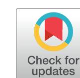
Science of the Total Environment 810 (2022) 152238

Contents lists available at ScienceDirect

Science of the Total Environment

journal homepage: www.elsevier.com/locate/scitotenv

Abundance and diversity of *n*-alkane and PAH-degrading bacteria and their functional genes – Potential for use in detection of marine oil pollution



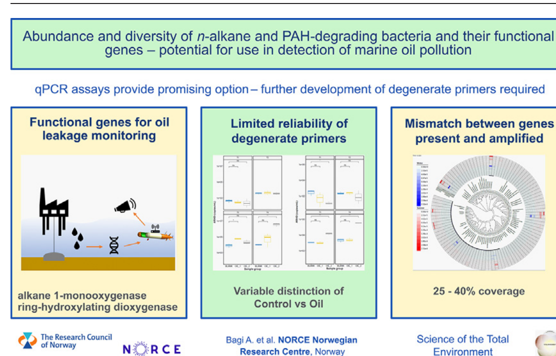
Andrea Bagi *, Kamila Knapik, Thierry Baussant

NORCE Norwegian Research Centre, Bergen, Norway

HIGHLIGHTS

- Quantifying functional genes from broad diversity of bacteria for monitoring oil pollution at sea
- qPCR revealed limited reliability of degenerate primers in signaling oil presence
- Mismatch between amplified and shotgun sequencing identified gene diversity

GRAPHICAL ABSTRACT



ARTICLE INFO

Article history:
Received 4 October 2021
Received in revised form 26 November 2021
Accepted 3 December 2021
Available online 9 December 2021

Editor: Frederic Coulon

Keywords:
Metagenomics
Marine monitoring
Alkane 1-monooxygenase
Aromatic ring-hydroxylating dioxygenase
qPCR
Oil

ABSTRACT

Monitoring environmental status through molecular investigation of microorganisms in the marine environment is suggested as a potentially very effective method for biomonitoring, with great potential for automation. There are several hurdles to that approach with regards to primer design, variability across geographical locations, seasons, and type of environmental pollution. Here, qPCR analysis of genes involved in the initial activation of aliphatic and aromatic hydrocarbons were used in a laboratory setup mimicking realistic oil leakage at sea. Seawater incubation experiments were carried out under two different seasons with two different oil types. Degenerate primers targeting initial oxygenases (alkane 1-monooxygenase; *alkB* and aromatic-ring hydroxylating dioxygenase; ARHD) were employed in qPCR assays to quantify the abundance of genes essential for oil degradation. Shotgun metagenomics was used to map the overall community dynamics and the diversity of *alkB* and ARHD genes represented in the microbial community. The amplicons generated through the qPCR assays were sequenced to reveal the diversity of oil-degradation related genes captured by the degenerate primers. We identified a major mismatch between the taxonomic diversity of *alkB* and ARHD genes amplified by the degenerate primers and those identified through shotgun metagenomics. More specifically, the designed primers did not amplify the *alkB* genes of the two most abundant alkane degraders that bloomed in the experiments, *Oceanobacter* and *Oleispira*. The relative abundance of *alkB* sequences from shotgun metagenomics and 16S rRNA-based *Oleispira*-specific qPCR assay were better signals for oil in water than the tested qPCR *alkB* assay. The ARHD assay showed a good agreement with PAHs degradation despite covering only 25% of the top 100 ARHD genes and missing several abundant *Cycloclasticus* sequences that were present in the metagenome. We conclude that further improvement of the degenerate primer approach is needed to rely on the use of oxygenase-related qPCR assays for oil leakage detection.

* Corresponding author at: Mekjarvik 12, 4070 Stavanger, Norway.
E-mail address: anba@norceresearch.no (A. Bagi).

http://dx.doi.org/10.1016/j.scitotenv.2021.152238
0048-9697/© 2021 The Authors. Published by Elsevier B.V. This is an open access article under the CC BY license (http://creativecommons.org/licenses/by/4.0/).

1. Introduction

Monitoring marine environmental status through analysis of microbial communities using molecular techniques is a potentially very sensitive, rapid and relatively easy-to-automate approach (Caruso et al., 2016; Jones et al., 2019). Implementing genomic tools into benthic biomonitoring of offshore oil and gas activities is well underway to complement or replace currently used time-consuming and costly standards (Aylagas et al., 2021; Mauffrey et al., 2020). However, little effort has been directed towards the automatization of the entire process from sediment collection to analysis (He et al., 2020). On the contrary, at least one commercially available fully integrated autonomous water sampling and processing device, the Environmental Sample Processor (ESP), has been developed for completely autonomous seawater analysis (Preston et al., 2011; Ussler et al., 2013). ESP can provide *in situ* and real-time information about the presence and quantity of a pre-defined set of microorganisms through an autonomous pipeline, which makes it an ideal candidate for environmental surveillance (Bagi et al., 2019). Detecting and tracking oil releases near offshore oil & gas installations using this technology has been evaluated as a potential application area in Norwegian waters (Krolicka et al., 2014).

Oil exposure is known to induce swift changes in microbial community composition with a transient dominance of obligate hydrocarbon-degrading and opportunistic genera including *Oleispira*, *Alcanivorax*, *Oleibacter*, *Colwellia*, and *Cycloclasticus* among others. (Brakstad et al., 2015; Catania et al., 2018; Catania et al., 2015; Chakraborty et al., 2012; Dubinsky et al., 2013; Hazen et al., 2010; Krolicka et al., 2017; Mason et al., 2012; Netzer et al., 2018; Ribicic et al., 2018a; Ribicic et al., 2018b; Tremblay et al., 2019). Quantifying these key indicator bacteria could therefore be used as a signal of oil presence in seawater (Krolicka et al., 2019). Challenges associated with this approach include the type of dominant bacteria varying significantly according to the geographical location, nutrient status, season and the hydrocarbon profile (Reunamo et al., 2013; Ribicic et al., 2018a; Ribicic et al., 2018c; Sun and Kostka, 2019). Consequently, a monitoring approach based on detecting specific hydrocarbon-associated microorganisms would either need to be location, season, and oil specific or it would need to target a comprehensive panel of hydrocarbon degrading microorganisms. The latter is challenging on the currently available ESP platform, where a maximum of 6 targets can be quantified simultaneously with qPCR. Nevertheless, this limitation could be circumvented by measuring a few selected functional genes essential in the processing of hydrocarbons and which are universally present in hydrocarbon-degraders (Bagi et al., 2019; Cébron et al., 2008; Miettinen et al., 2019; Viggor et al., 2015). Homologous genes present in a wide range of species can be simultaneously amplified with degenerate primer pairs, providing the possibility to detect an oil-specific yet broadly representative response with a limited number of qPCR assays (Jurelevicius et al., 2013; Kloos et al., 2006; Scoma et al., 2015; Wang et al., 2010).

The mono- and dioxygenase enzymes responsible for activating aliphatic and aromatic hydrocarbons have been proposed as specific markers of oil pollution (Fuentes et al., 2014). Without these enzymes, hydrocarbons cannot be metabolized and channeled into central metabolic pathways. In particular, qPCR assays designed for quantification of alkane 1-monooxygenase (*alkB*) and aromatic-ring-hydroxylating dioxygenase (ARHD) coding genes have been exploited to track oil pollution and to estimate biodegradative potential with varying success (Cébron et al., 2008; Miettinen et al., 2019; Reunamo et al., 2013; Shahsavari et al., 2016; Viggor et al., 2015). There is still limited knowledge regarding how well the degenerate primers amplify the gene diversity present in a particular sample. Large mismatches between the actual diversity and the fraction of the diversity amplified during the qPCR step could lead to false conclusions. Therefore, the objective of this study was (1) to assess the performance of a set of degenerate primers targeting functional genes (*alkB* and ARHD) for oil detection across two seasons and in response to two different oil types, (2) to compare their performance to a genus specific 16S rRNA assay designed and validated for quantification of *Oleispira* (Krolicka et al., 2014) and (3) compare the diversity of *alkB* and ARHD genes amplified by the degenerate qPCR primers with that obtained from shotgun metagenomics.

The main hypothesis was that under oil occurrence in water, there is a specific increase and signature of an array of microbial genes, and both *alkB* and ARHD gene abundance in the microbial community can be used to detect this change. The second hypothesis was that the designed degenerate qPCR primers can cover the diversity of microbes related to oil biodegradation, thereby enhancing detection sensitivity compared to the *Oleispira*-specific 16S rRNA gene assay. This was tested in a laboratory setup across two seasons (winter and summer) and with two different oil types (light and heavy crude oil), using natural seawater microbial community.

The long-term vision of this study is to use an array of robust microbial genes for use with ESP to automate their detection and quantification as a proxy of oil occurrence in seawater.

2. Materials and methods

2.1. Experimental setup

Two experiments were performed using natural seawater collected at the same site in two different seasons, winter (February) and late summer (September) in 2018. For both experiments, seawater was collected though a pipeline from approximately 80 m depth from the fjord nearby the NORCE facility (Byfjorden: 59.03 N, 5.63E, Norway) in autoclaved 10 L containers. The samples were immediately brought to climate rooms with air temperature set to the *in-situ* temperature of the fjord seawater (8 ± 1 °C in February and 10 ± 1 °C in September). Crude oil from a North Sea platform (OIL-1, density = 0.8832 g/cm^3, fraction of <C10 hydrocarbons (paraffins, naphthenes and aromatics combined) = 3.7 wt%) and a heavy crude (OIL-2, density = 1.004 g/cm^3, fraction of <C10 hydrocarbons (paraffins, naphthenes and aromatics combined) = 0.25 wt%) were added to the dedicated experimental bottles containing the 10 L sampled seawater (referred to as "LIVE" later in the text), using individual disposable syringes. The syringes were weighed before and after oil addition to determine the exact amount of oil added. The target nominal concentration was 10 mg/L in both experiments. Negative controls (referred to as "NEG" later in the text) for both oil types were prepared with autoclaved seawater in the same manner as described above for the test bottles. Experimental bottles were left open to allow for natural weathering from evaporation of the most volatile compounds as this would occur in the field. Control bottles containing only seawater (blank) were prepared without addition of oil and were placed in a separate climate room to avoid any cross contamination from the oil-exposed seawater containers. For each control and oil exposed treatment, triplicate bottles were prepared. In total, 15 bottles were set up, placed on magnetic stirrers (approximately 200 rpm to avoid formation of a vortex), and incubated in climate rooms under light cycles corresponding to the season for 14 days (8 h daylight in February and 13 h in September).

2.2. Bacterioplankton analysis

At the start of the experiment (T0), 2 L seawater samples were collected in autoclaved glass bottles ($n = 3$) and filtered immediately onto 25 mm diameter 0.22 μm pore size membrane filters (GSWP, Millipore). Seawater samples (1.0 L at T1 and 0.5 L subsequently) were collected from all experimental bottles at four time points thereafter on day 2, 4, 7, and 14 and filtered immediately upon collection. All filters were stored at −80 °C until extraction. Total DNA was extracted from the filters using the AllPrep Bacterial DNA/RNA/Protein Mini Kit (Qiagen) following the manufacturer's instructions. DNA quantity and quality were checked using NanoDrop, agarose gel and a Qubit instrument (Thermo Fisher Scientific).

2.2.1. Metagenomics

DNA from the triplicate bottles of blank, OIL-1 and OIL-2 from each season (summer: S and winter: W) was pooled at approximately equal DNA

amount for time points T0, T2, T3 and T4 for B and for time points T2, T3 and T4 for OIL-1 and OIL-2. One blank-negative pool (NB) and two oil-negative control pools (OIL-1N and OIL-2N) were also prepared using the T4 (14 days) DNA samples. Pooled DNA samples ($n = 26$ in total, 13 from each season) were sent to the sequencing facilities of Novogene Europe (Cambridgeshire, UK), where they were subjected to further quality control, library preparation and Illumina sequencing (paired-end, 150 bp, NovaSeq6000). Sequencing data (raw sequences without barcodes) was submitted to the Sequence Read Archive (SRA) of the National Center for Biotechnology Information (NCBI) under BioProject PRJNA556155. BioSample accessions are available in Supplementary Table S1A.

