Supplementary Information for

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

Maxim Rubin-Blum[1*], Chakkiath Paul Antony[1], Christian Borowski[1], Lizbeth Sayavedra[1], Thomas Pape[2], Heiko Sahling[2], Gerhard Bohrmann[2], Manuel Kleiner[3], Molly C. Redmond[4], David L. Valentine[5], Nicole Dubilier[1,2*]

[1]Max-Planck Institute for Marine Microbiology, Celsiusstrasse 1, 28359 Bremen, Germany

[2]MARUM, Center for Marine Environmental Sciences, University of Bremen, 28359 Bremen, Germany

[3]Department of Geoscience, University of Calgary, Calgary, 2500 University Drive Northwest, Alberta T2N 1N4, Canada

[4]Department of Biological Sciences, University of North Carolina at Charlotte, 9201 University City Blvd., Charlotte, NC 28223, USA

[5]Department of Earth Science, University of California at Santa Barbara, 1006 Webb Hall, Santa Barbara, CA 93106, USA

***Corresponding authors**:

Maxim Rubin-Blum

Phone: 0049 (0)421 2028 905
Fax: 0049 (0)421 2028 790
E-mail: mrubin@mpi-bremen.de

Nicole Dubilier

Phone: 0049 (0)421 2028 932
Fax: 0049 (0)421 2028 790
E-mail: ndubilie@mpi-bremen.de

Contents:

Supplementary Note 1

No evidence for degradation of the PAH naphthalene in *Cycloclasticus*-bearing *B. heckerae* gill tissues

Before our 2015 cruise to the Campeche Knolls, we assumed that the symbiotic *Cycloclasticus* of *B. heckerae* can degrade PAHs based on indications for this ability in *Cycloclasticus*-bearing *B. heckerae* we collected during our 2006 cruise to Campeche Knolls[1] . We therefore examined if symbiotic *Cycloclasticus* can use PAH during our 2015 cruise by incubating symbiont-bearing gill tissues from f *B. heckerae* individuals with isotopically labeled naphthalene. We chose this small PAH molecule because it can be easily degraded by marine *Cycloclasticus* that are known to grow on PAH[2]. Naphthalene mineralization was not detected in *B. heckerae* gill tissues under the experimental conditions described below (see Methods). Positive controls with *Cycloclasticus pugetii* ATCC 51542 respired approximately 50% of labeled naphthalene to CO_2 within 48 hours (Supplementary Fig. 7). Experiments with isotopically labeled short-chain alkanes were not performed because the genomic evidence for this metabolism was not discovered until months after the cruise when we analyzed the genomes and transcriptomes of the symbiotic *Cycloclasticus* and no longer had live mussels.

Supplementary Note 2

Genes involved in PAH degradation pathways could not be detected in metagenomes of *B. heckerae* and sponges.

To examine if the genes known to be used for PAH degradation were present in the symbiotic *Cycloclasticus*, we searched our unassembled metagenomic and metatranscriptomic libraries for genes involved in PAH degradation using BLAST and BBMap. To ensure that we could find genes in the symbiotic *Cycloclasticus* that were homologous to known PAH degradation genes, but might have only low sequence similarity, we used a very low stringency value of only 60% nucleotide identity in our searches. Given i) the high read depths, which ranged between 10 and 49 million in our metagenomic libraries, ii) the fact that we sequenced multiple host individuals, and iii) our low-stringency search parameters, we assume that we would have detected at least some of the genes involved in PAH degradation, if these were present in the symbiotic *Cycloclasticus*. In free-living *Cycloclasticus* that degrade PAH and whose genomes have been sequenced, such as *C. pugetii*, PAH degradation genes are present in multiple copies and are easily detected at even low sequencing depths.

In Supplementary Table 5, we show the number of normalized reads from metagenomic mussel and sponge libraries that had at least 60% nucleotide identity to key metabolic genes in *C. pugetii*. The table first lists three examples of genes common to *C. pugetti* and the symbiotic *Cycloclasticus*, and shows that these were found in high read numbers. The table then lists the 47 genes known to be involved in PAH metabolism in *C. pugetti*, and shows that for the vast majority of these genes, no reads were found in mussel and sponge metagenomes. The very few reads that did map to PAH genes most likely originated from environmental contaminants.

Supplementary Note 3

Genomic structure of the *Cycloclasticus* symbiont pHMO encoding genes

B. heckerae Cycloclasticus genomes contained two *pmoCAB* operons. In contrast to their contiguous arrangement in the *B. heckerae Cycloclasticus*, in the sponge *Cycloclasticus* symbionts the *pmoCAB* operon was split, i.e., the *pmoAB* and *pmoC* genes were located on two different contigs. All three genes from the split *pmoCAB* operon were highly expressed in the sponge *Cycloclasticus* symbionts, indicating that the operon is functional (Fig. 6). Indeed, an ancillary copy of *pmoC* plays an essential role in the methanotrophic growth of *Methylococcus capsulatus*[3].

Supplementary Note 4

No evidence for methane use by the *Cycloclasticus* symbionts

Genes encoding key enzymes in methane assimilation pathways such as the ribulose monophosphate pathway and the serine cycle were not found in the *Cycloclasticus* symbiont genomes. It is theoretically possible that the moderately expressed genes encoding serine hydroxymethyltransferase (SHMT) could direct methane-derived carbon into the amino-acids serine and glycine. However, SHMT is commonly used by bacteria for the biosynthesis of amino acids during growth on multicarbon compounds[4]. For example, SHMT expression levels remained unchanged irrespective of whether the methanotroph *Methylocystis* sp. SB2 was grown on methane or ethanol[5].

Supplementary Note 5

Incorporation of alkane-derived carbon by the *Cycloclasticus* symbionts

The genomes of the *Cycloclasticus* symbionts from all three host species encoded genes that enable them to fix carbon derived from short chain alkanes. Incorporation of alkane-derived carbon requires the synthesis of acyl-coenzyme A (CoA)[4]. Coding regions annotated as acetyl-CoA (C2), propionyl-CoA (C3) and butyryl-CoA (C4) synthetases were present in the genomes of all symbiotic *Cycloclasticus*. In the following, we discuss the pathways predicted to be involved in incorporating C2-C4 acyl-CoAs into biomass.

