Short-chain alkanes fuel mussel and sponge *Cycloclasticus* symbionts from deep-sea gas and oil seeps

Maxim Rubin-Blum, Chakkiath Paul Antony, Christian Borowski, Lizbeth Sayavedra, Thomas Pape, Heiko Sahling, Gerhard Bohrmann, Manuel Kleiner, Molly C. Redmond, David L. Valentine and Nicole Dubilier

*Cycloclasticus* bacteria are ubiquitous in oil-rich regions of the ocean and are known for their ability to degrade polycyclic aromatic hydrocarbons (PAHs). In this study, we describe *Cycloclasticus* that have established a symbiosis with *Bathymodiolus heckerae* mussels and poecilosclerid sponges from asphalt-rich, deep-sea oil seeps at Campeche Knolls in the southern Gulf of Mexico. Genomic and transcriptomic analyses revealed that, in contrast to all previously known *Cycloclasticus*, the symbiotic *Cycloclasticus* appears to lack the genes needed for PAH degradation. Instead, these symbionts use propane and other short-chain alkanes such as ethane and butane as carbon and energy sources, thus expanding the limited range of substrates known to power chemosynthetic symbioses. Analyses of short-chain alkanes in the environment of the Campeche Knolls symbioses revealed that these are present at high concentrations (in the μM to mM range). Comparative genomic analyses revealed high similarities between the genes used by the symbiotic *Cycloclasticus* to degrade short-chain alkanes and those of free-living *Cycloclasticus* that bloomed during the Deepwater Horizon oil spill. Our results indicate that the metabolic versatility of bacteria within the *Cycloclasticus* clade is higher than previously assumed, and highlight the expanded role of these keystone species in the degradation of marine hydrocarbons.

Fossil hydrocarbons, abundant in seafloor reservoirs as petroleum and natural gas deposits, fuel the global economy and play an important role in biogeochemical cycles. Natural seepage of fossil hydrocarbon has a pronounced effect on marine ecosystems and the atmosphere. The presence of oil and natural gas in the water column affects the microbial community, which responds rapidly to hydrocarbon infusions. These intrinsic microbes play an important role in the bioremediation of hydrocarbon pollution. For example, oil plume bacteria have the potential to mitigate persistent and toxic pollutants, such as polycyclic aromatic hydrocarbons (PAHs) and aromatic compounds, and stable carbon isotope signatures that have established a symbiosis with *Bathymodiolus heckerae* mussels and poecilosclerid sponges from asphalt-rich, deep-sea oil seeps at Campeche Knolls in the southern Gulf of Mexico. Genomic and transcriptomic analyses revealed that, in contrast to all previously known *Cycloclasticus*, the symbiotic *Cycloclasticus* appears to lack the genes needed for PAH degradation. Instead, these symbionts use propane and other short-chain alkanes such as ethane and butane as carbon and energy sources, thus expanding the limited range of substrates known to power chemosynthetic symbioses. Analyses of short-chain alkanes in the environment of the Campeche Knolls symbioses revealed that these are present at high concentrations (in the μM to mM range). Comparative genomic analyses revealed high similarities between the genes used by the symbiotic *Cycloclasticus* to degrade short-chain alkanes and those of free-living *Cycloclasticus* that bloomed during the Deepwater Horizon oil spill. Our results indicate that the metabolic versatility of bacteria within the *Cycloclasticus* clade is higher than previously assumed, and highlight the expanded role of these keystone species in the degradation of marine hydrocarbons.

© 2017 Macmillan Publishers Limited, part of Springer Nature. All rights reserved.
sponge species with a branching morphotype (one individual). Metabolic reconstruction of the *Cycloclasticus* genomes, complemented by transcriptomic and proteomic analyses, provided insights into the carbon and energy sources that fuel these symbionts. Analyses of gas and oil drops as well as asphalt and gas hydrates from the collection sites confirmed the presence of substrates predicted to play an important role in the metabolism of the *Cycloclasticus* symbionts. By identifying homologous key genes and transcripts in environmental data sets from the Deepwater Horizon (DWH) oil spill, we show that the carbon and energy pathways we discovered in the symbiotic *Cycloclasticus* may have also played a central role in the metabolism of free-living *Cycloclasticus* that bloomed en masse in the subsurface DWH hydrocarbon plume.

Results and discussion

*B. heckerae* mussels and two sponge species have host-specific *Cycloclasticus* 16S rRNA phylotypes. Eight *B. heckerae* individuals collected at Chapopote in 2015 were examined in this study. All had 16S rRNA sequences that were highly similar (>99.9%) to those of *Cycloclasticus* symbionts found in the two *B. heckerae* individuals sampled from Chapopote in 2006 (ref. 28) based on sequence analyses of PCR products (Supplementary Methods). Fluorescence in situ hybridization (FISH) with a probe targeting *Cycloclasticus* showed the intracellular location of the symbionts in *B. heckerae* gill bacteriocytes, where they co-occurred with previously described thiotrophic and methanotrophic symbionts (Fig. 2a,b). The relative abundance of *Cycloclasticus* in five *B. heckerae* individuals, based on estimates from Illumina metagenomic reads that mapped to the *Cycloclasticus* 16S rRNA gene, ranged from 5.1% to 10.7% (Supplementary Table 1). We also examined four *B. brooksi* individuals that co-occurred with *B. heckerae* in Chapopote using sequencing (15–49 million Illumina reads), FISH and PCR (Supplementary Methods). As in the two *B. brooksi* individuals collected at Chapopote in 2006, no *Cycloclasticus* were found in *B. brooksi* individuals collected in 2015.

