High Diversity of Anaerobic Alkane-Degrading Microbial Communities in Marine Seep Sediments Based on (1-methylalkyl)succinate Synthase Genes

Marion H. Stagars¹, S. Emil Ruff¹, ²†, Rudolf Amann¹ and Katrin Knittel¹ *

¹ Department of Molecular Ecology, Max Planck Institute for Marine Microbiology, Bremen, Germany, ² HGF MPG Joint Research Group for Deep-Sea Ecology and Technology, Max Planck Institute for Marine Microbiology, Bremen, Germany

Alkanes comprise a substantial fraction of crude oil and are prevalent at marine seeps. These environments are typically anoxic and host diverse microbial communities that grow on alkanes. The most widely distributed mechanism of anaerobic alkane activation is the addition of alkanes to fumarate by (1-methylalkyl)succinate synthase (Mas). Here we studied the diversity of MasD, the catalytic subunit of the enzyme, in 12 marine sediments sampled at seven seeps. We aimed to identify cosmopolitan species as well as to identify factors structuring the alkane-degrading community. Using next generation sequencing we obtained a total of 420 MasD species-level operational taxonomic units (OTU0.96) at 96% amino acid identity. Diversity analysis shows a high richness and evenness of alkane-degrading bacteria. Sites with similar hydrocarbon composition harbored similar alkane-degrading communities based on MasD genes; the MasD community structure is clearly driven by the hydrocarbon source available at the various seeps. Two of the detected OTU0.96 were cosmopolitan and abundant while 75% were locally restricted, suggesting the presence of few abundant and globally distributed alkane degraders as well as specialized variants that have developed under specific conditions at the diverse seep environments. Of the three MasD clades identified, the most diverse was affiliated with Deltaproteobacteria. A second clade was affiliated with both Deltaproteobacteria and Firmicutes likely indicating lateral gene transfer events. The third clade was only distantly related to known alkane-degrading organisms and comprises new divergent lineages of MasD homologs, which might belong to an overlooked phylum of alkane-degrading bacteria. In addition, masD geneFISH allowed for the in situ identification and quantification of the target guild in alkane-degrading enrichment cultures. Altogether, these findings suggest an unexpectedly high number of yet unknown groups of anaerobic alkane degraders and underline the need for comprehensive surveys of microbial diversity based on metabolic genes in addition to ribosomal genes.

Keywords: alkyl succinate synthase, MasD, AssA, sulfate reduction, anaerobic alkane oxidation, microbial diversity, geneFISH
INTRODUCTION

Alkanes are found throughout nature and belong to the most abundant organic compounds in the biogeoosphere (Wilkes et al., 2002). They are main components of crude oil and natural gas and are either formed by living organisms (Tissot and Welte, 1984) or through geological transformation of biomass (Claypool and Kvenvolden, 1983). In the marine environment, alkanes occur in gas hydrates, petroleum-rich hydrothermal sediments, seeps or areas anthropogenically contaminated such as occurred during the Exxon Valdez oil spill in 1989 or the Deep Water Horizon blowout in 2010. Marine seeps can differ remarkably in their hydrocarbon composition and concentration. Methane seeps mainly emit methane in micromolar to millimolar ranges m^{-2} d^{-1} (e.g., Hydrate Ridge in the NE Pacific, Suess et al., 1999), gas seeps emit a substantial amount of C2-C5 gasses in addition to methane (e.g., Mediterranean Amon mud volcano, Mastalerz et al., 2009) and hydrocarbon seeps emit a broad range of alkanes, alkenes and aromatics (e.g., at Guaymas Basin in the Gulf of California, Byrne and Emery, 1960; Simoneit and Welte, 1984) or through geological transformation of biomass (Claypool and Kvenvolden, 1983). In the marine environment, alkanes occur in gas hydrates, petroleum-rich hydrothermal sediments, seeps or areas anthropogenically contaminated such as occurred during the Exxon Valdez oil spill in 1989 or the Deep Water Horizon blowout in 2010. Marine seeps can differ remarkably in their hydrocarbon composition and concentration. Methane seeps mainly emit methane in micromolar to millimolar ranges m^{-2} d^{-1} (e.g., Hydrate Ridge in the NE Pacific, Suess et al., 1999), gas seeps emit a substantial amount of C2-C5 gasses in addition to methane (e.g., Mediterranean Amon mud volcano, Mastalerz et al., 2009) and hydrocarbon seeps emit a broad range of alkanes, alkenes and aromatics (e.g., at Guaymas Basin in the Gulf of California, Byrne and Emery, 1960; Simoneit and Lonsdale, 1982). Many microorganisms are able to utilize these hydrocarbons as their carbon source (Widdel et al., 2010). In seep sediments, degradation mainly takes place under anoxic conditions as oxygen is generally depleted within the first few millimeters of the sediment. A large fraction of sulfate reduction (SR) at gas and hydrocarbon seeps is fueled by the anaerobic oxidation of methane (AOM, Reeburgh, 2007). However, as indicated by a global median ratio of SR to AOM of 10.7, a major part of total SR is fueled by the oxidation of non-methane hydrocarbons, in particular the oxidation of alkanes (Bowles et al., 2011).

Microbial anaerobic oxidation of alkanes has been described for a large range of alkanes: gaseous (C2–C5) alkanes (Kniemeyer et al., 2007; Savage et al., 2010; Jaekel et al., 2012; Adams et al., 2013; Bose et al., 2013), mid-chain (C6–C12) alkanes (Ehrenreich et al., 2000; Davidova et al., 2006) and long-chain (C13–C20) alkanes (Ackersberg et al., 1991; So and Young, 1999; Zengler et al., 1999; Cravo-Laureau et al., 2004a). Cultivated anaerobic alkane degraders use sulfate (sulfate-reducing bacteria, SRB), nitrate, manganese or ferric iron Fe(III) as electron acceptors (Weelink et al., 2009; Widdel et al., 2010). Furthermore, hydrocarbon-degrading enrichment cultures have been established under methanogenic conditions (Zengler et al., 1999; Chang et al., 2006; Berdugo-Clavijo and Gieg, 2014; Embree et al., 2014). Isolated or enriched anaerobic alkane degraders belong to two phyla: Proteobacteria and Firmicutes (Figure 1). Only recently an archaeon, Archaeoglobus fulgidus, an alkylsuccinate synthase activity was hypothesized for pyruvate formate lyase (PFL) based on a higher similarity of PFLD with AssA compared with reference PFLs (Khelifi et al., 2014). Yet, there are only few environmental studies on AssA/MasD diversity in contaminated soils and groundwater (Wang et al., 2012; Zhou et al., 2012; von Netzer et al., 2013), contaminated river sediments and aquifers (Callaghan et al., 2010) or marine hydrocarbon-impacted sediments (Acosta-González et al., 2013; von Netzer et al., 2013; Johnson et al., 2015). Considering the high number of long branches in AssA/MasD phylogenetic trees these studies point to the existence of a broad diversity of microorganisms involved in the degradation of alkanes.

DSS clades that are active in the oxidation of short- and long-chain alkanes (Kleindienst et al., 2014). Cultivation-independent studies using clone libraries or next generation sequencing techniques (Teske et al., 2002; Orcutt et al., 2010; Ruff et al., 2015) and alkane incubation studies (Savage et al., 2010; Adams et al., 2013; Bose et al., 2013) showed a high diversity of SRB at hydrocarbon seeps suggesting potential for a much larger diversity of alkane degraders as is currently known from available isolates.

Several biochemical reactions have been described for alkane activation under anoxic conditions (Callaghan, 2013; Musat, 2015). The most well-described and particularly dominant pathway is the activation of alkanes to fumarate yielding alkylsuccinates. The enzyme involved in this initial activation step is a glycyl radical enzyme of the pyruvate formate lyase family, the (1-methylalkyl)succinate synthase, Mas (Grundmann et al., 2008), which has also been referred to as alkylsuccinate synthase, Ass (Callaghan et al., 2010). The putative catalytic subunit of Mas is subunit D (MasD), which is equivalent to Ass subunit A (AssA). Alkane activation by MasD has been described for a range of n-alkanes and cycloalkanes in sulfate-reducing and nitrate-reducing isolates and enrichment cultures (Kropp et al., 2000; Callaghan et al., 2006; Kniemeyer et al., 2007; Grundmann et al., 2008; Musat et al., 2010; Jaekel et al., 2015). Furthermore, fumarate addition has also been suggested for methanogenic alkane degradation based on the detection of assA/masD genes in enrichments (Davidova et al., 2011; Mbadinga et al., 2011; Zhou et al., 2012; Aitken et al., 2013; Cheng et al., 2013). Recently, an assABC operon was detected on a Smithella single cell genome (Tan et al., 2014). Considering the wide distribution of alkane activation via fumarate addition, MasD/AssA serves as valid biomarker for anaerobic alkane degradation. For alkane degradation by archaean Archaeoglobus fulgidus, an alkylsuccinate synthase activity was hypothesized for pyruvate formate lyase (PFL) based on a higher similarity of PFLD with AssA compared with reference PFLs (Khelifi et al., 2014). Yet, there are only few environmental studies on AssA/MasD diversity in contaminated soils and groundwater (Wang et al., 2012; Zhou et al., 2012; von Netzer et al., 2013), contaminated river sediments and aquifers (Callaghan et al., 2010) or marine hydrocarbon-impacted sediments (Acosta-González et al., 2013; von Netzer et al., 2013; Johnson et al., 2015). Considering the high number of long branches in AssA/MasD phylogenetic trees these studies point to the existence of a broad diversity of microorganisms involved in the degradation of alkanes.

