depth-resolved quantification of anaerobic toluene degraders and aquifer microbial community patterns in distinct redox zones of a tar oil contaminant plume. *Applied and Environmental Microbiology* **74:** 792-801.

Winderl C., Penning H., von Netzer F., Meckenstock R. U., Lueders T. (2010). DNA-SIP identifies sulfate-reducing *Clostridia* as important toluene degraders in tar-oil-contaminated aquifer sediment. *The ISME Journal* **4:** 1314-1325.

Wöhlbrand L., Kallerhoff B., Lange D., Hufnagel P., Thiermann J., Reinhardt R., Rabus R. (2007). Functional proteomic view of metabolic regulation in "*Aromatoleum aromaticum*" strain EbN1. *Proteomics* **7:** 2222-2239.

Xiong X., Yin X., Pei X., Jin P., Zhang A., Li Y., Gong W., Wang Q. (2012). Retrieval of glycoside hydrolase family 9 cellulase genes from environmental DNA by metagenomic gene specific multi-primer PCR. *Biotechnology Letters* **34:** 875-882.

Yachi S., Loreau M. (1999). Biodiversity and ecosystem productivity in a fluctuating environment: The insurance hypothesis. *Proceedings of the National Academy of Sciences of the United States of America* **96:** 1463-1468.

Yagi J. M., Suflita J. M., Gieg L. M., DeRito C. M., Jeon C.-O., Madsen E. L. (2010). Subsurface cycling of nitrogen and anaerobic aromatic hydrocarbon biodegradation revealed by nucleic acid and metabolic biomarkers. *Applied and Environmental Microbiology* **76:** 3124-3134.

Young L. Y., Phelps C. D. (2005). Metabolic biomarkers for monitoring in situ anaerobic hydrocarbon degradation. *Environmental Health Perspectives* **113:** 62-67.

Zamfirescu D., Grathwohl P. (2001). Occurrence and attenuation of specific organic compounds in the groundwater plume at a former gasworks site. *Journal of Contaminant Hydrology* **53:** 407-427.

Zedelius J., Rabus R., Grundmann O., Werner I., Brodkorb D., Schreiber F., Ehrenreich P., Behrends A., Wilkes H., Kube M., Reinhardt R., Widdel F. (2011). Alkane degradation under anoxic conditions by a nitrate-reducing bacterium with possible involvement of the electron acceptor in substrate activation. *Environmental Microbiology Reports* **3:** 125-135.

Zhou J., Bruns M. A., Tiedje J. M. (1996). DNA recovery from soils of diverse composition. *Applied and Environmental Microbiology* **62:** 316-322.

Zinder S. H. (2010). Anaerobic utilization of halohydrocarbons. In: Timmis K. N. (ed). *Handbook of Hydrocarbon and Lipid Microbiology*. Springer: Berlin Heidelberg. pp 2049-2064.

7. Publications and authorship clarification

Publications generated within the direct frame of this thesis:

1. **Winderl, C., H. Penning, F. von Netzer, R. U. Meckenstock, and T. Lueders.** 2010. DNA-SIP identifies sulfate-reducing *Clostridia* as important toluene degraders in tar-oil-contaminated aquifer sediment. *The ISME Journal* **4:**1314-1325.
2. **Pilloni, G., F. von Netzer, M. Engel, and T. Lueders.** 2011. Electron acceptor-dependent identification of key anaerobic toluene degraders at a tar-oil-contaminated aquifer by Pyro-SIP. *FEMS Microbiology Ecology* **78:**165-175.
3. **von Netzer, F. G. Pilloni, S. Kleindienst, M. Krüger, K. Knittel, F. Gründger, and T. Lueders.** 2013. Enhanced gene detection assays for fumarate-adding enzymes allow uncovering anaerobic hydrocarbon degraders in terrestrial and marine systems. *Applied and Environmental Microbiology* **79**: 543-552.
   Submitted to *Applied and Environmental Microbiology*, July 2012, accepted 31. October 2012.

Affiliation of novel, unidentified FAE gene clusters via stable isotope probing

Paper by Winderl *et al.* (2010): Tillmann Lueders, Christian Winderl designed the experiment. Christian Winderl conducted the experiment. Further analyses were provided by the PhD candidate and Holger Penning. Christian Winderl, Tillmann Lueders, Rainer Meckenstock and the graduate student wrote the paper. I performed the additional 16S rRNA and *bssA* gene T-RFLP analyses in order to trace the population development of *Desulfosporosinus* spp. and associated *bssA* homologous T-RFs in total DNA. I also wrote respective sections of the manuscript text. For this thesis, I calculated the phylogenetic trees depicted in Figure 8 with the sequences generated by Christian Winderl. Parts of the paper were used in the method section of this thesis.

Paper by Pilloni *et al.* (2011): Tillmann Lueders and Giovanni Pilloni developed the concept and designed the experiment. All measurements were performed by Giovanni Pilloni, except 16S rRNA gene analysis on inoculum DNA (Katrin Hörmann) and *bssA* T-RFLP (done by the author of this thesis). Technical pyrosequencing assistance was provided by Marion Engel. I performed the *bssA* T-RFLP in selected, fractionated DNA aliquots and analysed the respective data in order to identify the *bssA* genes on labelled degrader genomes. I also wrote

respective sections of the manuscript text. Parts of the paper were used in the method and results sections of this thesis.

Enhanced gene detection assay for fumarate-adding enzymes uncover anaerobic hydrocarbon degraders in terrestrial and marine systems

The concept for the experiments was developed by Tillmann Lueders. Experimental design was done by me and Tillmann Lueders. I conducted the experiments, measurements and data analysis. DNA was extracted and supplied by Giovanni Pilloni, Sara Kleindienst, Katrin Knittel, Martin Krüger and Friederike Gründger. These parts of this thesis are based on a manuscript written by me and Tillmann Lueders, submitted to Applied and Environmental Microbiology in July 2012, accepted in October 2012 and published in 2013 (Volume 79: pp 543-552). Parts of the paper were used in introduction, methods, results and discussion sections of this thesis.

Establishment of a pyrosequencing analysis pipeline for *bssA* amplicons

Tillmann Lueders and the PhD candidate designed the concept. Technical assistance was provided from Katrin Hörmann. DNA was provided from Giovanni Pilloni, Martin Krüger and Friederike Gründger. I produced the amplicons and performed the subsequent analyses. Technical pyrosequencing assistance was provided by Marion Engel. I wrote the respective sections of this thesis, which are to be developed into a manuscript in the near future.

Other publications with involvement of the candidate generated during this thesis, but with no direct contribution to the central objectives of this thesis:

1. **Schleheck, D., F. von Netzer, T. Fleischmann, D. Rentsch, T. Huhn, A. M. Cook, and H.-P. E. Kohler.** 2010. The missing link in linear alkylbenzenesulfonate surfactant degradation: 4-Sulfoacetophenone as a transient intermediate in the degradation of 3-(4-sulfophenyl)butyrate by *Comamonas testosteroni* KF-1. *Applied and Environmental Microbiology* **76:**196-202.
2. **Beckmann, S., T. Lueders, M. Kruger, F. von Netzer, B. Engelen, and H. Cypionka.** 2011. Acetogens and acetoclastic methanosarcinales govern methane formation in abandoned coal mines. *Applied and Environmental Microbiology* **77:**3749-3756.
3. **Kleindienst S., Herbst F.-A., von Netzer F., Amann R., Peplies J., von Bergen M., Seifert J., Musat F., Lueders T., Knittel K. (2012).** Specialists instead of generalists oxidize alkanes in anoxic marine hydrocarbon seep sediments. **In preparation.**

Appendix

A1 Selected supporting information

A1.1 Affiliation of novel, unidentified FAE gene clusters via stable isotope probing

Table 13. Complete *bssA* gene T-RF data, analysed in T-REX, as displayed in Figure 10.