*Cycloclasticus* were found in all three sponge individuals collected at Chapopote and Mictlan (Fig. 2c,d and Supplementary Table 1). Their abundances ranged between 4.1% and 8.0% based on the number of reads that mapped to the 16S rRNA gene of the sponge *Cycloclasticus* (Supplementary Table 1). The two encrusting sponges collected from Chapopote and Mictlan belonged to the same species based on 100% identical cytochrome c oxidase subunit I (COI) gene sequences, while the branching sponge individual belonged to a different species based on COI comparisons (10% nucleotide difference to COI of the encrusting sponge) and phylogenetic analyses (Supplementary Fig. 1). Both the encrusting and the branching sponge belong to the order Poecilosclerida based on their COI phylogeny (Supplementary Fig. 1).
Phylogenetic analyses of the *Cycloclasticus* 16S rRNA gene sequences from the Campeche Knoll *B. heckerae* mussels and sponges revealed that they belonged to a closely related clade (98% similarity) of cultivated and environmental *Cycloclasticus* often found in oil-contaminated habitats (Fig. 3). The mussel and sponge *Cycloclasticus* were phylogenetically distinct from cultivated *Cycloclasticus*, such as *C. pugetii*, which have all been shown to grow on PAHs (Fig. 3). In all three Campeche Knoll invertebrate species, the *Cycloclasticus* sequences were specific to the host species from which they originated, and phylogenetically distinct from each other. The closest relatives of the *Cycloclasticus* from the encrusting sponge were *Cycloclasticus* from the hydrocarbon seep sponge *Myxilla methanophila* (Gulf of Mexico), while the *Cycloclasticus* of *B. heckerae* was most closely related to uncultured *Cycloclasticus* in the oil and gas plume from the DWH well blowout (Fig. 3). The branching sponge *Cycloclasticus* were also related to bacteria from uncultured DWH *Cycloclasticus*, as well as bacteria from an ethane enrichment and from invertebrate hosts from chemosynthetic environments (Fig. 3).

Given that *B. heckerae* and the two sponge species each harboured a specific *Cycloclasticus* 16S rRNA phylotype that differed between these three host species, but was nearly identical between individuals of the same host species (including host individuals from the two collection sites Chapopote and Mictlan, separated by 25 km), we use the term ‘symbiotic’ for these host-specific bacteria.

**Symbiotic *Cycloclasticus* genotypes are specific to their host species.** We assembled draft *Cycloclasticus* genomes from the metagenomic sequencing of four *B. heckerae* individuals collected at Chapopote in 2015. The *B. heckerae* *Cycloclasticus* genomes had estimated sizes of 2.1–2.2 Mb, GC contents of 42%, and were 93–97% complete (Supplementary Table 3). The *Cycloclasticus* genomes from the four *B. heckerae* individuals were highly similar based on their average nucleotide identity (ANI, ≥ 99.95%, Supplementary Table 4) and phylogenomic analyses based on 11 single-copy genes (Supplementary Fig. 2).

Draft *Cycloclasticus* genomes from the two sponge species had estimated sizes of 1.6–2.3 Mb, GC contents of 43–44%, and were 90–95% complete (Supplementary Table 3). The *Cycloclasticus* genomes from the two encrusting sponge individuals were highly similar to each other (ANI = 99.8%), despite the fact that these hosts were collected at two different sites separated by 20 km (Chapopote and Mictlan, Fig. 1). Their genomes differed considerably from the *Cycloclasticus* genome of the branching sponge (ANI = 79.8%) and the *B. heckerae* mussels (ANI values <80%, Supplementary Table 4). Correspondingly, phylogenomic
analyses confirmed that the sponge *Cycloclasticus* symbionts differ from each other and those of *B. heckerae*. Taken together, ANI values, phylogenetic 16S rRNA and phylogenomic analyses all provide support for the conclusion that the three invertebrate species examined here, *B. heckerae* and two sponge species, harbour highly host-specific *Cycloclasticus* symbionts.
Genes for PAH degradation could not be detected in symbiotic *Cycloclasticus*. One of the key characteristics of *Cycloclasticus* is their ability to degrade PAHs, which they can use as their sole energy and carbon source. Correspondingly, the genomes of cultivated *Cycloclasticus* contain multiple clusters of genes involved in the degradation of PAHs (Supplementary Fig. 3 and Supplementary Table 5). Remarkably, none of the symbiotic *Cycloclasticus* genomes examined in this study, neither from the mussels nor the two sponge species, contained genes or transcripts attributable to PAH degradation (Supplementary Note 1 and 2 and Supplementary Fig. 3). These genomic and transcriptomic results were supported by incubation experiments on board with *Cycloclasticus*-bearing *B. heckerae* gill tissues; despite numerous attempts, we never observed oxidation of 14C-labelled naphthalene to 14CO2 by gill tissues, while control experiments with *Cycloclasticus* pugettii ATCC 51542 always showed naphthalene oxidation (Supplementary Note 1 and Supplementary Methods). These results contradict an earlier study that showed naphthalene oxidation (Supplementary Note 1 and 7) (transcriptomic sequencing of the branching sponge was not conducted). Genes encoding pHMOs (pmoCAB) in the *Cycloclasticus* symbionts of all three host species. These genes were genetically distant and phylogenetically distinct from the pMMO genes of the methane-oxidizing symbionts of these hosts and were most closely related to bacteria found in the DWH oil plume (Fig. 5). The *Cycloclasticus* symbiont pmoA sequences belonged to two groups: Group Z (ref. 31) and a clade of pMMOs related to those of short-chain alkane degrading *Nocardioides*32, and *Mycobacterium*.33 The latter clade forms a sister group to Group X pMMOs, which are suggested to have a high affinity to ethane34 or ethylene35 (Fig. 5). Hereafter, we refer to this clade as Group X-like pMMOs. Transcriptomic analyses of the symbiotic *Cycloclasticus* from four *B. heckerae* individuals and the two encrusting sponge individuals revealed that their pMMO genes were the most highly expressed genes compared to all other genes involved in carbon metabolism (Fig. 4 and Supplementary Tables 6 and 7) (transcriptomic sequencing of the branching sponge was not possible, as insufficient RNA was recovered). Proteomics revealed that subunit A of Group Z pMMO was among the most highly expressed *Cycloclasticus* proteins in three *B. heckerae* individuals (Supplementary Table 8). The second step in the oxidation of short-chain alkanes, after these have been oxidized to alcohol, is the formation of an aldehyde. In many bacteria that oxidize alcohols, this reaction is catalysed by pyrroloquinoline quinone-dependent alcohol dehydrogenases (PQQ-ADH). Indeed, we found genes encoding PQQ-ADHs in the symbiotic *Cycloclasticus* of all three host species, and these were also highly expressed in their transcriptomes (Fig. 4, Supplementary Tables 6 and 7). The membrane-bound PQQ-ADH was also found in the *Cycloclasticus* proteome of one *B. heckerae* individual (Supplementary Table 8). The symbiotic *Cycloclasticus* PQQ-ADH sequences were highly similar (91%) to PQQ-ADHs assumed to oxidize ethanol and higher alcohols, rather than methanol36 (Supplementary Fig. 4).