In this study we addressed the diversity of the alkane-degrading microbial community by massive parallel 454-tag sequencing of masD genes retrieved from 12 globally distributed marine seep sediments and correlate their characteristics with environmental parameters like hydrocarbon composition, water depth, temperature and SR rates. We hypothesize that in marine seep sediments the diversity of the anaerobic alkane-degrading microbial community is much higher than previously known, including new deeply branching taxonomic lineages. Due to the narrow range of substrate used by isolated alkane degraders (Musat, 2015) we hypothesize that the MasD-microbial community at methane and gas seeps differ significantly from
FIGURE 1 | Phylogenetic tree showing the affiliation of 16S rRNA gene sequences from isolated or enriched anaerobic hydrocarbon degraders to selected reference sequences of the domain Bacteria. Nitrate-reducing bacteria are printed in blue, sulfate-reducing bacteria in red, iron-reducing bacteria in purple, phototrophic bacteria in green, fermentative bacteria in orange and syntrophic bacteria degrading hydrocarbons in a consortium under methanogenic conditions in light blue. Georgfuchsia toluolica, printed in light green, has been shown to use Fe(III), Mn(IV), and nitrate as terminal electron acceptor for growth on aromatic compounds. Cd. Methylomirabilis oxyfera, printed in brown, can oxidize methane anaerobically by utilizing oxygen produced internally from dismutation of nitric oxide into nitrogen and oxygen gas. Substrate usage is given within parenthesis. The bar represents 10% estimated sequence divergence.
that at hydrocarbon seeps. As an integral part of the study, we optimized the geneFISH protocol to identify alkane-degrading communities in situ.

MATERIALS AND METHODS

Sampling Sites
Sediments were sampled from seven globally distributed marine seeps differing in their geographical, biological, chemical, and geological features (Table 1, Figure 2). Investigated methane seeps predominantly releasing methane were located on the Cascadian Margin at Hydrate Ridge (Pacific Ocean, station HR19), on the Hikurangi continental margin (New Zealand, Wairarapa, station NZ315) and in the central North Sea (Tommeliten, station Tomm). Gas seeps at the Mediterranean Amon mud volcano (AMV, stations AMV760, AMV825) release significant amounts of other gasses (C2–C5 alkanes) in addition to methane. Hydrocarbon seeps, in the northern (stations GoM4463 and GoM156), and southern Gulf of Mexico (Chapopote Asphalt Volcano: stations GoM140, GoM13, GoM17) and in the Guaymas Basin (stations GB4573 and GB4484) in the Gulf of California are characterized by seepage of complex hydrocarbons (Table 1).

| Table 1 | Description of sampling sites. |
|---|---|
| **Seep type** | **Sampling site** | **Station** | **Cruise; year** | **Latitude** | **Longitude** | **Water depth [m]** | **Hydrocarbon composition** | **Reference** |
| Hydrocarbon seeps | Chapopote Asphalt Volcano (Campeche Knolls, Southern Gulf of Mexico) | GoM140 | SO174; 2003 | 21°54.00′N | 93°26.40′W | 2902 | Mainly methane and C2–C4 alkanes, few C29–C32 alkanes | Orcutt et al., 2010; Wegener, unpublished |
| | GoM13 | M67/2; 2003 | 21°53.99′N | 93°26.18′W | 2908 | Mainly steranes, hopanes, few paraffins, alkylbenzenes, cycloalkanes | |
| Northern Gulf of Mexico | GoM17 | M67/2; 2003 | 21°53.94′N | 93°26.14′W | 2908 | Mainly asphaltenes (> C20) | |
| | GoM4463 | LeDEN; 2002 | 27°44.48′N | 91°19.04′W | 504 | Aromatic hydrocarbons and methane | |
| | GoM156 | SO174; 2003 | 27°46.95′N | 91°30.47′W | 550 | Methane, ethene, C15 alkanes, isoprenoids, naphthalene, toluene | |
| Guaymas Basin (Gulf of California) | GB4573 | AT15-56; 2009 | 27°00.69′N | 111°24.26′W | 2100 | C12–C18 alkanes, cycloalkanes, diverse aromatics | Bazylnski et al., 1988 |
| | GB4484 | AT15-56; 2009 | 27°00.64′N | 111°40.96′W | 2000 | C12–C18 alkanes, cycloalkanes, diverse aromatics | |
| **Gas seeps** | Amon Mud Volcano (Mediterranean Sea) | AMV760 | M70/2; 2006 | 32°22.129′N | 31°43E | 1122 | Mainly methane and C2–C4 alkanes | Mastalerz et al., 2009; Grünke et al., 2011 |
| | AMV825 | M70/2; 2006 | 32°22.128′N | 31°42E | 1122 | Mainly methane and C2–C4 alkanes | |
| Methane seeps | Hydrate Ridge (Cascadia Margin) | HR19 | SO148-1; 2000 | 44°34.18′N | 125°08.80′W | 777 | Mainly methane (>95%), a few other gaseous alkanes | Boertius et al., 2000 |
| | Wairarapa (New Zealand) | NZ315 | SO191-3; 2007 | 41°46.28′S | 175°25.78′E | 1058 | Mainly methane (>99%), a few other gaseous alkanes | Ruff et al., 2013 |
| | Tommeliten (North Sea) | Tomm | AL267; 2004 | 56°29.90′N | 02°59.80′E | 75 | Mainly methane (>99%), a few other gaseous alkanes | Wegener et al., 2008 |

Nucleic Acid Extraction
DNA was extracted by mechanical, chemical and enzymatic cell lysis steps according to the protocol of Zhou et al. (1996) from sediments frozen immediately after sampling at −20°C. Extracted nucleic acids were washed with ice-cold ethanol (80% v/v), centrifuged at 14,000 × g for 10 min at RT, dried for 15–30 min at RT and gently resuspended without pipetting in TE buffer (10 mM Tris-HCl, pH 8.0, 1 mM EDTA, pH 8.0) for 1 h at 4°C.

MasD Amplification and Pyrosequencing
The masD/assA gene was amplified using one of two primer pairs: 7757f-1-2 (TCG GAC GCG TGC AAC GMY CTG A; MasD amino acid position 395 in strain HxN1; accession number CAO03074)/8543r (TCG TCR TTG CCC CAY TTN GG; position 657 in HxN1) or primer pair 7766f (TGT AAC GGC ATG ACC ATT GCG CT; position 398 in HxN1)/8543r (TCG TCR TTG CCC CAY TTN GG) (von Netzer et al., 2013). Primers were barcoded and extended with an SfiI restriction site at the 5′ end for ligation with the 454-adapters. For each sample, eight replicate PCRs (20 μl volume) per primer pair were carried out containing, 0.5 μl primer each, 250 μl dNTPs, 0.3 μg μl−1 BSA, 1 × PCR buffer, 0.25 U Taq polymerase (5Prime, Germany) under the following conditions: initial denaturation...
at 95°C for 5 min, followed by 35 cycles of denaturation (96°C, 1 min), annealing (58°C, 1 min), elongation (72°C, 2 min), and a final elongation step (72°C, 10 min). Replicate PCR reactions of both primer pairs were pooled and the 800 bp-amplicons were then extracted from an agarose gel (1.5% w/v) and purified using the MiniElute PCR Purification Kit (Qiagen) according to the manufacturer’s recommendations. Massive parallel tag sequencing of the amplicons was carried out on a 454 Life Sciences GS FLX sequencer (Roche, Basel, Switzerland) at the MaxPlanck-Genome-Center, Cologne, Germany.

For pairwise comparison of 16S rRNA genes and MasD (see below) we amplified and Sanger sequenced masD from the following alkane-degrading strains and enrichments (Table 2): PF2803 (DSM16219, Cravo-Laureau et al., 2004b), LM2801 (DSM18843, Cravo-Laureau et al., 2007), and Propane60-GuB (Kniemeyer et al., 2007).

Sequence Processing
Raw reads were submitted to a rigorous quality control using a mothur version 1.29.1 routine (Schloss et al., 2009) including stringent quality filtering and chimera check using UCHIME (Edgar et al., 2011). Sequences were removed from the data set if they had ≥1 mismatch to the forward primer, were <200 bp in length, had >8 homopolymers, had an average quality score <20 (qthreshold = 20) or contained any ambiguities. Sequences were then translated in silico (RevTrans 1.4 Server) and screened for MasD based on the presence of the amino acid motive FECIR, FECIK, FECQR, FECVR, FDCIR or FDNIA. Ninety percent of MasD/AssA proteins in our database possess one of these motives at position 435 (HxN1). Sequences with stop codons were removed from the dataset. Sequences were further checked for the presence of the catalytically active cysteine (Grundmann et al., 2008) at position 477 of HxN1. We could not confirm the presence of the conserved motif (RVXG) that harbors the radical-storing glycine (Becker et al., 1999) and is characteristic for all glycy radical enzymes including MasD and BssA as it lies outside the region amplified by our primers (position 811; HxN1).

