Metagenomics and quantitative stable isotope probing offer insights into metabolism of polycyclic aromatic hydrocarbons degraders in chronically polluted seawater

Ella T. Sieradzki^{a*}, Michael Morando^a, Jed A. Fuhrman^{a#}

^aDepartment of Biological Sciences, University of Southern California, California, USA

Running title: Metabolism of PAH degrading bacteria in polluted water

#Address correspondence to fuhrman@usc.edu

*Current affiliation: Environmental Science, Policy and Management department, University of California, Berkeley, California, USA

Abstract

Bacterial biodegradation is a significant contributor to remineralization of polycyclic aromatic hydrocarbons (PAHs): toxic and recalcitrant components of crude oil as well as byproducts of partial combustion chronically introduced into seawater via atmospheric deposition. The Deepwater Horizon oil spill demonstrated the speed at which a seed PAH-degrading community maintained by low chronic inputs can respond to acute pollution. We investigated the diversity and functional potential of a similar seed community in the Port of Los Angeles, a chronically polluted site, using stable isotope probing with naphthalene, deep-sequenced metagenomes and carbon incorporation rate measurements at the port and in two sites further into the San Pedro Channel. We demonstrate the ability of a local seed community of degraders at the Port of LA to incorporate carbon from naphthalene, leading to a quick shift in the microbial community composition to be dominated by these normally rare taxa. We were able to directly show that assembled genomes belonged to naphthalene degraders by matching their 16S-rRNA gene with experimental stable isotope probing data. Surprisingly, we did not find a full PAH degradation pathway in those genomes and even when combining genes from the entire microbial community. We analyze metabolic pathways identified in 29 genomes whose abundance increased in the presence of naphthalene to generate metagenomic-based recommendations for future optimization of PAHs bioremediation.

Importance

Oil spills in the marine environment have a devastating effect on marine life and biogeochemical cycles. Oil-degrading bacteria occur naturally in the ocean, especially where they are supported

by chronic inputs of oil, and have a significant role in degradation of oil spills. The most recalcitrant and toxic component of oil is polycyclic aromatic hydrocarbons. Therefore, the bacteria who can break those molecules down are of particular importance. We identified such bacteria at the port of Los Angeles, one of the busiest ports worldwide, and characterized their metabolic capabilities. Based on those analyses we proposed chemical targets to stimulate the activity of these bacteria in case of an oil spill in the port of LA.

Introduction

Polycyclic aromatic hydrocarbons (PAHs) are recalcitrant, mutagenic and carcinogenic components of crude oil as well as byproducts of incomplete combustion (1). Microbial biodegradation has an important role in PAH remediation alongside physical weathering processes (2). Biodegradation of PAHs captured much scientific attention after the Deepwater Horizon (DWH) oil spill in the Gulf of Mexico in 2010. Several studies measured PAH degradation rates (3, 4) and showed enrichment of known PAH-degrading bacteria in beaches, surface water, the deep-sea plume and sediments even months after the spill began (5–9). Bacteria known to have the ability to utilize PAHs as a carbon source include strains of *Cycloclasticus*, *Colwellia*, *Pseudomonas*, *Alteromonas*, and others (6, 10–12). Many coastal sites worldwide experience chronic input of PAHs, mainly from atmospheric deposition and natural oil seeps. Recent studies show that chronic pollution supports a consistent “seed” of PAH degrading bacteria which can respond quickly to acute pollution such as an oil spill (6, 10–15). Stable isotope probing (SIP) is a well-established method for the identification of environmental bacteria utilizing targeted substrates, in this case PAHs (16, 17), in which ¹³C PAHs are added to

samples in order to make the DNA of PAH-utilizers heavy and thus capable of physical separation in a density gradient. A large-scale SIP study was performed on DWH surface and deep-plume water, revealing local strains of PAH-degrading bacteria that responded to the input of hydrocarbons (6). Some genomes of those bacteria were assembled from mesocosm metagenomes in order to further explore their PAH metabolism (18). The studies mentioned above, like many others, focus only on the high-density (i.e. most heavily ^{13}C -labeled) fractions under the assumption that the most heavily labeled organisms, and thus the main targets, will be found there. However, this strategy may lead to overlooking degraders with low-GC (i.e. naturally lower DNA density) genomes, whose DNA may not appear in the heaviest fractions even if they include moderate amounts of ^{13}C (19). For example, 50% enrichment of a genome with 30% GC using ^{13}C would result in a mean weighted density of 1.7175 g*ml^{-1} (20), whereas the center of gradient is commonly set at $1.71\text{-}1.72 \text{ g*ml}^{-1}$. Tag-SIP is a powerful and particularly sensitive extension of the standard SIP approach, in which DNA from both labeled samples and parallel unlabeled controls are separated into density fractions, and the 16S-rRNA gene is amplified from all density fractions of both samples for comparison. This approach, circumventing GC-based bias, allows us to track substrate incorporation by a single taxon demonstrated by an increase (shift) in its DNA density in the labeled samples compared to controls (21, 22).

One of the main motivations to study PAH-degrading organisms is to characterize their metabolic requirements. Understanding the suite of nutrients and cofactors those organisms require could potentially be applied towards bioremediation and biostimulation (23). The combination of SIP with metagenomics can help reveal metabolic dependencies within assembled genomes of PAH-degraders (2, 15, 18, 23). Additionally, it has been proposed that

PAH biodegradation is a community process rather than fully performed by a single taxon (18). Thus, it is important to investigate potential degradation using Tag-SIP and deep sequencing so as to identify not only the function of the organisms that initiate the degradation but also of the other organisms involved in later steps of degradation.

While the Port of LA (POLA) is not routinely monitored for PAHs, it is located in an area with natural oil seeps, it houses a ship-refueling station of high-aromatic-content marine diesel (ETC Canada, 11/22/2017) and it is surrounded by the second largest metropolitan area in the USA, which is likely a source of consistent atmospheric PAH deposition. The LA-Long Beach port is the busiest port in the United States and the 10th busiest in the world according to the international association of ports and harbors (IAPH, 11/26/2017). A report published in 2010 revealed detectable levels of many PAHs at various stations within the port and as far as outside the port entrance, corresponding to our sampling site (24, 25). Due to the high marine traffic at POLA, there is constant resuspension into the water column of sediment, which would normally (without such regular resuspension) serve as a sink for PAHs due to their tendency to attach to particles (24, 25).

Here we provide evidence for the existence of a rare seed PAH-degrading, free-living (0.2-1 μ) microbial community in the chronically polluted Port of Los Angeles. We show that this seed can “spring into action” as a response to high PAH input, and start a degradation process leading to a specialized community dominated by two genera. We then discuss metabolic characteristics of the local PAH-degraders using largely complete genomes of PAH degraders and propose targets for biostimulation experiments in this system.

Materials and Methods

Sample collection

Surface seawater was collected in July and October 2014 from three sites across the San Pedro Channel near Los Angeles, CA, USA (fig. 1): the port of Two Harbors, Santa Catalina Island (CAT), the San Pedro Ocean Time-series (SPOT) and the Port of Los Angeles (POLA). An additional sample was collected in May 2015 only from POLA. Water was collected by HDPE bucket into two 10 liter LDPE cubitainers and stored in a cooler in the dark until arrival to the lab.

Isotope addition and incubation

400 nM of either unlabeled (^{12}C) naphthalene or fully labeled ^{13}C -10-naphthalene (ISOTEC, Miamisburg, OH, USA) were added to each 10 L cubitainer of water collected without replication. This concentration is roughly 3 orders of magnitude lower than the solubility of naphthalene in water (30 mg/L = 234 μM). 1 μM ammonium-chloride was also added to each bottle to prevent nitrogen limitation of PAH degraders. Naphthalene-enriched seawater from July and October 2014 was incubated in 10% ambient light, which is comparable to the light at 5m depth at SPOT, at surface water temperature (17°C) for 24 hours, and from May 2015 in the dark at surface water temperature (20°C) for 88 hours. Incubation of POLA water in the dark was performed to simulate more closely conditions in that site, as light attenuation at POLA is much steeper compared to the other sites (sup. fig. S1). At the end of the incubation period the seawater was filtered through an 80 nM mesh and a glass fiber Acrodisc (Millipore-Sigma, St.

Louis, MO, USA) prefilter (pore size 1 μm) followed by a 0.2 μm polyethersulfone (PES) Sterivex filter (Millipore-Sigma) to capture only the free-living microbes. After filtration, 1.5ml Sodium-Chloride-Tris-EDTA (STE; 10mM Tris-HCl, 1mM EDTA and 100mM NaCl) buffer was injected into the Sterivex casing and the filters were promptly sealed and stored in -80°C .