In the third step of short-chain alkane degradation, the aldehydes are oxidized to carboxylic acids. This step could be carried out by the PQQ-ADHs of the symbiotic *Cycloclasticus*, as these enzymes have been shown37 to oxidize aldehydes to carboxylic acids in butane-degrading *Pseudomonas* spp. This third step could also be achieved via tungsten-containing aldehyde ferredoxin oxioreductases (AORs), which are known to use short-chain alkane derived aldehydes as their substrate38. AORs were highly expressed in the transcriptomes of *Cycloclasticus* symbionts of all three host species, and also present in the proteomes of *Cycloclasticus* from two *B. heckerae* individuals (Supplementary Tables 8 and 9).

The genomes of the *Cycloclasticus* symbionts from all three host species encoded genes that enabled them to fix carbon derived from C2–C4 alkanes (Fig. 4), and many of these genes were highly expressed. These pathways are discussed in detail in Supplementary Notes 7 and 8. While carbon derived from C2–C4 alkanes appears to play a central role in the metabolism of the symbiotic *Cycloclasticus*, it is unlikely that they can use methane or other C1 compounds. Although pMMOs and PQQ-ADHs are also highly expressed in metabolically active methanotrophs39,40, their genes are phylogenetically distant from those of the symbiotic *Cycloclasticus*, suggesting functional divergence. Moreover, genes encoding key enzymes in C1 assimilation pathways such as the ribulose monophosphate (RuMP) pathway, the serine cycle and the Calvin–Benson–Bassham (CBB) cycle were not present in the *Cycloclasticus* symbiont genomes.

In summary, given that genes involved in PAH degradation were not detected in our metagenomic, metatranscriptomic and metaproteomic data, and the high expression levels of genes involved in the use of short-chain alkanes such as pMMOs, PQQ-ADHs and AORs, the symbiotic *Cycloclasticus* investigated in this study most probably use gaseous non-aromatic short-chain hydrocarbons as energy and carbon sources. This is surprising, as all cultivated *Cycloclasticus* are able to degrade PAHs, but do not appear to use short-chain alkanes. The genomes of cultivated *Cycloclasticus* lack genes coding for the first two enzymes needed for their oxidation, pMMOs and PQQ-ADHs. AORs, the third group of enzymes, are present in the genomes of cultivated *Cycloclasticus*, but their protein sequences are only 70% similar to those of the symbiotic *Cycloclasticus* (Supplementary Fig. 5).

**DWH *Cycloclasticus* may also be able to degrade short-chain alkanes.** We found genes encoding the first three enzymes central to the oxidation of short-chain alkanes—phMO, PQQ-ADH and AOR—in genomes and transcriptomes from the DWH hydrocarbon plume that were highly similar to those of the symbiotic *Cycloclasticus* from Campeche Knoll hosts (Fig. 5, Supplementary Figs 4 and 5). The Group X-like pMMOs, which were highly expressed in the mussel *Cycloclasticus*, were also present in a single amplified genome, SAG AC281-P21, and transcripts from the DWH oil plume36. These sequences formed a highly supported monophyletic clade with the *Cycloclasticus* symbionts (Fig. 5). Correspondingly, the 16S rRNA sequence from the DWH SAG AC281-P21 belonged to the same clade as those from the *B. heckerae* *Cycloclasticus* symbionts (Fig. 3). The Group Z pMMOs from the Campeche Knoll mussel and sponge *Cycloclasticus* belonged to a monophyletic clade that included two DWH SAGs,
Figure 4 | Reconstruction of central carbon and energy metabolic pathways in symbiotic Cycloclasticus. a. *B. heckerae* Cycloclasticus symbiont. b. Encrusting sponge Cycloclasticus symbiont. Pathways used for the assimilation of C1–C4 alkanes, as well as for respiration and sulfur oxidation are shown (Supplementary Notes 7 and 10). Coloured arrows indicate C1–C4 assimilation pathways, and black arrows indicate general metabolic pathways. Enzyme background colours represent mean expression values in the *B. heckerae* transcriptomes (BHT1-4) or sponge transcriptomes (ST1-2), normalized to DNA gyrase subunit B (gyrB) expression values in respective samples. These values are described in detail in Supplementary Tables 5 and 6. Enzyme abbreviations: particulate hydrocarbon monooxygenase (pHMO, subunits a–c); PQQ-dependent alcohol dehydrogenase (ADH PQQ); aldehyde oxidoreductase (AOR); cytochrome c oxidase (Cyt. cox, coxA-C subunits); cytochrome c oxidase cbb3 type (Cyt. cbb3, coaV-O subunits); cytochrome c reductase (CYTB, RIP1, CYC1 subunits); citrate synthase (CS); isocitrate dehydrogenase (IDH); 2-oxoglutarate dehydrogenase complex (DLST, dihydrolipoamide succinyldihydrolactonase component E2; DLD, dihydrolipoamide dehydrogenase; OGDH, 2-oxoglutarate dehydrogenase E1); succinyl-CoA ligase (SCS, α- and β-chains); succinyl dehydrogenase (SdhA, iron–sulfur protein; SdhB, flavoprotein subunit; SdhC, succinate dehydrogenase cytochrome b-556 subunit; SdhD, succinate dehydrogenase hydrophobic membrane anchor protein); fumarate hydratase (FH); malate dehydrogenase (MD); isocitrate lyase (ICL); malate synthase (MS); 2-methylcitrate synthase (2-MCS); 2-methylcitrate dehydratase (2-MCD); 2-methylfumarate isomerase (2-MAH); methylisocitrate lyase (MICL); phosphoenolpyruvate carboxykinase (PEPCK); phosphoenolpyruvate synthase (PEPS); serine hydroxymethyltransferase (SHMT); pyruvate dehydrogenase (PD, e1 and e2 subunits); polyhydroxybutyrate (PHB); tetrahydrofolate (THF).
but was more distantly related to the symbiotic *Cycloclasticus* the genus sponge (Fig. 3). (The second SAG, AC281-N15, also belonged to *Cycloclasticus* sequences of one of the two SAGs (AC281-I03) and those of the relationship was supported by similarly close 16S rRNA gene

As with the pHMOs, the PQQ-ADH sequences of the *Cycloclasticus* symbionts were most closely related to PQQ-ADH sequences from DWH bacteria and formed a highly supported clade with SAG AC281-P21, as well as transcripts from the hydrocarbon plume, with the latter constituting 38% of the total PQQ-ADH transcripts in the plume (Supplementary Fig. 4).