Sample	Inoculum	A15	A03	A05	A08	A10
Substrate	no	13C	13C	13C	13C	13C
Incubation time (d)	0	1	8	29	64	86
51 bp	0.00	0.15	0.86	0.00	0.66	0.19
53 bp	16.03	9.67	9.13	7.83	9.87	2.25
55 bp	0.33	0.35	0.26	0.00	0.67	0.19
58 bp	0.00	0.11	0.19	0.00	0.00	0.00
60 bp	0.00	0.28	0.15	0.00	0.23	0.00
63 bp	1.44	2.75	1.02	2.32	0.99	1.19
68 bp	4.52	7.54	10.63	7.29	10.97	5.58
70 bp	0.29	0.00	0.40	0.00	0.00	0.00
75 bp	0.66	0.24	0.27	0.88	0.31	0.35
78 bp	1.09	0.18	0.65	0.00	0.54	0.21
81 bp	0.40	0.10	0.57	0.00	0.00	0.00
84 bp	0.00	0.11	0.15	0.00	0.00	0.00
89 bp	15.17	29.40	20.34	26.98	26.56	25.07
101 bp	0.23	0.00	0.21	0.00	0.00	0.00
112 bp	1.35	0.27	0.85	0.00	9.39	39.27
114 bp	1.86	1.27	1.53	1.67	0.00	0.00
119 bp	0.00	0.00	0.17	0.00	0.31	0.00
121 bp	0.27	0.15	0.00	0.00	0.00	0.00
125 bp	0.30	0.25	0.31	0.00	0.47	0.35
132 bp	0.00	0.14	0.27	0.00	0.00	0.00
150 bp	0.39	0.24	0.40	0.00	0.28	0.21
155 bp	0.72	0.54	0.60	0.42	0.73	0.40
164 bp	1.38	0.21	0.72	0.00	0.24	0.21
166 bp	0.00	0.11	0.27	0.00	0.42	0.00
169 bp	0.32	0.00	0.15	0.00	0.00	0.00
176 bp	0.56	2.31	1.55	1.14	2.74	1.00
178 bp	0.95	1.48	1.22	1.27	0.60	0.71
190 bp	0.34	0.00	0.17	0.00	0.00	0.00
192 bp	0.28	0.20	0.48	0.00	0.39	0.00
198 bp	1.61	1.26	1.30	1.37	1.31	1.27
199 bp	1.16	0.56	2.91	0.64	2.56	0.71
204 bp	4.69	5.57	6.65	2.98	3.43	1.51
206 bp	0.26	0.16	0.60	0.00	0.60	0.29
216 bp	0.35	0.26	0.71	0.00	0.50	0.25
220 bp	0.00	0.15	0.31	0.00	0.32	0.18
222 bp	1.81	0.21	0.71	0.54	1.09	0.68
224 bp	0.46	0.61	0.82	0.46	0.67	0.31
227 bp	0.63	0.00	0.30	0.00	0.00	0.00
230 bp	0.38	0.00	0.15	0.00	0.00	0.00
240 bp	27.27	19.09	16.81	34.13	15.36	13.16
242 bp	3.30	1.96	2.56	2.80	0.97	0.84
248 bp	0.00	0.00	0.00	0.00	0.65	1.08
252 bp	0.00	0.10	0.30	0.00	0.00	0.00
264 bp	0.00	0.25	0.20	0.00	0.23	0.33
295 bp	0.00	0.17	0.00	0.00	0.00	0.18
326 bp	0.37	0.26	0.26	0.56	0.36	0.22
329 bp	0.00	0.09	0.51	0.00	0.00	0.00
333 bp	0.00	0.30	0.18	0.00	0.22	0.00
T-cluster 364 bp	**8.50**	**9.89**	**8.73**	**5.44**	**3.01**	**1.17**
687 bp	0.00	0.00	0.00	0.00	0.32	0.18
738 bp	0.00	0.48	0.72	0.53	1.18	0.45

Table 14. 16S rRNA gene T-RF data, analysed in T-REX, as displayed in Figure 11

Sample	Inokulum	A15	A3	A5	A8	A10
Substrate	no	13C	13C	13C	13C	13C
Incubation time (d)	0	1	8	29	64	86
60 bp	0.56	0.21	0.15	0.21	0.32	0.21
62 bp	0.36	0.46	0.54	0.57	0.43	0.50
73 bp	0.56	1.26	0.68	0.53	0.40	0.26
78 bp	1.41	0.98	0.86	1.70	1.26	0.92
81 bp	0.21	0.10	0.10	0.16	0.09	0.00
84 bp	0.32	0.21	0.21	0.24	0.38	0.39
87 bp	0.59	0.32	0.24	0.46	0.42	0.39
90 bp	2.05	2.11	2.10	2.93	4.06	3.61
93 bp	0.00	0.14	0.13	0.16	0.19	0.20
103 bp	0.13	0.00	0.07	0.00	0.00	0.00
105 bp	0.04	0.00	0.05	0.00	0.00	0.00
110 bp	0.28	0.22	0.18	0.24	0.33	0.33
114 bp	0.14	0.20	0.22	0.15	0.15	0.16
119 bp	1.44	1.27	0.99	1.39	0.71	0.68
123 bp	0.70	1.37	0.86	0.85	1.19	1.02
127 bp	1.86	2.26	1.89	2.25	1.26	2.08
131 bp	0.55	0.86	0.43	0.24	0.40	0.53
137 bp	2.11	2.03	2.20	2.43	1.69	1.68
143 bp	6.74	16.72	14.16	8.81	9.67	7.15
146 bp	**6.25**	**6.05**	**5.47**	**6.20**	**9.30**	**7.74**
149 bp	0.87	0.90	1.05	1.22	0.66	1.26
156 bp	0.47	0.22	0.36	0.17	0.18	0.17
160 bp	5.58	3.91	2.92	2.96	2.89	4.14
163 bp	0.47	1.08	0.98	0.90	0.77	1.02
165 bp	0.90	1.08	0.84	0.73	0.31	0.20
170 bp	1.63	3.90	2.74	1.96	3.32	2.45
172 bp	0.03	0.00	0.06	0.00	0.00	0.00
175 bp	0.89	0.85	0.58	0.54	0.28	0.29
178 bp	**0.13**	**0.16**	**0.13**	**0.11**	**9.61**	**7.64**
181 bp	0.18	0.17	0.68	0.57	0.89	0.50
183 bp	0.26	0.56	0.36	0.37	0.64	0.65
188 bp	0.20	0.18	0.12	0.14	0.12	0.00
191 bp	0.35	0.33	0.36	0.34	0.27	0.25
194 bp	0.29	0.41	0.34	0.32	0.28	0.32
199 bp	0.21	0.22	0.26	0.83	0.23	0.29
202 bp	0.35	0.62	0.40	0.32	0.35	0.43
205 bp	0.33	0.30	0.24	0.27	0.85	1.12
208 bp	0.28	0.31	0.32	0.42	0.35	0.24
211 bp	1.21	0.23	0.18	0.48	0.19	0.41
214 bp	0.14	0.07	0.08	0.19	0.10	0.00
216 bp	0.05	0.00	0.07	0.00	0.00	0.00
222 bp	0.15	0.11	0.09	0.10	0.11	0.11
225 bp	0.39	0.55	0.46	0.71	0.69	0.59
229 bp	0.11	0.08	0.08	0.10	0.10	0.00
232 bp	0.18	0.18	0.14	0.17	0.22	0.17
237 bp	0.09	0.13	0.09	0.10	0.12	0.15
243 bp	0.11	0.11	0.09	0.00	0.00	0.12
247 bp	0.16	0.16	0.18	0.22	0.15	0.13
250 bp	0.28	0.19	0.19	0.21	0.21	0.23
253 bp	0.09	0.00	0.07	0.00	0.07	0.00
258 bp	0.10	0.09	0.08	0.11	0.00	0.10
260 bp	0.12	0.09	0.10	0.11	0.00	0.15
264 bp	0.26	0.19	0.17	0.20	0.14	0.14
267 bp	0.16	0.21	0.21	0.25	0.25	0.19
270 bp	0.23	0.17	0.17	0.20	0.13	0.11
273 bp	0.15	0.13	0.21	0.18	0.22	0.18
277 bp	0.50	0.17	0.28	0.16	0.25	0.16
281 bp	0.38	0.34	0.51	0.58	0.33	0.44
286 bp	0.53	0.37	0.38	0.35	0.32	0.27
290 bp	0.08	0.00	0.27	0.23	0.00	0.10
294 bp	0.14	0.17	0.14	0.13	0.19	0.15
297 bp	0.18	0.23	0.21	0.22	0.13	0.16
300 bp	0.09	0.00	0.00	0.00	0.12	0.14
304 bp	0.22	0.19	0.16	0.21	0.20	0.21
307 bp	0.18	0.16	0.19	0.23	0.13	0.13
311 bp	0.31	0.25	0.22	0.26	0.98	0.75
314 bp	0.09	0.08	0.16	0.11	0.15	0.10
317 bp	0.00	0.00	0.00	0.00	0.00	0.00
320 bp	0.09	0.09	0.10	0.11	0.08	0.00
323 bp	0.17	0.13	0.13	0.14	0.12	0.12
328 bp	0.09	0.08	0.09	0.00	0.00	0.00
331 bp	0.04	0.09	0.00	0.00	0.00	0.00
335 bp	0.04	0.08	0.05	0.00	0.00	0.00
343 bp	0.04	0.00	0.05	0.00	0.08	0.00
346 bp	0.00	0.00	0.05	0.00	0.07	0.00
352 bp	0.23	0.16	0.17	0.17	0.15	0.13
354 bp	0.22	0.13	0.19	0.17	0.15	0.19
361 bp	0.09	0.08	0.06	0.00	0.09	0.11
364 bp	0.04	0.00	0.00	0.00	0.00	0.00
372 bp	1.17	1.72	1.65	2.44	2.52	3.37
376 bp	0.03	0.00	0.00	0.00	0.00	0.00
379 bp	0.06	0.00	0.05	0.00	0.00	0.00
386 bp	0.06	0.00	0.06	0.00	0.00	0.00
396 bp	0.15	0.16	0.10	0.10	0.00	0.00
398 bp	0.68	0.91	0.85	0.92	0.73	1.43
402 bp	0.70	0.64	0.61	0.65	0.47	0.76
404 bp	0.50	0.39	0.44	0.51	0.38	0.37
412 bp	0.00	0.00	0.08	0.00	0.00	0.00
417 bp	0.17	0.12	0.10	0.15	0.10	0.10
421 bp	0.20	0.20	0.20	0.20	0.17	0.30
425 bp	0.10	0.10	0.11	0.13	0.09	0.00
429 bp	0.19	0.15	0.19	0.16	0.12	0.18
434 bp	0.89	0.97	1.09	0.97	0.84	1.53
438 bp	2.17	2.10	1.99	2.15	1.65	2.70
444 bp	6.31	3.67	4.30	4.31	2.90	3.06
447 bp	0.08	0.00	0.10	0.00	0.00	0.00
451 bp	0.31	0.28	0.28	0.24	0.24	0.22
453 bp	0.14	0.10	0.09	0.12	0.10	0.12
456 bp	0.69	0.55	0.48	0.48	0.45	0.43
460 bp	0.03	0.13	0.07	0.00	0.15	0.21
464 bp	0.86	0.51	0.72	0.70	0.36	0.54
468 bp	1.18	1.07	7.77	6.88	1.15	4.60
474 bp	2.22	0.88	0.82	2.10	0.85	0.76
476 bp	0.43	0.30	0.27	0.39	0.19	0.25
480 bp	0.04	0.00	0.00	0.00	0.00	0.00
485 bp	7.94	7.89	7.69	8.87	7.40	6.97
491 bp	17.57	10.75	11.76	11.58	7.44	6.04
496 bp	0.11	0.11	0.10	0.00	0.15	0.11
500 bp	0.07	0.08	0.06	0.00	0.00	0.00
502 bp	0.17	0.00	0.06	0.09	0.00	0.00
506 bp	0.66	0.44	0.31	0.43	0.49	0.51
510 bp	0.68	0.81	0.88	0.91	1.37	1.20
513 bp	0.69	0.95	0.63	0.86	1.05	0.90
518 bp	0.88	1.68	1.39	1.53	3.13	3.45
527 bp	1.62	1.82	1.47	1.33	1.17	1.30
535 bp	0.06	0.00	0.00	0.00	0.08	0.00
538 bp	0.10	0.09	0.06	0.00	0.00	0.10
543 bp	0.08	0.10	0.07	0.10	0.24	0.32
545 bp	0.10	0.14	0.11	0.12	0.13	0.17
548 bp	0.20	0.37	0.31	0.28	0.32	0.35
562 bp	0.11	0.16	0.17	0.15	0.18	0.20
564 bp	0.12	0.08	0.07	0.00	0.08	0.00
579 bp	0.21	0.30	0.44	0.52	0.77	0.96
585 bp	0.04	0.00	0.06	0.00	0.00	0.00
597 bp	0.25	0.30	0.27	0.27	0.29	0.27
601 bp	0.59	0.42	0.39	0.39	0.35	0.29
606 bp	0.03	0.00	0.00	0.00	0.00	0.00
645 bp	0.12	0.19	0.19	0.22	0.38	0.42
652 bp	0.11	0.10	0.09	0.00	0.10	0.00
684 bp	0.06	0.09	0.06	0.00	0.09	0.00
763 bp	0.03	0.10	0.10	0.12	0.21	0.21
864 bp	0.12	0.12	0.08	0.10	0.11	0.13
894 bp	0.18	0.20	0.23	0.20	0.28	0.29
900 bp	0.13	0.08	0.08	0.09	0.08	0.00
916 bp	0.23	0.16	0.16	0.13	0.17	0.14