Carbon incorporation rate measurement

Seawater from all three sites was incubated in quadruplicate 2-liter polycarbonate bottles with 400nM ^{13}C labeled naphthalene. 1 μM ammonium-chloride was also added to each bottle to prevent nitrogen limitation of PAH degraders. A single bottle was filtered immediately after amendment addition to establish a t_0 atom% ^{13}C of the particulate carbon at each site. The remaining triplicates were incubated in a temperature-controlled room (see above). Incubations were carried out for ~ 24 h and were terminated by filtration onto pre-combusted (~ 5 h at 400°C) 47mm GF/F filters (Whatman, Maidstone, United Kingdom). The filters were then dried at 60°C and kept in the dark until analysis. Isotopic enrichment was measured on an IsoPrime continuous flow isotope ratio mass spectrometer (CF-IRMS). IRMS data were corrected for both size effect and drift before being calculated as previously described (26).

DNA extraction

Total DNA was extracted from the Sterivex filters using a modified DNeasy Plant kit protocol (Qiagen, Hilden, Germany). The Sterivex filters containing STE buffer were thawed. 100 μL 0.1mm glass beads were added into the filter casing and put through two 10-minutes cycles of

bead beating in a TissueLyser (Qiagen, Hilden, Germany). The flow-through, pushed out using a syringe, was incubated for 30 minutes at 37°C with 2 mg/ml lysozyme followed by another 30 minutes incubation at 55°C with 1 mg/ml proteinase K and 1% SDS. The resulting lysate was loaded onto the DNeasy columns followed by the protocol as described in the kit instructions. Only samples from POLA 5/2015 yielded enough DNA for ultracentrifugation of both the labeled and unlabeled DNA. However, metagenomes were sequenced from all dates and sites except SPOT 10/2014.

Ultracentrifugation and density-fractions retrieval

Isopycnic ultracentrifugation and gradient fractionation were performed as described in previous work (22, 27). Briefly, DNA from the labeled and unlabeled samples was added into separate quick seal 5ml tubes (Beckman Coulter, Indianapolis, IN, USA) combined with 1.88 g/ml CsCl and gradient buffer for a final buoyant density of 1.725 g/ml. The tubes were sealed and centrifuged in a Beckman Optima L100 XP ultracentrifuge and near-vertical rotor NVT 65.2 at 44100 rpm (190,950 rcf) at 20°C for 64 hours.

The gradient was divided into 50 fractions of 100µl each. Refraction was measured using 10 µl of each fraction using a Reichert AR200 digital refractometer, and converted into buoyant density ($\rho = 10.927 \cdot n_c - 13.593$) (28). DNA in each fraction was preserved with 200 µl PEG and 1 µl glycogen, precipitated with ethanol, eluted in 50 µl Tris-EDTA (TE) buffer and quantified using PicoGreen (Invitrogen, Carlsbad, CA, USA).

Amplification of the 16S-rRNA V4-V5 hypervariable regions

175
176 PCR was performed on each fraction with detectable DNA. Each reaction tube contained 12 µl
177 5Prime Hot Master Mix (Quantabio, Beverly, MA, USA), 1 µl (10pg) barcoded 515F-Y forward
178 primer (5'-GTGYCAGCMGCCGCGGTAA), 1 µl (10pg) indexed 926R reverse primer (5'-
179 CCGYCAATTYMTTTRAGTTT), 1 ng of DNA and 10µl molecular-grade water.

180 Thermocycling conditions:

- 181 1. 3 min denaturation at 95°C
- 182 2. 30 cycles of denaturation at 95°C for 45 sec, annealing at 50°C for 45 sec, and elongation
183 at 68°C for 90 sec
- 184 3. final elongation at 68°C for 5 min

185 PCR products from each fraction were cleaned using 1x Agencourt AMPure XP beads (Beckman
186 Coulter), quantified by PicoGreen and diluted to 1 ng/µl. A pool of 1 ng of each uniquely-
187 barcoded product was cleaned and concentrated again with 0.8x Agencourt AMPure XP beads.
188 The pooled amplicons were sequenced on Illumina MiSeq (UC Davis, USA) for 600 cycles.
189 Each pool was also spiked with 1 ng of an even and a staggered mock community in order to
190 assess the sequencing run quality (29). All expected OTUs were found in the observed mock
191 communities and accounted in total for 99.5% of the reads, indicating that there was no
192 unexpected bias in the sequencing run (30).

193
194 *Amplicon data processing*

195
196 The raw reads were quality-trimmed using Trimmomatic (31) version 0.33 with parameters set to
197 LEADING:20 TRAILING:20 SLIDINGWINDOW:15:25 MINLEN:200 and merged with

Usearch version 7 (32) with a limit of maximum 3 differences in the overlapping region. The resulting merged-reads were analyzed in Mothur (33) and clustered at 99% identity following the MiSeq SOP (March 20th, 2016) (34) with one exception: we found that degapping the aligned sequences, aligning them again and dereplicating again fixed an artifact in which a few abundant OTUs were split due to the alignment, despite 100% identity, due to an additional terminal base between the OTU representative sequences. OTUs with a total of less than 10 reads over all fractions were removed from the analysis. The remaining 2366 OTUs were assigned taxonomy using the arb-silva SINA search and classify tool version 1.3.2 (35).

Detection of OTU enrichment due to substrate incorporation

Plots of normalized abundance as a function of density were generated in R (<https://www.r-project.org/>) for the top 200 most abundant OTUs in each sample. The abundance of each OTU was normalized to a sum of 1 across all fractions (eq. 1).

Equation 1: Normalization of OTU abundance per fraction where i=fraction number, Ri=relative abundance of the OTU in fraction i, Di=DNA quantity in fraction i, Ni=normalized abundance of the OTU per fraction i

$$Ni = \frac{Ri \times Di}{\sum (Ri \times Di)}$$

To detect enrichment, the weighted mean density of an OTU in the labeled and unlabeled samples was calculated and if the difference exceeded 0.005 g/ml the shift was determined

significant (19). As quantitative SIP (qSIP) is sensitive to OTU abundance (19, 36), OTUs were used further only if they were non-spurious, whose distribution could not be differentiated from a normal distribution in both enriched and control samples (Kolmogorov-Smirnov test, $\alpha=0.05$). With these criteria 180 OTUs were further analyzed.

Metagenomic library preparation

The original unfractionated DNA extracted from the Sterivex filters was sheared by Covaris m2 to a mean length of 800 bp. Libraries were prepared from 15 ng of sheared DNA using the Ovation Ultra-low DR Multiplex System v2 kit (Nugen, Redwood City, CA, USA) with 9 amplification cycles. The libraries were bead purified as described above and sequenced on Illumina MiSeq for 600 cycles (UC Davis, USA) or on Illumina HiSeq rapid run for 500 cycles (USC genome core). See sup. table S1 for a detailed list of sequenced metagenomes.

Metagenomic sequencing analysis

Reads were quality-trimmed using Trimmomatic version 0.33 as described above. Paired reads were assembled per sample in an iterative subsampling and assembly process as described in Hug et al. (37) but using metaSPAdes version 3.9.1 instead of IDBA-UD, followed by overlap assembly with minimus 2 with minimum overlap 200 bp and minimum identity 99%. Paired reads from all sequenced samples (sup. table S1) were mapped back to the contigs with BBmap (sourceforge.net/projects/bbmap/) requiring 95% identity. Binning was performed using two approaches: (1) binning the POLA 5/15 13C metagenome with CONCOCT (38) and bin

refinement in Anvi'o (39). (2) pooling contigs longer than 5kbp from four naphthalene-enriched metagenomes (POLA 5/15 12C, POLA 5/15 13C, POLA 7/14 12C, POLA 7/14 13C), dereplicating them with cd-hit at 99% id (40, 41) and binning them using a combination of MaxBin2 (42), CONCOCT (38) and MetaBAT2 (43). These bins were combined, refined and reassembled using the MetaWRAP pipeline (44). Genomic bins generated by both methods were dereplicated using dRep (45) and only bins that were at least 50% complete and under 10% redundant were analyzed.

Initial taxonomic assignment of metagenomic assembled genomes (MAGs) was performed using GTDB-Tk (46). We then improved the taxonomy by generating class-level phylogenomic trees with GToTree (47) using NCBI RefSeq complete genomes and placing the bins assigned to the class by GTDB-Tk within them.