2.2.2. qPCR

2.2.2.1. Quantification of alkB and ARHD genes. Genes involved in aliphatic and aromatic hydrocarbon activation, alkane 1-monooxygenase (*alkB*) and aromatic ring-hydroxylating dioxygenase alpha subunit (ARHD), were quantified in all samples (copy number/mL seawater). For this analysis, primers were designed using the on-line tool j-CODEHOP (Boyce et al., 2009). Alkane 1-monooxygenase protein sequences of *Marinobacter hydrocarbonoclasticus* VT8 (ABM17711), *Thalassomonas viridans* (KJE46164), *Shewanella waksmanii* (WP_028774418), *Colwellia sp.* Phe_37 (KXJ58097), *Alcanivorax sp.* 97CO-5 (EUC70824), *Thalassolituus oleivorans* (WP_015486581.1), unclassified *Oleiphilus* (WP_068514162) and *Oleispira antarctica* (WP_046011085) were used to design *alkB* primers, while aromatic ring-hydroxylating dioxygenase subunit alpha (ARHD) protein sequences of *Alteromonas naphthalenivorans* (WP_013785992), *Marinomonas profundimaris* (WP_024024133), *Pseudomonas stutzeri* B1SMN1 (EPL61971), *Pseudomonas balearica* DSM 6083 (AJE16202), *Burkholderia sp.* K24 (WP_035519969.1), *Cycloclasticus sp.* Phe_18 (KXJ47531) and *Pseudoalteromonas translucida* (WP_058372524.1) were used to design the degenerate ARHD primers. Sequences were downloaded from NCBI. The resulting primer sequences were as follows, alkB-degF: 5'-GCAGCTGACAT CTGCAAATTATATTgarcaytaygg-3′ (degenerate core coding for the peptide EHYG, degeneracy = 8), alkB-degR: 5′-TGATAAGAACGTGTAGGAT Tngcrtgrtgrtc-3′ (degenerate core coding for the peptide DHHA, degeneracy = 24), ARHD-α-degF: 5′- TCTGAATGTGTGTCGTcaymgngg −3′ (degenerate core coding for the peptide HRG, degeneracy = 16) and ARHD-α-degR: 5'-GCCGAAGGCccanccrtgrta-3′ (degenerate core coding for the peptide YHGW, degeneracy = 16). The primers generated products with 177 bp and 94 bp lengths, respectively. PCR conditions, *i.e.*, primer concentration and annealing temperature were optimized to maximize assay performance. Each qPCR reaction was carried out in 20 μL volume. The reaction contained 2 μL of 10-times-diluted DNA, 1.5 μM of each primer, 10 μL of PowerUp™ SYBR® Green Master Mix (Applied Biosystems), and nuclease-free water. PCR was performed using a StepOnePlus™ Real-Time PCR machine (Applied Biosystems). The PCR conditions included an initial step of 50 °C for 2 min and 95 °C for 2 min followed by 40 cycles of 95 °C for 15 s and an annealing/extension temperature of 58 °C for 30 s. The PCR reactions were performed in triplicate. The standard curves were prepared fresh from synthetic gene fragment or PCR product serially diluted 10-fold in water (DNA concentration range for *alkB* was 18.5-1,850,000 copies and for ARHD 37.9-3,790,000 copies). For *alkB*, the gBlocks® Gene Fragment (IDT) was designed based on sequence of alkane 1-monooxygenase from *Alcanivorax sp.* NBRC 101098 (GenBank accession number: BAP15615). For *ARHD*, PCR products were generated using primers that allowed to produce longer products that contained the binding sites for the qPCR primers, S-ARHD-F: 5′-TCGTCGGCAGCGTCAGATGTGTATAA GAGACAGCAYMGNGG-3′and S-ARHD-R: 5′-GTCTCGTGGGCTCGGAGAT GTGTATAAGAGACAGCCANCCRTGRTA-3′. Generated products were purified and quantified using Qubit (Thermo Fisher).

2.2.2.2. Quantification of Oleispira. *Oleispira* per mL seawater estimates were obtained using primers designed in a previous study, OLEA339F 3′-TGGA CGAAAGTCTGATGCAGCCATG-5′ and OLEA520R 3′-TCCGATTAACGCTT GCACCTTTAGT-5′ (Krolicka et al., 2014). PowerUp SYBR master mix (2X) was used according to the manufacturer's instructions, with 10-fold diluted DNA extracts as template (2 μL/reaction), 20 pmol of each primer per reaction (reaction volume = 20 μL) and an annealing temperature of 60 °C. The qPCR program was as follows: initial activation at 50 °C for 2 min, additional activation at 95 °C for 2 min, followed by 40 cycles of denaturation at 95 °C for 30 s and annealing and extension at 60 °C for 1 min. Standard curves were established using DNA extracted from pure culture of *Oleispira antarctica* RB-8. Standards were included on each plate and copy numbers were determined based on each individual standard curve for each individual plate. The highest standard contained 1.68 ng of *Oleispira* DNA, equivalent to approximately 344,000 *Oleispira* genomes (cells). A 6 point 10-fold dilution series were prepared each time, with the lowest standard containing DNA equivalent to approx. 3.44 *Oleispira* cells.

2.2.2.3. Quantification of total number of prokaryotes (TNP). TNP was quantified using degenerate 16S rRNA gene primers 515F (5′-GTGCCAGCMGCCG CGGTA-3′) and 806R (5′-GGACTACHVGGGTWTCTAAT-3′) recommended by the Earth Microbiome Project (Caporaso et al., 2011). PowerUp SYBR master mix (2X) was used according to the manufacturer's instructions, with 10-fold diluted DNA extracts as template (2 μL/reaction), 20 pmol of each primer per reaction (reaction volume = 20 μL) and an annealing temperature of 57 °C. The qPCR program was as follows: initial activation at 50 °C for 2 min, additional activation at 95 °C for 2 min, followed by 40 cycles of denaturation at 95 °C for 15 s and annealing at 57 °C for 30 s and extension at 72 °C for 45 min. Standard curves were established using full length 16S rRNA gene amplicons of *Oleispira antarctica* RB-8. Amplicons were generated using the primer set 27F and 1492R on DNA extracted from pure culture of the strain (Heuer et al., 1997). Standards were included on each plate and copy numbers were determined based on each individual standard curve for each individual plate. The highest standard contained 5,600,000 copies/μL from which a 5 point 10-fold dilution series were prepared each time.

2.3. Amplicon sequencing of alkB and ARHD

PCR products were generated using the same primer sets as used for qPCR analysis. Reactions were carried out in 100 μL total volume with Platinum HotStart 2X Master Mix (Thermo Fisher Scientific), 2 μL of 10-times-diluted DNA from each individual bottle, 1.5 μM of each primer, and nuclease-free water. The PCR conditions included an initial denaturation step of 94 °C for 2 min followed by 35 cycles of 94 °C for 15 s, 58 °C for 30 s and 72 °C for 15 s. Amplicon sequencing was carried out with Illumina paired-end strategy with read length of 150 bp (for ARHD amplicons) and 250 bp (for alkB amplicons) by Novogene Europe (Cambridgeshire, UK). Sequencing data (raw sequences without barcodes and primers) was submitted to the Sequence Read Archive (SRA) of the National Center for Biotechnology Information (NCBI) under BioProject PRJNA556155. BioSample accessions are available in Supplementary Table S1B.

2.4. Chemical analysis

Seawater was collected for chemical analysis simultaneously to seawater collection for bacterioplankton analysis. From all experimental bottles, 100 mL seawater was collected for liquid-liquid extraction followed by spectrofluorometric measurement of total aromatics on day 2, 4, 7 and 14. In order to estimate the amount of oil (aromatics) in the water fraction, the following procedure was carried out. Seawater samples (100 mL) were mixed with 5 mL cyclohexane in a Schott flask and shaken vigorously for 1 min. The organic fraction was then obtained after phase separation (30 min) in a separation funnel. To this, sodium-sulfate was added in order to adsorb any remaining water content. Extracts were analyzed immediately using a Lumina Spectrometer (Fisher Scientific) in “Wavescan” mode. Excitation wavelength was set to 270 nm and emission wavelength to 300–500 nm. Calibration curves for each of the two crude oils were obtained by dissolving the crude oil directly in cyclohexane at the concentrations expected in the experimental setup.

Pooled seawater samples (3 × 300 mL) for each condition were analyzed for total hydrocarbon content (THC) and polycyclic aromatic hydrocarbon (PAH) by GC-FID and GC–MS analysis, respectively. The analyses were performed by Intertek West Lab AS (Tananger, Norway) according to ISO 28540:2011. PAH analysis was performed for the 16 EPA (Environmental Protection Agency, USA) priority aromatic compounds (Keith, 2015) on day 2, 7 and 14 (T1, T3 and T4). THC analysis was performed on day 2, 4, 7, and 14. The limit of quantification for THC analysis was 0.4 $mg \cdot L^{-1}$ and for PAHs, this ranged from 0.01 to 0.02 $\mu g \cdot L^{-1}$.

2.5. Data analysis

2.5.1. Metagenomics

Processing of raw data including quality control, filtering, trimming, assembly, gene prediction, dereplication and mapping reads back onto the dereplicated genes have been carried out by the service provider (Novogene Europe, UK). These steps were performed as follows. Low quality bases ($Q \leq 38$) exceeding 40 bp were trimmed, reads with N nucleotides over 10 bp and reads which overlapped with adapter sequences over 15 bp were trimmed. Clean_Q20 values (the percentage of bases whose quality score is greater than 20, *i.e.*, error rate < 0.01) were summarized. Samples passing QC were initially assembled using a SOAPdenovo protocol. The resulting scaffolds were trimmed at "N" to obtain fragments without "N". These scaftigs (*i.e.*, continuous sequences within scaffolds without N) were used as fragments to map clean reads back to using SOAP 2.21. Unutilized PE reads were collected for all samples after the first round of assembly and a mixed assembly was then conducted on the unutilized reads with the same assembly parameters. The "NOVO_MIX" genes represent the mixed assembly result, while all other genes are named after the sample from which their scaftigs originated from. The scaftigs of each sample and mixed assembled scaftigs <500 bp were trimmed and remaining scaftigs were used for further analysis and gene prediction. Scaftigs (≥500 bp) were used for ORF (Open Reading Frame) prediction by MetaGeneMark. ORFs ≤100 nt were trimmed and dereplicated by CD-HIT to generate gene catalogues. Dereplicating was performed with default parameters: identity = 95%, coverage = 90% (CD-HIT parameters: -c 0.95, -G 0, -aS 0.9, -g 1, -d 0) (Li and Godzik, 2006). The longest gene was chosen as the representative gene (unigene). Clean reads were mapped to gene catalogue using SoapAligner with parameters: -m 200, -x 400, identity ≥95%. The resulting read count table was converted into relative counts for each gene (G_k) using the formula: $G_k = \frac{r_k}{L_k} \bullet \frac{1}{\sum \frac{r_i}{L_i}}$, where r_k is the read count of gene k and L_k is the length of gene k, r_i is the read count of the ith gene and L_i is the length if the ith gene.

Read count table, relative count table, dereplicated nucleotide and amino acid sequences were used in subsequent analysis. KEGG pathway mapping was performed by the service provider (Novogene Europe, UK) and the provided KEGG assignments and read count table (relative only) were used for downstream statistical analysis as obtained. Taxonomy assignment was performed as follows. Dereplicated protein sequences were searched against the NCBI nr database (downloaded in July 2019) using DIAMOND BLASTP with default parameters (e-value cut-off: 0.001) (Buchfink et al., 2015). BLAST results were imported into MEGAN and LCA was performed to assign taxonomy to each unique protein sequence (LCA parameters: MinScore = 50.0, MaxExpected = 0.001, TopPercent = 10.0, MinSupportPercent = 0.05, MinSupport = 896, LCA = naïve, mode = BlastP) (Huson et al., 2007). Taxonomy information, *i.e.*, taxonName and taxonPath were exported for further analysis.