A1.2 Enhanced gene detection assay for fumarate-adding enzymes uncover anaerobic hydrocarbon degraders in terrestrial and marine systems

Table 15. *bssA* gene T-RFLP data, analysed by T-REX and averaged from four independent DNA extractions, as displayed in Figure 13. The lower table shows the standard deviations.

		bssA 58°C			bssA 52°C			FAE-B			FAE-N		
		6.85 m	7.15 m	7.25 m	6.85 m	7.15 m	7.25 m	6.85 m	7.15 m	7.25 m	6.85 m	7.15 m	7.25 m
Average	73 bp				1.62	7.30	6.28						
	83 bp							3.32	0.00	7.69			
	117 bp				0.64	0.25	11.04						
	152 bp										25.74	0.00	0.00
	214 bp				0.43	0.61	0.00						
	242 bp	0.55	0.43	8.08	2.59	1.39	0.53						
	297 bp							0.83	0.75	1.53			
	324 bp							50.67	94.96	85.48	74.26	100.00	100.00
	364 bp	1.92	0.00	7.18	14.36	0.97	11.72	20.46	0.00	0.86			
	478 bp	97.54	99.57	84.74	61.42	60.18	61.44						
	687 bp				18.93	29.30	9.00						

		bssA 58°C			bssA 52°C			FAE-B			FAE-N		
		6.85 m	7.15 m	7.25 m	6.85 m	7.15 m	7.25 m	6.85 m	7.15 m	7.25 m	6.85 m	7.15 m	7.25 m
Standard deviation	73 bp				0.85	13.40	12.57						
	83 bp							2.80	0.00	6.05			
	117 bp				0.54	0.29	10.52						
	152 bp										9.72	0.00	0.00
	214 bp				0.50	0.42	0.00						
	242 bp	1.10	0.86	5.91	0.99	0.97	1.06						
	297 bp							0.73	1.49	1.05			
	324 bp							11.23	10.08	10.29	9.72	0.00	0.00
	364 bp	2.26	0.00	3.22	2.06	0.65	12.63	9.70	0.00	1.24			
	478 bp	3.15	0.85	8.16	3.52	9.76	28.21						
	687 bp				2.13	7.22	8.17						

Table 16. *bssA* gene T-RFLP data, analysed with T-REX, as displayed in Figure 16.

Primer	FAE-KM			FAE-N	
Time (d)	Inokulum	57	113	Inokulum	
51 bp	1.8	0.0	0.0	109 bp	3.0
53 bp	3.3	6.8	3.3	198 bp	1.3
54 bp	0.0	8.9	2.9	221 bp	16.2
67 bp	1.8	4.1	1.6	326 bp	6.5
68 bp	0.0	3.1	0.7	348 bp	73.2
71 bp	0.0	0.0	1.4		
77 bp	30.5	6.4	1.3		
79 bp	9.0	48.5	28.6		
111 bp	1.2	1.4	1.1		
147 bp	38.3	7.2	1.8		
156 bp	0.4	0.0	0.6		
222 bp	1.3	2.0	1.2		
438 bp	6.3	8.8	38.9		
788 bp	2.3	2.9	16.6		