Reads from additional unenriched seawater metagenomes from all three sites were also mapped to the dereplicated MAG set to detect MAGs whose abundance increased in the presence of naphthalene.

Metabolic analysis

Open reading frames (ORFs) in the final set of metagenomic assembled genomes (MAGs) and viral contigs were predicted using Prodigal (48) and annotated by assignment to clusters of orthologous groups (COGs) using the Anvi'o anvi-run-ncbi-cogs function. KEGG (Kyoto encyclopedia of genes and genomes) orthology for ORFs was assigned with Kofamscan using the prokaryote profile and its built-in thresholds (49). Kofamscan results were summarized using

KEGGdecoder (50) (sup. fig. S2). ORF taxonomy was determined by Kaiju (51) using the RefSeq database. Secondary metabolite clusters were identified using the antiSMASH bacterial version 5 online platform (52) with the “relaxed cutoffs” option. Iron-related transport and storage systems were identified using FeGenie (53). Additional functional analysis was also done with METABOLIC version 1.3 (54) and DRAM (55). Read recruitment from different samples to the MAGs and viral contigs was analyzed with Anvi'o (39) using the Q2Q3 setting. This setting ignores the 25% lowest covered and 25% highest covered positions within the MAG when calculating mean coverage to avoid bias due to islands or highly conserved genes.

Data availability

Metagenomic and amplicon raw reads from enrichment mesocosms can be found on EMBL-ENA under project PRJEB26952, samples ERS2512855-ERS2512864. The metagenomic library blank is under sample ERS2507713. Amplicon reads can be found under sample accession numbers ERS2507470-ERS2507679, PCR blanks and mock communities under samples ERS2507702-ERS2507712. Metagenomic t0 raw reads can be found under project PRJEB12234, samples ERS2512914-ERS2512919.

Results

Naphthalene uptake rates across the San Pedro Channel

Our main hypothesis was that at the Port of Los Angeles (POLA) there would be a seed of obligatory PAH-degrading bacteria due to chronic inputs. One indication to the existence of such a community would be measurable uptake of naphthalene-derived carbon. We also wanted to test whether this degradation potential extended out into the San Pedro Channel at the San Pedro Ocean Time-series (SPOT) and Two Harbors (CAT) (fig. 1). Indeed, isotopic enrichment measurements at POLA indicated a mean naphthalene uptake rate of 35 nM/d (standard deviation 19.53 nM/d) given a high input of 400nM naphthalene. However, naphthalene incorporation rates at SPOT and CAT were below detection.

Figure 1: map of the sampling sites across the San Pedro Channel. The sites are within a range of 40 km.

SIP-identified PAH-degrading taxa at POLA

303 ¹³C-naphthalene-enrichment of POLA seawater led to significant incorporation of labeled carbon
304 ($\geq 0.005 \text{ g} \cdot \text{ml}^{-1}$ buoyant density increase corresponding to 9 atom %excess) by 34 out of 180
305 OTUs (sup. dataset S1). After 88 hours of incubation *Colwellia* sp. and *Cycloclasticus* sp.
306 (Gammaproteobacteria) comprised 40% of the free-living (0.2-1 μ) microbial community. These
307 OTUs were enriched at 53 (*Colwellia*) and 47 (*Cycloclasticus*) atom %excess (56) (fig. 2 A,B).
308 These main naphthalene degraders were rare (<0.2% cumulative relative abundance of all OTUs
309 classified as *Colwellia* or *Cycloclasticus*) to non-detectable prior to enrichment (t_0) at all sites in
310 all dates. While both taxa were represented by multiple OTUs, there was always a dominant
311 OTU which matched the 16S-rRNA genes from the MAGs (see below). The most abundant OTU
312 accounted for 95% and 99% of *Colwellia* and *Cycloclasticus* amplicons, respectively. Additional
313 significantly enriched OTUs included Gammaproteobacteria (*Marinomonas*, *Neptuniibacter*,
314 *Porticoccus*, *Pseudoalteromonas*, SAR86 and *Vibrio*), Flavobacteriales (*Tenacibaculum*,
315 *Fluviicola*, *Polaribacter*, NS7 marine group and *Owenweeksia*), Sphingobacteriales,
316 Deferribacterales (Marine group A), and Rhodospirillales (fig. 2 C-F).
317

Figure 2: Density shifts demonstrating which taxa took up ^{13}C naphthalene. Distribution of labeled (^{13}C , red) and control (^{12}C , blue) normalized relative abundance as a function of buoyant density of (A) *Colwellia*, (B) *Cycloclasticus*, (C) *Neptuniibacter*, (D) *Owenweeksia*, (E)

Tenacibaulum and (F) Marine group A. Vertical lines represent the weighted mean of the distribution. The difference between the weighted mean density (WM) of the labeled (^{13}C) and control (^{12}C) (delWM) is noted on each plot.

Metagenome-assembled genomes (MAGs) from naphthalene-enriched water

We assembled and binned 43 dereplicated MAGs that were more than 50% complete (mean=88%, SD=11%) and less than 10% redundant (mean=4.6%, SD=2.6%) from naphthalene-enriched POLA water. We then mapped reads from naphthalene-enriched and unenriched (t_0) metagenomes to those bins in order to pinpoint potential degraders, under the assumption that potential degraders would be more abundant in mesocosm metagenomes compared to t_0 metagenomes (fig. 3). Interestingly, seven bins were enriched in the presence of naphthalene at SPOT and CAT: POLA0515-13_bin_43, POLA0515-13_bin_23, Bin_4_4, Bin_47_1, Bin_46_1, Bin_20_1_1 and Bin_1_2_1 (sup. fig. S3). Eight bins significantly increased in abundance only in naphthalene-enriched POLA water (mean coverage < 1 in all CAT and SPOT samples) (table 1).

Table 1: Taxonomy and genomic parameters of MAGs whose abundance increased at POLA upon naphthalene addition

MAG ID	Taxonomy	Completeness	Redundancy	Size (Mbp)
POLA0515-13_bin_34	<i>Colwellia</i>	91%	3.6%	3.5
Bin_42_1	<i>Cycloclasticus</i>	99%	2.9%	2.4
POLA0515-13_bin_16	Flavobacteria	99%	3.6%	2.7
POLA0515-13_bin_12	Cellvibrionales	89%	0.7%	2.3
POLA0515-13_bin_27	<i>Pseudohongiela</i>	93%	10%	2.8
POLA0515-13_bin_4	<i>Pseudohongiela</i>	61%	4.3%	3.8
Bin_39_1	Rhodospirillaceae	91%	5%	3.2

Figure 3: MAGs that respond positively to addition of naphthalene. Mean coverage of MAGs in naphthalene-enriched and unenriched seawater was normalized to sequencing depth and log-transformed. Each column represents a sample with collection month and year indicated (e.g. 0113 is January 2013). Enriched mesocosms have _12 (^{12}C -naphthalene added) or _13 (^{13}C -naphthalene added) next to the sample name and are surrounded by rectangles. Incubation time is denoted under the rectangles. The rest of the samples represent t0 (before naphthalene addition). MAG taxonomy by GToTree is displayed on the right. Bins in bold font had significantly higher coverage in naphthalene-enriched samples (p-value < 0.05). *16S-rRNA based taxonomy where its resolution was higher than GToTree taxonomy

Primary naphthalene degraders

Two MAGs of interest were classified as *Colwellia* sp. and *Cycloclasticus* sp. and contained 16S-rRNA genes that matched at 100% identity to the most abundant OTUs of PAH-carbon incorporators (fig. 2 A,B).

The *Colwellia* MAG had high coverage and breadth (portion of the bin that has at least 1x coverage) only in naphthalene-enriched POLA water from May 2015 and October 2014 (fig. 3). Its closest relative based on a phylogenomic tree of 117 single-copy genes (47) is *Colwellia* sp. PAMC 20917 isolated from the Mid-Atlantic Ridge cold, oxic seafloor aquifer (57) (sup. fig. S4; sup. dataset S2). It has a 78% average nucleotide identity (two-way ANI) and 63% average amino acid identity (two-way AAI) with the *Colwellia* MAG assembled from the Deepwater Horizon oil spill (18). This MAG contained subunits A and B of naphthalene dioxygenase enzyme (*nahAa*, *nahAb*), which is the first step in naphthalene degradation, but only parts of the remainder of the pathway known from pure cultures (e.g. *nahD*, *nahE*, salicylate hydroxylase large and small subunits, protocatechuate dioxygenase alpha and beta chains) (sup. dataset S3; S4). In addition, the MAG contained near-complete chemotaxis and flagella-assembly, near complete vitamin B6 and biotin biosynthesis, and a complete riboflavin biosynthesis pathway (sup. fig. S2). This organism has transporters for nitrite/nitrate, urea, phosphate, molybdate and heme. Finally, it can transport nitrite and potentially reduce it to ammonium via a dissimilatory pathway (*nirBD*, sup. dataset S3).