2.5.1.1. alkB and ARHD genes present in the metagenome. The alkane 1-monooxygenase (*alkB*) protein sequences (305 CDS) were retrieved from metagenomic data using KEGG pathway annotation (K00496). The *alkB* protein sequences were aligned using Clustal Omega and trimmed to the conserved regions targeted by primers (EHYG and DHHA). Sequences that did not overlap with the conserved region on both ends were removed (212 sequences remaining after filtering). Filtered and trimmed protein sequences were then clustered using CD-HIT with 90% similarity threshold (118 clusters). The normalized read counts were summed up for each cluster and assigned to the representative sequence from each cluster. The representative sequences were annotated using DIAMOND (Buchfink et al., 2015) BLASTP analysis (with minimum bitscore 40) against the NCBI's nr protein database (downloaded in May 2020). The representative sequences were assigned to the taxon if the identity was >90% and collapsed at the genus level.

Aromatic ring-hydroxylating dioxygenase (ARHD) protein sequences (1487 CDS) were retrieved from metagenomic data based on BLAST analysis (with minimum bitscore 50) against the HyDeg database (Scoma et al., 2015). The retrieved ARHD protein sequences were aligned using Clustal Omega (Madeira et al., 2019) and trimmed to the conserved regions targeted by primers (HRG and YHGW). Sequences that did not overlap with the conserved region on both ends were removed (1132 sequences remaining after filtering). Filtered and trimmed protein sequences were then clustered using CD-HIT with 90% similarity threshold (759 clusters). The normalized read counts were summed up for each cluster and assigned to the representative sequence from each cluster. The representative sequences were annotated using DIAMOND (Buchfink et al., 2015) BLASTP analysis (with minimum bitscore 40) against the NCBI's nr protein database (downloaded in May 2020). The representative sequences were assigned to the taxon if the identity was >90% and collapsed at the genus level.

2.5.2. Phylogenetic tree construction

The representative protein sequences of the top 100 most abundant clusters trimmed to the conserved regions targeted by primers were aligned using Clustal Omega. The phylogenetic tree was constructed with MEGA-X software version 10.7.1 (Kumar et al., 2018). The optimal trees were obtained using the Neighbor-Joining method and 1000 bootstrap replicates. The evolutionally distances were calculated with the p-distance method and pairwise *deletion*. The trees were visualized with iTOL v 6.1.1. online tool available at https://itol.embl.de/ (Letunic and Bork, 2007). The heatmap datasets were generated with iTOL heatmap tool. Datasets values for the heatmaps were based on the normalized read counts summed up for each cluster.

2.5.3. Amplicon sequencing – alkB and ARHD

As part of an initial filtering process, first the *alkB* and ARHD sequences were clustered into 'OTUs' with USEARCH (97% similarity) resulting in 1570 *alkB* and 17,862 ARHD OTUs. Next, presence of chimeric sequences was checked using UCHIME 6.0 in *de novo* mode implemented in the Ribosomal Database Project (RDP) functional gene (FunGene) pipeline (Fish et al., 2013) at http://fungene.cme.msu.edu/FunGenePipeline/ (no chimera sequences were detected). Finally, an abundance-based filtering was performed, where sequences with low read count (<50 reads across all samples) were removed. The processing of ARHD and *alkB* genes was performed in a slightly different manner in the subsequent annotation process. In case of the ARHD genes, the remaining OTUS (nucleotide sequences) were clustered with CD-HIT (Li and Godzik, 2006) using a 90% similarity threshold (1409 clusters) prior to translation into protein sequences. The *alkB* OTUs were translated without further clustering. ARHD and *alkB* sequences were translated into proteins with translate6frames.sh tool from BBMap package (https://sourceforge.net/projects/bbmap/) and manually frameshift corrected. Protein sequences were trimmed to the conserved regions targeted by primers and filtered to remove sequences that were incomplete, contained nonsense codons, or did not encode for alkane 1-monooxygenase or aromatic ring-hydroxylating dioxygense (1509 *alkB* and 1396 ARHD sequences remaining after filtering). The filtered *alkB* protein sequences were clustered at this stage using CD-HIT (Li and Godzik, 2006) at 90% similarity threshold (171 *alkB* clusters). The obtained ARHD and *alkB* representative protein sequences were finally annotated using DIAMOND BLASTP against the NCBI's nr protein database using best blast hit approach, curated manually, and assigned to taxon if identity was >90%. The read counts from OTU table generated by USEARCH were summed up for each cluster and assigned to the representative sequence from each cluster.

2.5.4. Statistical analysis

All statistical analysis and visualizations were carried out in RStudio, R version 4.1.0 "Camp Pontanezen" (R Core Team, 2021). PERMANOVA analysis was performed using the adonis function with 9999 permutations from the vegan package (Oksanen et al., 2020). Linear correlation between functional gene abundance and total number of bacteria were performed using Pearson correlation test implemented in base R, cor() function. Non-metric multidimensional scaling (NMDS) was performed using the metaMDS() function implemented in vegan, using Euclidean distance metric and visualized using the plot() function of base R. Heatmap representation of the Euclidean distances was drawn with heatmap() function in the stats package of base R. For comparing gene abundances between sample groups, we used two-tailed paired *t*-test.

3. Results

3.1. Chemical analysis of the oil exposed seawater

The results of the fluorescence spectrophotometry and GC–MS analysis are summarized in Supplementary Table S2A. Briefly, the concentration of total aromatics measured by fluorescence showed large variations between replicate flasks of oil-exposed seawater particularly at T3 and T4, while the negative controls were more consistent across replicates. A decrease in concentration occurred in both the LIVE samples and negative controls (NEG), with significantly lower concentrations in live compared to autoclaved controls observed only at T4. In case of OIL-1 in the summer, no significant difference was apparent between live and negative control. The GC–MS analysis showed that the natural seawater contained some of the naphthalene-compounds (C1-, C2- and C3-naphthalene) and traces of higher molecular weight PAHs (C1-, C2-, and C3-phenanthrenes/anthracenes, C2-dibenzothiophene and pyrene) in both seasons, albeit at very low concentrations near the limit of quantification (<0.1 μg/L, Practical Quantification Limit (PQL) of the method is 0.01 μg/L ± 30–40%) (Supplementary Table S2B). Among the 16 EPA PAHs, those larger than chrysene were below detection limit (PQL for chrysene is 0.02 μg/L) in all tested samples throughout the experiments except benzo(*a*)pyrene, which was present at very low concentrations in the winter samples of OIL-2 at T3 and T4. At T1 (after 2 days of oil exposure), the most abundant PAHs were naphthalene and its alkyl derivatives under all conditions (2.8–51 μg/L). The concentration of total 16 EPA PAHs was similar at the start of the experiment in both seasons, but with higher amounts detected in the OIL-2 exposed seawater at T1 (Fig. 1A). The concentration of PAHs

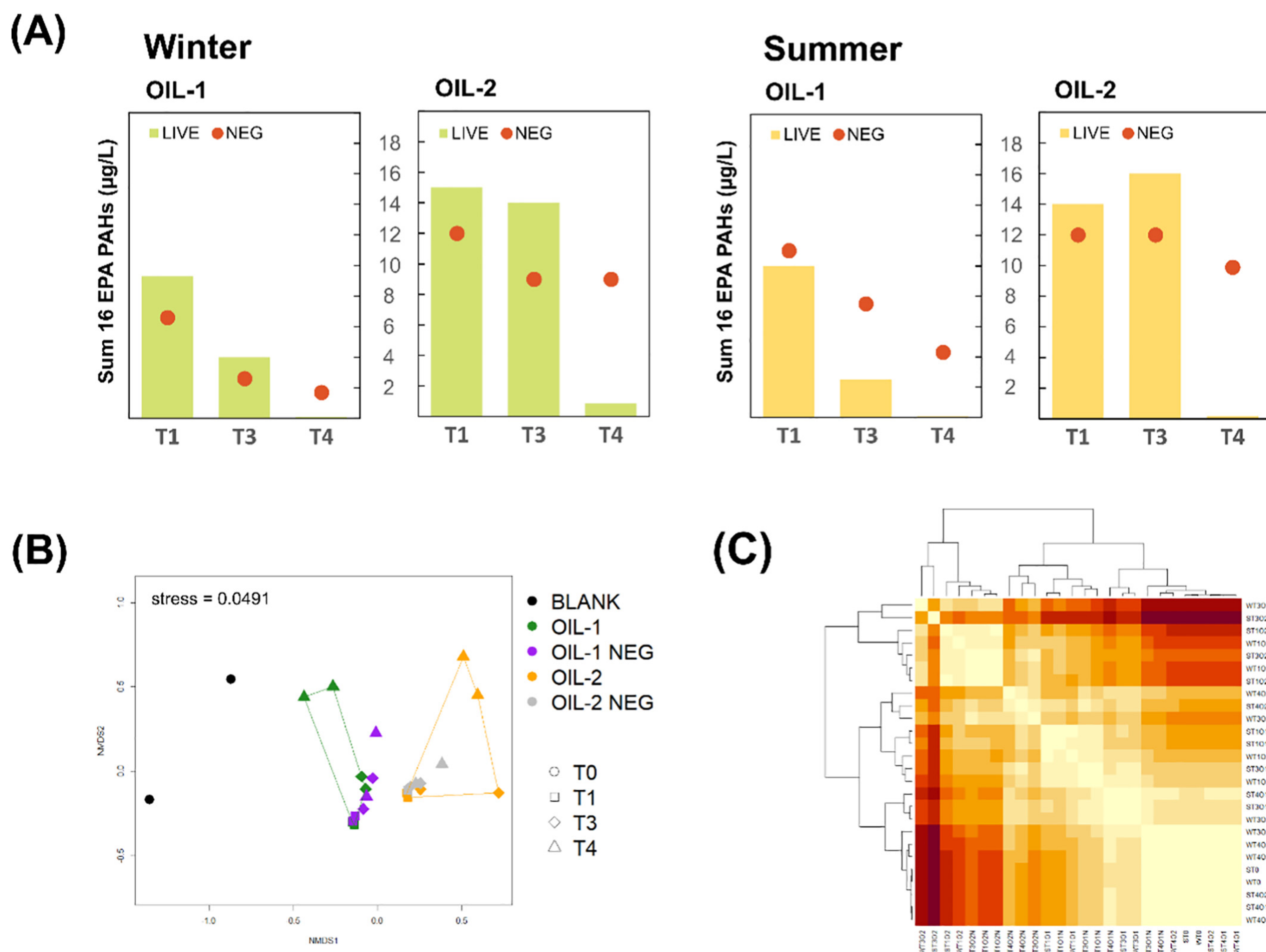


Fig. 1. PAH composition measured by gas chromatography mass spectrometry (GC–MS). **Panel A:** Concentration of the sum of 16 priority polycyclic aromatic hydrocarbons (PAHs) as defined by the Environmental Protection Agency (EPA) in seawater analyzed by GC–MS. **Panel B**: Non-metric multidimensional scaling (NMDS) representation of the PAH compositions (Euclidean distance was used on raw concentration values), green line connects the OIL-1 samples and orange line the OIL-2 samples. **Panel C**: Heatmap representation of the Euclidean distances (calculated using raw concentration values of the 16 EPA PAHs) between all samples. Red cell color indicates greater dissimilarity while yellow indicates greater similarity between samples. **Abbreviations**: W = winter and S = summer, T1 = day 2, T3 = day 7 and T4 = day 14 sampling time. NEG (N) = negative control (oil added to autoclaved seawater) and LIVE = exposed (oil added to seawater). O1 = OIL-1 and O2 = OIL-2.

generally decreased with time, to some extent also in the negative controls (Fig. 1A, "NEG"). Adsorption of the oil to the walls of all exposure bottles was visible. The decrease of PAHs appeared to be more gradual in the OIL-1 exposed seawater, while in case of OIL-2, the greatest decrease occurred between T3 and T4. There were no PAHs detected in OIL-1 exposed bottles at T4 and very low concentrations were measured in OIL-2 exposed ones.