Symbiotic, free-living and cultivated *Cycloclasticus* contain AORs (Supplementary Fig. 5). AOR sequences vary considerably within

---

**Figure 5 | Phylogeny of pHMO subunit A protein sequences.** The tree is drawn to scale, with branch lengths representing the number of substitutions per site. Bootstrap values below 50% are not shown. The analysis was based on 55 protein sequences and included 219 amino acid positions. pMMO, particulate methane monoxygenase; pHMO, particulate hydrocarbon monoxygenase.

as well as transcripts that accounted for 50% of all *pmaA* transcripts in a DWH plume transcriptome (Fig. 5). This close phylogenetic relationship was supported by similarly close 16S rRNA gene sequences of one of the two SAGs (AC281-103) and of the *Cycloclasticus* symbionts from *B. heckerae* and the encrusting sponge (Fig. 3). (The second SAG, AC281-N15, also belonged to the genus *Cycloclasticus* based on its 16S rRNA gene sequence, but was more distantly related to the symbiotic *Cycloclasticus.*)
In previous studies at Chapopote, high concentrations of C2 Knolls, we asked if these alkanes are present in their environment. Cycloclasticus free-living Cycloclasticus of short-chain alkanes by symbiotic Short-chain alkanes are abundant in natural gas and oil seeps. Hence, unlike the cultivated and symbiotic AC281-N15 also contained PAH-degradation genes, such as dioxygenases, similar to those of cultivated Cycloclasticus (Fig. 3). Hence, unlike the cultivated and symbiotic Cycloclasticus, some free-living Cycloclasticus may have the ability to use both PAHs and short-chain alkanes.

Short-chain alkanes are abundant in natural gas and oil seeps. Given the strong genomic and transcriptomic evidence for the use of short-chain alkanes by symbiotic Cycloclasticus from Campeche Knolls, we asked if these alkanes are present in their environment. In previous studies at Chapopote, high concentrations of C2–C4 alkanes (in the low mM range) were measured at 2–10 cm depth in surface asphalts, and the fractions of these compounds ranged between 3% and 33% of total light hydrocarbons11. In this study, we measured the concentrations and relative proportions of short-chain alkanes in the environment of the symbiont-bearing invertebrates at Campeche Knolls by sampling and analysing gas and oil bubbles a few centimetres above the invertebrate collection sites at Chapopote and Mictlan, and from a piece of surface asphalt with gas hydrate collected at Chapopote. Concentrations of ethane, propane, and butane in the asphalt were in the µM to low mM range, and their relative fractions ranged between 7% and 8% for C2, 1% and 4% for C3, and 0.5% and 1% for C4 isomers (Table 1). Gas and oil bubbles from Chapopote and Mictlan contained considerable fractions of C2 (1.5% and 16%), C3 (0.4% and 14.5%) and C4 (0.06% and 3.0%) (Table 1). For comparison, in the DWH plume, which was dominated by propane and ethane-degrading bacteria in the late stages of the plume10, the dissolved gas in the plume contained similar proportions and concentrations of ethane and propane (for C2, 0.01–0.03 mM, 8–9%; for C3, ~0.01–0.02 mM, 4–6%; for C4, 0.01–0.03 mM, 2–3%)12. These results show that in both anthropogenically induced and natural seafloor hydrocarbon seepage in the Gulf of Mexico, short-chain alkanes are present in considerable concentrations that are probably sufficient to support the growth of alkane-degrading bacteria.

Conclusions

Our study provides genomic and transcriptomic evidence for the use of short-chain alkanes by symbiotic and free-living Cycloclasticus. We thus provide a direct link between the phylogenetic identity of these bacteria and their metabolic function. Our results confirm earlier studies that found indirect evidence for the use of short-chain alkanes by Cycloclasticus based on their dominance in DWH plumes that had high rates of propane and ethane consumption, and stable isotope probing (SIP) with propane and ethane at natural hydrocarbon seeps10,34. These results show that metabolic versatility within the Cycloclasticus is higher than previously assumed and not limited to PAH degradation.

All cultivated Cycloclasticus are very closely related to each other (Fig. 3, top clade), and within this clade we found no evidence for genes involved in the use of short-chain alkanes. These bacteria appear to rely solely on PAH degradation. In contrast, SAGs from the DWH contained genes for the use of both PAHs and short-chain alkanes. Such a versatile metabolism could be advantageous for Cycloclasticus in their ephemeral environment. During the early development of hydrocarbon plumes, short-chain alkanes are abundant and their use costs less energy than that of PAHs. However, as plumes mature, the concentrations of short-chain alkanes can decrease and the plume Cycloclasticus can then switch to using PAHs (refs 10, 43, 44).

It is intriguing that the symbiotic Cycloclasticus are the only bacteria within this genus that lack the ability to degrade PAHs, despite the high concentrations of PAHs in Campeche Knoll asphalts35. (Some SAGs from DWH do not have genes for PAH degradation (Fig. 3), but these genomes are only 25–50% complete.) Unlike DWH oil plume Cycloclasticus, symbionts experience a consistent supply of short-chain alkanes, determined by the host’s location above a hydrocarbon source. The host can enhance the metabolic capacity of symbionts by providing a continuous supply of dissolved gases via active pumping. Moreover, given their protected environment within host tissues, the Cycloclasticus symbionts would not experience competition with free-living short-chain alkane degraders. Furthermore, the low aqueous solubility of PAH compared to short-chain alkanes, as well as diffusion barriers across the host cell membrane could limit the access of these symbionts to PAHs. The selective advantage in maintaining the large suite of genes needed for PAH degradation appears to no longer exist for symbiotic Cycloclasticus. Similarly, we could envision a similar selection-driven loss in free-living Cycloclasticus from environments with a continuous supply of short-chain alkanes, such as surface asphalt and sediments at hydrocarbon-rich seeps. Future molecular and physiological studies on both symbiotic and free-living Cycloclasticus are needed to test these hypotheses and gain a better understanding of how this group of bacteria contributes to the consumption of hydrocarbons.