Supplementary Note 6

Genes for C3 assimilation via the methylcitrate cycle were highly expressed in symbiotic *Cycloclasticus*

Propionyl-CoA synthetase is essential for the incorporation of carbon from propane and initiates the methylcitrate cycle, in which propionate is oxidized to pyruvate, which is then used for biosynthesis and energy production[6,7]. The genes encoding propionyl-CoA synthetase (*prpE*) were the most highly expressed acyl-CoA synthethases in the *Cycloclasticus* symbionts of *B. heckerae* and the encrusting sponge (Fig. 4, Supplementary Tables 6, 7). In the *Cycloclasticus* symbionts of *B. heckerae*, methylcitrate cycle genes were all highly expressed. In particular, expression levels of the gene encoding methylisocitrate lyase (*prpB*), the rate-limiting enzyme that catalyzes the last step of the cycle to form succinate and pyruvate, were nearly as high as those for pHMO (Fig. 4, Supplementary Tables 6, 7). In the two encrusting sponge individuals, genes involved in the methylcitrate cycle were also expressed, although not as highly as in the mussel symbionts. These results provide further support for the hypothesis that short-chain alkanes are metabolized by the *Cycloclasticus* symbionts. Furthermore, they suggest that propane

is a major source of carbon and energy for the symbiotic *Cycloclasticus* of mussels, and likely also for those of the encrusting sponge.

Supplementary Note 7

Assimilation of acetyl-CoA (C2) and butyryl-CoA (C4)

The assimilation of ethane (C2) via acetyl-CoA into biomass occurs through the glyoxylate bypass and requires two enzymes, isocitrate lyase (ICL) and malate synthase (MS)[8]. Genes encoding for ICL and MS were present in the genomes of all symbiotic *Cycloclasticus*. In the mussel *Cycloclasticus*, expression levels of these genes were considerably higher (ICL) or as high (MS) as the expression levels of single-copy housekeeping genes (Fig. 4, Supplementary Tables 5, 6).

For the assimilation of C2 compounds, isocitrate is cleaved to glyoxylate via ICL. Isocitrate is also a substrate for the catabolic tricarboxylic acid (TCA) cycle, in which it is oxidized to α-ketoglutarate by isocitrate dehydrogenase (IDH). Relative expression levels of ICL and IDH are therefore good indicators of carbon flux between the anabolic glyoxylate bypass and the catabolic TCA cycle. In the mussel *Cycloclasticus*, expression levels of ICL were higher than those of IDH, suggesting that C2 was routed into the assimilatory glyoxylate bypass (Fig. 4, Supplementary Tables 5, 6).

In the *Cycloclasticus* from the two encrusting sponge individuals, only one expressed ICL and MS genes, and at low levels (Fig. 6, Supplementary Tables 5, 6). Furthermore, expression ratios of ICL to IDH were much lower than those of the mussel *Cycloclasticus*. We

speculate that these differences reflect differences in the use of ethane by the mussel and sponge *Cycloclasticus*. Further support for this assumption is that Group X-like pHMOs, which are suggested to oxidize ethane[9], were highly expressed in the mussel *Cycloclasticus* but not present in the sponge *Cycloclasticus* genomes and transcriptomes.

Unlike cultivated *Cycloclasticus*, the symbiotic *Cycloclasticus* genomes contained genes encoding signature proteins involved in C4 metabolism, such as 3-hydroxybutyryl-CoA dehydratase, 3-hydroxybutyryl-CoA dehydrogenase and enoyl-CoA hydratase (Fig. 4, Supplementary Tables 5, 6). Although the expression of these enzymes was generally not higher than housekeeping genes, the incorporation of butane would be feasible in the symbiotic *Cycloclasticus* (Fig. 4).

Supplementary Note 8

The TCA cycle

While the expression levels for most TCA cycle enzymes were low to moderate (Fig. 5, Supplementary Tables 5, 6), the *Cycloclasticus* symbionts presumably derive the bulk of their cellular energy needs via the oxidation of C2-C4 substrates through the TCA cycle. A similar trend, with the TCA cycle enzymes constituting just 0.2-0.6 % of the total soluble protein, was observed when *Methylocella silvestris* was grown on propane as the sole carbon and energy source[10].

Supplementary Note 9

Short-chain alkanes can be converted to polyhydroxyalkanoate for carbon storage

Polyhydroxyalkanoates are carbon and energy rich compounds commonly stored by bacteria, and can play an important role in bacteria-invertebrate symbioses[11,12]. Polyhydroxybutyrate (PHB) is a common bacterial polyhydroxyalkanoate, and is generally synthesized from acetyl-CoA via condensation to acetoacetyl-CoA and its reduction to 3-hydroxybutyryl-CoA[13]. Finally, 3-hydroxybutyryl-CoA is polymerized to PHB via PHB synthase[13]. The expression of PHB synthase was substantial in the *Cycloclasticus* symbionts, suggesting similar to cultivated *Cycloclasticus*[14], the symbiotic *Cycloclasticus* actively synthesize PHB (Fig. 4, Supplementary Tables 5, 6).

Supplementary Note 10

An unusual sulfur oxidation (sox) operon in the symbiotic *Cycloclasticus*

Another unique feature found in the symbiotic *Cycloclasticus* genomes, yet absent in the PAH-degrading cultivated *Cycloclastiucs*, was the *sox* operon, which encodes periplasmic sulfur-oxidizing proteins[15].The *soxCDXYZAB* gene arrangement resembled the unusual *sox* operon that was described in *Congregibacter*/NOR5 clade[16,17]. The *Cycloclasticus* SoxB sequences clustered with the respective sequences from the *Congregibacter*/NOR5 clade (Supplementary Fig. 6). *Congregibacter* and *Cycloclasticus* have been found to co-exist in association with sea-surface oil aggregates[18], which can contain organic sulfur[19]. It was recently suggested that *Congregibacter* sulfur oxidation proteins degrade trace amounts of thiosulfate, possibly as a means of detoxification, without substantial energy yield[17]. Indeed, the poor expression of most of the *sox* genes in the symbiotic *Cycloclasticus*, excluding the substantially expressed *soxY* and

soxZ genes that encode a sulfur carrier complex[20], is atypical of metabolically active thiotrophs[21] (Supplementary Tables 5, 6).