Overall, season had little to no effect on the pattern of PAH degradation. The change in PAH composition over time showed significant differences only between oil types, time and the combination of the two variables (PERMANOVA, p-value = 0.0001 for both factors, 9999 permutations) (Oksanen et al., 2020). The successive PAH composition change of the two oil types is illustrated on an NMDS plot (Fig. 1B). The OIL-1 and OIL-2 samples, indicated by the line connecting the respective samples, grouped separately reflecting the distinct chemical characteristics of the two different oils. The NMDS plot also shows that the PAH composition of "LIVE" samples evolved differently from the "NEG" samples (autoclaved seawater amended with oil), distinguishing purely abiotic from the combination of biotic and abiotic changes. The hierarchical clustering representation placed all T4 samples (ST4O1, ST4O2, WT4O1 and WT4O2) in a single cluster, together with the T0 samples. This could indicate that the PAH composition was returning towards background levels existing in the natural seawater (Fig. 1C). As Fig. 1C shows, all T1 samples clustered together with their respective negative controls, suggesting that changes in PAH composition were purely due to abiotic processes until this time point.

3.2. Abundance of functional genes and 16S rRNA genes determined by qPCR

Measuring the abundances of two initial oxygenases allowed us to test the hypothesis whether oil-related functional assays targeting initial oxidases can be used for distinguishing oil exposed and non-exposed seawater communities under different seasons and with different oil types. The statistical analysis focused on comparing copy numbers of functional genes, *alkB* and ARHD, between control and oil-exposed samples over time.

3.2.1. Alkane 1-monooxygenase (alkB)

Abundance of *alkB* genes remained below 5000 copies/mL seawater in all samples in both seasons at all time points and showed a seasonal and oil-type dependent pattern (Fig. 2). In general, the T1 samples had very low number of *alkB* genes (<100 copies/mL seawater), particularly in the winter experiment (<50 copies/mL seawater). In this season, *alkB* copy numbers peaked at T3 in OIL-1 and where highest at T4 in OIL-2. On the same time points, *alkB* copy numbers were significantly higher in the oil exposed samples compared to the control, implying that the increase occurred as a result of oil presence. In the summer, the number of *alkB* genes in OIL-1 was not significantly higher than in the control, at any time points. In OIL-2, *alkB* abundance was significantly higher at T3 and T4 compared to control.

3.2.2. Aromatic ring-hydroxylating dioxygenase (ARHD)

In case of ARHD, much higher absolute values have been observed than for *alkB* in general (up to 4,000,000 copies/mL). The highest abundances were reached at the end of each experiment in OIL-2 samples (Fig. 3). During the winter experiment, significant differences between the blank and oil-exposed samples were observed at T3 for OIL-1 and at T4 for OIL-2. During the summer experiment, ARHD abundances showed no significant difference from the control on OIL-1, whereas there was a significant difference between OIL-2 exposed seawater and control at both T3 and T4 sampling times.

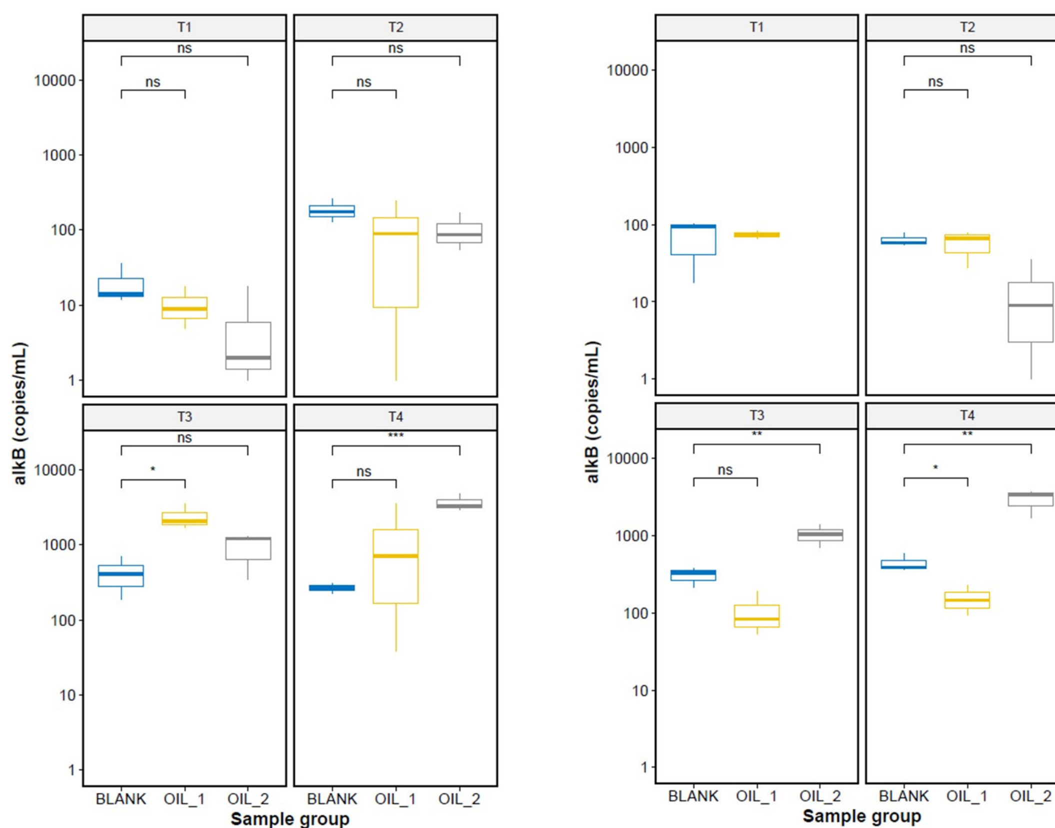


Fig. 2. Abundance of alkane 1-monooxygenase (*alkB*) genes determined by qPCR. **Left**: winter experiment and **Right**: summer experiment. **Abbreviations**: ns = non-significant according to paired t-test, * corresponds to: $0.01 < p < 0.05$ and ** corresponds to: $0.001 < p < 0.01$, T1 = day 2, T2 = day 4, T3 = day 7 and T4 = day 14. Sample group = experimental condition.

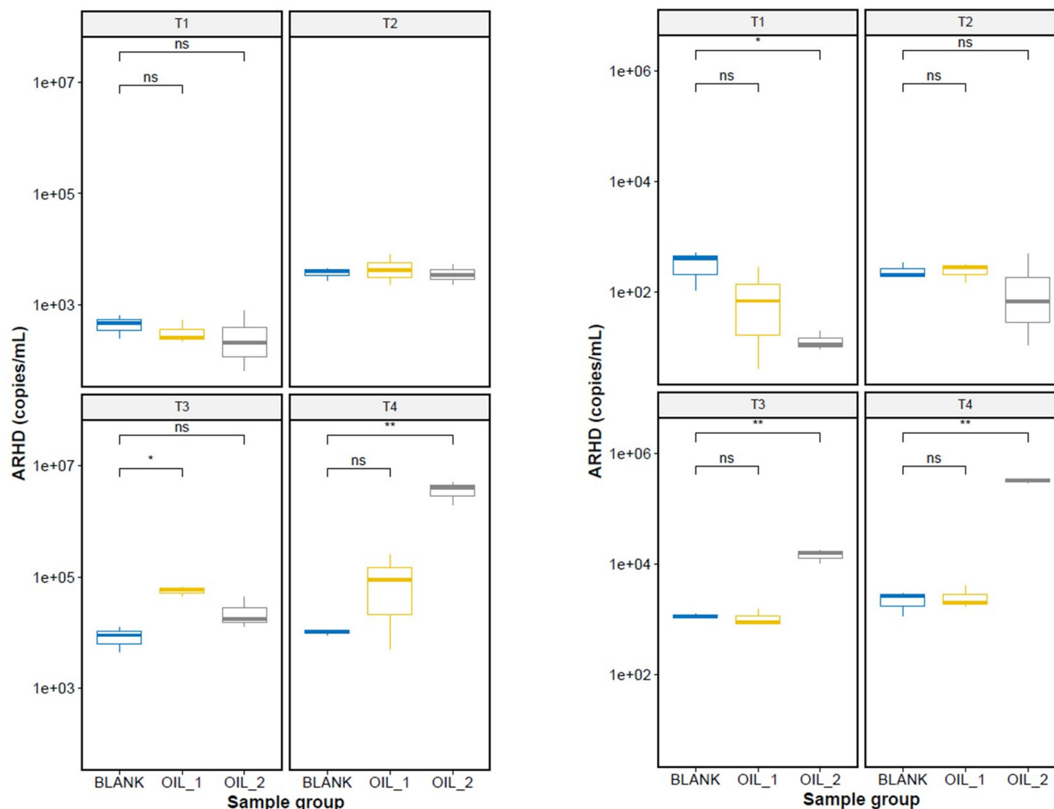


Fig. 3. Abundance of aromatic ring-hydroxylating dioxygenase (ARHD) genes determined by qPCR. **Left**: winter experiment and **Right**: summer experiment. **Abbreviations**: ns = non-significant according to paired *t*-test, * corresponds to: $0.01 < p < 0.05$ and ** corresponds to: $0.001 < p < 0.01$, T1 = day 2, T2 = day 4, T3 = day 7 and T4 = day 14. Sample group = experimental condition.

3.2.3. *Oleispira*

Oleispira was nearly undetectable in the natural seawater with abundances of 35 ± 19 (n = 3) per mL of seawater in the summer and < 10 per mL (n = 3) in the winter. During both seasons, *Oleispira* numbers increased in the control samples to some extent, reaching a maximum of 1988 ± 551 per mL in the summer (T3) and 1482 ± 653 per mL in the winter (T3). During the winter experiment, *Oleispira* numbers increased faster on OIL-1, peaking at T3 with 51,513–94,140 per mL in contrast to OIL-2 reaching maximum number of 3896–67,291 per mL at T4 (Fig. 4). During the summer experiment, *Oleispira* numbers again reached their maximum earlier on OIL-1 than on OIL-2, with peak abundances of 2998–18,470 per mL by T2 and 10,803–62,517 per mL by T3, respectively (Fig. 4). Mean *Oleispira* copy numbers appeared to be higher on OIL-1 compared to control already at T2 but significant differences between control and oil exposed samples were found only at T3 in both seasons.