Establishment of a Protein Database for Glycyl Radical Enzymes
We built a comprehensive protein database containing more than 10,000 sequences for MasD, BssA, NmsA and pyruvate formate lyase (Pfl). Sequences were retrieved from this study or from public databases GenBank, NCBI and DDBJ or from publically available complete or draft genomes, single cell genomes, and metagenomes accessible via the Integrated Microbial Genome (IMG) with microbiome samples (IMG/M) system at JGI1 (Markowitz et al., 2014). The mas operon sequence from strain HxN1 (AM748709) was used for a BlastX search versus the IMG/M database resulting in >1400 potential glycyl radical enzyme sequences. We then imported all amino acid sequences retrieved from these databases along with our own sequences retrieved from seep sediments into the software package ARB (Ludwig et al., 2004) and aligned them using MAFFT v7 (Katoh and Standley, 2013). Manual correction of frame shifts based on BLASTX was necessary for about 15 of our seep sequences due to insertions or deletions caused by 454-pyrosequencing. The amino acid alignment is provided in Supplementary information (Supplementary Table S1).

Definition of OTU Cut-Offs for MasD Amino Acid Sequences
Taxonomic units for MasD were defined based on 15 alkane-degrading pure cultures by pairwise comparison of their amino acid sequences with the corresponding 16S rRNA gene sequences. A list of strains used for this analysis is provided in Table 2. The sequence difference D and similarities S (S = 1 – D) were calculated. Finally, the similarity of

1http://img.jgi.doe.gov/m
TABLE 2 | Alkane-degrading strains used for calculation of MasD OTU thresholds.

| Strain/enrichment | e−-acceptor | Phylogenetic affiliation | MasD | 16S rRNA |
|-------------------|-------------|--------------------------|------|---------|
| Strain HxN1       | NO₃⁻        | Betaproteobacteria        | CAO03074 | AF331975 |
| Strain OcN1       | NO₃⁻        | Betaproteobacteria        | CBK27727 | AF331976 |
| Strain HdxN1      | NO₃⁻        | Gammaproteobacteria       | NC_014366 | AF331974 |
| Strain BuS5       | SO₄²⁻       | Deltaproteobacteria       | AXAM00000002 | EF077225 |
| Propane60GuB      | SO₄²⁻       | Firmicutes                | LN879422 | EF077227 |
| Butane12-Gme      | SO₄²⁻       | Deltaproteobacteria       | unpublished | EF077226 |
| Strain PL12       | SO₂⁻        | Deltaproteobacteria       | LC102219 | AB488588 |
| Desulfoglaeba alkanexedens (str. ALDC) | SO₄²⁻ | Deltaproteobacteria | GU453566 | DG030457 |
| Strain TD3        | SO₄²⁻       | Deltaproteobacteria       | Unpublished | XB0922 |
| Strain HxN3       | SO₂⁻        | Deltaproteobacteria       | unpublished | AF141881 |
| Desulfatibacillus alkenivorans (str. AK-01) | SO₂⁻ | Deltaproteobacteria | CP001322.1 | NR_074962 |
| Desulfatibacillus alkenivorans (str. PF2803) | SO₂⁻ | Deltaproteobacteria | LN879420 | NR_025795 |
| Desulfatibacillus aliphaticivorans (str. CV2803) | SO₂⁻ | Deltaproteobacteria | AUCT0100049 | AY184360 |
| Desulfatirera olefinivorans (str. LM2801) | SO₂⁻ | Deltaproteobacteria | LN879421 | DQ826724 |
| Strain PrN3       | SO₂⁻        | Deltaproteobacteria       | Unpublished | Y17501 |

MasD amino acid sequence pairs were plotted versus the similarity of the 16S rRNA gene sequence pairs of the same strains.

**Phylogenetic Tree Reconstruction**

The phylogenetic tree based on small subunit ribosomal RNA genes was calculated with nearly full-length sequences (>1350 bp) available in Arb Silva database release 111 (Pruesse et al., 2007) by neighbor-joining analysis in combination with filters which consider only 50% conserved regions of the 16S rRNA. Partial sequences were subsequently inserted into the reconstructed consensus tree by parsimony criteria, without allowing changes in the overall tree topology. MasD-based phylogenetic tree was constructed by maximum likelihood analysis (PhyML algorithm, Blosum 62 substitution model) considering 95 amino acid positions (position 436 to 506, strain HxN1; CAO03074) using 441 deduced amino acid sequences. Only one representative sequence per MasD family level OTU was shown in the final tree.

**Community Diversity Analysis**

Sequence abundance tables were generated by clustering the retrieved MasD seep sequences at 96% amino acid identity based on 120 amino acid positions (Pos. 398-500, HxN1) using a distance matrix in Mothur (Schloss et al., 2009) based on 120 amino acid positions (Pos. 398-500, HxN1) retrieved MasD seep sequences at 96% amino acid identity. Sequence abundance tables were generated by clustering the sequences. Community Diversity Analysis was calculated with nearly full-length sequences (>1350 bp) available in Arb Silva database release 111 (Pruesse et al., 2007) by neighbor-joining analysis in combination with filters which consider only 50% conserved regions of the 16S rRNA. Partial sequences were subsequently inserted into the reconstructed consensus tree by parsimony criteria, without allowing changes in the overall tree topology. MasD-based phylogenetic tree was constructed by maximum likelihood analysis (PhyML algorithm, Blosum 62 substitution model) considering 95 amino acid positions (position 436 to 506, strain HxN1; CAO03074) using 441 deduced amino acid sequences. Only one representative sequence per MasD family level OTU was shown in the final tree.

**Design of Probes for masD Detection**

To set-up a masD geneFISH assay, an enrichment culture with n-butane was established under sulfate-reducing conditions from hydrocarbon seep sediments of site GB4573.5 ml sediment slurry was made using a 1:1 mix of sediment from the upper 10 cm with artificial anoxic seawater (Widdel and Bak, 1992) and incubated in hungate tubes containing 5 ml anoxic media for anaerobic sulfate-reducing microorganisms (Widdel and Pfennig, 1982). As substrate, 1 bar n-butane gas (Messer Griesheim GmbH, Krefeld, Germany) was added to the headspace. All tubes were kept under N₂/CO₂ (90/10, v/v) atmosphere and were horizontally incubated at 28°C. When sulfide production reached 15 mM, which minimizes the total within-cluster variance. A network was built based on a presence absence matrix. The network vertices (nodes) were plotted using a Fruchterman and Reingold (1991) force-directed algorithm, which causes an increase in the nodes attraction to each other with increasing similarity between them; the more OTU shared between two samples, the closer they are in the network. Analyses were carried out with the R statistical environment and the packages vegan (Oksanen et al., 2012), ggplot2 (Wickham, 2009) and network (Butts, 2008) in addition to custom R scripts. In silico coverage (C) of MasD sequences was calculated per station according to the following equation

$$ C = \left(1 - \frac{n}{N}\right)^*100 $$

where n is the number of singletons (SSO_{abs} + SSO_{rel}) and N the total number of sequences analyzed.

**TABLE 2 | Alkane-degrading strains used for calculation of MasD OTU thresholds.**

*Stagars et al. MasD Diversity at Marine Seeps*
enrichment GB4573_14 and masD genes were amplified using primers ass/bss F and ass/bss R as described previously (Callaghan et al., 2010). Cloning and sequencing of 689 bp-amplicons was performed as described previously (Kleindienst et al., 2012). Based on 28 sequenced clones used to represent the breadth of diversity of masD in our sample, we designed a probe mix consisting of 9 masD-targeting dsDNA polynucleotide probes (316 bp; GC content 65%) using the PolyPro software (Moraru et al., 2010, 2011, Supplementary Table S2). The individual probes showed >75% nucleotide sequence identity to the target regions of all members of the masD gene clusters retrieved from the GB4573_14 enrichment and >72% nucleotide sequence identity to all retrieved MasD OTUs.

For probe synthesis, plasmid DNA was extracted from 9 selected clones using the SpinMiniprep Kit (Qiagen, Hilden, Germany) and dsDNA probes were synthesized by PCR with enzymatic incorporation of Dig-labeled nucleotides (dUTP) using the PCR Dig Probe Synthesis Kit (Roche, Diagnostics, Mannheim, Germany). As negative control, probe NonPolyPr350 was synthesized. Synthesized dsDNA probes were purified with the GeneClean Turbo kit (Q-Biogene).