The *Cycloclasticus* MAG (99% complete; 2.9% redundant, 2.4 Mbp) was detected in POLA naphthalene-enriched metagenomes from May 2015, July 2014 and October 2014 (fig. 3 B). This

MAG was most closely related to *Cycloclasticus zancles* 78-ME (sup. fig. S4; sup. dataset S2). It has a 78% average nucleotide identity (two-way ANI) and 81% average amino acid identity (two-way AAI) with the *Cycloclasticus* MAG assembled from the Deepwater Horizon oil spill (18). While it does not have the first steps of naphthalene degradation, it contains downstream genes of this pathway coding for 2-hydroxychromene-2-carboxylate isomerase (*nahD*), trans-o-hydroxybenzylidenepyruvate hydratase-aldolase (*nahE*) and catechol 2,3-dioxygenase (*xyIE*) (sup. dataset S5). This MAG contains a near-complete flagellar assembly pathway but not the chemotaxis pathway. It can incorporate nitrogen from cyanate and biosynthesize riboflavin and biotin. This organism can degrade the aromatic hydrocarbon cymene and specifically contains the ring-opening enzyme *cmtC*. Similar to the *Colwellia* MAG, this MAG also contains transporters for nitrite/nitrate, urea and phosphate, and can potentially reduce nitrite to ammonium via *nirBD* (sup. dataset S3). The *Colwellia* MAG contains secondary metabolite clusters of homoserine lactone (hserlactone) and lassoprotein which imply quorum sensing and a potential antimicrobial activity respectively (58, 59). The *Cycloclasticus* MAG also contains clusters of bacteriocin and non-ribosomal peptide synthase (NRPS) which may also point to a potential antibacterial activity.

Naphthalene-associated community metabolic characterization

In addition to *Colwellia* and *Cycloclasticus*, 27 MAGs were significantly more abundant in naphthalene-enriched samples (Wilcoxon rank test, p-value < 0.05). One of these MAGs, Bin_47_1 (*Porticoccus*) contained a 16S-rRNA gene that matched at 100% identity to an OTU that shifted significantly.

None of these MAGs contain the PAH or BTEX (benzene, toluene, ethylbenzene, xylene) degradation pathways defined by pure culture studies, but 12 of them had the KEGG pathway for degradation of the aromatic hydrocarbon cymene (sup. dataset S3). However, 18 MAGs included an aromatic ring-hydroxylating dioxygenase. Nine of these MAGs can assimilate nitrogen from nitroalkanes (fuel additives) or nitriles. These MAGs also contain secondary metabolite clusters of homoserine lactone (*Litoricola* Bin_27_2, Rhodobacteraceae POLA0714-12_bin_12), terpene (*Litoricola*, *Pseudohongiella*, Rhodobacteraceae, *Puniceispirillum*, SAR92, *Porticoccus*) and bacteriocin (Rhodobacteraceae). Finally, 24 out of these 29 MAGs contained at least one C1-oxidation enzyme, eight of the MAGs contain genes for oligosaccharide degradation such as alpha-L-rhamnosidase, beta-xylosidase and pullulanase and 23 MAGs contained tripartite ATP-independent periplasmic transporters (TRAP) (sup. dataset S3; sup. fig. S6).

Since several studies suggested that full degradation of PAHs may require bacterial consortia (18, 60), we also investigated the metabolic potential of the whole assembled community (open reading frames (ORFs) identified in contigs larger than 1kbp). The full naphthalene degradation pathway identified in pure cultures and incorporated into the KEGG database was not found even when combining all ORFs of the assembled MAGs.

Moreover, to confirm we did not miss additional naphthalene degradation genes because they did not assemble into contigs, we searched all forward reads from the POLA May 2015 ¹³C metagenome against all nucleotide sequences of naphthalene dehydrogenase ferredoxin subunit (*nahAc*, the first step of naphthalene degradation) from NCBI GenBank (188 references, September 2019). We found 248 reads (0.0008% of the total metagenome) that hit with a relaxed cutoff of e-value 10⁻³ (mean identity 97.8%, minimum bitscore 30). Using a conservative estimate that this gene comprises 0.1% of the metagenome (similar to 16S-rRNA), assuming one

copy per cell and that each read maps to one copy of the gene, this would imply that only 8 out of 10,000 cells carry the naphthalene dioxygenase *nahAc* gene. This fraction would decrease even further if we dropped the single-copy and one read per gene assumptions.

A cross-comparison of the 29 MAGs that were significantly more abundant in naphthalene-enriched water showed that while most MAGs had the ability to synthesize hydrophobic amino acids (leucine, isoleucine, valine and proline), some polar amino acids (cysteine, histidine, serine, threonine) as well as charged amino acid lysine and amphipathic amino acid tryptophan, only one had the biosynthetic pathway for methionine, and none had biosynthetic pathways for arginine, phenylalanine or tyrosine. However, 15 MAGs had a putative transport system for polar amino acids or branched amino acids, and nine had a transport system for L-amino acids. Only five MAGs contained a urea transport system, and while 34% naphthalene-enriched genomes had a urease gene, so did 43% of the genomes that did not respond to naphthalene-enrichment. In addition to *Colwellia* and *Cycloclasticus*, eight MAGs can degrade urea to ammonia via urease and four MAGs have genes coding for *nirBD* nitrite reductase. Five MAGs contain a phosphonate transport system whereas 20, including *Colwellia* and *Cycloclasticus*, can transfer phosphate.

Regarding vitamins, in addition to *Colwellia*, *Cycloclasticus*: four MAGs can synthesize biotin and 12 MAGs can synthesize pantothenate (vitamin B5). *Colwellia* and four additional MAGs can synthesize riboflavin (vitamin B2) (sup. fig. S2, sup. dataset S3).

Of these 29 genomes five are capable of anoxygenic photosynthesis (three Rhodobacteraceae and two *Halioglobus*), and 18 encode proteorhodopsins (sup. dataset S3).

Discussion

Naphthalene biodegradation potential does not extend outside the Port of Los Angeles

Naphthalene degradation rates were detectable only at POLA but not at SPOT or CAT. As there is tidal mixing across the San Pedro Channel, we would expect some potential for degradation at those sites by bacteria advected from the port. A possible explanation for the non-detectable rates at SPOT and CAT is that diminishing PAH inputs offshore (61) can support a very limited degrader seed community. The PAH degraders would then be out-competed for nutrients at those sites by microbes that are better adapted to oligotrophic conditions, and lose viability or be otherwise removed by processes like grazing and viral infection at rates faster than they are replenished by mixing or growth. Indeed, the abundance of MAGs of the primary degraders *Colwellia* and *Cycloclasticus* was extremely low at CAT and SPOT. More studies quantifying naphthalene degradation rates across San Pedro Channel are needed to determine whether this pattern consists throughout the year. The PAH carbon incorporation rate at POLA, however, indicated a minimum removal of ~10% of the initial concentration per day. This rate is roughly 3-fold higher than that measured in seawater from the Gulf of Mexico amended with crude oil (6, 62), although the incubation time was shorter compared to the Gulf of Mexico experiments, and the degradation rate may not be linear. The rate we measured is likely underestimated as measurement was performed using GF/F filters, which have a pore size leading to loss of roughly half of marine free-living prokaryotes. In addition, there could be degradation of PAH without incorporation of the labeled carbon into cells, which would not be accounted for by this measuring technique.

Naphthalene-associated community

Incubation of POLA water with naphthalene revealed a difference in PAH associated communities between the 24-hour and 88-hour incubations. First, it is noteworthy that many of the MAGs that were abundant in naphthalene-enriched cosms after 88 hours were already abundant after 24 hours. Some of them were common marine heterotrophs such as *Puniceispirillum* (SAR116) and Rhodobacteraceae. Both *Pseudohongiella* and OM182 MAGs were matched by 16S-rRNA gene to OTUs that did not incorporate carbon from naphthalene, indicating that their increased abundance in the presence of naphthalene is not due to them being primary degraders. We speculate that their increased abundance in enriched cosms is due to an early response to the presence of naphthalene, as their initial abundance was higher than that of the primary degraders *Colwellia* and *Cycloclasticus*. However, we also note a difference in response time between the primary degraders. In 3 out of 4 24 hours incubations, *Cycloclasticus* abundance was already comparable to its abundance after 88 hours, whereas *Colwellia* was only abundant in one 24 hours incubation, and its abundance was orders of magnitude lower compared to 88 hours. This may indicate a difference in growth rates and/or utilization of substrates and nutrients, as both taxa were undetectable in unenriched water. Gutierrez et al. (6) identified 4-5 orders of magnitude increase in abundance of *Colwellia* and *Cycloclasticus* in PAH enriched mesocosms after three days. It is notable that the same naphthalene degrading taxa appeared reproducibly in multiple incubations in different seasons. Thus, characterization of the metabolic requirements of these specific MAGs could be helpful in the future in case of an oil spill at the Port of LA.