3.2.4. Total number of prokaryotes (TNP)

The collected seawater contained 7743 ± 4382 and 735 ± 643 copies of the 16S rRNA gene per mL seawater in the summer and winter, respectively. During the experiments, TNP values showed season- and oil type-dependent trends (Supplementary Fig. S1). Within the first 2 days, TNP increased in the controls, and OIL-1 samples, while in case of OIL-2 samples,

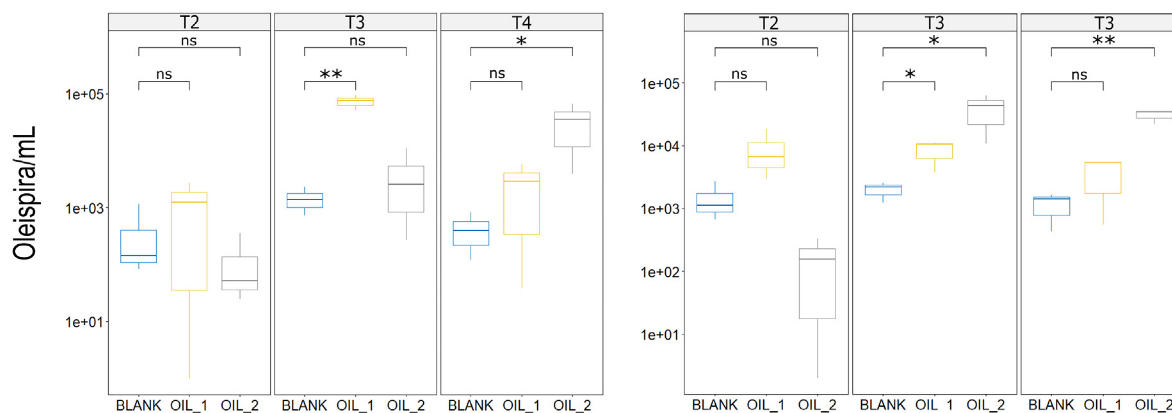


Fig. 4. Abundance of *Oleispira* (per mL seawater) determined by quantitative PCR (qPCR). **Left**: winter experiment and **Right**: summer experiment, ns = non-significant according to paired *t*-test, * = $0.01 < p < 0.05$ and ** = $0.001 < p < 0.01$. Sampling times: T2 = day 4, T3 = day 7 and T4 = day 14. Sample group = experimental condition.

an initial decrease (2 orders of magnitude) was observed in the summer experiment. Overall, the TNP followed the trends observed for *alkB* and ARDH genes for OIL1- and OIL-2 and in general, a greater microbial growth occurred on heavy oil than on the lighter Troll oil by the end of the incubation period. Growth on the Troll oil did not exceed that observed in the controls except for a peak at T3 in winter.

3.3. Community composition and functional diversity revealed by shotgun metagenomics

A total of 13.47 – 42.46 million reads were obtained for each sample with an average read count of 26.77 million across all samples. On the 4,193,933 scaftigs, 6,944,883 genes were predicted of which 2,660,591 were unique.

3.3.1. Taxonomy

The composition of the two natural seawater samples were very similar on the genus level, with unclassified Gammaproteobacteria, *Candidatus* Pelagibacter, *Candidatus* Thioglobus, environmental *Siphoviridae*, unclassified Chloroflexi, unclassified *Planctomycetaceae*, unclassified Euryarcheota, unclassified *Candidatus* Poseidoniales, unclassified Alphaproteobacteria, *Nitrosopumilus*, environmental viruses, SAR324 cluster, unclassified Flavobacterales and *Candidatus* Marinimicrobia being the most dominant genera in both seasons (Supplementary Fig. S2). These genera remained the most dominant throughout the 14 days in the control (blank) samples, with 4 additional ones appearing among the top 15, namely *Colwellia*, unclassified *Porticoccaceae*, *Oleispira* and *Oceanobacter*. The most notable change in the seawater community of the blank samples was the bloom of *Colwellia*, particularly in the winter, where its relative abundance reached 40% at T3 (Supplementary Fig. S3). Apart from this difference in the extent of the *Colwellia* bloom, the remaining abundant genera maintained fairly stable relative abundances in both seasons over time.

In the oil-exposed samples, the top 15 most abundant genera showed moderate season- and oil type-dependent changes, with time influencing the community structure most (PERMANOVA, $R^2 = 0.53$, p-value = 0.0006) and season having a weak and less significant effect (PERMANOVA, $R^2 = 0.14$, p-value = 0.044). As Fig. 5 shows, blooms of a few genera characterized the main changes in community structure. Like in the blank samples, *Colwellia* bloomed in the oil-exposed seawater during winter, dominating the OIL-2 samples at T3 with 50% relative abundance. Two obligate hydrocarbonoclastic bacterial (OHCB) genera, *Oleispira* and *Cycloclasticus* was clearly associated with the presence of both oil types. These two genera bloomed successively, with *Oleispira*'s increase in relative abundance preceding that of *Cycloclasticus*. This sequential pattern occurred earlier in the summer, than in the winter. Interestingly, *Oleispira* reached higher relative abundances during the summer than in the winter. *Cycloclasticus* showed an increase towards the end of the exposures, dominating the OIL-2 samples at t4 in both seasons with relative abundances reaching nearly 50%. The third major blooming genera, *Oceanobacter* accompanied *Oleispira* in both seasons and exceeded it in relative abundance during the winter. A third OHCB, *Alcanivorax* only appeared during the summer experiment in both oil types at T3 and T4, with very low relative abundances.

3.3.2. alkB and ARHD genes in the metagenome

Phylogenetic trees were constructed from genes identified in the metagenome as *alkB* or ARHD, to illustrate the taxonomic diversity of these sequences and the overlap with the obtained amplicons. The obtained trees are shown in Supplementary Tree S1 and Supplementary Tree S2 for *alkB* and *ARHD*, respectively and the iTOL projects are available through: https://itol.embl.de/shared/2hI5IxcpmUN3u and https://itol.embl.de/shared/1RCyg8cTIg6RI, for alkB and ARHD genes, respectively. The *alkB* sequences extracted from the metagenome represented a large diversity of bacteria, with the most abundant ones belonging to the genera *Cycloclasticus*, *Oleispira*, *Oleiphilus*, *Sphingopyxis*, *Paraperlucidibacteria*, SAR92 clade, *Pseudophaeobacter* and unclassified Gammaproteobacteria (Fig. 6). According to the *in silico* analysis, 2 out of 5 *Cycloclasticus*, 2 out of 5 *Oleispira*, all 4 *Alcanivorax*, and all 3 *Oleiphilus alkB* sequences could be amplified with the degenerate primers used in the qPCR analysis. In general, less than 40% of the *alkB* sequences present in the metagenome could be expected to be amplified, missing one abundant *Cycloclasticus* sequence and two abundant *Oleispira* sequences.

In case of the ARHD genes, a more limited taxonomic diversity was found in the metagenome, with *Cycloclasticus* clearly dominating (Fig. 7). Approximately one quarter of the top 100 genes was classified as *Cycloclasticus*. Other abundant taxa included *Pseudoalteromonas*, *Candidatus* Thioglobus, *Candidatus* Pelagibacter, *Oceanobacter*, SAR92 clade, and *Rhodospirillaceae*. When it comes to the coverage of the degenerate primers, mostly *Cycloclasticus* genes, unclassified Gammaproteobacteria, *Rhodospirillaceae* (although not the more abundant one), SAR92 clade, and a *Candidatus* Marinimicrobia genes would be expected to be amplified. Overall, less than 25% of the top 100 genes were covered according to the *in silico* analysis, missing several abundant *Cycloclasticus* sequences, an abundant *Pseudoalteromonas* sequence and the *Candidatus* Thioglobus sequences.

3.4. Functional gene diversity revealed by amplicon sequencing

The 171 alkane 1-monooxygenase sequences (clusters) represented 36 genera and 110 unclassified OTUs (Supplementary Table S3). The most abundant amplicons were from genus *Sphingopyxis*, *Pseudophaeobacter*, *Oleiphilus*, SAR92 clade bacterium, *Alcanivorax* and *Marinobacter*. *Oleispira*

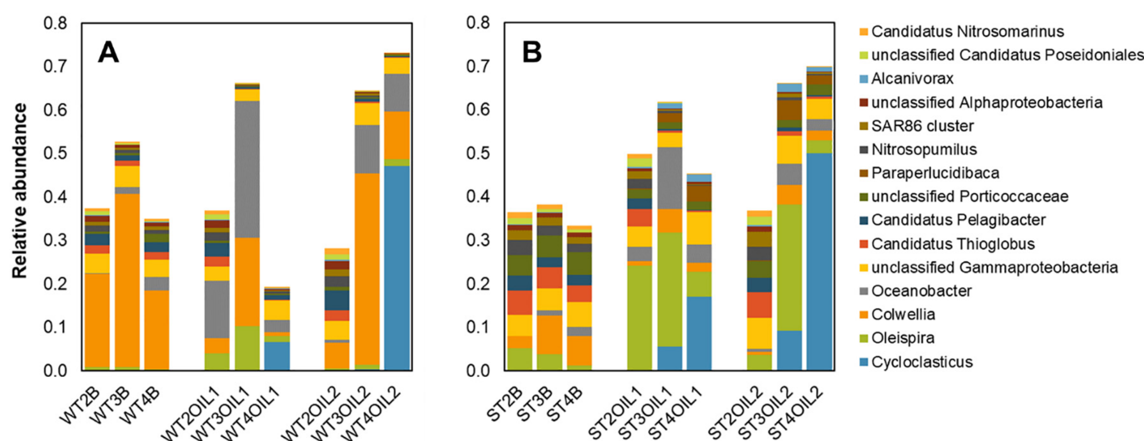


Fig. 5. Top 15 genera detected in the metagenomes of the winter experiment (**A**) and the summer experiment (**B**). **Abbreviations**: W = winter and S = summer, T2 = day 4, T3 = day 7, and T4 = day 14, B = blank.

Fig. 6. Phylogenetic tree of alkane-1-monooxygenase sequences (K00496) extracted from the metagenome. Thick black lines mark those genes which were *in silico* covered by the primer sequences. Winter relative abundance values are shown in blue while summer relative abundance values are shown in red color. Grey circles represent bootstrap value >0.5.

did appear among the taxa, ranked as number 21 based on average relative abundance. No *Cycloclasticus* sequence was found among the amplicons. A few seasonal differences could be observed in the taxonomic composition of the *alkB* amplicons. The collected seawater contained a higher relative abundance of *Pseudophaeobacter* and *Arenicella* in the winter, while SAR92 clade bacterium, Gammaproteobacteria bacterium, *Alcanivorax*, and *Porticoccaceae* bacterium showed higher relative abundances in the summer. In the oil exposed samples, *Sphyngopyxis* increased in relative abundance during the winter while *Oleiphilus* and *Alcanivorax* during the summer.

The 1396 ARHD sequences (clusters) represented 39 genera and 288 unclassified OTUs. The two most abundant taxa were *Cycloclasticus* and *Gammaproteobacteria* bacterium (Supplementary Table S4). Only minor differences can be seen between the two seasons (controls and oil exposed alike) and nearly none between the two oil types. In the winter, *Cycloclasticus* and *Gammaproteobacteria* bacterium ARHD clusters dominated and over time, *Cycloclasticus* amplicons dominated in all oil exposed samples.

A comparison of the top ten best represented taxa in the metagenomes *vs* qPCR amplicons is shown in Table 1. Despite the large mismatch between the existing diversity of these functional genes and the breadth of the diversity expected to be amplified by the degenerate primers, there were several overlapping taxa.