Fluorescence In Situ Hybridization of masD (geneFISH)
One ml was subsampled from the enrichment culture GB4573_14 at three different time points, fixed with formaldehyde (2% final concentration in 1x PBS (pH 7.4) for 1 h at RT) and an aliquot was filtered on polycarbonate filters (GTTP, pore size: 0.2 μm). Filters were incubated in 0.01 M HCl for 10 min at room temperature to inactivate endogenous peroxidases, followed by incubation in 10 mg/ml lysozyme for 1 h at 37°C to permeabilize cell walls. CARD-FISH targeting 16S rRNA with probe DS5658 (Manz et al., 1998) specific for the deltaproteobacterial clade of Desulfosarcina/Desulfococcus was performed as described earlier (Ishii et al., 2004) using Alexa488-labeled tyramide for signal amplification. Following hybridization of 16S rRNA, inactivation of the probe-coupled horseradish peroxidase enzymes (HRP) was achieved by incubations of the filters in 3% H2O2 in 1xPBS for 30 min at RT followed by incubation in 0.1 M HCl for 10 min. RNAs in the cells were digested by incubating the filters in RNase solution (0.5 U μl−1 RNase I, Ambion), 30 μg ml−1 RNase A (Sigma), 0.1 M Tris-HCl pH 8 for 4 – 5 h at 37°C. Filter sections, with either the masD probe mix or negative control probe NonPolyPr350 were incubated in hybridization buffer containing 45% formamide as calculated by the PolyPro software. After initial denaturation at 85°C for 25 min, hybridization lasted for 18 – 22 h at 50°C followed by binding of the anti-Dig HRP-conjugated antibody (Fab fragments) and signal amplification with a Alexa 594-labeled tyramide. Filter sections were embedded in SlowFadeGold antifade reagent (Invitrogen), containing 1 μg ml−1 4′,6-diamidino-2-phenylindole (DAPI). Microscopy was performed on an epifluorescence microscope (Axioplan, Carl Zeiss). To correct for false positive signals, i.e., extracellular signals and non-specific probe binding to particles or filter matrix, we subtracted negative control counts from the total gene counts.

Nucleotide Sequence Accession Numbers
MasD clone sequences from enrichment culture GB4573_14 have been deposited in the EMBL, GenBank and DDBJ nucleotide sequence database under accession numbers LN610408 to LN610424 and HG764719 to HG764728 and MasD sequences from strains PF2803 (DSM 16219), LM2801 (DSM 18843) and the phylotype Propane60GuB under numbers LN879420-LN879422. Raw masD pyrotag sequences have been stored in the sequence read archive under SRA bioproject number 278019.

RESULTS

Definition of OTU Cut-Offs for masD
Thresholds for OTU clustering were calculated by linear correlation of 15 cultured alkane-degrading strains with both 16S rRNA genes and MasD sequence information (Table 2). The taxonomic threshold for microbial species is 98.7% for 16S rRNA genes (Stackebrandt and Ebers, 2006; Yarza et al., 2014), which corresponded to a threshold of 96% for MasD (OTU 0.96; Figure 3). The 16S rRNA gene genus-level (94.5%) and family level thresholds (86.5%; Yarza et al., 2014) corresponded to cutoff values for MasD of 86% (OTU 0.86) and 72% (OTU 0.72), respectively.

Diversity of MasD Community in Seep Sediments
The diversity of anaerobic alkane-degrading bacteria was studied by pyrosequencing of masD present in sediments from 12 globally distributed stations at methane, gas or hydrocarbon seep sites. We retrieved a total of 12,745 raw sequences. Strict quality filtering of raw sequences was an important part of our data analysis as sequencing errors and chimeras affect OTU clustering of protein sequences more severely than clustering of nucleic acid sequences and would result in an overestimation of environmental MasD diversity. After filtering, 40% of raw reads (5,131 masD sequences, 652 nmsA sequences and 79 bssA sequences) were included in further analyses. After translation and clustering the protein sequences on species-level (OTU 0.96) we obtained 420 MasD OTU 0.96, 1 NmsA OTU 0.96, and 1 BssA OTU 0.96 (Table 3). The retrieval of nmsA and bssA sequences from our seep sites was a result of unspecific binding of our masD primers. It is likely then that diversity of bssA and nmsA is not fully covered by these primers and is omitted from further analysis.

MasD Richness and Evenness
Observed richness of alkane degraders ranged between 13 and 93 MasD species-level OTUs (Table 3). Coverage ranged between 83 and 94% indicating sufficient sampling effort (Table 3, Supplementary Figure S1) except for the site with the highest inverse Simpson index, GoM4463, for which the coverage was only 74%. OTU abundance for other taxonomic levels is provided in Supplementary Table S3. After subsampling to standardize sequencing effort, OTU 0.96 diversity differed by a maximum
FIGURE 3 | Pairwise comparison of MasD amino acid and 16S rRNA gene sequence identity for 15 strains capable of anaerobic alkane degradation. Accession numbers of used strains are given in Table 2. The intersection of vertical dashed lines and the regression line gives the cutoff values for defining a species (96%), genus (86%), and family (72%) based on MasD identity corresponding to the proposed values based on 16S rRNA genes (Yarza et al., 2014).

factor of 2.25 between sites. Diversity was highest at GoM4463 (27 OTU0.96) and GoM156 (21 OTU0.96) but clearly lower at GB4484 and all methane seep sites (13–18 OTU0.96). Inverse Simpson diversity index (1/D), which takes into account both OTU richness and evenness, ranged between 3 and 9.

Similarity of MasD Communities at Different Seep Sites

The MasD community structure at different seep sites was visualized by NMDS based on amino acid sequences clustered at species level (OTU0.96; Figure 4A). Sites were grouped according to hydrocarbon type at the individual sites, i.e., methane (=methane seeps), gaseous hydrocarbons (=gas seeps), or diverse hydrocarbons (=hydrocarbon seeps) and connected to the weighted centroid of the within-group distances. MasD diversity of samples within these three groups showed a high similarity with a high shared proportion of taxa (31%). Dissimilarity was most pronounced between methane seeps and all other sites. Methane seep sites HR19, NZ315 and Tomm revealed quite similar community structures (43% shared taxa) but were clearly different from those at gas and diverse hydrocarbon seeps (<6% shared taxa). AMV gas seeps host MasD communities partially overlapping with those of the hydrocarbon seeps GB and GoM. The frequency of co-occurrence was highest between the hydrocarbon seeps GoM13 and GoM156, between hydrocarbon and gas seeps GB4484, AMV825 and GoM17 and between methane seeps Tomm and HR19.

Hierarchical clustering of co-occurrence frequencies identified two main clusters with cluster A consisting of only gas and hydrocarbon seeps (GoM13, GB4573, GB4484, AMV825, GoM17) and cluster B including all seep types (AMV760, GoM140, NZ315, GoM4463, Tomm, HR19; Figure 4C).

TABLE 3 | Mas subunit D (MasD) diversity at investigated hydrocarbon seep sites.

| Sample  | Raw reads | Quality reads | OTU0.96 no. | Relative SSO* [%] | Absolute SSO** [%] | Coverage [%] | OTU0.96 no.*** | Inverse Simpson [1/D] |
|---------|-----------|---------------|-------------|------------------|-------------------|-------------|----------------|---------------------|
| GoM140  | 535       | 430           | 58          | 36               | 43                | 89          | 17              | 6                   |
| GoM13   | 751       | 487           | 47          | 47               | 43                | 92          | 16              | 4                   |
| GoM17   | 892       | 226           | 27          | 63               | 26                | 89          | 15              | 3                   |
| GoM4463 | 1419      | 154           | 44          | 55               | 36                | 74          | 27              | 9                   |
| GoM156  | 1078      | 803           | 87          | 32               | 47                | 91          | 21              | 4                   |
| GB4573  | 1255      | 989           | 93          | 15               | 49                | 94          | 17              | 7                   |
| GB4484  | 834       | 572           | 66          | 29               | 55                | 90          | 14              | 4                   |
| AMV760  | 837       | 134           | 21          | 48               | 43                | 89          | 12              | 3                   |
| AMV825  | 637       | 376           | 50          | 56               | 32                | 88          | 19              | 3                   |
| HR19    | 708       | 65            | 13          | 62               | 23                | 83          | 13              | 3                   |
| NZ315   | 2043      | 722           | 65          | 28               | 52                | 93          | 16              | 3                   |
| Tomm    | 1756      | 173           | 33          | 58               | 24                | 84          | 18              | 6                   |
| Total   | 12745     | 5131          | 420         | 16               | 61                |             |                 |                     |

*SSOrel., relative single sequence OTU0.96: sequences occurring only once in at least one sample but may occur more often in other samples, given in percent of OTU0.96 per station.

**SSOabs., absolute single sequence OTU0.96: sequences occurring only once in the whole dataset, given in percent of OTU0.96 per station.

***Standardized numbers of OTU0.96 based on resampling of 65 sequences without replacement.
**Cosmopolitan and Locally Restricted OTUs**

Two out of the 420 species-level OTU\(_{0.96}\) were cosmopolitan as defined by their presence in at least 10 of the 12 stations investigated (Figures 4B and 5). These two OTU comprised the majority of MasD sequences. Fifteen percent of total MasD sequences were assigned to OTU\#1 that dominated at almost all sites with 5–55% of quality sequences retrieved from the individual sites. OTU\#1 was rare only at methane seep sites NZ315 and Tomm with <2%. Conversely, OTU\#2 dominated these two methane seep sites with 53% of total quality reads at NZ315 and 22% at Tomm, in addition to 52% at HR19. The third most abundant OTU\#3, which was present at six sites (Supplementary Table S4), strongly dominated in GoM13 (44% of the reads), GoM156 (26%) and AMV825 (14%) sediments. In general, the next ten abundant OTU\#4 to \#13 were only occasionally abundant and found to be dominant at a single seep site (Supplementary Table S4). For example, OTU\#11 comprised only 2.8% of total sequences and were found at five sites but dominated GoM17 (22%) and AMV760 (65%).