A1.3 Establishment of a pyrosequencing analysis pipeline for *bssA* amplicons

A1.3.1 Sequences used for creating a FAE BLASTN database

Accession	gene	Full name
JX219271	*bssA*	Amon mud volcano clone BAA01
JX219272	*bssA*	Amon mud volcano clone BAA03
JX219273	*bssA*	Amon mud volcano clone BAA11
JX219274	*bssA*	Amon mud volcano clone BAA17
JX219275	*bssA*	Amon mud volcano clone BAA19
BX682953	*bssA*	*Aromatoleum aromaticum* EbN1
GU477530	*bssA*	*Aromatoleum* Marmara Sea Sediment clone bssAMSS6
GU477525	*bssA*	*Azoarcus* Marmara Sea Sediment clone bssAMSS1
GU477526	*bssA*	*Azoarcus* Marmara Sea Sediment clone bssAMSS2
AB285034	*bssA*	*Azoarcus* sp. DN11
AY032676	*bssA*	*Azoarcus* sp. T
FN564032	*bssA*	Banisveld aquifer bssA clone 39b_2004(2F)
FN564033	*bssA*	Banisveld aquifer bssA clone P3B3B
FN564034	*bssA*	Banisveld aquifer bssA clone P3B3G
FN564035	*bssA*	Banisveld aquifer bssA clone P3Bb1A
FN564036	*bssA*	Banisveld aquifer bssA clone P3B4G
FN564037	*bssA*	Banisveld aquifer bssA clone P3C1F
FN564038	*bssA*	Banisveld aquifer bssA clone P3C1G
FN564039	*bssA*	Banisveld aquifer bssA clone P3C2B
FN564040	*bssA*	Banisveld aquifer bssA clone BBS075
FN564041	*bssA*	Banisveld aquifer bssA clone P7C10A
FN564042	*bssA*	Banisveld aquifer bssA clone P5A8F
FN564043	*bssA*	Banisveld aquifer bssA clone P5A8H
FN564043	*bssA*	Banisveld aquifer bssA clone P5B1G
FN564044	*bssA*	Banisveld aquifer bssA clone BBS050
FN564045	*bssA*	Banisveld aquifer bssA clone P5A9F
FN564046	*bssA*	Banisveld aquifer bssA clone P5A9G
FN564047	*bssA*	Banisveld aquifer bssA clone P5A10C
FN564048	*bssA*	Banisveld aquifer bssA clone P5A11D
FN564049	*bssA*	Banisveld aquifer bssA clone P5A12E
FN564050	*bssA*	Banisveld aquifer bssA clone P5A10F
FN564053	*bssA*	Banisveld aquifer bssA clone P5B3B
FN564054	*bssA*	Banisveld aquifer bssA clone P5B4D
FN564055	*bssA*	Banisveld aquifer bssA clone P5B6G
FN564056	*bssA*	Banisveld aquifer bssA clone 68a_2004(9G)
FN564057	*bssA*	Banisveld aquifer bssA clone BBS130
FN564058	*bssA*	Banisveld aquifer bssA clone P7B
FN564059	*bssA*	Banisveld aquifer bssA clone P7B3A
FN564060	*bssA*	Banisveld aquifer bssA clone P7C7A
FN564061	*bssA*	Banisveld aquifer bssA clone P7C7C
FN564062	*bssA*	Banisveld aquifer bssA clone P7C7E
FN564063	*bssA*	Banisveld aquifer bssA clone P7C8C
FN564064	*bssA*	Banisveld aquifer bssA clone P7C8F
FN564065	*bssA*	Banisveld aquifer bssA clone P7C9E
FN564066	*bssA*	Banisveld aquifer bssA clone P7C9H
FN564067	*bssA*	Banisveld aquifer bssA clone P7C10C
FN564068	*bssA*	Banisveld aquifer bssA clone P7C11A
FN564069	*bssA*	Banisveld aquifer bssA clone P7C11B
FN564070	*bssA*	Banisveld aquifer bssA clone P7C11C
FN564071	*bssA*	Banisveld aquifer bssA clone P7C11E
FN564072	*bssA*	Banisveld aquifer bssA clone P7C11G
FN564073	*bssA*	Banisveld aquifer bssA clone P7C12F
FN564074	*bssA*	Banisveld aquifer bssA clone P7C12H
FN564075	*bssA*	Banisveld aquifer bssA clone 6b_2004(4F)
FN564076	*bssA*	Banisveld aquifer bssA clone P1C11H
FN564077	*bssA*	Banisveld aquifer bssA clone P4B1C
FN564078	*bssA*	Banisveld aquifer bssA clone P4B1G
FN564079	*bssA*	Banisveld aquifer bssA clone BBS035
FN564080	*bssA*	Banisveld aquifer bssA clone P4B3H
FN564081	*bssA*	Banisveld aquifer bssA clone P4B5G
FN564082	*bssA*	Banisveld aquifer bssA clone 48a_1999(2C)
FN564082	*bssA*	Banisveld aquifer bssA clone P4F12C
FN564083	*bssA*	Banisveld aquifer bssA clone BBS020
FN564084	*bssA*	Banisveld aquifer bssA clone 48a_1999(5B)
FN564085	*bssA*	Banisveld aquifer bssA clone P6A7B
FN564086	*bssA*	Banisveld aquifer bssA clone P6A7D
FN564087	*bssA*	Banisveld aquifer bssA clone P6A7E
FN564088	*bssA*	Banisveld aquifer bssA clone P6A8A
FN564089	*bssA*	Banisveld aquifer bssA clone BBS090
FN564090	*bssA*	Banisveld aquifer bssA clone P6A8F
FN564091	*bssA*	Banisveld aquifer bssA clone P6A9C
FN564092	*bssA*	Banisveld aquifer bssA clone P6A12A
FN564093	*bssA*	Banisveld aquifer bssA clone P6F110C
FN564094	*bssA*	Banisveld aquifer bssA clone P8B1B
FN564095	*bssA*	Banisveld aquifer bssA clone P8B1H
FN564096	*bssA*	Banisveld aquifer bssA clone P8B2D
FN564097	*bssA*	Banisveld aquifer bssA clone 89a_1999(10D)
FN564098	*bssA*	Banisveld aquifer bssA clone P11F111E
EF190465	*bssA*	Banisveld toluene-degrading enrichment GW3Aa
EF190466	*bssA*	Banisveld toluene-degrading enrichment GW3Ab
Lueders *et al* . n.p.	*bssA*	Brand aquifer clone A14bss01
Lueders *et al* . n.p.	*bssA*	Brand aquifer clone A14bss03
Lueders *et al* . n.p.	*bssA*	Brand aquifer clone A14bss05
Lueders *et al* . n.p.	*bssA*	Brand aquifer clone A14bss06
Lueders *et al* . n.p.	*bssA*	Brand aquifer clone A14bss07
Lueders *et al* . n.p.	*bssA*	Brand aquifer clone A14bss08
Lueders *et al* . n.p.	*bssA*	Brand aquifer clone A14bss09
Lueders *et al* . n.p.	*bssA*	Brand aquifer clone A14bss10
Lueders *et al* . n.p.	*bssA*	Brand aquifer clone A14bss11
Lueders *et al* . n.p.	*bssA*	Brand aquifer clone A14bss12
Lueders *et al* . n.p.	*bssA*	Brand aquifer clone A14bss13
Lueders *et al* . n.p.	*bssA*	Brand aquifer clone A14bss14
Lueders *et al* . n.p.	*bssA*	Brand aquifer clone A14bss15
Lueders *et al* . n.p.	*bssA*	Brand aquifer clone A14bss16
Lueders *et al* . n.p.	*bssA*	Brand aquifer clone A14bss17
Lueders *et al* . n.p.	*bssA*	Brand aquifer clone A14bss18
Lueders *et al* . n.p.	*bssA*	Brand aquifer clone A14bss19
Lueders *et al* . n.p.	*bssA*	Brand aquifer clone A14bss28
Lueders *et al* . n.p.	*bssA*	Brand aquifer clone A14bss31
Lueders *et al* . n.p.	*bssA*	Brand aquifer clone A14bss35
Lueders *et al* . n.p.	*bssA*	Brand aquifer clone A14bss39
Lueders *et al* . n.p.	*bssA*	Brand aquifer clone A14bss48
Lueders *et al* . n.p.	*bssA*	Brand aquifer clone A14bss50
Lueders *et al* . n.p.	*bssA*	Brand aquifer clone A14bss65
Lueders *et al* . n.p.	*bssA*	Brand aquifer clone A14bss67
Lueders *et al* . n.p.	*bssA*	Brand aquifer clone A14bss73
Lueders *et al* . n.p.	*bssA*	Brand aquifer clone A14bss78
Lueders *et al* . n.p.	*bssA*	Brand aquifer clone A14bss81
Lueders *et al* . n.p.	*bssA*	Brand aquifer clone A14bss96
Lueders *et al* . n.p.	*bssA*	Brand aquifer clone L21bss01
Lueders *et al* . n.p.	*bssA*	Brand aquifer clone L21bss02
Lueders *et al* . n.p.	*bssA*	Brand aquifer clone L21bss03
Lueders *et al* . n.p.	*bssA*	Brand aquifer clone L21bss04
Lueders *et al* . n.p.	*bssA*	Brand aquifer clone L21bss06
Lueders *et al* . n.p.	*bssA*	Brand aquifer clone L21bss07
Lueders *et al* . n.p.	*bssA*	Brand aquifer clone L21bss08
Lueders *et al* . n.p.	*bssA*	Brand aquifer clone L21bss09
Lueders *et al* . n.p.	*bssA*	Brand aquifer clone L21bss10
Lueders *et al* . n.p.	*bssA*	Brand aquifer clone L21bss11
Lueders *et al* . n.p.	*bssA*	Brand aquifer clone L21bss12
Lueders *et al* . n.p.	*bssA*	Brand aquifer clone L21bss13
Lueders *et al* . n.p.	*bssA*	Brand aquifer clone L21bss14
Lueders *et al* . n.p.	*bssA*	Brand aquifer clone L21bss15
Lueders *et al* . n.p.	*bssA*	Brand aquifer clone L21bss16
Lueders *et al* . n.p.	*bssA*	Brand aquifer clone L21bss17
Lueders *et al* . n.p.	*bssA*	Brand aquifer clone L21bss19
Lueders *et al* . n.p.	*bssA*	Brand aquifer clone L21bss20
Lueders *et al* . n.p.	*bssA*	Brand aquifer clone L21bss23
Lueders *et al* . n.p.	*bssA*	Brand aquifer clone L21bss26
Lueders *et al* . n.p.	*bssA*	Brand aquifer clone L21bss42
Lueders *et al* . n.p.	*bssA*	Brand aquifer clone L21bss55
Lueders *et al* . n.p.	*bssA*	Brand aquifer clone L21bss60
Lueders *et al* . n.p.	*bssA*	Brand aquifer clone L21bss65
Lueders *et al* . n.p.	*bssA*	Brand aquifer clone L21bss67
Lueders *et al* . n.p.	*bssA*	Brand aquifer clone L21bss74
Lueders *et al* . n.p.	*bssA*	Brand aquifer clone L21bss76
Lueders *et al* . n.p.	*bssA*	Brand aquifer clone L21bss77
Lueders *et al* . n.p.	*bssA*	Brand aquifer clone L21bss82
Lueders *et al* . n.p.	*bssA*	Brand aquifer clone L21bss83
Lueders *et al* . n.p.	*bssA*	Brand aquifer clone L21bss84
Lueders *et al* . n.p.	*bssA*	Brand aquifer clone L21bss89
GU453668	*bssA*	Casper aquifer OTU 1
GU453669	*bssA*	Casper aquifer OTU 2
GU453671	*bssA*	Casper aquifer OTU 4
GU453673	*bssA*	Casper aquifer OTU 5
GU453670	*bssA*	Casper aquifer OTU 6
GU453674	*bssA*	Casper aquifer OTU 6
GU477533	*bssA*	*Deltaproteobacteria* Marmara Sea Sediment clone bssMSS9
von Netzer *et al*. n.p.	*bssA*	*Desulfitobacterium aromaticivorans* UKTL bssA2
EF123662	*bssA*	*Desulfobacterium cetonicum*
GU477528	*bssA*	*Desulfobacula* Marmara Sea Sediment clone bssAMSS4
GU477532	*bssA*	*Desulfobacula* Marmara Sea Sediment clone bssAMSS8
CU466268	*bssA*	*Desulfobacula toluolica* Tol2
EF123663	*bssA*	*Desulfobacula toluolica* Tol2
Winderl *et al.* n.p.	*bssA*	*Desulfosporosinus* sp. SIP clone GL10BS03
Winderl *et al.* n.p.	*bssA*	*Desulfosporosinus* sp. SIP clone GL10BS26
Winderl *et al.* n.p.	*bssA*	*Desulfosporosinus* sp. SIP clone GL10BS43
EF123678	*bssA*	Flingern aquifer clone D1203
EF123679	*bssA*	Flingern aquifer clone D1208
EF123680	*bssA*	Flingern aquifer clone D1209
EF123681	*bssA*	Flingern aquifer clone D1210
EF123682	*bssA*	Flingern aquifer clone D1212
EF123683	*bssA*	Flingern aquifer clone D1218
EF123684	*bssA*	Flingern aquifer clone D1223
EF123685	*bssA*	Flingern aquifer clone D1231
EF123686	*bssA*	Flingern aquifer clone D1256
EF123687	*bssA*	Flingern aquifer clone D1264
JX219277	*bssA*	Flingern aquifer DS03 clone F3A02