4. Discussions

4.1. Chemistry

Chemistry analysis showed that despite the addition of a relatively high nominal oil concentration (10 mg/L) the levels of dissolved PAHs were low (on the order of magnitude of 10 μg/L) and aliphatic compounds were below detection limit (<100 μg/L). Such low concentrations are typically

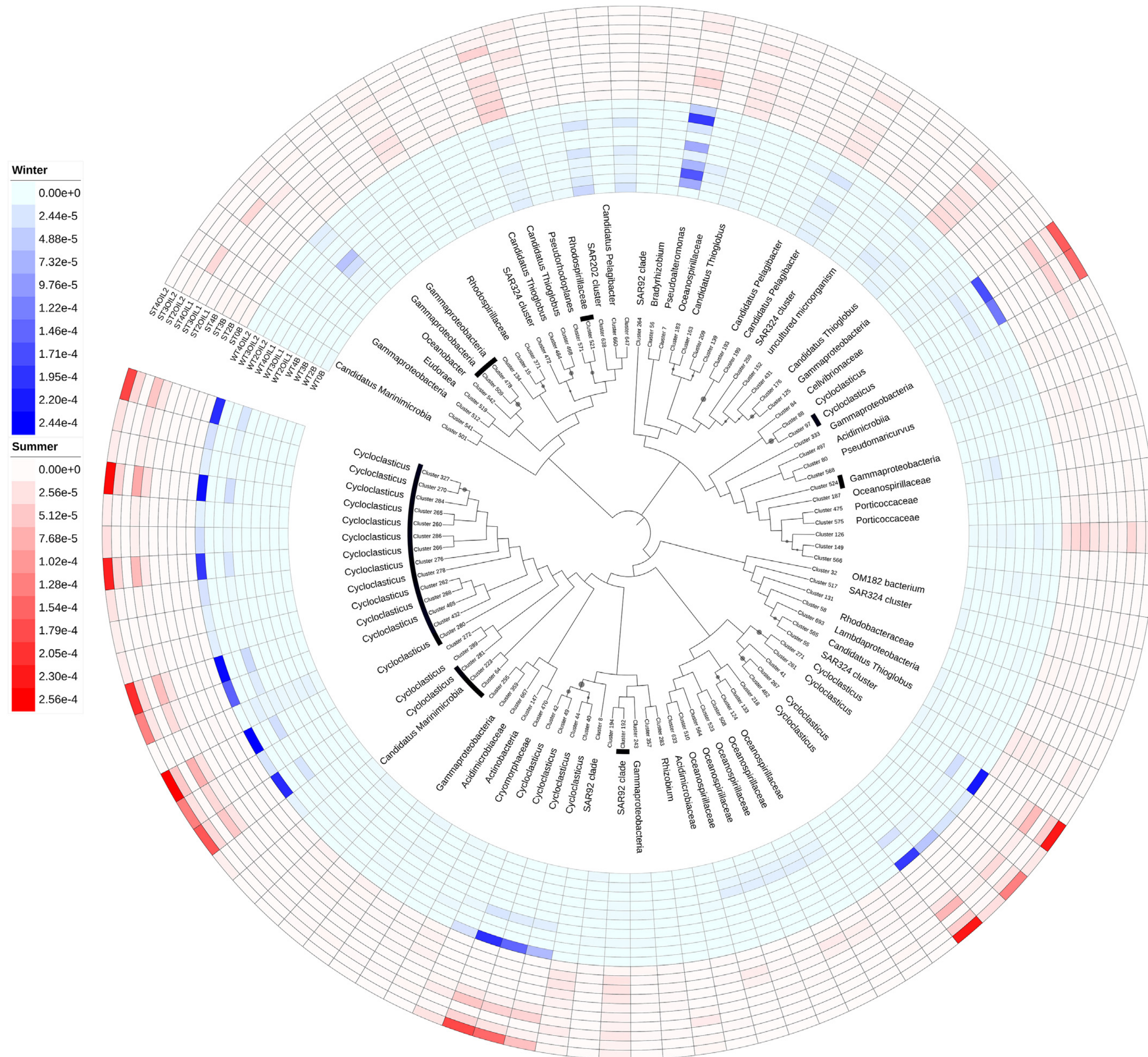


Fig. 7. Phylogenetic tree of aromatic ring-hydroxylating dioxygenases extracted from the metagenome. Thick black lines mark those genes which were *in silico* covered by the primer sequences. Winter relative abundance values are shown in blue while summer relative abundance values are shown in red color. Grey circles represent bootstrap value >0.5.

Table 1
Top 10 taxa identified from the metagenomics-derived sequences of *alkB* and ARHD genes and the amplicon sequences of *alkB* and ARHD generated by the degenerate qPCR primers.

alkB genes		ARHD genes	
Metagenome	Amplicons	Metagenome	Amplicons
Cycloclasticus	*Sphingopyxis*	*Cycloclasticus*	*Cycloclasticus*
Oleispira	*Pseudophaeobacter*	Candidatus Thioglobus	Gammaproteobacteria
Thalassolituus	OTU1	Gammaproteobacteria	Porticoccaceae
Gammaproteobacteria	*Oleiphilus*	*Pseudoalteromonas*	Rhodobacteraceae
Oleiphilus	OTU4	Rhodobacteraceae	Rhodospirillaceae
ST3OIL1_124315	SAR92 clade	*Actinobacteria*	Alphaproteobacteria
WT4OIL1_100072	*Alcanivorax*	Oceanospirillaceae	Myxococcales
SAR92 clade	*Marinobacter*	Rhodospirillaceae	Dehalococcoidia
Paraperlucidibacteria	*Arenicella*	Acidimicrobiaceae	Pseudomonadales
Sphingopyxis	Gammaproteobacteria	Candidatus Pelagibacter	OTU285

a result of low energy mixing systems without dispersants, like the gentle magnetic stirring employed in this study, besides the overall low solubility of alkanes and PAHs in seawater at low temperatures (Brakstad and Faksness, 2000; Brakstad et al., 2018; Faksness et al., 2008; Whitehouse, 1984). We chose this nominal concentration to mimic a realistic scenario in the field with a relatively small leakage, rather than a massive oil spill (*e.g.* >10 g/L oil in seawater). Similar nominal oil concentrations (up to 100 mg/L) are used in studies where water accommodated fractions (WAF) of oil are prepared and used in subsequent biodegradation experiments (Brakstad and Faksness, 2000; Hansen et al., 2018). The low hydrocarbon concentrations in the water, however, may lead to limited growth, hampering the detection of hydrocarbonoclastic bacteria. For example, genus specific qPCR assays targeting bacteria known to bloom in the presence of oil have previously revealed an concentration-dependent growth of *Oleispira*, *Cycloclasticus* and *Alcanivorax* in the range of 30–2000 μg/L oil (Krolicka et al., 2019).

4.2. Seasonal and oil-type dependent community changes

Shotgun metagenomics revealed a season and oil type dependent community change over the 14 days with sequential blooms of hydrocarbon-degradation associated genera as previously observed (Head et al., 2006; Ribicic et al., 2018a; Ribicic et al., 2018c; Thomas et al., 2021; Tremblay et al., 2019). A transient increase of alkane-degrading *Oleispira* and *Oceanobacter* was followed by the bloom of PAH-degrading *Cycloclasticus,* as it would be expected in case of a primary degradation of aliphatic hydrocarbons, followed by that of the aromatic compounds. Although the chemical analysis covered only PAHs, the increase in alkane-degrading bacteria indicates that aliphatic hydrocarbons were present and consumed.

The two major seasonal differences in this study were the higher relative abundance of *Colwellia* and *Oceanobacter* in the colder season. *Colwellia* genus contains cold-adapted opportunistic microorganisms associated with degradation of a wide range of organic matter (*e.g.* agar from macroalgae and collagen from bones) and also oil compounds (Boccadoro et al., 2018; Borchert et al., 2021; Campeao et al., 2019; Gui Yuanyuan, 2020; Gutierrez et al., 2013; Krolicka et al., 2019; Peña-Montenegro et al., 2020). Despite its widespread occurrence in oil-related environments, there is still little direct evidence for its ability to activate aliphatic or aromatic hydrocarbons. Here, we did identify both an alkane 1-monooxygenase and an aromatic ring-hydroxylating dioxygenase from *Colwellia* suggesting its potential in such processes. Nevertheless, to associate the presence of *Colwellia* with biodegradation of aliphatic or aromatic hydrocarbons remains challenging due to its versatile metabolic profile and its tendency to bloom in control samples under laboratory conditions (Krolicka et al., 2017). Neither the *alkB* nor the ARHD genes of *Colwellia* evident from the metagenome data were amplified by the degenerate primers. The other season-specific genus, *Oceanobacter*, belongs to *Oceanospirillaceae*, a family known to contain several species with unique metabolic traits, including ability to degrade hydrocarbons such as *Oleispira* (Satomi and Fujii, 2014). Indeed, *Oceanobacter* was found to degrade aliphatic hydrocarbons equally well as *Alcanivorax borkumensis* SK2 with the advantage of not requiring nutrient addition to do so (Teramoto et al., 2009). Although these results were established under tropical conditions, it is likely that there exists a cold-adapted species within this genus, able to perform similar metabolic processes under cold conditions. The *Oceanobacter* in our dataset did possess both *alkB* and ARHD genes, confirming this possibility. Nevertheless, neither of these were present in the amplicon data (*i.e.* not capture by the qPCR assays). Due to its specific growth in oil-exposed seawater, *Oceanobacter* should be further evaluated for its inclusion in the repertoire of taxa suitable for detecting oil presence in the sea.

4.3. Abundance patterns and their implication for oil detection

Based on *alkB* or ARHD gene abundances, none of the samples collected within the first 4 days of the experiment showed signs of oil presence (*i.e.*, no significant difference between control and oil exposed samples). Both functional gene assays successfully signaled the presence of both oils during the winter experiment, while neither the *alkB* nor the ARHD assay showed significantly higher abundances on OIL-1 compared to control during the summer. The assays signaled oil presence only for OIL-2 exposure during the summer. According to the sequencing results, there was a clear discrepancy in the taxonomic diversity captured by the degenerate primers and that revealed by the shotgun metagenomic approach. The *in silico* coverage analysis revealed that one of the more abundant *Oleispira alkB* genes present in the metagenome was not captured by the *alkB* qPCR assay. Indeed, *Oleispira alkB* sequences were not among the top 10 best amplified, despite *Oleispira* being the most abundant alkane degrader in these samples. The same was observed for the other abundant alkane-degrader, *Oceanobacter*. Instead, the *alkB* assay appeared to preferentially amplify *Sphingopyxis*, *Pseudophaeobacter*, unclassified OTU1 and *Oleiphilus* genes across all season and sampling times. In general, the taxonomic composition of amplified *alkB* and also ARHD genes was very similar between oil-exposed and control samples. Both assays mainly amplified genes from taxa that were present at fairly stable low relative abundances (except for ARHD, *Cycloclasticus*) and largely missed those that bloomed in response to oil. This led to the observed low copy numbers and small differences between sample types and that the assays could not detect significantly higher *alkB* or ARHD abundances in the oil-exposed samples compared to control even when the taxonomy clearly indicated oil-related changes. The ARHD assay was more successful due to its ability to capture several of the abundant *Cycloclasticus* genes on T3 and T4. As a result, its abundance changes showed good agreement with PAH degradation patterns (Supplementary Fig. S7).

Considering the abundance patterns of functional genes in the light of the total prokaryotic abundance (TNP), it was evident that *alkB* and ARHD copy numbers followed the changes in overall prokaryotic abundance to some extent. Pearson test revealed a strong correlation between the abundance of functional genes and TNP values ($R = 0.86$, $p < 2.2E\text{-}16$ and $R = 0.89$, $p < 2.2E\text{-}16$ for *alkB* and ARHD, respectively), meaning that the overall bacterial abundance change dictated the abundance trends of the functional genes. Similar results were reported in a study of Baltic Sea seawater and sediment, where it was found that rather than correlating with hydrocarbon concentration, *alkB* and PAH-oxygenase abundance correlated with 16S rRNA copy numbers (Miettinen et al., 2019). These authors studied chronically polluted samples, most likely harboring stable microbial communities well adapted to the presence of hydrocarbons. Such well-established communities may have a stable proportion of hydrocarbon exposure related microorganisms containing *alkB* and PAH-oxygenases. Since our approach was to mimic an acute oil leakage event, which is followed by the rearrangement of community composition, we hypothesized that an elevated *alkB* abundance in the oil exposed sample in comparison to control should be expected, due to the transient bloom of HCBs (*i.e.* a change in the proportion of the community related to hydrocarbon exposure). This assumption overlooked the fact that initially, oil exposure can lead to a decrease in prokaryotic abundance due to toxic effects which in turn results in an overall decrease in functional gene abundances as well (Heipieper and Martínez, 2010). However, an overall biomass and diversity loss could also be expected to result in that the fraction of oil-tolerant and oil-degrading microbes increases selectively already in the early stages. The Ecological Index of Hydrocarbon Exposure (EIHE) relies on this idea (Lozada et al., 2014). In fact, initial EIHEs in our experiments were 1.5 and 2.0% for the natural seawater in the summer and in the winter experiment, respectively. Following an initial increase, EIHEs declined over time in the blank samples. On the contrary, a continuous increase in EIHE was observed in case of OIL-2 (maximum of 52 and 56% in winter and summer, respectively at T4) and a transient increase with maximum at T3 (18 and 39% in winter and summer, respectively) in case of OIL-1 in both seasons (Supplementary Fig. S4). Hence, even if the absolute *alkB* numbers did not provide a reliable signal of oil presence in our study, the percentage of *alkB* genes (*alkB* to TNP ratio) could have shown a difference between control and oil exposed samples. Unfortunately, this was not the case for the ratios calculated from the *alkB* and TNP copy numbers, as they generally remained below the ratios of the blank samples (Supplementary Fig. S5). The relative abundances of *alkB* genes present in the metagenome

(K00496), however, showed the anticipated trend with an early increase in OIL-1, already at T2. (Supplementary Fig. S6). The ARHD to TNP ratio again performed better than *alkB* to some degree (Supplementary Fig. S5).