Rare OTU\(_{0.96}\) appearing only once in the whole data set (i.e., absolute single sequence OTUs; SSO\(_{abs}\)) are referred to as ‘locally restricted.’ A majority of OTU\(_{0.96}\) (61%) were assigned to SSO\(_{abs}\). Conditionally rare organisms are those appearing only once in a given sample but more often in one or more of the other samples (i.e., relative single sequence OTU\(_{0.96}\), SSO\(_{rel}\), Gobet et al., 2012). The conditionally rare microbial MasD community comprised 16% of total OTU\(_{0.96}\) of which >50% were represented by a low read number (<10 reads per OTU).

**Phylogenetic Affiliation of MasD**

For phylogenetic analysis, MasD was clustered on a proposed family level of 72% similarity (OTU\(_{0.72}\)). In total, 83 family level OTUs were detected. Representative sequences formed three clusters that shared <60% sequence similarity (Figure 6). The three clusters might comprise organisms from different phyla because the inter-cluster identity of <60% is close to the estimated MasD phylum level OTU threshold at 52% (according to the phylum threshold of 75% for 16S rRNA genes; Yarza et al., 2014). The proposed three phyla are mostly consistent...
Strains within cluster I all belong to Proteobacteria, cluster II does not contain any sequences from cultured strains and cluster III includes sequences from Proteobacteria (strain BuS5, Deltaproteobacteria) but also from Firmicutes.

OTU richness was highest in cluster I: 372 species-level OTU (OTU\textsubscript{0.96}) were identified comprising 4099 sequences (80\% of total) with 72 total family level OTU (OTU\textsubscript{0.72}). In cluster II, 47 species-level OTU\textsubscript{0.96} were identified comprising 1022 sequences (20\% of total) with 10 family level OTU\textsubscript{0.72}. And finally cluster III contained only 10 sequences clustered into a single OTU on species, genus and family level (Figure 6). For cluster I we defined seven monophyletic subclades: cluster Ia to Ig. Clusters Ia, Ib, and Ig only contain MasD from uncultivated organisms. Clusters Ia and Ib are comprised of a particularly high number of family level OTU\textsubscript{0.72} (18 and 16, respectively). Cluster Ib also included the AssA found in a draft genome from Smithella SCADC, a syntrophic deltaproteobacterium derived from different methanogenic alkane-degrading enrichment cultures (Embree et al., 2014; Tan et al., 2014), and Candidatus Magnetoglobus multicellularis (ATBP010000388 on IMG/M). The only cultivated relative in cluster Id is the long-chain alkane- and alkene-degrading Desulfatibacillum aliphaticivorans. This cluster includes the two most abundant and cosmopolitan OTU\textsubscript{0.96} #1 and #2. Cluster II is quite diverse and comprised of MasD from betaproteobacterial nitrate reducers and deltaproteobacterial sulfate reducers but also from medium- or short-chain alkane degraders (Figure 6).

**In Situ Identification of Alkane-Degrading Bacteria by masD geneFISH**

A geneFISH assay for masD was established to identify the alkane-degrading bacterial community in the environment. A sediment-free culture from GB4573 sediments grown under sulfate-reducing conditions with \textit{n}-butane as sole carbon source was used to optimize the protocol. The enrichment is dominated by alkane-degrading SRB of the SCA2 clade belonging to the deltaproteobacterial \textit{Desulfooccus/Desulfosarcina} (DSS) branch (Kleindienst et al., 2014). As a basis for probe design, a masD gene library was constructed from this enrichment culture. The obtained sequences fall into all three defined clusters (Figure 6). Because similarity between clusters I, II, and III was too low on the DNA level to allow the design of a single dsDNA polynucleotide probe covering the whole diversity of the gene (Figure 6, Supplementary Table S1), clusters were targeted independently. For cluster II, it was possible to cover all sequences with one dsDNA polynucleotide probe. For cluster I and III, we designed probe mixes of 4 dsDNA polynucleotide probes each for complete coverage. The DNA-based sequence similarity between MasD OTU\textsubscript{0.72} from the 12 seeps and the
individual probes was above the threshold of 72%, which has been reported to be detected by a 350 bp probe (Moraru et al., 2010). The dsDNA masD probes were synthesized separately and applied as mix on the GB enrichment culture. Hybridization of the enrichment culture showed masD gene signals co-localized with signals of probe DSS658 (Figure 7A), which targets the Desulfosarcina/Desulfococcus branch of Deltaproteobacteria, indicating a successful in situ identification of alkane-degrading bacteria. Members of the DSS made up between 73 and 90% of total DAPI-stained cells. Of these cells, 49–58% showed positive masD gene signals with the masD probe mix (Figure 7B).

**DISCUSSION**

**Diversity of Alkane-Degrading Bacteria**

To date, only about 20 bacterial strains or phylotypes in enrichments have been found to degrade alkanes anaerobically. They belong to a total of ten families within the phyla Proteobacteria and Firmicutes (Figure 1). In contrast, a quite high bacterial 16S rRNA gene diversity has been reported for hydrocarbon seeps in the Guaymas Basin and Gulf of Mexico, in particular for deltaproteobacterial SRB (Teske et al., 2002; Dhillon et al., 2003; Orcutt et al., 2010; Wegener and Knittel, unpublished data), suggesting that there are likely many more SRB and other bacteria than currently known that thrive anaerobically on alkanes. Our data clearly support this hypothesis as we found a total of 420 MasD species-level OTU0.96 and 83 family level OTU0.72 from the 12 investigated hydrocarbon-impaired sites. Based on rarefaction curves, diversity was not fully covered; therefore increased sequencing efforts might reveal even more diversity. Furthermore, we also might have missed new lineages which did not have the sequence motifs we were searching for. For example, we excluded about 70% of reads retrieved from GoM4463 and HR19 from the analysis only due to the absence of these motifs. The high overall MasD diversity can be explained either by the presence of a microbial seed bank, which comprises dormant alkane-degrading organisms that are resuscitated following environmental changes (Lennon and Jones, 2011) or by the presence of many niches as alkanes are abundant in nature.

The oily sediment site from the northern Gulf of Mexico (GoM4463) had the highest MasD α-diversity (Inverse Simpson Index), followed by site GB4573 in the Guaymas Basin, indicating that these habitats are species-rich because of the presence of very diverse hydrocarbons. This high diversity might also enable the microbes to withstand certain environmental changes, like the strong temperature gradients found at Guaymas Basin. In contrast, the asphaltic sample GoM17 was the least diverse, supporting the notion that in an extreme environment with a rather limited buffet of degradable hydrocarbons, small changes would have serious impact on the microbial communities where few new species accumulate.

**Cosmopolitan and Rare Alkane-Degrading Bacteria**

Two abundant cosmopolitan species-level MasD OTU0.96 were detected in this study indicating that only a small number of dominant alkane degraders are globally distributed. These OTU were assigned to cluster Id, which included Desulfotabicullum alphaticivorum str. CV2803, a sulfate-reducing long-chain alkane- (C13–C18) and alkene- (C7–C23) degrading bacterium isolated from hydrocarbon-polluted sediments in the Gulf of Fos (France, Cravo-Laureau et al., 2004a). Their closest relatives were MasD sequences (e.g., Guayamas clone LCA_537) retrieved from an enrichment with dodecane (Kleindienst et al., 2014) suggesting long chain alkanes as substrates for the organisms representing OTU#1 and #2. Their cosmopolitan presence suggests an as of yet unknown environmental importance of members in this group for the degradation of long-chain alkanes and alkenes in marine seep sediments. Cosmopolitan distribution was also shown for other seep-associated organisms, such as certain ANME that have been found in seep sediments worldwide (Ruff et al., 2015).

A tremendous amount of species-level single sequence OTU0.96 (SSOabs and SSOrel) was retrieved from the 12 different sediments. SSOabs are locally restricted, permanently rare species (Gobet et al., 2012). In contrast, SSOrel are organisms that are rare in one ecosystem, but very common or even dominant in another. These organisms may start growth when the conditions change, which was shown for pelagic communities where rare organisms became abundant after disturbance (Sjöstedt et al., 2012) or showed seasonal patterns (Hugoni et al., 2013). Guaymas Basin site 4573 had the most SSOrel and SSOabs supporting the uniqueness of this habitat with strong geochemical gradients, in particular temperature gradients, previously reported (Teske et al., 2002; Teske and Sørensen, 2008). These minor MasD variants have developed under specialized circumstances that may be linked to the diversity and the structure of the alkanes in that environment.
Factors Driving the Alkane-Degrading Community Structure

The MasD community structure was so clearly driven by the hydrocarbon source at the various seeps that the pressure of other factors such as water depth has little to no effect. MasD OTU from methane, gas and hydrocarbon seep sites was most similar within the habitat type as shown by non-metric multidimensional scaling. Thus, the range of available alkanes seems to strongly influence the diversity of MasD-carrying microbes. This is consistent with the narrow substrate range of cultured strains. For example, strain BuS5 oxidizes only propane and butane (Kniemeyer et al., 2007), Azoarcus sp. str. HxN1 oxidizes C_6–C_8 (Ehrenreich et al., 2000) and Desulfatibacillum aliphaticivorans str. CV2803 oxidizes C_{13}–C_{18} (Cravo-Laureau et al., 2004a). It has been shown that seep-associated anaerobic methanotrophs (ANME clades) are structured by sediment depth and sediment temperature, (Ruff et al., 2015) while sulfate reducers (SEEP-SRB clades) are structured more by faunal activity and thus biogeochemistry (Felden et al., 2014). In all studies, water depth was not identified as an influencing factor.