n.p. not published. Continued on the following page

Accession	gene	Full name
JX219278	*bssA*	Flingern aquifer DS03 clone F3A03
JX219279	*bssA*	Flingern aquifer DS03 clone F3A06
JX219280	*bssA*	Flingern aquifer DS03 clone F3A07
JX219281	*bssA*	Flingern aquifer DS03 clone F3A08
JX219282	*bssA*	Flingern aquifer DS03 clone F3A10
JX219283	*bssA*	Flingern aquifer DS03 clone F3A11
JX219284	*bssA*	Flingern aquifer DS03 clone F3A14
JX219285	*bssA*	Flingern aquifer DS03 clone F3A15
JX219286	*bssA*	Flingern aquifer DS03 clone F3A19
JX219287	*bssA*	Flingern aquifer DS03 clone F3A20
JX219288	*bssA*	Flingern aquifer DS03 clone F3A29
JX219289	*bssA*	Flingern aquifer DS03 clone F3B22
JX219290	*bssA*	Flingern aquifer DS04 clone F4A01
JX219291	*bssA*	Flingern aquifer DS04 clone F4A02
JX219292	*bssA*	Flingern aquifer DS04 clone F4A03
JX219293	*bssA*	Flingern aquifer DS04 clone F4A05
JX219294	*bssA*	Flingern aquifer DS04 clone F4A06
JX219295	*bssA*	Flingern aquifer DS04 clone F4A07
JX219296	*bssA*	Flingern aquifer DS04 clone F4A08
JX219297	*bssA*	Flingern aquifer DS04 clone F4A09
JX219298	*bssA*	Flingern aquifer DS04 clone F4A10
JX219299	*bssA*	Flingern aquifer DS04 clone F4A11
JX219301	*bssA*	Flingern aquifer DS04 clone F4A13
JX219302	*bssA*	Flingern aquifer DS04 clone F4A14
JX219303	*bssA*	Flingern aquifer DS04 clone F4A21
JX219304	*bssA*	Flingern aquifer DS04 clone F4A22
JX219305	*bssA*	Flingern aquifer DS04 clone F4A28
JX219306	*bssA*	Flingern aquifer DS04 clone F4A29
JX219307	*bssA*	Flingern aquifer DS05 clone F5A03
JX219308	*bssA*	Flingern aquifer DS05 clone F5A05
JX219309	*bssA*	Flingern aquifer DS05 clone F5A07
JX219310	*bssA*	Flingern aquifer DS05 clone F5A08
JX219311	*bssA*	Flingern aquifer DS05 clone F5A10
JX219312	*bssA*	Flingern aquifer DS05 clone F5A14
JX219313	*bssA*	Flingern aquifer DS05 clone F5A16
JX219314	*bssA*	Flingern aquifer DS05 clone F5A19
JX219315	*bssA*	Flingern aquifer DS05 clone F5A24
JX219316	*bssA*	Flingern aquifer DS05 clone F5A29
GU453672	*bssA*	Fort Lupton aquifer OTU1
GU453675	*bssA*	Fort Lupton aquifer OTU2
NC_011979	*bssA*	*Geobacter daltonii* FRC-32 bssA1
NC_011979	*bssA*	*Geobacter daltonii* FRC-32 bssA2
EF123664	*bssA*	*Geobacter grbiciae* TACP
GU477534	*bssA*	*Geobacter* Marmara Sea Sediment clone bssAMSS10
GU477529	*bssA*	*Geobacter* Marmara Sea Sediment clone bssAMSS5
AF441130	*bssA*	*Geobacter metallireducens* GS-15
EF123666	*bssA*	*Geobacter toluenoxydans* TMJ1
FN562993	*bssA*	*Georgfuchsia toluolica* G5G6
FJ810622	*bssA*	Glenn Falls aquifer clone BssAC01
FJ810623	*bssA*	Glenn Falls aquifer clone BssAC02
FJ810624	*bssA*	Glenn Falls aquifer clone BssAC03
FJ810625	*bssA*	Glenn Falls aquifer clone BssAC04
FJ810626	*bssA*	Glenn Falls aquifer clone BssAC05
FJ810627	*bssA*	Glenn Falls aquifer clone BssAC06
FJ810628	*bssA*	Glenn Falls aquifer clone BssAC07
FJ810629	*bssA*	Glenn Falls aquifer clone BssAC08
FJ810630	*bssA*	Glenn Falls aquifer clone BssAC09
FJ810631	*bssA*	Glenn Falls aquifer clone BssAC10
FJ810632	*bssA*	Glenn Falls aquifer clone BssAC11
FJ810633	*bssA*	Glenn Falls aquifer clone BssAC12
FJ810634	*bssA*	Glenn Falls aquifer clone BssAC13
FJ810636	*bssA*	Glenn Falls aquifer clone BssAC15
FJ810637	*bssA*	Glenn Falls aquifer clone BssAC16
FJ810638	*bssA*	Glenn Falls aquifer clone BssAC17
FJ810639	*bssA*	Glenn Falls aquifer clone BssAC18
FJ810640	*bssA*	Glenn Falls aquifer clone BssAC19
FJ810641	*bssA*	Glenn Falls aquifer clone BssAC20
FJ810642	*bssA*	Glenn Falls aquifer clone BssAC21
FJ810643	*bssA*	Glenn Falls aquifer clone BssAD01
FJ810644	*bssA*	Glenn Falls aquifer clone BssAD02
FJ810645	*bssA*	Glenn Falls aquifer clone BssAD03
FJ810646	*bssA*	Glenn Falls aquifer clone BssAD06
FJ810647	*bssA*	Glenn Falls aquifer clone BssAD07
FJ810648	*bssA*	Glenn Falls aquifer clone BssAD10
JX219276	*bssA*	Gulf of Mexico oil sediment 161 clone BG2A13
GU477531	*bssA*	*Magnetospirillum* Marmara Sea Sediment clone bssAMSS7
AB167725	*bssA*	*Magnetospirillum* sp. TS-6 bssA
von Netzer *et al.* n.p.	*bssA*	Marmara microcosm bssA clone T3N03
von Netzer *et al.* n.p.	*bssA*	Marmara microcosm bssA clone T3N06r
von Netzer *et al.* n.p.	*bssA*	Marmara microcosm bssA clone T3N24r
von Netzer *et al.* n.p.	*bssA*	Marmara microcosm bssA clone T3N44r
JF754547	*bssA*	methanogenic alkane degrader enrichment clone bssA40
JF754549	*bssA*	methanogenic oily sludge enrichment clone bssA12
JF754548	*bssA*	methanogenic oily sludge enrichment clone bssA16