Second, some MAGs were abundant in enriched water from SPOT and/or CAT even though there was no measurable degradation of naphthalene. They could be benefitting from toxicity of naphthalene to other bacteria or from byproducts of naphthalene oxidation. In the case of byproducts utilization we would have expected OTUs associated with these MAGs to shift to a heavier density. Only one of these MAGs representing *Porticoccus* contained a 16S-rRNA gene, and the matching OTU indeed shifted.

Incubation conditions were not identical between the 24-hour cosms from all three sites in July and October 2014 and the 88-hour cosm (POLA only, May 2015). Incubation in the dark (88 hours) would promote heterotrophy while incubation in the light should still allow cyanobacteria and photosynthetic picoeukaryotes to compete over nutrients with potential naphthalene-associated bacteria. However, if light had affected the results, we would have expected it to be reflected in MAG abundance. As it is not, we suggest that under naphthalene enrichment these photoautotrophs are not good competitors.

The duration of the 88-hour SIP experiment may have contributed to the accumulation of labeled carbon in bacterial DNA. In order to observe pronounced enrichment, the PAH degraders have to replicate at least once, and DNA density increases with every replication as the new strand is made with labeled nucleotides. Additionally, it is methodologically difficult to observe enrichment in rare taxa (36). Since the seed degrader community is made up of rare organisms, their abundance must increase substantially before their enrichment can be tracked. The other side of the coin is that a longer incubation increases the chance of cross-feeding, namely labeling of organisms that incorporate intermediate- or end-products of naphthalene degradation. Cross-feeding of intermediate products of naphthalene degradation is a vital part of the complete remineralization of this hydrocarbon which appears to have been often overlooked. In this study,

there were 24 OTUs which did not belong to the main naphthalene degraders (*Colwellia*, *Cycloclasticus* and *Neptuniibacter*) but were still significantly enriched (9-15%). While two of them (i.e. *Marinobacter* and *Pseudoalteromonas*) may be involved directly in degradation of naphthalene but either grow more slowly or were simply outcompeted by *Colwellia* and *Cycloclasticus* (63, 64), the rest are more generalist heterotrophs which are also associated with degradation of algal blooms in the San Pedro Channel (e.g. Flavobacteria, Marine group A, SAR86) (65).

Metabolism of putative PAH-degraders and associated taxa

In order to gain more insight into metabolic requirements of naphthalene degrading bacteria at POLA, we examined metabolic pathways and key transporters and enzymes within the annotated proteins in our metagenomic assembled genomes (MAGs). To date, only one published study characterized metabolic pathways in assembled genomes of marine oil degrading bacteria, using metagenomes from naphthalene- and phenanthrene-enriched seawater from the Deepwater Horizon (DWH) oil spill (6, 18). Within the DWH mesocosms, the prominent naphthalene degraders belonged to the genera *Cycloclasticus* and *Alteromonas*, and phenanthrene degraders included *Neptunomonas*, *Cycloclasticus* and *Colwellia*, whereas we found *Cycloclasticus* and *Colwellia* to be the primary naphthalene degraders at POLA.

Both *Colwellia* and *Cycloclasticus* were non-detectable in both 16S-rRNA amplicons and MAG coverage before enrichment. However, after 88 hours of incubation they were the most dominant taxa in the mesocosms, and exhibited very significant enrichment indicating incorporation of naphthalene-derived carbon into their DNA over a substantial number of replication cycles.

The *Colwellia* MAG was the only one assembled here that contained annotated subunits of naphthalene 1,2-dioxygenase (*nahAa*, *nahAb*). As this is the first step of the pathway, requiring investment of reducing power (66), it is not surprising that some of the downstream meta-cleavage of catechol is also present in the MAG. In comparison to the naphthalene degradation enzymes found in the *Colwellia* genome from the Gulf of Mexico, this MAG also had subunit A of naphthalene dioxygenase (*nahAa*) as well as *nahD*. It had an additional subunit (*nahAb*) of naphthalene dioxygenase but lacked the next enzyme encoded by *nahB*.

Cycloclasticus genes dominated the metagenome of the phenanthrene-enriched DWH mesocosm (18), highlighting its importance as an obligate PAH degrader (67). Unlike the *Colwellia* MAG, our *Cycloclasticus* MAG contained hardly any part of the KEGG-defined aerobic naphthalene or phenanthrene degradation pathways which are based on pure cultures. This is in stark contrast to the *Cycloclasticus* genome assembled from naphthalene-enriched water from the Gulf of Mexico (18). However, it did include several ring-hydroxylating dioxygenases. While these enzymes are best known to degrade single aromatic hydrocarbons, previous studies demonstrated that some single-ring aromatic degrading enzymes are capable of degrading naphthalene (two aromatic rings) efficiently (67–70). Additionally, the *Cycloclasticus* MAG has a sigma-54-dependent transcription regulator with a potential hydrocarbon-binding domain as identified in *Cycloclasticus zancles* 78-ME (GenBank accession AGS40441.1), which could control transcription of hydrocarbon degradation genes. Dioxygenases require iron and an iron-binding domain, such as ferredoxin which can be shared by multiple enzymes (71, 72), which we found in this MAG. As we know that our *Cycloclasticus* strain can take up naphthalene-derived carbon as its main carbon source based on the SIP results, and has genes that can participate in similar pathways but none of the culture-defined pathway, we posit that non-traditional dioxygenases

with affinity to naphthalene were utilized to degrade naphthalene. It is reasonable to expect that generalist aromatics-degrading organisms might “most-economically” possess multi-functional pathways that do not fully optimize degradation of a single substrate but work adequately with several aromatic substrates, without the need to carry the “extra baggage” needed to optimize each specific pathway. To further support this idea, metatranscriptomes of the microbial community within the oil plume of the Deepwater Horizon spill revealed low to nonexistent transcription of known PAH-degradation genes despite their presence in metagenomes (73). Moreover, the use of stable isotope probing indicated that both *Colwellia* and *Cycloclasticus* had the ability to degrade PAHs, whereas if only metagenomics was used, we might have concluded that they did not.

The presence of secondary metabolite clusters in both genomes suggests that part of the success of these organisms may be attributed to antimicrobial activity in addition to the ability to incorporate carbon from naphthalene.

Interestingly, most enriched MAGs also had Tripartite ATP-independent periplasmic (TRAP) transporters. TRAP transporters are common in marine bacteria (74) have been shown to participate in chlorobenzoate (75) and lignin-derived aromatic degradation (76), suggesting that they may participate in uptake of PAHs or their byproducts in this system. The presence of C1-oxidizing enzymes in most naphthalene-enriched MAGS also implies that even those who cannot directly degrade naphthalene may benefit greatly from degradation byproducts.

MAG-generated hypotheses for future biostimulation experiments at POLA

Previous studies revealed that in many systems bioremediation of crude oil can be enhanced by addition of nitrogen and phosphorous (15, 77). While oil degrading bacteria are found in many marine systems (78), their metabolic requirements may vary by system, depending on limitations *in-situ*. The significant genomic dissimilarity between the strains of *Colwellia* and *Cycloclasticus* found at POLA and the ones found at the Gulf of Mexico implies that we cannot assume similarity in metabolic requirements between systems. As oil degradation requires a considerable amount of nitrogen (79), choosing the correct substrate could be crucial.

Both *Colwellia* and *Cycloclasticus* displayed a potential for using a variety of nitrogen sources to different ends, with a full dissimilatory nitrite reduction to ammonium (DNRA) pathway and nitrite/nitrate transporters (*focA*, *narK*) as well as a urea transporter and the ammonium-assimilating glutamate synthase pathway. Nitrate, nitrite and ammonium are always detectable at POLA in surface seawater (sup. fig. S5), and thus are not limiting nutrients in this site. DNRA is an anaerobic process, which appears to be in conflict with the aerobic nature of surface seawater. However, PAHs are hydrophobic and tend to attach to particles, and particles can provide anaerobic microniches in their interior (80). It is possible that PAH degradation occurs in large part on particles at POLA, which would explain the presence of this pathway within the MAGs, and that in these taxa nitrite is sometimes used as an oxidizer whereas ammonia and/or urea are used as nitrogen sources. *Colwellia* and *Cycloclasticus* strains have also been isolated from sediments, where conditions may become anaerobic and support DNRA (8, 10).