Phylogeny of Anaerobic Alkane-Degrading Bacteria

For 16S rRNA genes there are widely accepted thresholds for the definition of a species, genus or family (Rosselló-Móra and Amann, 2015). With the clustering of protein-coding genes, however, we must consider the high variation in sequence conservation. To establish taxonomic levels based on MasD we used cultivated alkane-degrading species and correlated their 16S rRNA nucleotide sequences with their MasD amino acid sequences. This approach has also been implemented for methane-oxidizing bacteria (PmoA, Degelmann et al., 2010) and nitrogen-fixing bacteria (NifH, Bowen et al., 2013). We determined the following threshold values for MasD: 96% amino acid identity for species, 87% for genus, and 72% for families. The species threshold of 96% that we defined for MasD was higher than those determined for PmoA (93%) and NifH (88%), which can in part be explained by the recent re-evaluation and implementation of a new 16S rRNA gene species threshold of 98.7% (Yarza et al., 2014) used in our study compared to the 97% identity threshold used in the above previous studies.

Phylogenetic analysis of MasD resulted in three clusters of which the first was dominant, the second regularly present and the third very rare. The 60% amino acid identity between these clusters is likely in the range of a phylum level threshold for MasD. The only available non-deltaproteobacterial MasD is from phylotype Propane60GuB (cluster III) and showed a maximum of 56% amino acid identity to any sequence of cluster I and cluster II.

Based on the MasD protein sequences alone we cannot exclude that the respective enzymes have similar substrate spectra or affinities. In the MasD phylogenetic tree (Figure 6) there was also no obvious separation between sequences retrieved from freshwater (Callaghan et al., 2010) and those marine habitats (Kleindienst et al., 2012; Acosta-González et al., 2013). Thus, cultivation attempts as well as enzyme characterizations are necessary to evaluate phylum-level functional diversity of MasD.

The dominant cluster I was most diverse including several isolated deltaproteobacterial and betaproteobacterial sulfate- and nitrate-reducing bacteria. It was present at all sites suggesting that these microbes are able to adapt to the local environmental features which differ remarkably between seeps. Within subcluster Ib, eight OTU_{8-72} cluster tightly with MasD from methanogenic enrichments and Smithella sp. We hypothesize that these MasD belong to syntrophic methanogenic hydrocarbon-degrading communities that are relevant for alkane degradation at marine seeps. However, so far studies on methanogenic alkane-degradation are limited to enrichment cultures (Zengler et al., 1999; Chang et al., 2006; Berdugo-Clavijo and Gieg, 2014; Embree et al., 2014). In situ studies that aim to quantify the responsible syntrophic bacteria and archaea directly in their environment have not been conducted, although the importance of crude oil biodegradation via methanogenesis has been proposed for subsurface oil reservoirs (Jones et al., 2008). The anaerobic biodegradation of octacosane (C_{28}), a solid paraffin, under methanogenic conditions (Davidova et al., 2011; Callaghan, 2013), could be one process of interest for future in situ studies.

MasD of the only cultured short-chain alkane-degrading strain BuS5 (Musat, 2015) grouped into cluster III. Although
this strain is found in numbers of ca. 2% of total cell counts at
AMV and Guaymas Basin (Kleindienst and Knittel, unpublished
data), we did not retrieve any OTU related to BuS5-MasD. Their
absence is most likely due to a substantial primer mispairing; the
forward primers had 11 (7757f1-f2, 22mer) and 13 (7766f, 23mer)
mismatches, respectively, to the BuS5-masD sequence retrieved
from the isolate genome (JGI gene ID 2513990058). In addition
to deltaproteobacterial BuS5, cluster III is also comprised of two
Firmicutes, i.e., a single cell from family Peptococcaceae and a
phylotype from the enrichment Propane60GuB, indicative of a
lateral gene transfer event. We also report a second potential
instance of lateral gene transfer in MasD cluster I, which contains
both Beta- and Deltaproteobacteria in subclusters.

Despite the broad diversity of cluster II (10 family level
OTU0.72 from this study), this cluster is comprised of only two
other sequences from a metagenomic study from Santa Barbara
oil seep sediments (Hess, JGI metagenome, IMG submission ID
26744). To date few environmental studies have been conducted
and most have investigated fresh water environments (Callaghan
et al., 2010; Cheng et al., 2013) and few marine environments
(Acosta-González et al., 2013; Kimes et al., 2013; von Netzer et al.,
2013; Johnson et al., 2015). All of them implement clone libraries
with very limited sequencing depth.

Considering the high overall diversity found within MasD,
the range of substrates currently known to be activated by
MasD/AssA might expand with further studies. For example,
ethane is the most abundant short-chain non-methane alkane
at several of our sites (Bazylnski et al., 1988; Boetius et al.,
2000; Orcutt et al., 2008; Felden et al., 2013). However,
isolation of anaerobic ethane-degrading organisms is lacking.
Recently, sequences retrieved from a Gulf of Mexico batch
reactor with ethane and sulfate have been shown to be
closely related to Desulfosarcina variabilis and strain BuS5
(Bose et al., 2013). Nevertheless, cultivation or enrichment
is necessary in order to get detailed insights into the
metabolic pathways and substrate spectrum. Large metagenomic
studies with systematic screening for masD will help to
further unravel MasD diversity and improve current primer
sets for future studies. Another approach that could prove
useful in future masD marker gene surveys is oligotyping,
which uses only the most informative nucleotide positions
to define phylogenetically distinct oligotypes (Eren et al.,
2013).

**In Situ Identification of Alkane-Degrading Communities**

In this study we also used a comparative sequence analysis of
masD for the development of a polynucleotide probe set. We demonstrated that these probes could be used for
the in situ identification of bacteria catalyzing the anaerobic
degradation of alkanes by geneFISH. The individual geneFISH
probes designed in this study could also be used to target
subpopulations, i.e., MasD cluster I, cluster II, and cluster III
(Figure 6). geneFISH was originally invented for *pmoA* and
applied on *Escherichia coli* clones in order to link function
with phylogeny (Moraru et al., 2010). Later, the method was
applied on thin sections of the hydrothermal vent mussel
*Bathymodiolus* to detect uptake hydrogenases (*hupL*) in its
symbionts (Petersen et al., 2011). Here, we showed that direct
application of geneFISH to sediment samples from Guaymas
Basin was also possible but challenging due to high background
fluorescence.

The in situ quantification of masD will help to identify
the size of alkane-degrading communities in the environment
independent of cultivation and enrichment and independent
from *a priori* knowledge regarding which taxa are capable of
alkane degradation. This method is and will be a major step
forward to begin to describe the global relevance and diversity
of alkane degradation at natural hydrocarbon seeps and in oil-
polluted environments.

**CONCLUSION**

As alkane is the predominant component of crude oil,
our findings point to a so far overlooked high potential
of marine benthic microbes to react to natural changes in
hydrocarbon seepage or to massive hydrocarbon input as
encountered during anthropogenic oil spills. The number
of MasD OTU0.72 that we detected in this study indicated
that there may be eight times more family level bacterial
classes that are capable of hydrocarbon degradation than
previously assumed. This observed high diversity of MasD
might allow for higher rates of survival and adaptability
within a microbial community exposed to such unstable and
changing environmental conditions. Exploring the prevalence
and diversity of MasD can help to identify novel lineages of
alkane degraders as well as to differentiate closely related
phylotypes. Several MasD clusters were discovered that do not
yet have cultured representatives. Therefore, one major objective
for future studies should be enrichment and isolation of these
lineages for further characterization of substrate specialization.
Further methodological approaches could include metagenomics
and single cell genomics of FACS sorted cells carrying masD
genes.

Furthermore, this study represents a first step toward
establishing comparative MasD sequencing and masD
geneFISH as powerful tools for targeted investigation of the
diversity and abundance of alkane-degrading bacteria in anoxic
environments.

**ACKNOWLEDGMENTS**

We greatly acknowledge Antje Boetius for providing samples
from expeditions with the research vessels SONNE, METEOR,
ALKOR, and SEWARD JOHNSON II. Samples were taken in
the framework of the GEOTECHNOLOGIEN programs MUMM
I and II (grants 03G0554A and 03G0608A) funded by the
German Ministry of Education and Research (BMBF) and the
German Research foundation and by the EU 5th FP METROL
and 6th FP HERMES. Andreas Teske and Marc Mussmann are
acknowledged for providing samples from an expedition with
ATLANTIS. Special thanks go to Sara Kleindienst who initiated the enrichment cultures, Florian Musat for fruitful discussions and sharing unpublished MasD sequences, Jimena Barrero Canosa for introduction and technical support on geneFISH, Hanno Teeling for help with 454 tag sequencing analysis as well as Manabu Fukui and Hisayaki Kojima for providing the MasD sequence of strain PL12. We are grateful to Meghan Chafee for critical reading, valuable suggestions and language editing. SER was supported by the Leibniz program awarded to Antje Boetius. Further support was provided by the Max Planck Society, Germany.