Supplementary Methods

Incubations with isotopically labeled naphthalene

Gill tissue from four *B. heckerae* mussels was dissected immediately upon retrieval (the tissues from these four animals were also used for transcriptome analyses, Supplementary Table 2). Incubations were done with intact gill pieces that displaced approximately 0.5 ml seawater, as well as with 0.5 ml of gill homogenate (from different individuals). As a positive control we used *Cycloclasticus pugetii* ATCC 51542, obtained from the American Type Culture Collection (ATCC) and inoculated onboard in Marine Agar 2216 amended with a small crystal of unlabeled naphthalene. The culture was washed with naphthalene-free 0.2 μm filtered seawater and incubated at approximately 10^5 cells ml^{-1}. Negative controls included only naphthalene incubations, as well as incubations with formaldehyde-fixed tissue and cultures, non-symbiotic foot tissues and gill tissues from two *B. brooksi* individuals.

Naphthalene, uniformly labeled with ^{14}C [ring-14C (U)] (ca. 25 mCi $mmol^{-1}$), as well as ^{14}C labeled sodium bicarbonate (ca. 50 mCi $mmol^{-1}$) was obtained from Hartmann Analytic (Braunschweig, Germany Germany). The incubations were carried out by shaking samples at 4 °C, in 10 ml of 0.2 μm filtered seawater in 20 ml glass vials crimped with Teflon caps, following addition of 6.6 – 10 kBq (400000 – 600000 DPM) of ^{14}C naphthalene diluted in 0.1 ml of 20% ethanol. All treatments were conducted in triplicate. At each time point (0 h, 24 h and 48 h), formaldehyde was added to yield a 5% v/v solution to stop active metabolism and fix the

samples. Following 30 minutes fixation at room temperature, 10 µl of concentrated HCl was added and the vials were vigorously vortexed. The acidified vials were kept at room temperature for two hours. The headspace gas was pushed into an air-tight syringe by simultaneously injecting saturated NaCl brine into the vial. The gas was then injected into a 12 ml glass exetainer containing 1:7 phenylethilamine:2-methoxyethanol solution. Following vigorous vortexing, the solution was subsampled into scintillation vials containing Permafluor E+ (PerkinElmer) and the ^{14}C activity (DPM) from mineralized CO_2 was counted with a liquid scintillation analyzer (Packard, Tri-carb 2900TR, USA). The efficiency of this method was 99%, and was evaluated by acidifying, trapping and counting alkaline sodium bicarbonate with an initial activity of 10 kBq under experimental conditions identical to those of naphthalene incubations.

PCR amplification of the *Cycloclasticus* 16S rRNA gene

Primers that amplify a 484 bp fragment of the *Cycloclasticus* 16S rRNA gene, CycF346 (5`- ggaggcagcagtggggaata-3`) and CycR830 (5`- cggaaacccgcccaaca-3`), were designed using Primer-BLAST[1] and verified with Silva TestPrime[2]. The PCR conditions were as follows: denaturation at 95 °C for 30 s, annealing at 55 °C for 45 s and extension at 72 °C for 45 s, for 30 cycles. Bands of the appropriate size were present in all eight *B. heckerae* and three sponge DNA individuals, but not in the four *B. brooksi* individuals we examined. PCR products were purified using the QIAquick PCR Purification Kit (Qiagen, Hilden, Germany) and sequenced directly to verify the specificity of the PCR reaction. Sequencing was performed using the BigDye terminator v3.1 Cycle Sequencing Kit with the Genetic Analyzer ABI PRISM 3130 (Applied

Bio- systems). The amplified 16S rRNA gene sequences, determined by direct sequencing, were identical to the 16S rRNA sequences assembled from the *Cycloclasticus* genomes from the respective hosts (Figure 3).

Phylogenomic analyses

Publicly available *Cycloclasticus* genomes from isolates and SAGs, as well as methylotrophs from the *Piscirickettsiaceae* and *Methylococcaceae* families of the Gammaproteobacteria were retrieved from GenBank and JGI-IMG. Only two out of six *Cycloclasticus* SAGs from the JGI-IMG database had enough marker genes for phylogenomic analyses. Phylogenomic treeing was performed using scripts available at phylogenomics-tools (DOI:10.5281/zenodo.46122). Marker proteins that are universally conserved across the bacterial domain were extracted from genomes using the AMPHORA2 pipeline[22]. Eleven markers that were identified to occur in single copy in all genomes were used for alignment on MUSCLE. An alignment mask was generated using Zorro[23]. Poorly aligned regions or misaligned regions were inspected visually and removed from the alignments. The marker alignments were concatenated into a single partitioned alignment and the best protein substitution model was predicted using the concat_align.pl script (phylogenomics-tools). The best tree with SH-like aLRT support values was calculated on RAxML[24] using the tree_calculations.pl script available at phylogenomics-tools.

Supplementary References:

1. Raggi, L., Schubotz, F., Hinrichs, K.-U., Dubilier, N. & Petersen, J. M. Bacterial symbionts of *Bathymodiolus* mussels and *Escarpia* tubeworms from Chapopote, an

asphalt seep in the Southern Gulf of Mexico. *Environ. Microbiol.* **15,** 1969–1987 (2013).

2. Gutierrez, T. *et al.* Hydrocarbon-degrading bacteria enriched by the Deepwater Horizon oil spill identified by cultivation and DNA-SIP. *ISME J.* **7,** 2091–2104 (2013).