Accession	gene	Full name
EF123690	*bssA*	Pasing aquifer clone Pb11209
EF123688	*bssA*	Pasing aquifer clone Pb11204
EF123689	*bssA*	Pasing aquifer clone Pb11208
EF123691	*bssA*	Pasing aquifer clone Pb11210
EF123692	*bssA*	Pasing aquifer clone Pb11213
EF123693	*bssA*	Pasing aquifer clone Pb11215
EF123694	*bssA*	Pasing aquifer clone Pb11216
EF123695	*bssA*	Pasing aquifer clone Pb11219
EF123696	*bssA*	Pasing aquifer clone Pb11240
EF123697	*bssA*	Pasing aquifer clone Pb11272
EF123698	*bssA*	Pasing aquifer clone Pb11285
EF123702	*bssA*	Pasing aquifer clone Pb31235
EF123699	*bssA*	Pasing aquifer clone Pb31207
EF123700	*bssA*	Pasing aquifer clone Pb31216
EF123701	*bssA*	Pasing aquifer clone Pb31225
EF123703	*bssA*	Pasing aquifer clone Pb31280
AM887516	*bssA*	petroleum microbial mat clone 05
AM887514	*bssA*	petroleum microbial mat clone 07
AM887517	*bssA*	petroleum microbial mat clone 08
AM887512	*bssA*	petroleum microbial mat clone 10
AM887515	*bssA*	petroleum microbial mat clone 12
AM887513	*bssA*	petroleum microbial mat clone 13
EU780921	*bssA*	Strain PROTL1
EF123667	*bssA*	Strain TRM1
GU133297	*bssA*	Testfeld Süd SIP clone LA07bs08
Winderl *et al.* n.p.	*bssA*	Testfeld Süd SIP clone GL10BS02
Winderl *et al.* n.p.	*bssA*	Testfeld Süd SIP clone GL10BS05
Winderl *et al.* n.p.	*bssA*	Testfeld Süd SIP clone GL10BS18
Winderl *et al.* n.p.	*bssA*	Testfeld Süd SIP clone GL10BS21
Winderl *et al.* n.p.	*bssA*	Testfeld Süd SIP clone GL10BS23
Winderl *et al.* n.p.	*bssA*	Testfeld Süd SIP clone GL10BS24
Winderl *et al.* n.p.	*bssA*	Testfeld Süd SIP clone GL10BS25
Winderl *et al.* n.p.	*bssA*	Testfeld Süd SIP clone GL10BS27
Winderl *et al.* n.p.	*bssA*	Testfeld Süd SIP clone GL10BS29
Winderl *et al.* n.p.	*bssA*	Testfeld Süd SIP clone GL10BS33
Winderl *et al.* n.p.	*bssA*	Testfeld Süd SIP clone GL10BS35
Winderl *et al.* n.p.	*bssA*	Testfeld Süd SIP clone GL10BS46
GU133291	*bssA*	Testfeld Süd SIP clone LA07bs01
GU133292	*bssA*	Testfeld Süd SIP clone LA07bs02
GU133293	*bssA*	Testfeld Süd SIP clone LA07bs03
GU133294	*bssA*	Testfeld Süd SIP clone LA07bs04
GU133295	*bssA*	Testfeld Süd SIP clone LA07bs05
GU133296	*bssA*	Testfeld Süd SIP clone LA07bs07
GU133298	*bssA*	Testfeld Süd SIP clone LA07bs09
GU133299	*bssA*	Testfeld Süd SIP clone LA07bs11
GU133300	*bssA*	Testfeld Süd SIP clone LA07bs12
GU133301	*bssA*	Testfeld Süd SIP clone LA07bs13
GU133302	*bssA*	Testfeld Süd SIP clone LA07bs14
GU133303	*bssA*	Testfeld Süd SIP clone LA07bs15
GU133304	*bssA*	Testfeld Süd SIP clone LA07bs16
GU133305	*bssA*	Testfeld Süd SIP clone LA07bs17
GU133306	*bssA*	Testfeld Süd SIP clone LA07bs20
GU133307	*bssA*	Testfeld Süd SIP clone LA07bs21
GU133308	*bssA*	Testfeld Süd SIP clone LA07bs22
GU133309	*bssA*	Testfeld Süd SIP clone LA07bs24
Winderl *et al.* n.p.	*bssA*	Testfeld Süd SIP clone S4bs46r2
Winderl *et al.* n.p.	*bssA*	Testfeld Süd SIP clone S4bs46r3
AJ001848	*bssA*	*Thauera aromatica* K127
AF113168	*bssA*	*Thauera aromatica* T1
GU477527	*bssA*	*Thauera* Marmara Sea Sediment clone bssAMSS3
AB066263	*bssA*	*Thauera* sp. DNT-1
JN210553	*bssA*	toluene-degrading Fort Lupton enrichment clone MTOL_1
JQ039472	*bssA*	toluene-degrading microcosm
JQ039471	*bssA*	toluene-degrading microcosm (*Desulfobulbaceae*)
JQ039470	*bssA*	toluene-degrading microcosm clone 1 (*Thauera* sp.)
JQ039468	*bssA*	toluene-degrading microcosm clone 2 (*Thauera* sp.)
JQ039466	*bssA*	toluene-degrading microcosm clone 3 (*Thauera* sp.)
JQ039467	*bssA*	toluene-degrading microcosm clone 5 (*Thauera* sp.)
JQ039469	*bssA*	toluene-degrading microcosm clone 6 (*Thauera* sp.)
GU123598	*bssA*	toluene-degrading Zeitz enrichment clone B1B6
GU123600	*bssA*	toluene-degrading Zeitz enrichment clone B2B2
GU123599	*bssA*	UFZ enrichment clone JBEB1C1
FJ889136	*bssA*	xylene-degrading Zeitz aquifer enrichment clone ZZoX12
FJ889137	*bssA*	xylene-degrading Zeitz aquifer enrichment clone ZZoX76
GU357866	*bssA*-homologue	benzene-degrading *Peptococcaceae* culture BF contig
von Netzer *et al.* n.p.	*bssA*-homologue	*Desulfosporosinus* sp. enrichment bssA1
JX219317	*bssA*-homologue	Flingern aquifer DS03 clone F3B04
JX219318	*bssA*-homologue	Flingern aquifer DS03 clone F3B05
JX219319	*bssA*-homologue	Flingern aquifer DS03 clone F3B11
JX219320	*bssA*-homologue	Flingern aquifer DS03 clone F3B25
JX219300	*bssA*-homologue	Flingern aquifer DS04 clone F4A12
JX219321	*bssA*-homologue	Flingern aquifer DS05 clone F5A04
JX219322	*bssA*-homologue	Flingern aquifer DS05 clone F5A09
JX219323	*bssA*-homologue	Flingern aquifer DS05 clone F5A30