Methods

Sample collection. Mussels and sponges were collected with the remotely operated vehicle (ROV) MARUM-QUEST 4000m during the RV Meteor M114-2 cruise to the Campeche Knolls in March 2015 (Fig. 1a). B. heckeri and B. brooki mussels were collected during three dives from the Chapopote ‘bubble’ site (21°54’ N; 93°26’ W) at a water depth of 2.925 m. This site is characterized by the presence of fresh asphalts, exposed gas hydrates and flourishing faunal communities (Fig. 1b). An asphalt piece inhabited by both branching and encrusting sponges was also recovered from the ‘bubble’ site. Another encrusting sponge was collected at Mictlan invertebrate collection sites (Fig. 1) ND, not detected.

| Sample location on asphalt/hydrate piece | C1 | C2 | C3 | i-C4 | n-C4 | n-C5 | C4/(C2+C3) | C4/C2 | C4/C3 |
|------------------------------------------|----|----|----|------|------|------|-------------|-------|-------|
| Outside microbial mat                    | 21,473.2 | 1,026.5 | 8.8 | ND | 9.7 | 4.4 | 20.5 | 20.9 | 2,440.1 |
| Inside microbial mat                     | 5,244.3 | 396.9 | 67.9 | 23.2 | 73.9 | 28.6 | 8.9 | 13.2 | 77.2 |
| Below microbial mat                      | 8,598.7 | 839.4 | 376.7 | 57.1 | 411.6 | 128.7 | 4.7 | 10.2 | 22.8 |

*Concentrations of low-molecular-weight alkanes (in µmol l−1 bulk asphalt) and resulting C4/(C2+C3) ratios in gas samples prepared from an asphalt/gas hydrate piece from one of the Chapopote invertebrate collection sites (GeoB19329-5). Relative composition of light hydrocarbons (in mol %) of dissolved gases via active pumping. Moreover, given their protected environment within host tissues, the Cycloclasticus symbionts would not experience competition with free-living short-chain alkane degraders. Furthermore, the low aqueous solubility of PAH compared to short-chain alkanes, as well as diffusion barriers across the host cell membrane could limit the access of these symbionts to PAHs. The selective advantage in maintaining the large suite of genes needed for PAH degradation appears to no longer exist for symbiotic Cycloclasticus. Similarly, we could envision a similar selection-driven loss in free-living Cycloclasticus from environments with a continuous supply of short-chain alkanes, such as surface asphalt and sediments at hydrocarbon-rich seeps. Future molecular and physiological studies on both symbiotic and free-living Cycloclasticus are needed to test these hypotheses and gain a better understanding of how this group of bacteria contributes to the consumption of hydrocarbons.
Knolls from a site at 3,106 m water depth that was characterized by the presence of fresh asphalt and considerable gas and oil bubbling (22'-1° 4' N, 96° 31' W, Fig. 1c). A detailed description of the collection sites is available elsewhere22-23.

Mussels were identified on board based on their morphology24. Their correct morphological identification was later confirmed based on 100% similarity of their COI genes to those collected from Chapopotte in 2006 (ref. 28) and the published COI sequences for these species27 (see section ‘Genome annotation’). The symbiont-bearing gills of the mussels were dissected and fixed immediately after retrieval, as described at the end of this section. The sponges appeared intact immediately after collection, with no visible evidence of tissue damage. Encrusting sponge tissue was carefully removed from the underlying asphalt with a scalpel. The excised sponge tissue was rinsed immediately upon retrieval. The Chapopotte sponge was kept for several hours (at 4 °C) on the asphalt piece to which they were attached in a bucket containing water collected together with the sample, until processed.

Samples for transcriptomic and metagenomic analyses were fixed in RNAlater (Sigma) according to the manufacturer’s instructions and stored at −80 °C. Samples for microscopy were fixed in 2% paraformaldehyde in 1× PBS for (at most) 12 h at 4 °C, rinsed three times in 1× PBS and stored at 4 °C in 0.5× PBS/50% ethanol.

Gas sampling and analysis. Gas bubbles (sample GeoB19323-13 from Chapopotte, dive 354) and oil drops (GeoB19336-6 from Moclan, dive 357) were collected with the MARUM-QUEST ROV several centimetres above mussel beds with the pressure-tight Gas Bubble Sampler48. Additionally, three gas samples were prepared from the MARUM-QUEST ROV several centimetres above mussel beds with the pressure-tight Gas Bubble Sampler48. The samples were prepared at the edge of the mussels and one mussel assembly combined from four B. heckerae transcriptomic libraries. Blast49 was used to search for candidate genes, such as aromatic hydrocarbon mono- and dioxygenases.

Cycloclasticus cyclopdoma transcriptor and protein preparation. We prepared tryptic digests of gill samples from three B. heckerae individuals also used for transcriptomic analyses, following the filter-aided sample preparation (FASP) protocol50 with some small modifications51. Peptides were not desalted. Approximate peptide concentrations were determined using the Pierce Micro BCA assay (Thermo Scientific Pierce) following the manufacturer’s instructions.

Proteome analysis. Protein extraction and peptide preparation. For protein identification, a database was created using all protein sequences predicted from the B. heckerae Cycloclasticus symbiont genome published in this study (PRJNA318571), from preliminary SAG sequencing, assembly, contamination removal from transcriptome reads with BBMap using a minimum identity value of 0.98. Mapped reads were assigned to genomic features with featureCounts53. To compare the transcriptome libraries of each individual, a normalization factor was estimated with calcNormFactors based on the trimmed mean of M-values (TMM) implemented in the edgeR package48. The TMM normalized read counts were converted to reads per kilobases of exon per million reads mapped (RPKM) with the Rsubread R package (http://www.bioconductor.org). The expression values were normalized to expression of housekeeping genes with high general expression stability, such as those encoding the DNA gyrase subunit A and B, and RecA (ref. 65). Sponge and mussel metatranscriptomes were assembled with Trinity54 to verify the absence of transcripts involved in PAH degradation pathways (two individual sponge assemblies and one mussel assembly combined from four B. heckerae transcriptomic libraries). Blast55 was used to search for candidate genes, such as aromatic hydrocarbon mono- and dioxygenases.

Cycloclasticus SAG generation and analysis. Samples were collected on 15 June 2010 from the deep-sea hydrocarbon plume created by the DWH (ref. 10). The water column was characterized by high dissolved oxygen content and elevated hydrocarbon mono- and dioxygenase activities. The water column was dominated by the Cycloclasticus genus, which has been shown to degrade PAHs and other hydrocarbons. The Cycloclasticus SAG was generated using all protein sequences predicted from the Cycloclasticus genome published in this study (PRJNA318571), from preliminary SAG sequencing, assembly, contamination removal from transcriptome reads with BBMap using a minimum identity value of 0.98. Mapped reads were assigned to genomic features with featureCounts53. To compare the transcriptome libraries of each individual, a normalization factor was estimated with calcNormFactors based on the trimmed mean of M-values (TMM) implemented in the edgeR package48. The TMM normalized read counts were converted to reads per kilobases of exon per million reads mapped (RPKM) with the Rsubread R package (http://www.bioconductor.org). The expression values were normalized to expression of housekeeping genes with high general expression stability, such as those encoding the DNA gyrase subunit A and B, and RecA (ref. 65). Sponge and mussel metatranscriptomes were assembled with Trinity54 to verify the absence of transcripts involved in PAH degradation pathways (two individual sponge assemblies and one mussel assembly combined from four B. heckerae transcriptomic libraries). Blast55 was used to search for candidate genes, such as aromatic hydrocarbon mono- and dioxygenases.