Based on the functional pathways and ABC transporters found in MAGs of naphthalene incorporators, we propose specific targets for future experiments on enhancement of PAH bioremediation at the Port of Los Angeles. While it appears both local strains of *Colwellia* and *Cycloclasticus* can utilize urea, when considering the remaining MAGs enriched in the presence

of naphthalene, addition of ammonia rather than urea may be more beneficial in our system. Similarly, phosphate is more likely to augment bioremediation than phosphonate. To supply iron for dioxygenase synthesis, adding heme/hemoproteins should be superior to addition of Fe(II) or Fe(III) as 27/29 enriched MAGs contain heme transporters. Marine bacterioplankton have been shown to be able to incorporate iron from heme groups (81, 82). Finally, Polar amino acids for which we did not find any biosynthetic pathways but did find transporters in enriched MAGs could also potentially augment bioremediation.

In the current climate of excessive use of fossil fuels, chronic deposition of toxic and recalcitrant polycyclic aromatic hydrocarbons into the coastal ocean is inevitable. PAH-degrading bacteria may provide some control over the remineralization of these inputs and could serve as targets for bioremediation technologies. Identification of naturally-occurring biodegraders is a crucial first step, but optimization of the degradation process requires knowledge of the metabolic requirements of local organisms (23). Moreover, their genomic information remains an available resource should other hypotheses for biostimulation arise. The combination of stable isotope probing with metagenomics could provide a source for genomically-generated hypotheses and experimental design for future bioremediation experiments.

Acknowledgements

We extend our gratitude to Prof. Douglas Capone (USC) and Troy Eric Gunderson for the uptake rate measurements and for comments on this manuscript. We would like to thank Dr. Ben Tully for assistance with analysis of functional pathways. This research was made possible with the support of NSF grants 1136818 and 1737409, Gordon and Betty Moore Foundation Marine

Microbiology Initiative grant GBMF3779 and The Wrigley Institute for Environmental Studies
Norma and Jerol Sonosky fellowships to E.T.S.

Competing interests

The authors declare no competing interests.

References

1. Kanaly RA, Harayama S. 2000. Biodegradation of High-Molecular-Weight Polycyclic Aromatic Hydrocarbons by Bacteria. *Journal of Bacteriology*.
2. Head IM, Jones DM, Rling WFM. 2006. Marine microorganisms make a meal of oil. *Nat Rev Microbiol* 4:173–182.
3. Bacosa HP, Erdner DL, Liu Z. 2015. Differentiating the roles of photooxidation and biodegradation in the weathering of Light Louisiana Sweet crude oil in surface water from the Deepwater Horizon site. *Mar Pollut Bull* 95:265–272.
4. Liu Z, Liu J, Gardner WS, Christopher Shank G, Ostrom NE. 2016. The impact of Deepwater Horizon oil spill on petroleum hydrocarbons in surface waters of the northern Gulf of Mexico. *Deep Sea Research Part II: Topical Studies in Oceanography*.
5. Kostka JE, Prakash O, Overholt WA, Green SJ, Freyer G, Canion A, Delgardio J, Norton N, Hazen TC, Huettel M. 2011. Hydrocarbon-degrading bacteria and the bacterial community response in gulf of Mexico beach sands impacted by the deepwater horizon oil spill. *Appl Environ Microbiol* 77:7962–7974.

6. Gutierrez T, Singleton DR, Berry D, Yang T, Aitken MD, Teske A. 2013. Hydrocarbon-degrading bacteria enriched by the Deepwater Horizon oil spill identified by cultivation and DNA-SIP. *ISME J* 7:2091–2104.
7. Dubinsky EA, Conrad ME, Chakraborty R, Bill M, Borglin SE, Hollibaugh JT, Mason OU, M Piceno Y, Reid FC, Stringfellow WT, Tom LM, Hazen TC, Andersen GL. 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.
8. Mason OU, Scott NM, Gonzalez A, Robbins-Pianka A, Bælum J, Kimbrel J, Bouskill NJ, Prestat E, Borglin S, Joyner DC, Fortney JL, Jurelevicius D, Stringfellow WT, Alvarez-Cohen L, Hazen TC, Knight R, Gilbert JA, Jansson JK. 2014. Metagenomics reveals sediment microbial community response to Deepwater Horizon oil spill. *ISME J* 8:1464–1475.
9. Kleindienst S, Grim S, Sogin M, Bracco A, Crespo-Medina M, Joye SB. 2016. Diverse, rare microbial taxa responded to the Deepwater Horizon deep-sea hydrocarbon plume. *ISME J* 10:400–415.
10. Dyksterhouse SE, Gray JP, Herwig RP, Lara JC, Staley JT. 1995. *Cycloclasticus pugetii* gen. nov., sp. nov., an aromatic hydrocarbon-degrading bacterium from marine sediments. *Int J Syst Bacteriol* 45:116–123.
11. Niepceron M, Portet-Koltalo F, Merlin C, Motelay-Massei A, Barray S, Bodilis J. 2010. Both *Cycloclasticus* spp. and *Pseudomonas* spp. as PAH-degrading bacteria in the Seine estuary (France). *FEMS Microbiol Ecol* 71:137–147.
12. Jin HM, Kim JM, Lee HJ, Madsen EL, Jeon CO. 2012. *Alteromonas* as a key agent of polycyclic aromatic hydrocarbon biodegradation in crude oil-contaminated coastal sediment. *Environ Sci Technol* 46:7731–7740.

13. Hazen TC, Dubinsky EA, DeSantis TZ, Andersen GL, Piceno YM, Singh N, Jansson JK, Probst A, Borglin SE, Fortney JL, Stringfellow WT, Bill M, Conrad ME, Tom LM, Chavarria KL, Alusi TR, Lamendella R, Joyner DC, Spier C, Baelum J, Auer M, Zemla ML, Chakraborty R, Sonnenthal EL, D'haeseleer P, Holman H-YN, Osman S, Lu Z, Van Nostrand JD, Deng Y, Zhou J, Mason OU. 2010. Deep-sea oil plume enriches indigenous oil-degrading bacteria. *Science* 330:204–208.
14. Valentine DL, Mezić I, Maćešić S, Črnjarić-Žic N, Ivić S, Hogan PJ, Fonoberov VA, Loire S. 2012. Dynamic autoinoculation and the microbial ecology of a deep water hydrocarbon irruption. *Proc Natl Acad Sci U S A* 109:20286–20291.
15. Bargiela R, Herbst F-A, Martínez-Martínez M, Seifert J, Rojo D, Cappello S, Genovese M, Crisafi F, Denaro R, Chernikova TN, Barbas C, von Bergen M, Yakimov MM, Ferrer M, Golyshin PN. 2015. Metaproteomics and metabolomics analyses of chronically petroleum-polluted sites reveal the importance of general anaerobic processes uncoupled with degradation. *Proteomics* 15:3508–3520.
16. Radajewski S, Ineson P, Parekh NR, Murrell JC. 2000. Stable-isotope probing as a tool in microbial ecology. *Nature* 403:646–649.
17. Jeon CO, Park W, Padmanabhan P, DeRito C, Snape JR, Madsen EL. 2003. Discovery of a bacterium, with distinctive dioxygenase, that is responsible for in situ biodegradation in contaminated sediment. *Proc Natl Acad Sci U S A* 100:13591–13596.
18. Dombrowski N, Donaho JA, Gutierrez T, Seitz KW, Teske AP, Baker BJ. 2016. Reconstructing metabolic pathways of hydrocarbon-degrading bacteria from the Deepwater Horizon oil spill. *Nat Microbiol* 1:16057.
19. Sieradzki ET, Koch BJ, Greenlon A, Sachdeva R, Malmstrom RR, Mau RL, Blazewicz SJ, Firestone MK, Hofmockel KS, Schwartz E, Hungate BA, Pett-Ridge J. 2020. Measurement Error and Resolution in Quantitative Stable Isotope Probing: Implications for Experimental Design. *mSystems*