3. Stolyar, S., Costello, A. M., Peeples, T. L. & Lidstrom, M. E. Role of multiple gene copies in particulate methane monooxygenase activity in the methane-oxidizing bacterium *Methylococcus capsulatus* Bath. *Microbiology* **145,** 1235–1244 (1999).

4. Anthony, C. *The biochemistry of methylotrophs*. *London: Academic* **75,** (Academic Press, 1983).

5. Vorobev, A. *et al.* Genomic and transcriptomic analyses of the facultative methanotroph *Methylocystis* sp. strain SB2 grown on methane or ethanol. *Appl. Environ. Microbiol.* **80,** 3044–3052 (2014).

6. Brämer, C. O. & Steinbüchel, A. The methylcitric acid pathway in *Ralstonia eutropha*: new genes identified involved in propionate metabolism. *Microbiology* **147,** 2203–2214 (2001).

7. Brock, M., Maerker, C., Schütz, A., Völker, U. & Buckel, W. Oxidation of propionate to pyruvate in *Escherichia coli*: Involvement of methylcitrate dehydratase and aconitase. *Eur. J. Biochem.* **269,** 6184–6194 (2002).

8. Kornberg, H. & Krebs, H. Synthesis of cell constituents from C2-units by a modified tricarboxylic acid cycle. *Nature* **179,** 988–991 (1957).

9. Redmond, M. C., Valentine, D. L. & Sessions, A. L. Identification of novel methane-, ethane-, and propane-oxidizing bacteria at marine hydrocarbon seeps by stable isotope probing. *Appl. Environ. Microbiol.* **76,** 6412–22 (2010).

10. Patel, N. A. *et al.* Comparison of one- and two-dimensional liquid chromatography approaches in the label-free quantitative analysis of *Methylocella silvestris*. *J. Proteome Res.* **11,** 4755–4763 (2012).

11. Kleiner, M. *et al.* Metaproteomics of a gutless marine worm and its symbiotic microbial community reveal unusual pathways for carbon and energy use. *Proc. Natl. Acad. Sci. U.S.A.* **109,** 1173-1182 (2012).

12. Kim, J. K. *et al.* Polyester synthesis genes associated with stress resistance are involved in an insect-bacterium symbiosis. *Proc. Natl. Acad. Sci. U.S.A.* **110,** E2381–E2389 (2013).

13. Verlinden, R. A. J., Hill, D. J., Kenward, M. A., Williams, C. D. & Radecka, I. Bacterial synthesis of biodegradable polyhydroxyalkanoates. *J. Appl. Microbiol.* **102,** 1437–1449 (2007).

14. Messina, E. *et al.* Genome sequence of obligate marine polycyclic aromatic hydrocarbons-degrading bacterium *Cycloclasticus* sp. 78-ME, isolated from petroleum deposits of the sunken tanker Amoco Milford Haven, Mediterranean Sea. *Mar. Genomics* **25**, 11-13

(2015).

15. Friedrich, C. G., Rother, D., Bardischewsky, F., Quentmeier, A. & Fischer, J. Oxidation of reduced inorganic sulfur compounds by bacteria: emergence of a common mechanism? *Appl. Environ. Microbiol.* **67,** 2873–2882 (2001).

16. Fuchs, B. M. *et al.* Characterization of a marine gammaproteobacterium capable of aerobic anoxygenic photosynthesis. *Proc. Natl. Acad. Sci. U.S.A.* **104,** 2891–2896 (2007).

17. Spring, S. Function and evolution of the sox multienzyme complex in the marine gammaproteobacterium *Congregibacter litoralis*. *ISRN Microbiol.* **2014,** 597418 (2014).

18. Arnosti, C., Ziervogel, K., Yang, T. & Teske, A. Oil-derived marine aggregates – hot spots of polysaccharide degradation by specialized bacterial communities. *Deep Sea Res. Part II* **129,** 179–186 (2014).

19. Payzant, J. D., Montgomery, D. S. & Strausz, O. P. Sulfides in petroleum. *Org. Geochem.* **9,** 357–369 (1986).

20. Sauvé, V., Bruno, S., Berks, B. C. & Hemmings, A. M. The SoxYZ complex carries sulfur cycle intermediates on a peptide swinging arm. *J. Biol. Chem.* **282,** 23194–23204 (2007).

21. Stewart, F. J., Dmytrenko, O., Delong, E. F. & Cavanaugh, C. M. Metatranscriptomic analysis of sulfur oxidation genes in the endosymbiont of *Solemya velum*. *Front. Microbiol.* **2,** 134 (2011).

22. Wu, M. & Scott, A. J. Phylogenomic analysis of bacterial and archaeal sequences with AMPHORA2. *Bioinformatics* **28,** 1033–1034 (2012).

23. Wu, M., Chatterji, S. & Eisen, J. A. Accounting for alignment uncertainty in phylogenomics. *PLoS One* **7,** 1–10 (2012).

24. Stamatakis, A. RAxML version 8: A tool for phylogenetic analysis and post-analysis of large phylogenies. *Bioinformatics* **30,** 1312–1313 (2014).

Supplementary tables

Supplementary Table 1 | The relative abundance (coverage) of methanotrophic, thiotrophic and *Cycloclasticus* symbionts (%), based on unambiguous read mapping to the respective 16S rRNA gene sequences. SD is standard deviation.

	B.heckerae B	*B.heckerae D*	*B.heckerae E*	*B.heckerae O*	*B.heckerae P*	*Average*	*SD*
Methanotroph	60.6	58.0	61.6	66.4	69.9	63.3	4.8
Thiotroph	28.2	33.9	32.6	28.3	17.1	28.0	6.6
Cycloclasticus	10.7	5.8	5.5	5.1	9.2	7.3	2.5
Methylotroph	0.6	2.4	0.3	0.2	3.8	1.5	1.6

	Encrusting sponge Mictlan	*Encrusting sponge Chapopote*	*Branching sponge Chapopote*	*Average*	*SD*
Methanotroph	75.8	66.5	60.1	67.5	7.9
Thiotroph	4.8	5.5	10.9	7.0	3.4
Cycloclasticus	8.0	4.1	4.8	5.6	2.1
Other	11.4	24.0	24.2	19.8	7.3

Supplementary Table 2 | Samples used for genomic and transcriptomic analyses.
Sequencing depth for respective DNA and RNA libraries and genome accession numbers are shown.