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: http://journal.frontiersin.org/article/10.3389/fmicb.2015.01511

REFERENCES

Acosta-González, A., Rosselló-Móra, R., and Marqués, S. (2013). Characterization of the anaerobic microbial community in oil-polluted subtidal sediments: aromatic biodegradation potential after the Prestige oil spill. Environ. Microbiol. 15, 77–92. doi: 10.1111/1462-2920.2012.02782.x

Adams, M. M., Hoarfrost, A. L., Boge, A., Joyce, S. B., and Girguis, P. R. (2013). Anaerobic oxidation of short-chain alkanes in hydrothermal sediments: potential influences on sulfur cycling and microbial diversity. Front. Microbiol. 4:110. doi: 10.3389/fmicb.2013.00110

Aeckersberg, F., Bak, F., and Widdel, F. (1991). Anaerobic oxidation of saturated-hydrocarbons to CO2 by a new type of sulfate-reducing bacterium. Arch. Microbiol. 156, 5–14. doi: 10.1007/BF01418180

Aitken, C. M., Jones, D. M., Maguire, M. J., Gray, N. D., Sherry, A., Bowler, B. F. J., Adams, M. M., Hoarfrost, A. L., Boe, A., Joyce, S. B., and Girguis, P. R. (2013). Evidence that crude oil alkane activation proceeds by different mechanisms under sulfate-reducing and methanogenic conditions. Geochim. Cosmochim. Acta 109, 162–174. doi: 10.1016/j.gca.2013.01.031

Bazylinski, D. A., Farrington, J. W., and Jannasch, H. W. (1988). Hydrocarbons in surface sediments from a Guaymas Basin hydrothermal vent site. Org. Geochem. 12, 547–558. doi: 10.1016/0146-6380(88)90165-5

Becker, A., Fritz-Wolf, K., Kabsch, W., Knappe, J., Schultz, S., and Wagner, A. F. V. (1999). Structure and mechanism of the glycyl radical enzyme pyruvate formate-lyase. Nat. Struct. Biol. 6, 969–975. doi: 10.1038/13341

Berdugo-Clavijo, C., and Gieg, L. M. (2014). Conversion of crude oil to methane by a microbial consortium enriched from oil reservoir production waters. Front. Microbiol. 5:197. doi: 10.3389/fmicb.2014.00197

Boetius, A., Ravenskragh, K., Schubert, C. J., Rickert, D., Widdel, F., Gieseke, A., et al. (2000). A marine microbial consortium apparently mediating anaerobic oxidation of methane. Nature 407, 623–626. doi: 10.1038/35036572

Boe, A., Rogers, D. R., Adams, M. M., Joyce, S. B., and Girguis, P. R. (2013). Geomicrobiological linkages between short-chain alkane consumption and sulfate reduction rates in seep sediments. Front. Microbiol. 4:386. doi: 10.3389/fmicb.2013.00386

Bowen, J. L., Byrnes, J. E. K., Weisman, D., and Colaneri, C. (2013). Functional gene pyrosequencing and network analysis: an approach to examine the response of denitrifying bacteria to increased nitrogen supply in salt marsh sediments. Front. Microbiol. 4:342. doi: 10.3389/fmicb.2013.00342

Bowles, M. W., Samarkin, V. A., Bowles, K. M., and Joyce, S. B. (2011). Weak coupling between sulfate reduction and the anaerobic oxidation of methane in methane-rich seafloor sediments during ex situ incubation. Geochim. Cosmochim. Acta 75, 500–519. doi: 10.1016/j.gca.2010.09.043

Bray, J. R., and Curtis, J. T. (1957). An Ordination of the upland forest communities of Southern Wisconsin. Ecol. Monogr. 27, 325–349. doi: 10.2307/1942268

Butts, C. T. (2008). Network: a package for managing relational data in R. J. Stat. Softw. 24, 1–36. doi: 10.18637/jss.v024.i02

Byrne, J. V., and Emery, K. O. (1960). Sediments of the gulf of California. Geol. Soc. Am. Bull. 71, 983–1010. doi: 10.1130/0016-7066(1960)71(983:STSOGC)2.0.CO;2

Callaghan, A. V. (2013). Enzymes involved in the anaerobic oxidation of n-alkanes: from methane to long-chain paraffins. Front. Microbiol. 4:89. doi: 10.3389/fmicb.2013.00089