Proteome analysis. Protein extraction and peptide preparation. We prepared tryptic digests of gill samples from three B. heckerae individuals also used for transcriptomic analyses, following the filter-aided sample preparation (FASP) protocol50 with some small modifications51. Peptides were not desalted. Approximate peptide concentrations were determined using the Pierce Micro BCA assay (Thermo Scientific Pierce) following the manufacturer’s instructions.

One-dimensional liquid chromatography—tandem mass spectrometry (1D-LC-MS/MS). Samples were analysed by 1D-LC-MS/MS. For each mussel sample, three technical replicates were run. Two wash runs and one blank run were carried out between samples to reduce carry over. For each run, 2-7 μg of peptide was loaded onto a 5 mm, 300 μm I/D C18 Acclaim PepMap100 pre-column (Thermo Fisher Scientific) using an UltiMateTM 3000 RSLCnano liquid chromatograph (Thermo Fisher Scientific) and desalted on the pre-column. After desalting the peptides, the pre-column was switched in line with a 50 cm × 75 μm analytical EASY-Spray column packed with PepMap RSLC C18, 2 μm particle size (Thermo Fisher Scientific), which was heated to 45 °C. The analytical column was connected via an Easy-Spray source to a Q Exactive Plus hybrid quadrupole-Orbitrap mass spectrometer (Thermo Fisher Scientific). Peptides were separated on the analytical column, and mass spectra were acquired in the Orbitrap, as described previously56. Roughly 650,000 MS/MS spectra were acquired per sample (three technical replicates combined).

Protein identification and quantification. For protein identification, a database was created using all protein sequences predicted from the B. heckerae Cycloclasticus symbiont genome published in this study (PRJNA318571), from preliminary SAG sequencing, assembly, contamination removal from transcriptome reads with BBMap using a minimum identity value of 0.98. Mapped reads were assigned to genomic features with featureCounts53. To compare the transcriptome libraries of each individual, a normalization factor was estimated with calcNormFactors based on the trimmed mean of M-values (TMM) implemented in the edgeR package48. The TMM normalized read counts were converted to reads per kilobases of exon per million reads mapped (RPKM) with the Rsubread R package (http://www.bioconductor.org). The expression values were normalized to expression of housekeeping genes with high general expression stability, such as those encoding the DNA gyrase subunit A and B, and RecA (ref. 65). Sponge and mussel metatranscriptomes were assembled with Trinity54 to verify the absence of transcripts involved in PAH degradation pathways (two individual sponge assemblies and one mussel assembly combined from four B. heckerae transcriptomic libraries). Blast55 was used to search for candidate genes, such as aromatic hydrocarbon mono- and dioxygenases.

Cycloclasticus SAG generation and analysis. Samples were collected on 15 June 2010 from the deep-sea hydrocarbon plume created by the DWH (ref. 10). The water column was characterized by high dissolved oxygen content and elevated hydrocarbon mono- and dioxygenase activities. The water column was dominated by the Cycloclasticus genus, which has been shown to degrade PAHs and other hydrocarbons. The Cycloclasticus SAG was generated using all protein sequences predicted from the Cycloclasticus genome published in this study (PRJNA318571), from preliminary SAG sequencing, assembly, contamination removal from transcriptome reads with BBMap using a minimum identity value of 0.98. Mapped reads were assigned to genomic features with featureCounts53. To compare the transcriptome libraries of each individual, a normalization factor was estimated with calcNormFactors based on the trimmed mean of M-values (TMM) implemented in the edgeR package48. The TMM normalized read counts were converted to reads per kilobases of exon per million reads mapped (RPKM) with the Rsubread R package (http://www.bioconductor.org). The expression values were normalized to expression of housekeeping genes with high general expression stability, such as those encoding the DNA gyrase subunit A and B, and RecA (ref. 65). Sponge and mussel metatranscriptomes were assembled with Trinity54 to verify the absence of transcripts involved in PAH degradation pathways (two individual sponge assemblies and one mussel assembly combined from four B. heckerae transcriptomic libraries). Blast55 was used to search for candidate genes, such as aromatic hydrocarbon mono- and dioxygenases.
screening and annotation were performed by the DOE Joint Genome Institute (JGI) following standard JGI protocols for microbial single-cell genomes.

**Phylogenetic analyses.** The evolutionary history of protein and 16S rRNA gene sequences was inferred by using the maximum likelihood method based on the Le Gascuel 2008 model and Kimura 2-parameter model for amino acid and nucleotide sequences, respectively. The trees with the highest log likelihood were chosen. The percentage of trees in which the associated taxa clustered together was determined based on 200 and 1,000 bootstrap resamples for amino acid and nucleotide sequences, respectively. Phylogenetic analyses were conducted in MEGAS (ref. 77).

**Data availability.** COI gene sequences were submitted to GenBank under accession nos KU659139 (B. heckeri) and KU659136–8 (sponges). Genomes with curated metadata are available through the Integrated Microbial Genomes (IMG) database under IMG taxon IDs 2599185276, 2599185294, 2599185283, 2602042074, 2599185280 and 2599185270. The mass-spectrometry proteomics data and the protein sequence database have been deposited in the ProteomeXchange Consortium via the PRIDE partner repository with data set identifier PXD005351.

**Received 2 September 2016; accepted 9 May 2017; published 19 June 2017

**References**

1. Bolin, B. in *Oceanography* (ed. Brewer, P. G.) 305–326 (Springer, 1983).

2. Dando, P. R. & Howland, M. Environmental effects of submarine seeping natural gas. *Cont. Shelf Res.* 12, 1197–1207 (1992).

3. Ciccone, R. J. & Oremland, R. Biogeological aspects of atmospheric methane. *Biogeochemistry* 2, 229–327 (1988).

4. Collins, W. J., Derwent, R. G., Johnson, C. E. & Stevenson, D. S. The oxidation of organic compounds in the troposphere and their global warming potentials. *Chim. Change 52, 453–479 (2002)."