Host	*Collection site*	*Collection date*	*DNA library*	*Genome reads (million)*	*Cycloclasticus genome GOLD ID*	*RNA library*	*Transcriptome reads (million)*	*Proteome*
B. heckerae	Chapopote	20-Mar-15	B	20.5	Ga0105567			
B. heckerae	Chapopote	14-Mar-15	D	30.6	Ga0097688			
B. heckerae	Chapopote	14-Mar-15	E	12.6	Ga0105568			
B. heckerae	Chapopote	10-Mar-15	O*	17.7		BHT3	24.8	BHP3
B. heckerae	Chapopote	10-Mar-15	P	33.0	Ga0097674	BHT4	28.7	
B. heckerae	Chapopote	10-Mar-15	Q	10.4				
B. heckerae	Chapopote	20-Mar-15				BHT1	24.7	BHP1
B. heckerae	Chapopote	20-Mar-15				BHT2	20.6	BHP2
B. brooksi	Chapopote	14-Mar-15		46.6				
B. brooksi	Chapopote	20-Mar-15		46.2				
B. brooksi	Chapopote	20-Mar-15		49.1				
B. brooksi	Chapopote	14-Mar-15		15.0				
Encrusting sponge	Mictlan	11-Mar-15	L	27.7	Ga0105570	ST1	30.0	
Branching sponge	Chapopote	20-Mar-15	M	42.6	Ga0097675			
Encrusting sponge	Chapopote	20-Mar-15	N	29.4	Ga0097676	ST2	38.0	

* Due to low coverage, a *Cycloclasticus* genome of satisfactory quality could not be assembled from library O.

Supplementary Table 3 | Assembly quality assessment for *B. heckerae* and sponge *Cycloclasticus* genomes. General assembly parameters as assessed on Quast, and draft genome completeness as estimated using CheckM.

Bin	*B.heckerae B*	*B.heckerae D*	*B.heckerae E*	*B.heckerae P*	*Encrusting sponge Mictlan*	*Encrusting sponge Chapopote*	*Branching sponge Chapopote*
# contigs	121	194	199	123	67	77	244
Largest contig	220606	186995	209425	183234	201656	312680	109505
Total length	2141540	2189493	2212916	2130426	1616205	1916434	2321044
GC (%)	42.15	42.22	42.27	42.21	43.05	42.74	43.72
N50	84355	43544	119959	82558	82730	99813	27724
N75	35482	20252	43418	45939	27266	66376	9409
L50	8	14	7	9	6	5	26
L75	18	31	15	17	15	11	62
# N's per 100 kbp	0	4.34	0	0	0	4.64	4.35
Completeness (%)	94.7	93.3	94.9	97.1	89.9	91.5	95.6

Supplementary Table 4 | Pairwise two-way average nucleotide identities (ANI) of the symbiotic *Cycloclasticus* draft genomes. SD is standard deviation of ANI values.

Genomes compared	*2 way ANI (%)*	*SD (%)*
B. heckerae B vs B. heckerae D	99.97	0.08
B. heckerae B vs B. heckerae E	99.98	0.09
B. heckerae B vs B. heckerae P	99.97	0.66
B. heckerae D vs B. heckerae E	99.96	0.15
B. heckerae D vs B. heckerae P	99.95	0.62
B. heckerae E vs B. heckerae P	99.99	0.04
Encrusting sponge Chapopote *vs B. heckerae P*	79.13	4.23
Encrusting sponge Mictlan *vs B. heckerae P*	79.08	4.24
Branching sponge Chapopote *vs B. heckerae P*	78.64	4.56
Branching sponge *Chapopote vs* encrusting sponge Mictlan	79.81	5.01
Branching sponge *Chapopote vs* encrusting sponge Chapopote	79.80	4.99
Branching sponge *Mictlan vs* branching sponge Chapopote	99.84	1.17
B. heckerae vs C. pugetii	77.10	3.53
Branching sponge Chapopote *vs C. pugetii*	76.23	3.33
Encrusting sponge Chapopote *vs C. pugetii*	76.32	3.01
Encrusting sponge Mictlan *vs C. pugetii*	76.35	2.96

Supplementary Table 5 | Genes homologous to those involved in PAH degradation pathways in *Cycloclasticus pugetii* could not be detected in metagenomes of *B. heckerae* and sponges. Values in this table show the number of Illumina reads from *B. heckerae* and sponge metagenomic libraries that mapped to selected gene clusters in *C. pugetii* genome (GenBank accession number: NZ_ARVU01000001), using a tolerant minimum sequence identity of 0.6. These values are normalized to gene length and sequencing depth using the following formula: $\frac{N_x * 10^6}{L_x * \sum_{i=1}^{n} \frac{N_i}{L_i}}$, where N_x is number of Illumina reads mapped to the gene of interest, L_x is length (bp) of the gene of interest, $\sum_{i=1}^{n} \frac{N_i}{L_i}$ is the sum of length normalized read counts from a metagenomic library of interest for all n genes in *C. pugetii* genome. We have multiplied the values by 10^6 to yield comfortably-read numerals. Each group of rows separated by a blank row represents a gene cluster. First two groups, those of isocitrate lyase/malate synthase and methycitrate cycle genes are an example of gene clusters conserved among *C. pugetii*, B. heckerae and sponge *Cycloclasticus*. Gene clusters that are conserved among *C. pugetii* and sponge, but not *B. heckerae Cycloclasticus* are presented in the third cluster (urea ABC transporter genes). The very few reads that did map to PAH genes most likely originated from environmental contaminants. The remaining rows in this table represent all genes coding for enzymes that catalyze the central reactions in PAH metabolism of *C. pugetii*. * This gene is highly conserved not only among *Cycloclasticus*, but also among other bacteria that are present in the metagenomes, hence the exceptionally high coverage values. ** Fragments from these genes were amplified by PCR in *B. heckerae* samples collected in 2006.