Callaghan, A. V., Davodova, I. A., Savage-Ashtock, K., Parisi, V. A., Gieg, L. M., Sullita, J. M., et al. (2010). Diversity of benzyl- and allylsuccinate synthase genes in hydrocarbon-impacted environments and enrichment cultures. Environ. Sci. Technol. 44, 7287–7294. doi: 10.1021/es1020232
using 16S rRNA gene data. *Methods Ecol. Evol.* 4, 1111–1119. doi: 10.1007/1091-210X-12114
Feldén, J., Lichtschlag, A., Wenzhöfer, F., De Beer, D., Fessek, T., Pop Ristova, P., et al. (2013). Limitations of microbial hydrocarbon degradation at the Amon mud volcano (Nile deep-sea fan). *Biogeosciences* 10, 3269–3283. doi: 10.5194/bg-10-3269-2013.
Feldén, J., Ruff, S. E., Ertelai, T., Inagaki, F., Hinrichs, K. U., and Wenzhöfer, F. (2014). Anaerobic methanotrophic community of a 5346-m-deep vesicomyid clam colony in the Japan Trench. *Geobiology* 12, 183–199. doi: 10.1111/gbi.12078
Fruchterman, T. M. J., and Reingold, E. M. (1991). Graph drawing by force-directed placement. *Softw. Pract. Exp.* 21, 1129–1164. doi: 10.1002/spe.4380211102
Gobet, A., Boer, S. I., Huse, S. M., Van Beusekom, J. E. E., Quince, C., Sogin, M. L., et al. (2012). Diversity and dynamics of rare and of resident bacterial populations in coastal sands. *ISME J.* 6, 542–553. doi: 10.1038/ismej.2011.132.
Good, I. J. (1953). The population frequencies of species and the estimation in situ abundance of sulfate-reducing bacteria in diverse marine hydrocarbon seeps. *Front. Microbiol.* 7, doi:10.1038/nrmicro2504
Ludwig, W., Strunk, O., Westram, R., Richter, L., Meier, H., Yadhukumar, et al. (2004). AAR: a software environment for sequence data. *Nucleic Acids Res.* 32, 1363–1371. doi: 10.1093/nar/gkh293
Manz, W., Eisenbrecher, M., Neu, T. R., and Szewzyk, U. (1998). Abundance and spatial organization of gram-negative sulfate-reducing bacteria in activated sludge investigated by in situ probing with specific 16S rRNA-targeted oligonucleotides. *FEMS Microbiol. Ecol.* 25, 43–61. doi: 10.1111/j.1574-6941.1998.tb00459.x
Markowitz, V. M., Chen, I.-M., Chu, K., Szeto, E., Palaniappan, K., Pillay, M., et al. (2007). IMG/M 4 version of the integrated metagenome comparative analysis system. *Nucleic Acids Res.* 42, D568–D573. doi: 10.1093/nar/gkl919
Mastalerz, V., De Lange, G. J., and Dähnmann, A. (2009). Differential and anaerobic oxidation of hydrocarbon gases discharged at mud volcanoes in the Nile deep-sea fan. *Geochim. Cosmochim. Acta* 73, 3849–3863. doi: 10.1016/j.gca.2008.12.030
Mbadinga, S. M., Wang, L.-Y., Zhou, L., Liu, J.-F., Gu, J.-D., and Mu, B.-Z. (2011). Microbial communities involved in anaerobic degradation of alkanes. *Int. Biodeter. Biodegradation* 65, 1–13. doi: 10.1016/j.ibiod.2010.11.009
Moraru, C., Lam, P., Fuchs, B. M., Kuyper, M. M. M., and Amann, R. (2010). GeneFISH – an in situ technique for linking gene presence and cell identity in environmental microorganisms. *Environ. Microbiol.* 12, 3057–3073. doi: 10.1111/j.1462-2920.2010.02281.x
Moraru, C., Moraru, G., Fuchs, B. M., and Amann, R. (2011). Concepts and software for a rational design of polynucleotide probes. *Environ. Microbiol. Rep.* 3, 69–78. doi: 10.1111/j.1758-2229.2010.00189.x
Musat, F. (2015). The anaerobic degradation of gaseous, nonmethane alkanes: from in situ processes to microorganisms. *Comput. Struct. Biotechnol. J.* 13, 222–228. doi: 10.1016/j.csbj.2015.03.002
Musat, F., Wilkes, H., Behrends, A., Woebcken, D., and Widdel, F. (2007). Microbial nitrate-dependent cyclohexane degradation coupled with anaerobic ammonium oxidation. *ISME J.* 4, 1290–1301. doi: 10.1038/ismej.2010.50
Oksanen, J., Blanchet, F. G., Kindt, R., Legendre, P., O’Hara, R. B., Simpson, G. L., et al. (2012). *vegan: Community Ecology Package*. Available at: http://CRAN.R-project.org/package=vegan.
Orcutt, B., Samarkin, V., Boetius, A., and Joyce, S. B. (2008). On the relationship between methane production and oxidation by anaerobic methanotrophic communities from cold seeps of the Gulf of Mexico. *Environ. Microbiol.* 10, 1108–1117. doi: 10.1111/j.1462-2920.2007.01528.x
Orcutt, B. N., Joyce, S. B., Kleindienst, S., Knittel, K., Ramette, A., Reitz, A., et al. (2010). Impact of natural oil and higher hydrocarbons on microbial diversity, distribution, and activity in Gulf of Mexico cold-seep sediments. *Deep Sea Res. Part II* 57, 2008–2021. doi: 10.1016/j.dsr2.2010.05.014
Petersen, J. M., Zielinski, F. U., Pape, T., Seifert, R., Moraru, C., Amann, R., et al. (2011). Hydrogen is an energy source for hydrothermal vent symbioses. *Nature* 476, 176–180. doi: 10.1038/nature10325
Pruesse, E., Quast, C., Knittel, K., Fuchs, B. M., Ludwig, W., and Peplies, J. (2007). * Silva*: a comprehensive online resource for quality checked and aligned ribosomal RNA sequence data compatible with ARB. *Nucleic Acids Res.* 35, 7188–7196. doi: 10.1093/nar/gkm864
Reeburgh, W. S. (2007). Oceanic methane biogeochemistry. *Chem. Rev.* 107, 486–513. doi: 10.1021/cr050362v
Tan, B., Nesbo, C., and Foght, J. (2014). Re-analysis of omics data indicates von Netzer, F., Pilloni, G., Kleindienst, S., Krüger, M., Knittel, K., Gründger, F., Treude, T., Boetius, A., Knittel, K., Wallmann, K., and Jørgensen, B. B. (2003). Teske, A., and Sørensen, K. B. (2008). Uncultured archaea in deep marine Savage, K. N., Krumholz, L. R., Ryabin, T., Hall, J. R., Hartmann, M., Hollister, E. B., et al. (2009). Introducing mothur: open-source, platform-independent, community-supported software for describing and comparing microbial communities. Appl. Environ. Microbiol. 75, 7537–7541. doi: 10.1128/AEM.01541-09 Simonet, B. R. T., and Lonsdale, P. F. (1982). Hydrothermal petroleum in mineralized mounds at the seabed of Guaymas Basin. Nature 295, 198–202. doi: 10.1038/295198a0 Sjöstedt, J., Koch-Schmidt, P., Pontarp, M., Canbäck, B., Tunlid, A., Lundberg, P., et al. (2012). Recruitment of members from the rare biosphere of marine bacterioplankton communities after an environmental disturbance. Appl. Environ. Microbiol. 78, 1361–1369. doi: 10.1128/AEM.05542-11 So, C. M., and Young, L. Y. (1999). Initial reactions in anaerobic alkane degradation by a sulfate reducer, strain AK-01. Appl. Environ. Microbiol. 65, 5532–5540. Stackebrandt, E., and Ebers, J. (2006). Taxonomic parameters revisited: tarnished gold standards. Microbiol. Today 8, 6–9. Suess, E., Torres, M., Bohrmann, G., Collier, R. W., Greinert, J., Linke, P., et al. (1999). Gas hydrate destabilization: enhanced dewatering, benthic material turnover and large methane plumes at the Cascadia convergent margin. Earth Planet. Sci. Lett. 170, 1–15. doi: 10.1016/S0012-821X(99)00092-8 Tan, B., Nesbo, C., and Vogt, J. (2014). Re-analysis of omics data indicates Smithella may degrade alkanes by addition to fumarate under methanogenic conditions. ISME J. 8, 2353–2356. doi: 10.1038/ismej.2014.87 Teske, A., Hinrichs, K.-U., Edgcomb, V., De Vera Gomez, A., Kysela, D., Sylva, S. P., et al. (2002). Microbial diversity of hydrothermal sediments in the Guayas Basin: evidence for anaerobic methanotrophic communities. Appl. Environ. Microbiol. 68, 1994–2007. doi: 10.1128/AEM.68.4.1994-2007.2002 Teske, A., and Sørensen, K. B. (2008). Uncultured archaea in deep marine subsurface sediments: have we caught them all? ISME J. 2, 3–18. doi: 10.1038/ismej.2007.90 Tissot, B. P., and Welse, D. H. (1984). Petroleum Formation and Occurrence: A New Approach to Oil and Gas Exploration. Berlin: Springer Verlag. Treude, T., Boetius, A., Knittel, K., Wallmann, K., and Jørgensen, B. B. (2003). Anaerobic oxidation of methane above gas hydrates at Hydrate Ridge, NE Pacific Ocean. Mar. Ecol. Prog. Ser. 264, 1–14. doi: 10.3354/meps264001 von Netzer, F., Pilloni, G., Kleindienst, S., Krüger, M., Knittel, K., Gründger, F., et al. (2013). Enhanced gene detection assays for fumarate-adding enzymes allow uncovering of anaerobic hydrocarbon degraders in terrestrial and marine systems. Appl. Environ. Microbiol. 79, 543–552. doi: 10.1128/AEM.02362-12 Wang, L.-Y., Li, W., Mbadinga, S. M., Liu, J.-F., Gu, J.-D., and Mu, B.-Z. (2012). Methanogenic microbial community composition of oil sludge and its enrichment amended with alkanes incubated for over 500 days. Geomicrobiol. J. 29, 716–726. doi: 10.1080/01490451.2011.619634 Ward, J. H. (1963). Hierarchical grouping to optimize an objective function. J. Am. Stat. Assoc. 58, 236–244. doi: 10.1080/01621459.1963.10500845 Weelink, S. A. B., Van Doesburg, W., Saia, F. T., Rijpstra, W. I. C., Rüling, W. F. M., Smidt, H., et al. (2009). A strictly anaerobic betaproteobacterium Georgioplaca toluolica gen. nov., sp. nov. degrades aromatic compounds with Fe(III), Mn(IV) or nitrate as an electron acceptor. FEMS Microbiol. Ecol. 70, 575–585. doi: 10.1111/j.1574-6941.2009.00778.x Wegener, G., Shovitin, M., Knittel, K., Niemann, H., Hövland, M., and Boetius, A. (2008). Biogeochemical processes and microbial diversity of the Gulffaks and Tømmeliten methane seeps (Northern North Sea). Biogeochemistry 5, 1129–1144. doi: 10.5194/bg-5-1127-2008 Wickham, H. (2009). ggplot2: Elegant Graphics for Data Analysis. New York: Springer-Verlag. Widl, F., and Böck, F. (1992). “Gram-negative mesophilic sulfate-reducing bacteria,” in The Prokaryotes, 2nd Edn, eds A. Balows, H. G. Truper, M. Dworkin, W. Harder, and K. H. Schleifer (New York, NY: Springer-Verlag). Widl, F., Knittel, K., and Galushko, A. (2010). “Anaerobic hydrocarbon-degrading microorganisms: an overview,” in Handbook of Hydrocarbon and Lipid Microbiology, eds K. N. Timmis, T. McGonigey, J. R. Van Der Meer, and V. De Lorenzo (Heidelberg: Springer). Widl, F., and Pfenning, N. (1982). Studies on dissimilatory sulfate-reducing bacteria that decompose fatty acids II. Incomplete oxidation of propionate by Desulfobolus propionicus gen. nov., sp. nov. Arch. Microbiol. 131, 360–365. doi: 10.1007/BF00411187 Wilkes, H., Rabus, R., Fischer, T., Armstroff, A., Behrendts, A., and Widl, F. (2002). Anaerobic degradation of n-hexane in a denitrifying bacterium: further degradation of the initial intermediate (1-methylpentyl)succinate via C-skeleton rearrangement. Arch. Microbiol. 177, 235–243. doi: 10.1007/s00203-001-0381-3 Yarza, P., Yilmaz, P., Pruesse, E., Glockner, F. O., Ludwig, W., Schleifer, K.-H., et al. (2014). Uniting the classification of cultured and uncultured bacteria and archaea using 16S RNA gene sequences. Nat. Rev. Microbiol. 12, 635–645. doi: 10.1038/nrmicro3330 Zengler, K., Richnow, H. H., Rosselló-Móra, R., Michaelis, W., and Widl, F. (1999). Methane formation from long-chain alkane by anaerobic microorganisms. Nature 401, 266–269. doi: 10.1038/ 45777 Zhou, J., Bruns, M. A., and Tiedje, J. M. (1996). DNA recovery from soils of diverse composition. Appl. Environ. Microbiol. 62, 316–322. Zhou, L., Li, K.-P., Mbadinga, S., Yang, S.-Z., Gu, J.-D., and Mu, B.-Z. (2012). Analyses of n-alkanes degrading community dynamics of a high-temperature methanogenic consortium enriched from production water of a petroleum reservoir by a combination of molecular techniques. Ecotoxicology 21, 1680–1691. doi: 10.1007/s10646-012-0949-5 Conflict of Interest Statement: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest. Copyright © 2016 Stagars, Ruff, Ammann and Knittel. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) or licensor are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.