Despite *Oleispira* not being the single most abundant alkane degrader in the winter experiment, the *Oleispira*-specific qPCR assay was successful in distinguishing oil-exposed sample from control where the *alkB* assay failed to show oil presence. Moreover, the ratio of *Oleispira vs* TNP performed better in signaling oil presence earlier than the absolute copy numbers. Unlike for the functional assays, there was a good agreement between the relative abundances of *Oleispira* calculated from qPCR data (Oleispira/TNP ratio) and the *Oleispira* relative abundances obtained from the metagenome ($R^2 = 0.9578$ and 0.9862 in winter and summer, respectively). This observation suggests that relative abundances could serve as better signals of oil presence in the field rather than absolute abundances. It is therefore warranted to measure total prokaryotic abundance simultaneously with any targets, specific taxa or degenerate primer-based functional assay and aim at defining a threshold for relative abundance (Krolicka et al., 2019).

4.4. Challenges associated with designing degenerate primers

Several primer sets have been previously designed and used to explore abundance and diversity of *alkB* and ARHD genes in various environments and reported approximate coverages of the amplified products based on cloning and sequencing (Jurelevicius et al., 2013; Meynet et al., 2015; Paisse et al., 2011; Wang et al., 2010). Designing degenerate primers for the detection and quantification of homologous genes has a number of inherent challenges and has proven to be difficult for oil-degradation related enzymes so far (Linhart and Shamir, 2002). In addition, highly degenerate primers (>100) could also lead to significant quantification bias (Gaby and Buckley, 2017). Here, we employed the Consensus-Degenerate Hybrid Oligonucleotide Primer (CODEHOP) approach, which allows for designing primers based on a short (3–4 amino acid) degenerate 3′ core and a 5′ non-degenerate (consensus) clamp. This approach allows for designing primers with less degeneracy (thus reduced quantification bias), while in theory still capturing the intended diversity of sequences. As for all PCR primer design, CODEHOP uses multiple sequence alignment of known protein members of a target family to detect the core amino acid motifs. Then a set of primers containing all possible codons for these motifs are generated. Unlike other approach, each CODEHOP in the pool receives an identical 5′ consensus clamp derived from the most probable nucleotide at each position encoding the conserved amino acids flanking the targeted motif (Boyce et al., 2009; Rose et al., 2003). The main purpose of this clamp is to facilitate a robust amplification during the later cycles of the PCR reaction. Once a pool of templates has been “captured” by their matching degenerate core containing primer, the PCR products can be further amplified by all the different CODEHOPs thanks to all of them having the same complementary flanking nucleotide sequences. The resulting *alkB* primers designed in this study were similar to some of the previously published primer and probe sequences, *e.g.*, to the hybridization probe alk-f5 and the reverse primer alk-r5 published but targeted a shorter amino acid motif (Kloos et al., 2006; Miettinen et al., 2019; Pérez-de-Mora et al., 2011; Powell et al., 2006; Smits et al., 1999; Viggor et al., 2015; Wang et al., 2010). While most of the previously published *alkB* primer sets generate relatively long products in the range of 300 to 600 bp we aimed to amplify shorter sequences which could be better suited for qPCR analysis (Meynet et al., 2015). The designed ARHD primers were similar to those published by Chadhain et al. but again targeted shorter peptide sequences, *i.e.*, HRG instead of CR***HRG*** in case of the forward primer and YHGW instead of CS***YHGW*** in case of the reverse primer (Ní Chadhain et al., 2006). Despite limited degeneracy, amplification of a short product, targeting a shorter degenerate core and the use of a consensus clamp, the primers reported here did not recover (amplify) the taxonomic diversity of *alkB* and ARHD genes present in the metagenome. To our knowledge, this is the first study to examine diversity of generated PCR products using next-generation sequencing technology (amplicon sequencing) and compare the diversity of the amplicons with that observed from shotgun metagenomic analysis. Further work is required to explore the obtained sequences and re-iterate the primer design to improve on the apparent mismatch.

5. Conclusions

The goal of this project was to test if oil occurrence in water can be detected by an increase in abundance of microbial functional genes involved in oil biodegradation. This could be useful particularly for small leakages, which can remain undetected using a chemical approach. To our knowledge, this study is the first to thoroughly examine the coverage of *alkB* and ARHD gene-targeting assays for detecting oil in seawater through a combination of amplicon and shotgun sequencing of marine microbial communities. We identified a major mismatch between the diversity of *alkB* and ARHD genes amplified by the degenerate primers and those identified by shotgun metagenomics. More specifically, the designed primers did not amplify the *alkB* genes of the two most abundant alkane degraders that bloomed after addition of oil in the experiments, *Oceanobacter* and *Oleispira*. This was a major issue that has affected the performance and reliability of the tested qPCR assays. In the winter exposure, where *Oceanobacter* dominated the earlier oil exposed samples, the *Oleispira* assay did not accurately reflect the overall increase in alkane-degrading bacterial abundance. It is precisely in such situations that the degenerate functional approach would be most helpful. However, for that, further improvement of the degenerate primer approach is needed to ensure comprehensive coverage of the functional gene diversity present in various marine environments. To reliably use either one of the assay types (specific taxon or broad spectrum functional) for biomonitoring, one line of future research is to develop several genus specific primers to cover biogeographical gradients in the oceans. As this can be rather tedious, another alternative is to direct efforts towards establishing *in situ* automated metabarcoding systems and circumvent the challenges of targeted detection.

Supplementary data to this article can be found online at https://doi.org/10.1016/j.scitotenv.2021.152238.

Funding

This work was supported by the Research Council of Norway; Genomape project, grant number 255494.

CRediT authorship contribution statement

Andrea Bagi: Conceptualization, Investigation, Formal analysis, Visualization, Writing – original draft. **Kamila Knapik:** Conceptualization, Investigation, Formal analysis, Visualization, Writing – review & editing. **Thierry Baussant:** Conceptualization, Supervision, Funding acquisition, Project administration, Writing – review & editing.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgements

We would like to thank Kjell Birger Øysæd for his assistance and support in the chemical analysis of the seawater samples, Adriana Królicka for performing the USEARCH analysis and Emily Lyng for writing assistance and proofreading the article.

References

Aylagas, E., Atalah, J., Sánchez-Jerez, P., Pearman, J., Casado, N., Asensi, J., et al., 2021. A step towards the validation of bacteria biotic indices using DNA metabarcoding for benthic monitoring. Mol. Ecol. Resour. 21 (6), 1889–1903. https://doi.org/10.1111/1755-0998.13395.

Bagi, A., Knapik, K., Nilsen, M.M., Krolicka, A., Boccadoro, C., Baussant, T., 2019. Environmental surveillance of marine systems using genosensors: application to offshore activities. In: Skovhus, T.L., Whitby, C. (Eds.), Oilfield Microbiol.CRC Press, pp. 235–246

Boccadoro, C., Krolicka, A., Receveur, J., Aeppli, C., Le Floch, S., 2018. Microbial community response and migration of petroleum compounds during a sea-ice oil spill experiment in Svalbard. Mar. Environ. Res. 142, 214–233.

Borchert, E., García-Moyano, A., Sanchez-Carrillo, S., Dahlgren, T.G., Slaby, B.M., GEK, Bjerga, et al., 2021. Deciphering a marine bone-degrading microbiome reveals a complex community effort. mSystems 6, e01218–e01220.

Boyce, R., Chilana, P., Rose, T.M., 2009. iCODEHOP: a new interactive program for designing COnsensus-DEgenerate hybrid oligonucleotide primers from multiply aligned protein sequences. Nucleic Acids Res. 37, W222–W228.

Brakstad, O.G., Throne-Holst, M., Netzer, R., Stoeckel, D.M., Atlas, R.M., 2015. Microbial communities related to biodegradation of dispersed macondo oil at low seawater temperature with norwegian coastal seawater. Microbiol. Biotechnol. 8, 989–998.

Buchfink, B., Xie, C., Huson, D.H., 2015. Fast and sensitive protein alignment using DIAMOND. Nat. Methods 12, 59–60.

Campeao, M.E., Swings, J., Silva, B.S., Otsuki, K., Thompson, F.L., Thompson, C.C., 2019. "Candidatus colwellia aromaticivorans" sp. Nov., "Candidatus halocyntiibacter alkanivorans" sp. Nov., and "Candidatus ulvibacter alkanivorans" sp. NovGenome Sequences. Microbiol Resour Announc 8.

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

Caruso, G., La Ferla, R., Azzaro, M., Zoppini, A., Marino, G., Petochi, T., et al., 2016. Microbial assemblages for environmental quality assessment: knowledge, gaps and usefulness in the european marine strategy framework directive. Crit. Rev. Microbiol. 42, 883–904.

Catania, V., Santisi, S., Signa, G., Vizzini, S., Mazzola, A., Cappello, S., et al., 2015. Intrinsic bioremediation potential of a chronically polluted marine coastal area. Mar. Pollut. Bull. 99, 138–149.

Catania, V., Cappello, S., Di Giorgi, V., Santisi, S., Di Maria, R., Mazzola, A., et al., 2018. Microbial communities of polluted sub-surface marine sediments. Mar. Pollut. Bull. 131, 396–406.

Cébron, A., Norini, M.-P., Beguiristain, T., Leyval, C., 2008. Real-time PCR quantification of PAH-ring hydroxylating dioxygenase (PAH-RHDα) genes from gram positive and gram negative bacteria in soil and sediment samples. J. Microbiol. Methods 73, 148–159.

Chakraborty, R., Borglin, S.E., Dubinsky, E.A., Andersen, G.L., Hazen, T.C., 2012. Microbial response to the MC-252 oil and corexit 9500 in the Gulf of Mexico. Front. Microbiol. 3, 357.

Dubinsky, E.A., Conrad, M.E., Chakraborty, R., Bill, M., Borglin, S.E., Hollibaugh, J.T., et al., 2013. Succession of hydrocarbon-degrading bacteria in the aftermath of the Deepwater horizon oil spill in the gulf of Mexico. Environ. Sci. Technol. 47, 10860–10867.

Fish, J.A., Chai, B., Wang, Q., Sun, Y., Brown, C.T., Tiedje, J.M., et al., 2013. FunGene: the functional gene pipeline and repository. Front. Microbiol. 4, 291.

Fuentes, S., Méndez, V., Aguila, P., Seeger, M., 2014. Bioremediation of petroleum hydrocarbons: catabolic genes, microbial communities, and applications. Appl. Microbiol. Biotechnol. 98, 4781–4794.

Gaby, J.C., Buckley, D.H., 2017. The use of degenerate primers in qPCR analysis of functional genes can cause dramatic quantification bias as revealed by investigation of nifH primer performance. Microb. Ecol. 74, 701–708.

Gui Yuanyuan GXLJZP, 2020. Biodiversity analysis of cultured agar-degrading bacteria from surfaces of Antarctic macroalgae. Chin. J. Polar Res. 32, 504–511.

Gutierrez, T., Singleton, D.R., Berry, D., Yang, T., Aitken, M.D., Teske, A., 2013. Hydrocarbon-degrading bacteria enriched by the Deepwater horizon oil spill identified by cultivation and DNA-SIP. ISME J. 7, 2091–2104.

Hazen, T.C., Dubinsky, E.A., DeSantis, T.Z., Andersen, G.L., Piceno, Y.M., Singh, N., et al., 2010. Deep-sea oil plume enriches indigenous oil-degrading bacteria. Science 330, 204–208.