Library	*B.heckerae* B	*B.heckerae* D	*B.heckerae* E	*B.heckerae* P	Encrusting sponge Mictlan	Encrusting sponge Chapopote	Branching sponge Chapopote
Isocitrate lyase	806	680	916	989	798	555	702
Malate synthase	770	666	701	779	407	309	608
2-methylaconitate cis-trans isomerase	322	208	338	326	468	330	226
2-methylcitrate dehydratase	996	665	925	873	916	700	452
2-methylcitrate synthase	982	847	858	880	755	510	600
Methylisocitrate lyase	1286	937	945	1133	952	619	762
Acetyl-coenzyme A synthetase	816	510	613	776	719	564	588
Urea ABC transporter, urea binding protein*	0	0	0	3	2892	3201	0
Urea ABC transporter, permease protein UrtB	1	0	0	0	678	690	0
Urea ABC transporter, permease protein UrtC	0	1	3	0	736	745	1

Urea ABC transporter, ATPase protein UrtD	0	1	0	3	731	824	9
Urea ABC transporter, ATPase protein UrtE	0	0	0	3	321	443	6
Small subunit naph/bph dioxygenase	0	0	0	0	0	0	0
Large subunit naph/bph dioxygenase	0	0	0	0	1	1	0
2,3-dihydroxybiphenyl 1,2-dioxygenase	0	0	0	0	0	2	0
Small subunit naph/bph dioxygenase	0	0	0	0	0	3	0
Large subunit naph/bph dioxygenase	0	0	0	0	0	1	0
2,3-dihydroxybiphenyl 1,2-dioxygenase	0	0	0	0	0	2	0
Ortho-halobenzoate 1,2-dioxygenase alpha-ISP protein OhbB	0	0	0	0	1	0	0
2-hydroxy-6-oxo-6-phenylhexa-2,4-dienoate hydrolase	0	0	0	0	1	4	0
Ortho-halobenzoate 1,2-dioxygenase beta-ISP protein OhbA	0	0	0	0	0	0	0
Vanillate O-demethylase oxygenase subunit	0	0	0	0	0	0	0
Benzoate 1,2-dioxygenase alpha subunit	0	0	0	0	0	0	0
Benzoate 1,2-dioxygenase beta subunit	0	0	0	0	0	0	0
2-hydroxy-6-oxo-6-phenylhexa-2,4-dienoate hydrolase	0	0	0	3	0	0	0
Large subunit naph/bph dioxygenase	0	0	0	0	0	1	0
Small subunit naph/bph dioxygenase	6	7	0	4	0	0	0
Large subunit naph/bph dioxygenase	0	0	0	0	0	1	0
Small subunit naph/bph dioxygenase	0	0	0	0	0	0	0
Aromatic-ring-hydroxylating dioxygenase, beta subunit	0	0	0	0	0	0	0
Large subunit naph/bph dioxygenase	0	0	0	0	0	2	1
2,3-dihydroxybiphenyl 1,2-dioxygenase	0	0	0	0	0	0	0
Small subunit naph/bph dioxygenase	0	0	0	0	0	3	0
Large subunit naph/bph dioxygenase **KC007443-4/AGC82205-6	0	0	0	0	0	0	0
Large subunit toluate/benzoate dioxygenase	0	0	0	0	0	0	0
4-hydroxythreonine-4-phosphate dehydrogenase	0	0	0	0	0	0	0
hypothetical protein	0	0	0	0	0	0	0

Ortho-halobenzoate 1,2-dioxygenase beta-ISP protein OhbA	0	0	0	0	0	0	0
Ortho-halobenzoate 1,2-dioxygenase alpha-ISP protein OhbB	0	0	0	0	0	0	0
Large subunit, aromatic oxygenase	0	0	0	0	0	0	0
Ortho-halobenzoate 1,2-dioxygenase beta-ISP protein OhbA	0	0	0	0	0	0	0
2,3-dihydroxy-2,3-dihydro-phenylpropionate dehydrogenase	0	0	0	0	1	0	0
Small subunit naph/bph dioxygenase	0	0	0	0	2	0	0
Large subunit naph/bph dioxygenase	0	0	0	0	0	9	0
1,2-dihydroxycyclohexa-3,5-diene-1-carboxylate dehydrogenase	0	0	0	0	0	0	0
Benzoate 1,2-dioxygenase alpha subunit	0	0	0	0	0	1	1
Benzoate 1,2-dioxygenase beta subunit	0	0	0	0	0	0	0
2,3-dihydroxybiphenyl 1,2-dioxygenase	0	0	0	0	0	2	0
Ortho-halobenzoate 1,2-dioxygenase alpha-ISP protein OhbB	0	0	0	0	1	0	0
2-hydroxy-6-oxo-6-phenylhexa-2,4-dienoate hydrolase	0	0	0	0	1	4	0
Ortho-halobenzoate 12C2-dioxygenase beta-ISP protein OhbA	0	0	0	0	0	0	0
Vanillate O-demethylase oxygenase subunit	0	0	0	0	0	0	0
Benzoate 1,2-dioxygenase alpha subunit	0	0	0	0	0	0	0
Benzoate 1,2-dioxygenase beta subunit	0	0	0	0	0	0	0
2-hydroxy-6-oxo-6-phenylhexa-2,4-dienoate hydrolase	0	0	0	3	0	0	0
Ortho-halobenzoate 1,2-dioxygenase alpha-ISP protein OhbB	0	0	0	0	0	0	0
Ortho-halobenzoate 1,2-dioxygenase beta-ISP protein OhbA	0	0	0	0	0	0	0
Aromatic-ring-hydroxylating dioxygenase, beta subunit	0	0	0	0	0	0	0
Benzoate 1,2-dioxygenase alpha subunit	0	0	0	0	1	0	0
Scale	0	250	500	1000	1500	2000	3000

Supplementary Table 6 | Expression profiles of genes involved in carbon and energy metabolism in *Bathymodiolus* and sponge *Cycloclasticus* symbionts. Trimmed mean of M-values (TMM) and size (RPKM) normalized raw counts are shown. BHT1-4 and ST1-2 are transcriptome replicates for *B. heckerae* mussels and encrusting sponges, respectively.