5. Katzenstein, A. S., Doezema, L. A., Simpson, I. J., Blake, D. R. & Rowland, F. S. Transcriptome reads that mapped to the *Cycloclasticus* symbionts were deposited in the European Nucleotide Archive under the accession no. PRJEB12576. SAG assembly statistics and annotations are available through the Integrated Microbial Genome (IMG) database under IMG taxon IDs 2599185276, 2599185294, 2599185283, 2602042074, 2599185280 and 2599185270. The mass-spectrometry proteomics data and the protein sequence database have been deposited in the ProteomeXchange Consortium via the PRIDE partner repository with data set identifier PXD005351.

6. Received 2 September 2016; accepted 9 May 2017; published 19 June 2017

**References**

1. Bolin, B. in *Oceanography* (ed. Brewer, P. G.) 305–326 (Springer, 1983).

2. Dando, P. R. & Howland, M. Environmental effects of submarine seeping natural gas. *Cont. Shelf Res.* 12, 1197–1207 (1992).

3. Ciccone, R. J. & Oremland, R. Biogeological aspects of atmospheric methane. *Biogeochemistry* 2, 229–327 (1988).

4. Collins, W. J., Derwent, R. G., Johnson, C. E. & Stevenson, D. S. The oxidation of organic compounds in the troposphere and their global warming potentials. *Chim. Change 52, 453–479 (2002)."

5. Katzenstein, A. S., Doezema, L. A., Simpson, I. J., Blake, D. R. & Rowland, F. S. Transcriptome reads that mapped to the *Cycloclasticus* symbionts were deposited in the European Nucleotide Archive under the accession no. PRJEB12576. SAG assembly statistics and annotations are available through the Integrated Microbial Genome (IMG) database under IMG taxon IDs 2599185276, 2599185294, 2599185283, 2602042074, 2599185280 and 2599185270. The mass-spectrometry proteomics data and the protein sequence database have been deposited in the ProteomeXchange Consortium via the PRIDE partner repository with data set identifier PXD005351.

6. Received 2 September 2016; accepted 9 May 2017; published 19 June 2017

**References**

1. Bolin, B. in *Oceanography* (ed. Brewer, P. G.) 305–326 (Springer, 1983).

2. Dando, P. R. & Howland, M. Environmental effects of submarine seeping natural gas. *Cont. Shelf Res.* 12, 1197–1207 (1992).

3. Ciccone, R. J. & Oremland, R. Biogeological aspects of atmospheric methane. *Biogeochemistry* 2, 229–327 (1988).

4. Collins, W. J., Derwent, R. G., Johnson, C. E. & Stevenson, D. S. The oxidation of organic compounds in the troposphere and their global warming potentials. *Chim. Change 52, 453–479 (2002)."

5. Katzenstein, A. S., Doezema, L. A., Simpson, I. J., Blake, D. R. & Rowland, F. S. Transcriptome reads that mapped to the *Cycloclasticus* symbionts were deposited in the European Nucleotide Archive under the accession no. PRJEB12576. SAG assembly statistics and annotations are available through the Integrated Microbial Genome (IMG) database under IMG taxon IDs 2599185276, 2599185294, 2599185283, 2602042074, 2599185280 and 2599185270. The mass-spectrometry proteomics data and the protein sequence database have been deposited in the ProteomeXchange Consortium via the PRIDE partner repository with data set identifier PXD005351.
46. Gustason, R. G., Turner, R. D., Lutz, R. A. & Vrijenhoek, R. C. A new genus and five new species of mussels (Bivalvia, Mytilidae) from deep-sea sulfide/hydrocarbon seeps in the Gulf of Mexico. *Malacologia* **40**, 63–112 (1998).

47. Faure, B., Schaeffer, S. W. & Fisher, C. R. Species distribution and population connectivity of deep-sea mussels at hydrocarbon seeps in the Gulf of Mexico. *PLoS ONE* **10**, e0118460 (2015).

48. Pape, T. et al. Molecular and isotopic partitioning of low-molecular-weight hydrocarbons during migration and gas hydrate precipitation in deposits of a high-flux seepage site. *Chem. Geol.* **269**, 350–363 (2010).

49. Maryama, A. et al. Dynamics of microbial populations and strong selection for *Cycloclasticus pugetii* following the Nakhodka oil spill. *Microb. Ecol.* **46**, 442–453 (2003).

50. Zhou, J., Bruns, M. A. & Tiedje, J. M. DNA recovery from soils of diverse applications to single-cell sequencing. *Nat. Biotechnol.* **24**, 453–458 (2006).

51. Camacho, C. *EggNOG*: a new protein family classification for comparative genomics. *Nucleic Acids Res.* **38**, 415–421 (2010).

52. Vizcaíno, J. A.,ing et al. *ProteomeXchange* provides globally coordinated proteomics data submission and dissemination. *Nat. Biotechnol.* **32**, 223–228 (2014).

53. Albertsen, M. *et al.* DNA recovery from soils of diverse composition. *Appl. Environ. Microbiol.* **62**, 316–322 (1996).

54. Peng, Y., Leung, H. C. M., Yiu, S. M. & Chin, F. Y. L. *LD UDA*: a de novo assembler for single-cell and metagenomic sequencing data with highly uneven depth. *Bioinformatics* **28**, 1420–1428 (2012).

55. Seah, B. K. B. & Gruber-Vodicka, H. R. *Gbtools*: interactive visualization of metagenome bins in *R*. *Front. Microbiol.* **6**, 1451 (2015).

56. Albertsen, M. *et al.* Genome sequences of rare, uncultured bacteria obtained by differential coverage binning of multiple metagenomes. *Nat. Biotechnol.* **31**, 533–538 (2013).

57. Nurk, S., Bankevich, A. & Antipov, D. Assembling genomes and mini-metagenomes from highly chimeric reads. *Res. Comput. Mol. Biol.**10**, 158–170 (2013).

58. Bankevich, A. *et al.* SPAdes: a new genome assembly algorithm and its applications to single-cell sequencing. *J. Comput. Biol.* **19**, 455–477 (2012).

59. Gurevich, A., Saveliev, V., Vyahhi, N. & Tesler, G. *QUAST*: quality assessment tool for genome assemblies. *Bioinformatics* **29**, 1072–1075 (2013).

60. Parks, D. H., Imelfort, M., Skennerton, C. T., Hugenholtz, P. & Tyson, G. W. *CheckM*: assessing the quality of microbial genomes recovered from isolates, single cells, and metagenomes. *Genome Res.* **25**, 1043–1055 (2015).