n.p. not published. Continued on the following page

APPENDIX

Accession	gene	Full name
Washer *et al.* n.p.	*bssA*-homologue	methanogenic consortium clone MCbssA3
EF134965	*bssA*-homologue	Pensacola methanogenic consortium clone MCbssA2
JN651976	*bssA*-homologue	Shengli oil reservoir enrichment SL35assA-5
EF123668	*bssA*-homologue	Testfeld Süd aquifer clone B49bs012
EF123669	*bssA*-homologue	Testfeld Süd aquifer clone B49bs020
EF123670	*bssA*-homologue	Testfeld Süd aquifer clone B49bs021
EF123671	*bssA*-homologue	Testfeld Süd aquifer clone B49bs087
EF123672	*bssA*-homologue	Testfeld Süd aquifer clone B49bs088
EF123673	*bssA*-homologue	Testfeld Süd aquifer clone B49bs098
EF123674	*bssA*-homologue	Testfeld Süd aquifer clone B49bs100
EF123677	*bssA*-homologue	Testfeld Süd aquifer clone B49bs105
EF123676	*bssA*-homologue	Testfeld Süd aquifer clone B49bs110
EF123677	*bssA*-homologue	Testfeld Süd aquifer clone B49bs112
EF134966	*bssA*-homologue	toluene-degrading Pensacola enrichment clone MCbssA1
FJ889141	*bssA*-homologue	xylene-degrading Zeitz aquifer enrichment clone ZZpX01
FJ889142	*bssA*-homologue	xylene-degrading Zeitz aquifer enrichment clone ZZpX04
FJ889143	*bssA*-homologue	xylene-degrading Zeitz aquifer enrichment clone ZZpX04b
FJ889144	*bssA*-homologue	xylene-degrading Zeitz aquifer enrichment clone ZZpX06
FJ889145	*bssA*-homologue	xylene-degrading Zeitz aquifer enrichment clone ZZpX06b
FJ889146	*bssA*-homologue	xylene-degrading Zeitz aquifer enrichment clone ZZpX08
FJ889147	*bssA*-homologue	xylene-degrading Zeitz aquifer enrichment clone ZZpX80
FJ889148	*bssA*-homologue	xylene-degrading Zeitz aquifer enrichment clone ZZpX88
JX219324	*bssA*-homologue (OX39)	Amon mud volcano clone BAA23
von Netzer *et al.* n.p.	*bssA*-homologue (OX39)	*Desulfitobacterium aromaticivorans* UKTL bssA1
von Netzer *et al.* n.p.	*bssA*-homologue (OX39)	*Desulfosporosinus* sp. enrichment bssA2
EF123665	*bssA*-homologue (OX39)	*Desulfotomaculum* sp. Ox39 bssA homologue
JX219360	*bssA*-homologue (OX39)	Flingern aquifer DS03 clone F3N20
JX219325	*bssA*-homologue (OX39)	Gulf of Mexico oil sediment 161 clone BG2A01
JX219326	*bssA*-homologue (OX39)	Gulf of Mexico oil sediment 161 clone BG2A02
JX219327	*bssA*-homologue (OX39)	Gulf of Mexico oil sediment 161 clone BG2A05
JX219328	*bssA*-homologue (OX39)	Gulf of Mexico oil sediment 161 clone BG2A06
JX219329	*bssA*-homologue (OX39)	Gulf of Mexico oil sediment 161 clone BG2A14
JX219330	*bssA*-homologue (OX39)	Gulf of Mexico oil sediment 161 clone BG2A18
JX219331	*bssA*-homologue (OX39)	Gulf of Mexico oil sediment 161 clone BG2A21
JX219332	*bssA*-homologue (OX39)	Gulf of Mexico oil sediment 161 clone BG2A23
JX219333	*bssA*-homologue (OX39)	Gulf of Mexico oil sediment 161 clone BG2B03
FJ889139	*bssA*-homologue (OX39)	xylene-degrading Zeitz aquifer enrichment clone ZZoX08
FJ889138	*bssA*-homologue (OX39)	xylene-degrading Zeitz aquifer enrichment clone ZZoX16
FJ889140	*bssA*-homologue (OX39)	xylene-degrading Zeitz aquifer enrichment clone ZZoX16b
FJ889135	*bssA*-homologue (OX39)	xylene-degrading Zeitz aquifer enrichment clone ZZoX56
GU080100	*nmsA*	*Deltaproteobacterium* N47
CU466266	*nmsA*	*Deltaproteobacterium* NaphS2
CU466267	*nmsA*	*Deltaproteobacterium* NaphS3
CU466269	*nmsA*	*Deltaproteobacterium* NaphS6
JX219348	*nmsA*	Flingern aquifer DS03 clone F3B02
JX219349	*nmsA*	Flingern aquifer DS03 clone F3B07
JX219350	*nmsA*	Flingern aquifer DS03 clone F3B13
JX219351	*nmsA*	Flingern aquifer DS03 clone F3B17
JX219352	*nmsA*	Flingern aquifer DS03 clone F3B23
JX219353	*nmsA*	Flingern aquifer DS03 clone F3B26
JX219354	*nmsA*	Flingern aquifer DS03 clone F3B28
JX219355	*nmsA*	Flingern aquifer DS03 clone F3N01
JX219356	*nmsA*	Flingern aquifer DS03 clone F3N03
JX219357	*nmsA*	Flingern aquifer DS03 clone F3N04
JX219358	*nmsA*	Flingern aquifer DS03 clone F3N16
JX219359	*nmsA*	Flingern aquifer DS03 clone F3N19
JX219361	*nmsA*	Flingern aquifer DS03 clone F3N21
JX219362	*nmsA*	Flingern aquifer DS03 clone F3N22
JX219363	*nmsA*	Flingern aquifer DS03 clone F3N23
JX219364	*nmsA*	Flingern aquifer DS03 clone F3N24
JX219365	*nmsA*	Flingern aquifer DS03 clone F3N28
JX219366	*nmsA*	Flingern aquifer DS03 clone F3N29
JX219340	*nmsA*	Guaymas oil seep clone BGN05
JX219341	*nmsA*	Guaymas oil seep clone BGN06
JX219342	*nmsA*	Guaymas oil seep clone BGN07
JX219343	*nmsA*	Guaymas oil seep clone BGN08
JX219344	*nmsA*	Guaymas oil seep clone BGN12
JX219345	*nmsA*	Guaymas oil seep clone BGN15
JX219346	*nmsA*	Guaymas oil seep clone BGN21
JX219347	*nmsA*	Guaymas oil seep clone BGN23
JX219334	*nmsA*	Gulf of Mexico oil sediment 161 clone BG2N02
JX219335	*nmsA*	Gulf of Mexico oil sediment 161 clone BG2N03
JX219336	*nmsA*	Gulf of Mexico oil sediment 161 clone BG2N05
JX219337	*nmsA*	Gulf of Mexico oil sediment 161 clone BG2N14
JX219338	*nmsA*	Gulf of Mexico oil sediment 161 clone BG2N15
JX219339	*nmsA*	Gulf of Mexico oil sediment 161 clone BG2N18
FN675935	*assA*	*Aromatoleum* sp. OcN1
GU453666	*assA*	Arthur Kill OTU 1
GU453641	*assA*	Arthur Kill OTU 2
GU453642	*assA*	Arthur Kill OTU 3
GU453639	*assA*	Arthur Kill OTU 4
GU453640	*assA*	Arthur Kill OTU 5
AM748709	*assA*	*Azoarcus* sp. HxN1
Lueders *et al*. n.p.	*assA*	Brand aquifer clone A14bss02

Accession	gene	Full name
Lueders *et al*. n.p.	*assA*	Brand aquifer clone A14bss22
Lueders *et al*. n.p.	*assA*	Brand aquifer clone A14bss36
Lueders *et al*. n.p.	*assA*	Brand aquifer clone A14bss42
Lueders *et al*. n.p.	*assA*	Brand aquifer clone A14bss49
Lueders *et al*. n.p.	*assA*	Brand aquifer clone A14bss57
Lueders *et al*. n.p.	*assA*	Brand aquifer clone A14bss61
DQ826035	*assA*	*Desulfatibacillum alkenivorans* AK01 assA1
DQ826036	*assA*	*Desulfatibacillum alkenivorans* AK01 assA2
GU453656	*assA*	*Desulfoglaeba alkanexedens* assA
GU453644	*assA*	enrichment culture SDB OTU 2
GU453645	*assA*	enrichment culture SDB OTU 3
GU453662	*assA*	enrichment culture SDB OTU 4
GU453658	*assA*	enrichment culture SDB OTU 5
GU453643	*assA*	Fort Lupton aquifer OTU1
GU453663	*assA*	Fort Lupton aquifer OTU2
GU453664	*assA*	Fort Lupton aquifer OTU3
GU453659	*assA*	Gowanus Canal OTU 1
GU453646	*assA*	Gowanus Canal OTU 2
GU453660	*assA*	Gowanus Canal OTU 3
GU453648	*assA*	Gowanus Canal OTU 4
GU453653	*assA*	Gowanus Canal OTU 5
JX219367	*assA*	Guaymas oil seep clone BGM02
JX219368	*assA*	Guaymas oil seep clone BGM24
HQ704430	*assA*	hexadecane-degrading methanogenic consortium clone assA10
HQ704426	*assA*	hexadecane-degrading methanogenic consortium clone assA05
HQ704425	*assA*	hexadecane-degrading methanogenic consortium clone assA06
HQ704427	*assA*	hexadecane-degrading methanogenic consortium clone assA07
HQ704428	*assA*	hexadecane-degrading methanogenic consortium clone assA08
HQ704429	*assA*	hexadecane-degrading methanogenic consortium clone assA09
HQ704417	*assA*	hexadecane-degrading methanogenic consortium clone assA11
HQ704418	*assA*	hexadecane-degrading methanogenic consortium clone assA12
HQ704419	*assA*	hexadecane-degrading methanogenic consortium clone assA14
HQ704420	*assA*	hexadecane-degrading methanogenic consortium clone assA21
HQ704421	*assA*	hexadecane-degrading methanogenic consortium clone assA22
HQ704422	*assA*	hexadecane-degrading methanogenic consortium clone assA32
HQ704423	*assA*	hexadecane-degrading methanogenic consortium clone assA33
HQ704424	*assA*	hexadecane-degrading methanogenic consortium clone assA36
HQ704431	*assA*	hexadecane-degrading methanogenic consortium clone assA37
HQ704432	*assA*	hexadecane-degrading methanogenic consortium clone assA52
HQ704433	*assA*	hexadecane-degrading methanogenic consortium clone assA56
HQ704434	*assA*	hexadecane-degrading methanogenic consortium clone assA57
HQ704435	*assA*	hexadecane-degrading methanogenic consortium clone assA58
HQ704436	*assA*	hexadecane-degrading methanogenic consortium clone assA59
HQ704437	*assA*	hexadecane-degrading methanogenic consortium clone assA60
HQ704438	*assA*	hexadecane-degrading methanogenic consortium clone assA62
HQ704416	*assA*	hexadecane-degrading methanogenic consortium clone assA63
GU477516	*assA*	Marmara Sea Sediment clone assAMSS1
GU477517	*assA*	Marmara Sea Sediment clone assAMSS2
GU477518	*assA*	Marmara Sea Sediment clone assAMSS3
GU477519	*assA*	Marmara Sea Sediment clone assAMSS4
GU477520	*assA*	Marmara Sea Sediment clone assAMSS5
GU477521	*assA*	Marmara Sea Sediment clone assAMSS6
GU477522	*assA*	Marmara Sea Sediment clone assAMSS7
GU477523	*assA*	Marmara Sea Sediment clone assAMSS8
GU477524	*assA*	Marmara Sea Sediment clone assAMSS9
JF754509	*assA*	methanogenic alkane degrader enrichment clone assA55
JF754507	*assA*	methanogenic alkane degrader enrichment clone assA65
JF754510	*assA*	methanogenic oily sludge enrichment 1 (UNVERIFIED)
JF754508	*assA*	methanogenic oily sludge enrichment clone assA57
JF754511	*assA*	methanogenic oliy sludge enrichment 2 (UNVERIFIED)
GU453652	*assA*	Newtown Creek sediment OTU1 clone
GU453665	*assA*	Newtown Creek sediment OTU2 clone
GU453661	*assA*	Paraffin-degrading San Diego enrichment SDB OTU1 clone
GU453655	*assA*	Passaic River OTU 1
GU453657	*assA*	Passaic River OTU 2
GU453651	*assA*	Passaic River OTU 3
GU453637	*assA*	Passaic River OTU 4
GU453638	*assA*	Passaic River OTU 5
GU453647	*assA*	Passaic River OTU 6
GU453649	*assA*	Passaic River OTU 7
GU453650	*assA*	Passaic River OTU 8
GU453654	*assA*	Passaic River OTU 9
GU453667	*assA*	Propane-degrading Zodletone enrichment OTU1 clone
JN652006	*assA*	Shengli oil reservoir enrichment SL34assA-1
JN652007	*assA*	Shengli oil reservoir enrichment SL34assA-2
JN652008	*assA*	Shengli oil reservoir enrichment SL34assA-3
JN652009	*assA*	Shengli oil reservoir enrichment SL34assA-4
AE001003	*pfl*	*Archaeoglobus fulgidus* pflD2
23114478	*pfl*	*Desulfitobacterium hafniense* DCB-2 fCat
ZP_00099983	*pfl*	*Desulfitobacterium hafniense* DCB-2 pfl
23473769	*pfl*	*Desulfovibrio desulfuricans* G20 fCat
ZP_00130812	*pfl*	*Desulfovibrio desulfuricans* G20 pfl

n.p. not published. Tillmann Lueders, Cheryl Washer, Elizabeth Edward and Christian Winderl are acknowledged for giving access to sequences prior to publication respective unpublished sequences.