	BHT1	*BHT2*	*BHT3*	*BHT4*	*ST1*	*ST2*
Total reads mapped to coding sequences (CDS)	62910	26269	175665	47264	108351	13938
pHMO A-subunit (Group Z)	1485	4405	5562	3618	19752	4806
pHMO B-subunit (Group Z)	1225	1503	2972	3879	25894	5274
pHMO C-subunit (Group Z)	1082	437	1403	3437	8003	2404
pHMO A-subunit (Group X-like)	426	780	1627	546		
pHMO B-subunit (Group X-like)	253	691	1184	633		
pHMO C-subunit (Group X-like)	1279	644	3777	2006		
Serine hydroxymethyltransferase	38	207	21	50	137	92
Quino(hemo)protein alcohol dehydrogenase PQQ-dependent	187	444	194	783	3331	1096
Tungsten-containing aldehyde:ferredoxin oxidoreductase	542	397	519	1515	4752	1984
Acetyl-coenzyme A synthetase	136	278	267	434	591	119
Isocitrate lyase	796	650	218	789	208	0
Malate synthase	0	25	34	186	123	0
Citrate synthase	0	77	62	65	46	89
Aconitate hydratase	74	400	155	170	369	111
Isocitrate dehydrogenase	0	72	36	57	109	156
Dihydrolipoamide succinyltransferase component (E2)	19	596	137	186	482	139
Dihydrolipoamide dehydrogenase	33	125	150	88	120	0
2-Oxoglutarate dehydrogenase E1 component	35	149	108	52	120	184
Succinyl-CoA ligase [ADP-forming] beta chain	20	35	169	126	317	396
Succinyl-CoA ligase [ADP-forming] alpha chain	0	46	113	96	91	0
Succinate dehydrogenase iron-sulfur protein	0	0	34	107	130	0
Succinate dehydrogenase flavoprotein subunit	41	278	179	176	134	0
Succinate dehydrogenase cytochrome b-556 subunit	0	0	0	53	145	146
Succinate dehydrogenase hydrophobic membrane anchor protein	0	0	0	55	0	454
Fumarate hydratase	31	14	29	0	60	0
Malate dehydrogenase	49	674	73	65	517	0
Propioniyl-coenzyme A synthetase *PrpE*	177	400	640	742	1045	182
2-Methylcitrate synthase *PrpC*	320	1328	531	634	693	102
2-Methylcitrate dehydratase FeS dependent *PrpD*	185	748	638	252	767	467
2-Methylaconitate isomerase *prpF acnB*	0	356	38	71	88	0
Methylisocitrate lyase *PrpB*	378	2746	1802	1042	542	845
Butyryl-CoA synthetase	74	222	83	338	795	107

Trans-2-enoyl-CoA reductase (NAD+)	80	16	37	0	77	192
Butyryl-CoA dehydrogenase	0	82	0	0	11	0
Enoyl-CoA hydratase/isomerase	0	0	0	0	151	75
3-Hydroxybutyryl-CoA dehydratase	211	151	180	53	250	139
Polyhydroxyalkanoic acid (PHB) synthase	56	197	747	232	1945	67
Polyhydroxyalkanoic acid (PHB) synthase copy 2	0	57	8	0	11	0
3-Hydroxybutyryl-CoA dehydrogenase	85	94	32	0		
3-Hydroxyacyl-CoA dehydrogenase	42	43	12	45	28	0
Acetoacetyl-CoA reductase	0	321	383	0	1720	927
Acetoacetyl-CoA reductase copy 2	0	163	85	228	1073	703
Acetyl-CoA acetyltransferase	405	673	293	107	1681	243
ATP synthase alpha chain	156	1334	794	218	387	112
ATP synthase beta chain	18	910	582	441	348	125
ATP synthase gamma chain	195	416	507	121	1010	467
ATP synthase delta chain	0	524	301	434	366	108
ATP synthase epsilon chain	0	288	22	252	56	139
ATP synthase F0 sector subunit a	0	0	0	0	141	0
ATP synthase F0 sector subunit b	0	333	745	219	742	120
ATP synthase F0 sector subunit c	0	169	528	0	429	241
Ubiquinol-cytochrome C reductase iron-sulfur subunit	0	68	46	0	39	196
Ubiquinol-cytochrome c reductase cytochrome B subunit	0	82	73	0	18	47
Ubiquinol cytochrome C oxidoreductase cytochrome C1	32	0	121	0	77	0
Cytochrome c oxidase subunit *CcoN*	33	142	215	29		
Cytochrome c oxidase subunit *CcoO*	39	66	191	0		
Cytochrome c oxidase subunit *CcoQ*	0	115	0	0		
Cytochrome c oxidase subunit *CcoP*	0	22	50	47		
Cytochrome c oxidase polypeptide I *coxA*	43	159	190	260	756	329
Cytochrome c oxidase polypeptide II *coxB*	0	25	62	160	468	102
Cytochrome c oxidase polypeptide III *coxC*	0	24	135	219	81	0
SoxA	0	48	11	50	0	0
SoxB	0	0	36	0	0	0
SoxC	0	0	14	0	7	0
SoxD	0	21	29	0	11	0
SoxX	0	57	103	0	0	326
SoxY	314	87	156	138	25	251
SoxZ	79	66	89	139	39	1143
Gyrase B	119	124	22	104	352	96
Gyrase A	19	39	52	49	195	89
RecA	23	135	155	41	570	441

Supplementary Table 7 | Relative expression of key genes in carbon and energy metabolism of the symbiotic *Cycloclasticus* as inferred from transcriptomic analysis. The values presented here were normalized to *gyrA* (**a**), *gyrB* (**b**) and *recA* (**c**), following trimmed mean of M-values (TMM) and size (RPKM) normalizations. BHT1-4 are transcriptome replicates from four *B. heckerae* individuals, and ST1-2 from the two encrusting sponge individuals. Empty cells indicate absence of the gene in sponge *Cycloclasticus* genomes.