- 695 5.
- 696 20. Schildkraut CL, Marmur J, Doty P. 1962. Determination of the base composition of
697 deoxyribonucleic acid from its buoyant density in CsCl. *J Mol Biol* 4:430–443.
- 698 21. Connelly TL, Baer SE, Cooper JT, Bronk DA, Wawrik B. 2014. Urea Uptake and Carbon Fixation
699 by Marine Pelagic Bacteria and Archaea during the Arctic Summer and Winter Seasons. *Applied and*
700 *Environmental Microbiology*.
- 701 22. Morando M, Capone DG. 2016. Intracelld Heterogeneity in Nitrogen Utilization by Marine
702 Prokaryotes Revealed Using Stable Isotope Probing Coupled with Tag Sequencing (Tag-SIP). *Front*
703 *Microbiol* 7:1932.
- 704 23. Mapelli F, Scoma A, Michoud G, Aulenta F, Boon N, Borin S, Kalogerakis N, Daffonchio D. 2017.
705 Biotechnologies for Marine Oil Spill Cleanup: Indissoluble Ties with Microorganisms. *Trends*
706 *Biotechnol* 35:860–870.
- 707 24. Sabin LD, Maruya KA, Lao W, Diehl D, Tsukada D, Stolzenbach KD, Schiff KC. 2010. Exchange
708 of polycyclic aromatic hydrocarbons among the atmosphere, water, and sediment in coastal
709 embayments of southern California, USA. *Environ Toxicol Chem* 29:265–274.
- 710 25. Passow U, Ziervogel K, Asper V, Diercks A. 2012. Marine snow formation in the aftermath of the
711 Deepwater Horizon oil spill in the Gulf of Mexico. *Environmental Research Letters*.
- 712 26. Dugdale RC, Goering JJ. 1967. Uptake of new and regenerated forms of nitrogen in primary
713 productivity 1. *Limnol Oceanogr* 12:196–206.
- 714 27. Neufeld JD, Dumont MG, Vohra J, Murrell JC. 2007. Methodological considerations for the use of
715 stable isotope probing in microbial ecology. *Microb Ecol* 53:435–442.
- 716 28. Colin Murrell J, Whiteley AS. 2010. Stable Isotope Probing and Related Technologies. *American*

Society for Microbiology Press.

29. Parada AE, Needham DM, Fuhrman JA. 2016. Every base matters: assessing small subunit rRNA primers for marine microbiomes with mock communities, time series and global field samples. *Environ Microbiol* 18:1403–1414.
30. Yeh Y-C, Needham DM, Sieradzki ET, Fuhrman JA. 2018. Taxon Disappearance from Microbiome Analysis Reinforces the Value of Mock Communities as a Standard in Every Sequencing Run. *mSystems*.
31. Bolger AM, Lohse M, Usadel B. 2014. Trimmomatic: a flexible trimmer for Illumina sequence data. *Bioinformatics* 30:2114–2120.
32. Edgar RC. 2010. Search and clustering orders of magnitude faster than BLAST. *Bioinformatics* 26:2460–2461.
33. Schloss PD, Westcott SL, Ryabin T, Hall JR, Hartmann M, Hollister EB, Lesniewski RA, Oakley BB, Parks DH, Robinson CJ, Sahl JW, Stres B, Thallinger GG, Van Horn DJ, Weber CF. 2009. Introducing mothur: open-source, platform-independent, community-supported software for describing and comparing microbial communities. *Appl Environ Microbiol* 75:7537–7541.
34. Kozich JJ, Westcott SL, Baxter NT, Highlander SK, Schloss PD. 2013. Development of a dual-index sequencing strategy and curation pipeline for analyzing amplicon sequence data on the MiSeq Illumina sequencing platform. *Appl Environ Microbiol* 79:5112–5120.
35. Pruesse E, Peplies J, Glöckner FO. 2012. SINA: accurate high-throughput multiple sequence alignment of ribosomal RNA genes. *Bioinformatics* 28:1823–1829.
36. Youngblut ND, Barnett SE, Buckley DH. 2018. SIPSIm: A Modeling Toolkit to Predict Accuracy and Aid Design of DNA-SIP Experiments. *Front Microbiol* 9:570.

37. Hug LA, Thomas BC, Sharon I, Brown CT, Sharma R, Hettich RL, Wilkins MJ, Williams KH, Singh A, Banfield JF. 2016. Critical biogeochemical functions in the subsurface are associated with bacteria from new phyla and little studied lineages. *Environ Microbiol* 18:159–173.
38. Alneberg J, Bjarnason BS, de Bruijn I, Schirmer M, Quick J, Ijaz UZ, Lahti L, Loman NJ, Andersson AF, Quince C. 2014. Binning metagenomic contigs by coverage and composition. *Nat Methods* 11:1144–1146.
39. Eren AM, Murat Eren A, Esen ÖC, Quince C, Vineis JH, Morrison HG, Sogin ML, Delmont TO. 2015. Anvi'o: an advanced analysis and visualization platform for 'omics data. *PeerJ*.
40. Li W, Godzik A. 2006. Cd-hit: a fast program for clustering and comparing large sets of protein or nucleotide sequences. *Bioinformatics*.
41. Fu L, Niu B, Zhu Z, Wu S, Li W. 2012. CD-HIT: accelerated for clustering the next-generation sequencing data. *Bioinformatics* 28:3150–3152.
42. Wu Y-W, Simmons BA, Singer SW. 2016. MaxBin 2.0: an automated binning algorithm to recover genomes from multiple metagenomic datasets. *Bioinformatics* 32:605–607.
43. Kang DD, Li F, Kirton E, Thomas A, Egan R, An H, Wang Z. 2019. MetaBAT 2: an adaptive binning algorithm for robust and efficient genome reconstruction from metagenome assemblies. *PeerJ* 7:e7359.
44. Uritskiy GV, DiRuggiero J, Taylor J. 2018. MetaWRAP-a flexible pipeline for genome-resolved metagenomic data analysis. *Microbiome* 6:158.
45. Olm MR, Brown CT, Brooks B, Banfield JF. 2017. dRep: a tool for fast and accurate genomic comparisons that enables improved genome recovery from metagenomes through de-replication. *ISME J* 11:2864–2868.

46. Parks DH, Chuvochina M, Waite DW, Rinke C, Skarshewski A, Chaumeil P-A, Hugenholtz P. 2018. A standardized bacterial taxonomy based on genome phylogeny substantially revises the tree of life. *Nat Biotechnol* 36:996–1004.
47. Lee MD. 2019. GToTree: a user-friendly workflow for phylogenomics. *Bioinformatics* <https://doi.org/10.1093/bioinformatics/btz188>.
48. Hyatt D, Chen G-L, Locascio PF, Land ML, Larimer FW, Hauser LJ. 2010. Prodigal: prokaryotic gene recognition and translation initiation site identification. *BMC Bioinformatics* 11:119.
49. Aramaki T, Blanc-Mathieu R, Endo H, Ohkubo K, Kanehisa M, Goto S, Ogata H. KofamKOALA: KEGG ortholog assignment based on profile HMM and adaptive score threshold.
50. Graham ED, Heidelberg JF, Tully BJ. 2018. Potential for primary productivity in a globally-distributed bacterial phototroph. *ISME J* 12:1861–1866.
51. Menzel P, Ng KL, Krogh A. 2016. Fast and sensitive taxonomic classification for metagenomics with Kaiju. *Nat Commun* 7:11257.
52. Blin K, Shaw S, Steinke K, Villebro R, Ziemert N, Lee SY, Medema MH, Weber T. 2019. antiSMASH 5.0: updates to the secondary metabolite genome mining pipeline. *Nucleic Acids Res* 47:W81–W87.
53. Garber AI, Nealson KH, Okamoto A, McAllister SM, Chan CS, Barco RA, Merino N. 2020. FeGenie: A Comprehensive Tool for the Identification of Iron Genes and Iron Gene Neighborhoods in Genome and Metagenome Assemblies. *Front Microbiol* 11:37.
54. Zhou Z, Tran P, Liu Y, Kieft K, Anantharaman K. 2019. METABOLIC: A scalable high-throughput metabolic and biogeochemical functional trait profiler based on microbial genomes. *bioRxiv*.
55. Shaffer M, Borton MA, McGivern BB, Zayed AA, La Rosa SL, Solden LM, Liu P, Narrowe AB,