He, S., Peng, Y., Jin, Y., Wan, B., Liu, G., 2020. Review and analysis of key techniques in marine sediment sampling. Chin.<span/>J.<span/>Mech.<span/>Eng. 33, 66.

Head, I.M., Jones, D.M., Röling, W.F., 2006. Marine microorganisms make a meal of oil. Nat. Rev. Microbiol. 4, 173–182.

Heipieper, H.J., Martínez, P.M., 2010. Toxicity of hydrocarbons to microorganisms. In: Timmis, K.N. (Ed.), Handbook of Hydrocarbon and Lipid Microbiology. Springer, Berlin Heidelberg, Berlin, Heidelberg, pp. 1563–1573.

Heuer, H., Krsek, M., Baker, P., Smalla, K., Wellington, E.M., 1997. Analysis of actinomycete communities by specific amplification of genes encoding 16S rRNA and gel-electrophoretic separation in denaturing gradients. Appl. Environ. Microbiol. 63, 3233–3241.

Huson, D.H., Auch, A.F., Qi, J., Schuster, S.C., 2007. MEGAN analysis of metagenomic data. Genome Res. 17, 377–386.

Jones, D.O.B., Gates, A.R., Huvenne, V.A.I., Phillips, A.B., Bett, B.J., 2019. Autonomous marine environmental monitoring: application in decommissioned oil fields. Sci. Total Environ. 668, 835–853.

Jurelevicius, D., Alvarez, V.M., Peixoto, R., Rosado, A.S., Seldin, L., 2013. The use of a combination of alkB primers to better characterize the distribution of alkane-degrading bacteria. PLoS One 8, e66565.

Keith, L.H., 2015. The source of U.S. EPA's sixteen PAH priority pollutants. Polycycl. Aromat. Compd. 35, 147–160.

Kloos, K., Munch, J.C., Schloter, M., 2006. A new method for the detection of alkane-monooxygenase homologous genes (alkB) in soils based on PCR-hybridization. J. Microbiol. Methods 66, 486–496.

Krolicka, A., Boccadoro, C., Mæland, M., Preston, C.M., Birch, J., Scholin, C., et al., 2014. Detection of oil leaks by quantifying hydrocarbonoclastic bacteria in cold marine environments using the environmental sample processor. Proceedings of the 37th AMOP Technical Seminar on Environmental Contamination and Response, Canada, pp. 791–807.

Krolicka, A., Boccadoro, C., Nilsen, M.M., Baussant, T., 2017. Capturing early changes in the marine bacterial community as a result of crude oil pollution in a mesocosm experiment. Microbes Environ. 32, 358–366.

Krolicka, A., Boccadoro, C., Nilsen, M.M., Demir-Hilton, E., Birch, J., Preston, C., et al., 2019. Identification of microbial key-indicators of oil contamination at sea through tracking of oil biotransformation: an Arctic field and laboratory study. Sci. Total Environ. 696, 133715.

Kumar, S., Stecher, G., Li, M., Knyaz, C., Tamura, K., 2018. MEGA X: molecular evolutionary genetics analysis across computing platforms. Mol. Biol. Evol. 35, 1547–1549.

Letunic, I., Bork, P., 2007. Interactive tree of life (iTOL): an online tool for phylogenetic tree display and annotation. Bioinformatics 23, 127–128.

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

Linhart, C., Shamir, R., 2002. The degenerate primer design problem. Bioinformatics 18, S172–S181.

Lozada, M., Marcos, M.S., Commendatore, M.G., Gil, M.N., Dionisi, H.M., 2014. The bacterial community structure of hydrocarbon-polluted marine environments as the basis for the definition of an ecological index of hydrocarbon exposure. Microbes Environ. 29, 269–276.

Madeira, F., Park, Y.M., Lee, J., Buso, N., Gur, T., Madhusoodanan, N., et al., 2019. The EMBL-EBI search and sequence analysis tools APIs in 2019. Nucleic Acids Res. 47, W636–W641.

Mason, O.U., Hazen, T.C., Borglin, S., Chain, P.S., Dubinsky, E.A., Fortney, J.L., et al., 2012. Metagenome, metatranscriptome and single-cell sequencing reveal microbial response to Deepwater horizon oil spill. ISME J. 6, 1715–1727.

Mauffrey, F., Cordier, T., Apothéloz-Perret-Gentil, L., Cermakova, K., Merzi, T., Delefosse, M., et al., 2020. Benthic monitoring of oil and gas offshore platforms in the North Sea using environmental DNA metabarcoding. Mol. Ecol. n/a.

Meynet, P., Head, I.M., Werner, D., Davenport, R.J., 2015. Re-evaluation of dioxygenase gene phylogeny for the development and validation of a quantitative assay for environmental aromatic hydrocarbon degraders. FEMS Microbiol. Ecol. 91.

Miettinen, H., Bomberg, M., Nyyssonen, M., Reunamo, A., Jorgensen, K.S., Vikman, M., 2019. Oil degradation potential of microbial communities in water and sediment of Baltic Sea coastal area. PLoS One 14, e0218834.

Netzer, R., Henry, I.A., Ribicic, D., Wibberg, D., Bronner, U., Brakstad, O.G., 2018. Petroleum hydrocarbon and microbial community structure successions in marine oil-related aggregates associated with diatoms relevant for Arctic conditions. Mar. Pollut. Bull. 135, 759–768.

Ní Chadhain, S.M., Norman, R.S., Pesce, K.V., Kukor, J.J., Zylstra, G.J., 2006. Microbial dioxygenase gene population shifts during polycyclic aromatic hydrocarbon biodegradation. Appl. Environ. Microbiol. 72, 4078–4087.

Oksanen, J., Blanchet, F.G., Friendly, M., Kindt, R., Legendr, P., et al., McGlinn, D., 2020. vegan: Community Ecology Package. R package version 2.5-7.

Paisse, S., Duran, R., Coulon, F., Goni-Urriza, M., 2011. Are alkane hydroxylase genes (alkB) relevant to assess petroleum bioremediation processes in chronically polluted coastal sediments? Appl. Microbiol. Biotechnol. 92, 835–844.

Peña-Montenegro, T.D., Kleindienst, S., Allen, A.E., Eren, A.M., JP, McCrow, Sánchez-Calderón, J.D., et al., 2020. *Colwellia* and *Marinobacter* metapangenomes reveal species-specific responses to oil and dispersant exposure in deepsea microbial communities. bioRxiv https://doi.org/10.1101/2020.09.28.317438.

Pérez-de-Mora, A., Engel, M., Schloter, M., 2011. Abundance and diversity of n-alkane-degrading bacteria in a forest soil co-contaminated with hydrocarbons and metals: a molecular study on alkB homologous genes. Microb. Ecol. 62, 959–972.

Powell, S.M., Ferguson, S.H., Bowman, J.P., Snape, I., 2006. Using real-time PCR to assess changes in the hydrocarbon-degrading microbial community in Antarctic soil during bioremediation. Microb. Ecol. 52, 523–532.

Preston, C.M., Harris, A., Ryan, J.P., Roman, B., Marin, R., Jensen, S., et al., 2011. Underwater application of quantitative PCR on an ocean mooring. PLoS ONE 6, e22522.

Reunamo, A., Riemann, L., Leskinen, P., Jørgensen, K.S., 2013. Dominant petroleum hydrocarbon-degrading bacteria in the Archipelago Sea in south-West Finland (Baltic Sea) belong to different taxonomic groups than hydrocarbon degraders in the oceans. Mar. Pollut. Bull. 72, 174–180.

Ribicic, D., McFarlin, K.M., Netzer, R., Brakstad, O.G., Winkler, A., Throne-Holst, M., et al., 2018a. Oil type and temperature dependent biodegradation dynamics - combining chemical and microbial community data through multivariate analysis. BMC Microbiol. 18, 83.

Ribicic, D., Netzer, R., Hazen, T.C., Techtmann, S.M., Drablos, F., Brakstad, O.G., 2018b. Microbial community and metagenome dynamics during biodegradation of dispersed oil reveals potential key-players in cold norwegian seawater. Mar. Pollut. Bull. 129, 370–378.

Ribicic, D., Netzer, R., Winkler, A., Brakstad, O.G., 2018c. Microbial communities in seawater from an Arctic and a temperate norwegian fjord and their potentials for biodegradation of chemically dispersed oil at low seawater temperatures. Mar. Pollut. Bull. 129, 308–317.

Rose, T.M., Henikoff, J.G., Henikoff, S., 2003. CODEHOP (COnsensus-DEgenerate hybrid oligonucleotide Primer) PCR primer design. Nucleic Acids Res. 31, 3763–3766.

Satomi, M., Fujii, T., 2014. The family oceanospirillaceae. In: Rosenberg, E., DeLong, E.F., Lory, S., Stackebrandt, E., Thompson, F. (Eds.), The Prokaryotes: Gammaproteobacteria. Springer, Berlin Heidelberg, Berlin, Heidelberg, pp. 491–527.

Scoma, A., Hernandez-Sanabria, E., Lacoere, T., Junca, H., Boon, N., Pieper, D.H., et al., M., T., N., B., 2015. Primers: Bacterial Genes Encoding Enzymes for Aerobic Hydrocarbon Degradation. In: T., K. (Ed.), Hydrocarbon and Lipid Microbiology Protocols. Springer Protocols Handbooks. Springer, Berlin Heidelberg, pp. 23–37.

Shahsavari, E., Aburto-Medina, A., Taha, M., Ball, A.S., 2016. A quantitative PCR approach for quantification of functional genes involved in the degradation of polycyclic aromatic hydrocarbons in contaminated soils. MethodsX 3, 205–211.

Smits, T.H., Röthlisberger, M., Witholt, B., van Beilen, J.B., 1999. Molecular screening for alkane hydroxylase genes in gram-negative and gram-positive strains. Environ. Microbiol. 1, 307–317.

Sun, X., Kostka, J.E., 2019. Hydrocarbon-degrading microbial communities are site specific, and their activity is limited by synergies in temperature and nutrient availability in surface ocean waters. Applied and Environmental Microbiology 85 e00443-19.

Team, R.Core, 2021. R: A Language and Environment for Statistical Computing. R Foundation for Statistical Computing, Vienna, Austria.

Teramoto, M., Suzuki, M., Okazaki, F., Hatmanti, A., Harayama, S., 2009. Oceanobacter-related bacteria are important for the degradation of petroleum aliphatic hydrocarbons in the tropical marine environment. Microbiology (Reading) 155, 3362–3370.

Thomas, G.E., Brant, J.L., Campo, P., Clark, D.R., Coulon, F., Gregson, B.H., et al., 2021. Effects of dispersants and biosurfactants on crude-oil biodegradation and bacterial community succession. Microorganisms 9.

Tremblay, J., Fortin, N., Elias, M., Wasserscheid, J., King, T.L., Lee, K., et al., 2019. Metagenomic and metatranscriptomic responses of natural oil degrading bacteria in the presence of dispersants. Environ. Microbiol. 21, 2307–2319.

Ussler, W., Preston, C., Tavormina, P., Pargett, D., Jensen, S., Roman, B., et al., 2013. Autonomous application of quantitative PCR in the Deep Sea. In situ surveys of aerobic methanotrophs using the Deep-Sea environmental sample processor. Environ. Sci. Technol. 47, 9339–9346.

Viggor, S., Joesaar, M., Vedler, E., Kiiker, R., Parnpuu, L., Heinaru, A., 2015. Occurrence of diverse alkane hydroxylase alkB genes in indigenous oil-degrading bacteria of Baltic Sea surface water. Mar. Pollut. Bull. 101, 507–516.

Wang, W., Wang, L., Shao, Z., 2010. Diversity and abundance of oil-degrading bacteria and alkane hydroxylase (alkB) genes in the subtropical seawater of Xiamen Island. Microb. Ecol. 60, 429–439.