A1.3.2 Classification pipeline

Preparation of BLASTN results

(Microsoft Excel 2010, in square brackets function names in German user interface)

1. **Import** BLASTN results to Excel.
2. **Sort** BLASTN results **largest to smallest** according to *Score.*
3. **Remove Duplicates (**menu tab **Data).**
4. **Copy** and **insert** *sequence read length* from Greengenes trimming information files. **Control** with **=cell=cell** is necessary as there are some reads without a BLASTN result.
5. **Calculate** *Identities/Query (%)* from *Identities* and *sequence read length.*
6. **Sort** according *Identities/Query (%)* **largest to smallest.**
7. **Cut-off** at 90 % *Identities/Query* : copy BLASTN results over that value into a new table.
8. Insert column *Phylogeny* for classification: =**MID(*Identifier*;4;10)** [=TEIL(*Identifier*;4;10)].
9. **Sort** according to *Identifier* for the correct phylogenetic order.

Classification according to *Species*:

1. Copy values of column *Identifier* and delete duplicates, call column header *Species#.*
2. **=COUNTIF(*Identifier*;*Species#*)** [=ZÄHLENWENN(*Identifier*;*Species#*)].

Classification according to *Phylogeny:*

1. Copy values of column *Phylogeny* and delete duplicates, call column header *Phylogeny #.*
2. **=COUNTIF(*Phlogeny*;*Phylogeny#*)** [=ZÄHLENWENN(*Phlogeny*;*Phylogeny#*)].

Using the 'FAE-DB name attributor' file

For assigning the complete phylogenetic information to the *Identifier* values from the BLASTN results, a separate Excel-file called ‚FAE-DB name attributor' is required. The assignment is done with the function =**VLOOKUP** [=SVERWEIS]. The matrix consists in the first column of the *Identifier* values, in the second column, of the value *Identifier* and phylogeny (see *Table 7*), combined by a ‚=' (or ARB name and fullname for classification of FAE gene species level).

1. Insert *Identifier* results from excel file with BLASTN results.
2. Copy results of VLOOKUP-function as values to the inserted *Identifier* BLASTN results.
3. Separate the assigned names with the function **Text to Columns** (in menu tab **Data**) and copy the results into the file with the BLASTN results.

A1.3.3 Supporting data

Table 17. Data and format used for calculation of rarefaction shown in Figure 17 with the package 'vegan' in the statistical program environment R.

		Clone libary						
	No. of OTUs	1	2	3	4	5	6	7
Samples	F3A	1	6	0	0	0	0	5
	F4A	0	9	0	0	0	1	7
	F5A	0	7	1	0	0	3	2
	F3B	1	0	0	0	8	3	0

		Pyrosequencing libary										
	OTU	|2-3	|2	|1-2	|1-2-4	|1-1-3-6	|1-1-2	|1-1-1-5	|1-1-1-4	|1-1-1-3	|1-1-1-2	|1-1-1-1
Samples	F3A	0	0	0	0	255	0	408	715	20	3	173
	F3B	0	0	0	462	230	0	1	43	0	4	344
	F4A	0	0	0	0	31	0	433	1328	9	0	101
	F5A	0	0	0	0	183	19	189	777	16	70	132
	G3-52	0	0	0	0	4	309	0	12	22	0	0
	G3-58	0	0	0	0	0	967	20	71	2	0	0

Table 18. Data used for Figure 19: Pyrosequencing reads, T-RF abundance (%) and number of clones.

Pyrosequencing				
OTU	F3A	F4A	F5A	F3B
|1-2-4				462
|1-1-3-6	255	31	183	230
|1-1-2			19	
|1-1-1-5	131	31	115	1
|1-1-1-5-2	277	402	74	
|1-1-1-4	3	6	16	18
|1-1-1-4-1	712	1322	761	25
|1-1-1-3	20	9	16	
|1-1-1-2	3		70	4
|1-1-1-1	173	101	132	344

T-RFLP				
T-RF	F3A	F4A	F5A	F3B
73 bp	2	7	6	
83 bp				3
117 bp	1	0	11	
152 bp				
214 bp	0	1	0	
242 bp	3	1	1	
297 bp				1
324 bp				51
364 bp	14	1	12	20
478 bp	61	60	61	
687 bp	19	29	9	

Clones					
Lineage	T-RF (bp)	F3A	F4A	F5A	F3B
Strain N47	328	0	0	0	8
T-Cluster	364	0	1	3	3
F1-Cluster	476	5	7	2	0
Deltaproteobacteria	115	0	0	1	0
F2-Cluster	682	3	9	4	0
Clostridia	52	3	0	3	0
Betaproteobacteria	34/37	1	0	0	1

Table 19. Pyrosequencing read data used for Figure 20B.

OTU	58 °C	52 °C
|1-1-3-6		4
|1-1-2	967	309
|1-1-1-5-2	20	
|1-1-1-4	70	7
|1-1-1-4-1	1	5
|1-1-1-3	2	22

A2 Figure index

A3 Table index

A4 Lebenslauf – Frederick von Netzer

20. Juni 1985	Geboren in Wuppertal-Elberfeld
2004	Abitur Otto-Hahn-Gymnasium Furtwangen
2004-2008	Studium der Biologie, Universität Konstanz
2008	Diplomarbeit „Charakterisierung des Abbaus von 3-(4-Sulfophenyl)butyrat in *Comamonas testosteroni*“, Arbeitsgruppe Prof. Ph.D. A. M. Cook, Universität Konstanz
2009-2012	Doktorarbeit in der Arbeitsgruppe molekulare Ökologie am Institut für Grundwasserökologie (Helmholtz Zentrum München), Neuherberg, betreut durch Dr. Tillmann Lueders und Prof. Dr. Rainer U. Meckenstock.

The work presented in this thesis was funded by the DFG priority programme 1319 „Biologische Umsetzungen von Kohlenwasserstoffen in Abwesenheit von Sauerstoff: Von molekularer zu globaler Ebene“ (grant LU 1188/4-1) and conducted at the Institute of Groundwater Ecology at the Helmholtz Zentrum Munich.

This work would not have been possible without the scientific as well as moral support from my supervisors, cooperation partners and colleagues.

At first, I would like to thank Dr. Tillmann Lüders. Not only for giving me the opportunity to work with this exciting project in the frame of the aforementioned DFG priority programme 1319 and for supervising me, but also for his sedulous help, advice, patience and trust. He never gave up on me in my darkest moments, when I was in doubt about my results and me. Thank you very much!

Of course I thank also Prof. Dr. Rainer U. Meckenstock, not only for accepting me as a PhD at the IGÖ but also for his interest, scientific input, discussions and advice, which was not limited to the thesis committees. I thank also Prof. Dr. Matthias Boll, for joining and actively participating in my thesis committee despite spatial and temporal difficulties and showing interest in my work.

I am also thankful that Prof. Dr. Wolfgang Liebl accepted the work-intensive part as second reviewer at such short notice. In this spirit I want also to thank Prof. Dr. Siegfried Scherer for taking the chair of the examination committee.

The thesis would not have so much content without the different fruitful cooperation projects within and outside of the DFG priority program. My thanks go out to Giovanni Pilloni, Sara Kleindienst, Katrin Knittel, Sabrina Beckmann, Friederike Gründger und last but not least Martin Krüger for sharing samples and enlightening discussions.

My thanks go out to all the colleagues who helped by reading and correcting various drafts of this thesis. I wish also to thank all my active and former colleagues of the molecular ecology group for their support in matters of daily routine work and the excellent working atmosphere. This applies of course for the entire Institute for Groundwater Ecology, too!

Therefore, I thank also the usual suspects of the „Early Lunch Group”, all helpful and well-disposed fellow PhD candidates, PostDocs, technicians, group leaders and especially the comrades of office 130 (and recently also of office 12) for scientific discussions, convivial gatherings within and outside work.

Last, but not at all least, I thank my parents, relatives and friends for their on-going remote moral support (“Ah, come on, it’s just your hobby!”).