61. Goris, J. *et al.* DNA–DNA hybridization values and their relationship to whole-genome sequence similarities. *Int. J. Syst. Evol. Microbiol.* **57**, 81–91 (2007).

62. Petkau, A., Stuart-Edwards, M., Stothard, P. & van Domselaar, G. Interactive microbial genome visualization with GView. *Stand. Genomic Sci.* **8**, 3123–3126 (2010).

63. Li, W. & Godzik, A. *Cd-hit*: a fast program for clustering and comparing large genome sequences of rare, uncultured bacteria obtained by *spades*: a new genome assembly algorithm and its applications to single-cell sequencing. *Bioinformatics* **30**, 932–940 (2014).

64. Robinson, M. D., McCarthy, D. J. & Smyth, G. K. *edgeR*: a Bioconductor package for differential expression analysis of digital gene expression data. *Bioinformatics* **26**, 139–140 (2010).

65. Rocha, D. J. P., Santos, C. S. & Pacheco, L. G. C. Bacterial reference genes for gene expression studies by RT–qPCR: survey and analysis. *Antonie Van Leeuwenhoek* **106**, 685–693 (2013).

66. Grubherr, M. G. *et al.* Full-length transcriptome assembly from RNA-Seq data without a reference genome. *Nat. Biotechnol.* **29**, 644–652 (2011).

67. Cassava, C. *et al.* BLAST+: architecture and applications. *BC Bioinformatics* **10**, 421 (2009).

68. Winiewski, J. R., Zougman, A., Nagaraj, N. & Mann, M. Universal sample preparation method for proteome analysis. *Nat. Methods* **6**, 359–362 (2009).

69. Hammar, E. *et al.* Environmental Breviata halophila multilocus Artocystobacter ephippion. *Nature* **534**, 254–258 (2016).

70. Petersen, J. M. *et al.* Chemosynthetic symbionts of marine invertebrate animals are capable of nitrogen fixation. *Nat. Microbiol.* **2**, 16195 (2016).

71. Bettencourt, R. *et al.* High-throughput sequencing and analysis of the gill tissue transcriptome from the deep-sea hydrothermal vent mussel *Bathybiomphalina azoricus*. *Microb. Genomics* **11**, 685 (2019).

72. Li, W. & Godzik, A. Cds-hit: A fast program for clustering and comparing large sets of protein or nucleotide sequences. *Bioinformatics* **22**, 1658–1659 (2006).

73. Flores, L. *et al.* Analyzing chromatin remodeling complexes using shotgun proteomics and normalized spectral abundance factors. *Methods* **46**, 303–311 (2006).

74. Swan, B. K. *et al.* Potential for chemolithoheterotrophy among ubiquitous bacteria lineages in the deep ocean. *Science* **333**, 1296–1300 (2011).

75. Le, S. Q. & Gascuel, O. An improved general amino acid replacement matrix. *Mol. Biol. Evol.* **25**, 1307–1320 (2008).

76. Kimura, M. A simple method for estimating evolutionary rate of base substitutions through comparative studies of nucleotide sequences. *J. Mol. Evol.* **16**, 111–120 (1980).

77. Kumar, S., Stecher, G. & Tamura, K. *MEGA7*: Molecular Evolutionary Genetics Analysis version 7.0 for bigger datasets. *Mol. Biol. Evol.* **33**, 1870–1874 (2016).

78. Markowitz, V. M. *et al.* IMG 4 version of the integrated microbial genomes comparative analysis system. *Nucleic Acids Res.* **42**, 560–567 (2014).

Acknowledgements

The authors thank all who helped during the RV *Meteor* research cruise M114, including onboard technical and scientific personnel, the captain and crew, and the ROV MARUM QUEST 4000m team. The authors thank A. Crombie, University of East Anglia, UK, for discussions on propane metabolism and A. Rivers, DOE Joint Genome Institute, for sharing DWH plume PQQX-ADH transcript sequences. The authors thank M. Strous for access to proteomics equipment and E. Thorson for technical assistance with the determination of peptide concentrations. The purchase of the proteomics equipment was supported by a grant of the Canadian Foundation for Innovation to M. Strous. The authors acknowledge the Max Planck-Centre-Genome-Centre Cologne (http://mpg.mpimpz.mpg.de/de/home/) for generating the metagenomic and metatranscriptomic data used in this study. *Cycloclasticus SAG* sequencing was supported by the Joint Genome Institute’s Community Science Program. The work conducted by the US Department of Energy Joint Genome Institute is supported by the Office of Science, Biological and Environmental Research Program of the US Department of Energy and by the University of California, Lawrence Berkeley National Laboratory under contract no. DE-AC02-05CH11231, Lawrence Livermore National Laboratory under contract no. DE-AC52-07NA27344 and Los Alamos National Laboratory under contract no. DE-AC02-08NA25396. The Campiche Knoll cruise was funded by the German Research Foundation (DFG, Deutsche Forschungsgemeinschaft). Additional support was provided through the MARUM DFG-Research Center/Excellence Cluster ‘The Ocean in the Earth System’ at the University of Bremen. The authors acknowledge the Mexican authorities for granting permission to conduct this research in the southern Gulf of Mexico (permission of DGOPA: 025/4014 from 5th November 2014). This study was funded by the Max Planck Society and the MARUM DFG-Research Center/Excellence Cluster ‘The Ocean in the Earth System’ at the University of Bremen. Further support was provided by an ERC Advanced Grant (BathyBio, 340355) and a Gordon and Betty Moore Foundation Marine Microbial Initiative Investigator Award to N.D. (grant no. GBMF3881). D.L.V. and M.R.-B. received funding from the US National Science Foundation grants OCE-1155855 and OCE-1046144. C.P.A. was supported by a postdoctoral fellowship from the Humboldt Foundation. M.K. was supported by a NSERC Banting Postdoctoral Fellowship.

Author contributions

M.R.-B., C.R., C.P.A. and N.D. conceived the study. M.R.-B. and N.D. analysed their genomes. M.R.-B. and N.D. wrote the manuscript with contributions from all co-authors.

Additional information

Supplementary information is available for this article.

Reprints and permissions information is available at www.nature.com/reprints.

Correspondence and requests for materials should be addressed to M.R.-B. and N.D.

How to cite this article: Rubin-Blum, M. *et al.* Short-chain alkanes fuel mussel and *Cycloclasticus* symbionts from deep-sea gas and oil seeps. *Nat. Microbiol.* **2**, 17093 (2017).

Publisher’s note: Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Competing interests

The authors declare no competing financial interests.