- Rodríguez-Ramos J, Bolduc B, Gazitúa MC, Daly RA, Smith GJ, Vik DR, Pope PB, Sullivan MB, Roux S, Wrighton KC. 2020. DRAM for distilling microbial metabolism to automate the curation of microbiome function. *Nucleic Acids Res* <https://doi.org/10.1093/nar/gkaa621>.
56. Hungate BA, Mau RL, Schwartz E, Caporaso JG, Dijkstra P, van Gestel N, Koch BJ, Liu CM, McHugh TA, Marks JC, Morrissey EM, Price LB. 2015. Quantitative microbial ecology through stable isotope probing. *Appl Environ Microbiol* 81:7570–7581.
57. Tully BJ, Wheat CG, Glazer BT, Huber JA. 2018. A dynamic microbial community with high functional redundancy inhabits the cold, oxic subseafloor aquifer. *ISME J* 12:1–16.
58. Zhao N, Pan Y, Cheng Z, Liu H. 2016. Lasso peptide, a highly stable structure and designable multifunctional backbone. *Amino Acids* 48:1347–1356.
59. Jayaraman A, Wood TK. 2008. Bacterial quorum sensing: signals, circuits, and implications for biofilms and disease. *Annu Rev Biomed Eng* 10:145–167.
60. Gallego S, Vila J, Tauler M, Nieto JM, Breugelmans P, Springael D, Grifoll M. 2014. Community structure and PAH ring-hydroxylating dioxygenase genes of a marine pyrene-degrading microbial consortium. *Biodegradation*.
61. Duran R, Cravo-Laureau C. 2016. Role of environmental factors and microorganisms in determining the fate of polycyclic aromatic hydrocarbons in the marine environment. *FEMS Microbiol Rev* 40:814–830.
62. Bacosa HP, Erdner DL, Rosenheim BE, Shetty P, Seitz KW, Baker BJ, Liu Z. 2018. Hydrocarbon degradation and response of seafloor sediment bacterial community in the northern Gulf of Mexico to light Louisiana sweet crude oil. *ISME J* 12:2532–2543.
63. Hedlund BP, Staley JT. 2006. Isolation and characterization of *Pseudoalteromonas* strains with

divergent polycyclic aromatic hydrocarbon catabolic properties. *Environ Microbiol* 8:178–182.

64. Gao W, Cui Z, Li Q, Xu G, Jia X, Zheng L. 2013. *Marinobacter nanhaiticus* sp. nov., polycyclic aromatic hydrocarbon-degrading bacterium isolated from the sediment of the South China Sea. *Antonie Van Leeuwenhoek* 103:485–491.

65. Needham DM, Fuhrman JA. 2016. Pronounced daily succession of phytoplankton, archaea and bacteria following a spring bloom. *Nature Microbiology*.

66. Baboshin MA, Golovleva LA. 2012. Aerobic bacterial degradation of polycyclic aromatic hydrocarbons (PAHs) and its kinetic aspects. *Microbiology*.

67. Yakimov MM, Timmis KN, Golyshin PN. 2007. Obligate oil-degrading marine bacteria. *Curr Opin Biotechnol* 18:257–266.

68. Kuhm AE, Stolz A, Knackmuss HJ. 1991. Metabolism of naphthalene by the biphenyl-degrading bacterium *Pseudomonas paucimobilis* Q1. *Biodegradation* 2:115–120.

69. Lessner DJ, Johnson GR, Parales RE, Spain JC, Gibson DT. 2002. Molecular characterization and substrate specificity of nitrobenzene dioxygenase from *Comamonas* sp. strain JS765. *Appl Environ Microbiol* 68:634–641.

70. Suen WC, Haigler BE, Spain JC. 1996. 2,4-Dinitrotoluene dioxygenase from *Burkholderia* sp. strain DNT: similarity to naphthalene dioxygenase. *J Bacteriol* 178:4926–4934.

71. Ensley BD, Gibson DT. 1983. Naphthalene dioxygenase: purification and properties of a terminal oxygenase component. *J Bacteriol* 155:505–511.

72. Nikolaivits E, Dimarogona M, Fokialakis N, Topakas E. 2017. Marine-Derived Biocatalysts: Importance, Accessing, and Application in Aromatic Pollutant Bioremediation. *Front Microbiol* 8:265.

73. Mason OU, Hazen TC, Borglin S, Chain PSG, Dubinsky EA, Fortney JL, Han J, Holman H-YN, Hultman J, Lamendella R, Mackelprang R, Malfatti S, Tom LM, Tringe SG, Woyke T, Zhou J, Rubin EM, Jansson JK. 2012. Metagenome, metatranscriptome and single-cell sequencing reveal microbial response to Deepwater Horizon oil spill. *ISME J* 6:1715–1727.
74. Rosa LT, Bianconi ME, Thomas GH, Kelly DJ. 2018. Tripartite ATP-Independent Periplasmic (TRAP) Transporters and Tripartite Tricarboxylate Transporters (TTT): From Uptake to Pathogenicity. *Front Cell Infect Microbiol* 8:33.
75. Chae J-C, Zylstra GJ. 2006. 4-Chlorobenzoate uptake in *Comamonas* sp. strain DJ-12 is mediated by a tripartite ATP-independent periplasmic transporter. *J Bacteriol* 188:8407–8412.
76. Salmon RC, Cliff MJ, Rafferty JB, Kelly DJ. 2013. The CouPSTU and TarPQM transporters in *Rhodopseudomonas palustris*: redundant, promiscuous uptake systems for lignin-derived aromatic substrates. *PLoS One* 8:e59844.
77. Ron EZ, Rosenberg E. 2014. Enhanced bioremediation of oil spills in the sea. *Curr Opin Biotechnol* 27:191–194.
78. Miller JI, Techtman S, Joyner D, Mahmoudi N, Fortney J, Fordyce JA, GaraJayeva N, Askerov FS, Cravid C, Kuijper M, Pelz O, Hazen TC. 2020. Microbial Communities across Global Marine Basins Show Important Compositional Similarities by Depth. *MBio* 11.
79. Atlas RM, Ronald M. Atlas, Bartha R. 1972. Degradation and mineralization of petroleum in sea water: Limitation by nitrogen and phosphorous. *Biotechnology and Bioengineering*.
80. Alldredge AL, Cohen Y. 1987. Can microscale chemical patches persist in the sea? Microelectrode study of marine snow, fecal pellets. *Science* 235:689–691.
81. Hogle SL, Barbeau KA, Gledhill M. 2014. Heme in the marine environment: from cells to the iron

cycle. Metallomics 6:1107–1120.

82. Hogle SL, Brahamsha B, Barbeau KA. 2017. Direct Heme Uptake by Phytoplankton-Associated Bacteria. mSystems 2.

Supplementary figures titles:

Figure S1: Relative photosynthetically active radiation (PAR) profiles at the port of Los Angeles (POLA), the San Pedro Ocean Time-series (SPOT) and Santa Catalina Island (CAT) in spring (April), winter (January) and fall (October)

Figure S2: Basic metabolic pathways in MAGs represented as fraction of the enzymes present. *Colwellia* and *Cycloclasticus* MAGs are labeled with a star.

Figure S3: Mean coverage of MAGs with increased abundance in naphthalene-enriched water from SPOT/CAT

Figure S4: Phylogenomic tree showing the placement of our *Colwellia* and *Cycloclasticus* MAGs. The MAGs are labeled in purple. The tree is based on 117 single-copy genes of Gammaproteobacteria and was built in GToTree. Complete reference genomes were downloaded from RefSeq on 8/20/2019.

Supplementary table titles:

Table S1: Sample details – number of metagenomics (MG) reads and 16S-rRNA amplicons. * Number of amplicons in enriched mesocosms are combined across all density fractions.

870

871 Supplementary dataset titles:

872 Dataset S1: p-values of Kolmogorov-Smirnov normality test of ¹²C- and ¹³C-naphthalene enriched OTU
873 distribution as a function of density, density shift, corresponding atom% excess and SILVA-assigned
874 taxonomy of 180 OTUs analyzed from POLA 5/15.

875 Dataset S2: General information on metagenomic assembled genomes (% completion, % redundancy,
876 length), taxonomy determined by GTDB-Tk, Anvi'o, GToTree and 16S-rRNA and mean coverage in
877 metagenomes (second and third quartile – Anvi'o setting Q2Q3)

878 Dataset S3: Results of METABOLIC analysis of the 29 MAGs whose abundance increased in the
879 presence of naphthalene. Each column represents a MAG, and each row represent a Kegg module.

880 Dataset S4: Results of DRAM hits to open reading frames of the *Colwellia* MAG and *Cycloclasticus*
881 MAG

882 Dataset S5: Results of DRAM hits to TRAP systems of the 29 MAGs whose abundance increased in the
883 presence of naphthalene.