a)

	BHT1	*BHT2*	*BHT3*	*BHT4*	*ST1*	*ST2*
pHMO A-subunit (Group Z)	78.16	112.95	106.96	73.84	101.29	54.00
pHMO B-subunit (Group Z)	64.47	38.54	57.15	79.16	132.79	59.26
pHMO C-subunit (Group Z)	56.95	11.21	26.98	70.14	41.04	27.01
pHMO A-subunit (Group X-like)	22.42	20.00	31.29	11.14		
pHMO B-subunit (Group X-like)	13.32	17.72	22.77	12.92		
pHMO C-subunit (Group X-like)	67.32	16.51	72.63	40.94		
Serine hydroxymethyltransferase	2.00	5.31	0.40	1.02	0.70	1.03
Quino(hemo)protein alcohol dehydrogenase PQQ-dependent	9.84	11.38	3.73	15.98	17.08	12.31
Tungsten-containing aldehyde:ferredoxin oxidoreductase	28.53	10.18	9.98	30.92	24.37	22.29
Acetyl-coenzyme A synthetase	7.16	7.13	5.13	8.86	3.03	1.34
Isocitrate lyase	41.89	16.67	4.19	16.10	1.07	0.00
Malate synthase	0.00	0.64	0.65	3.80	0.63	0.00
Citrate synthase	0.00	1.97	1.19	1.33	0.24	1.00
Aconitate hydratase	3.89	10.26	2.98	3.47	1.89	1.25
Isocitrate dehydrogenase	0.00	1.85	0.69	1.16	0.56	1.75
Dihydrolipoamide succinyltransferase component (E2)	1.00	15.28	2.63	3.80	2.47	1.56
Dihydrolipoamide dehydrogenase	1.74	3.21	2.88	1.80	0.62	0.00
2-Oxoglutarate dehydrogenase E1 component	1.84	3.82	2.08	1.06	0.62	2.07
Succinyl-CoA ligase [ADP-forming] beta chain	1.05	0.90	3.25	2.57	1.63	4.45
Succinyl-CoA ligase [ADP-forming] alpha chain	0.00	1.18	2.17	1.96	0.47	0.00
Succinate dehydrogenase iron-sulfur protein	0.00	0.00	0.65	2.18	0.67	0.00
Succinate dehydrogenase flavoprotein subunit	2.16	7.13	3.44	3.59	0.69	0.00
Succinate dehydrogenase cytochrome b-556 subunit	0.00	0.00	0.00	1.08	0.74	1.64
Succinate dehydrogenase hydrophobic membrane anchor protein	0.00	0.00	0.00	1.12	0.00	5.10
Fumarate hydratase	1.63	0.36	0.56	0.00	0.31	0.00

Malate dehydrogenase	2.58	17.28	1.40	1.33	2.65	0.00
Propioniyl-coenzyme A synthetase *PrpE*	9.32	10.26	12.31	15.14	5.36	2.04
2-Methylcitrate synthase *PrpC*	16.84	34.05	10.21	12.94	3.55	1.15
2-Methylcitrate dehydratase FeS dependent *PrpD*	9.74	19.18	12.27	5.14	3.93	5.25
2-Methylaconitate isomerase *prpF acnB*	0.00	9.13	0.73	1.45	0.45	0.00
Methylisocitrate lyase *PrpB*	19.89	70.41	34.65	21.27	2.78	9.49
Butyryl-CoA synthetase	3.89	5.69	1.60	6.90	4.08	1.20
Trans-2-enoyl-CoA reductase (NAD+)	4.21	0.41	0.71	0.00	0.39	2.16
Butyryl-CoA dehydrogenase	0.00	2.10	0.00	0.00	0.06	0.00
Enoyl-CoA hydratase/isomerase	0.00	0.00	0.00	0.00	0.77	0.84
3-Hydroxybutyryl-CoA dehydratase	11.11	3.87	3.46	1.08	1.28	1.56
Polyhydroxyalkanoic acid (PHB) synthase	2.95	5.05	14.37	4.73	9.97	0.75
Polyhydroxyalkanoic acid (PHB) synthase copy 2	0.00	1.46	0.15	0.00	0.06	0.00
3-Hydroxybutyryl-CoA dehydrogenase	4.47	2.41	0.62	0.00	0.00	0.00
3-Hydroxyacyl-CoA dehydrogenase	2.21	1.10	0.23	0.92	0.14	0.00
Acetoacetyl-CoA reductase	0.00	8.23	7.37	0.00	8.82	10.42
Acetoacetyl-CoA reductase copy 2	0.00	4.18	1.63	4.65	5.50	7.90
Acetyl-CoA acetyltransferase	21.32	17.26	5.63	2.18	8.62	2.73
ATP synthase alpha chain	8.21	34.21	15.27	4.45	1.98	1.26
ATP synthase beta chain	0.95	23.33	11.19	9.00	1.78	1.40
ATP synthase gamma chain	10.26	10.67	9.75	2.47	5.18	5.25
ATP synthase delta chain	0.00	13.44	5.79	8.86	1.88	1.21
ATP synthase epsilon chain	0.00	7.38	0.42	5.14	0.29	1.56
ATP synthase F0 sector subunit a	0.00	0.00	0.00	0.00	0.72	0.00
ATP synthase F0 sector subunit b	0.00	8.54	14.33	4.47	3.81	1.35
ATP synthase F0 sector subunit c	0.00	4.33	10.15	0.00	2.20	2.71
Ubiquinol-cytochrome C reductase iron-sulfur subunit	0.00	1.74	0.88	0.00	0.20	2.20
Ubiquinol-cytochrome c reductase cytochrome B subunit	0.00	2.10	1.40	0.00	0.09	0.53
Ubiquinol cytochrome C oxidoreductase cytochrome C1	1.68	0.00	2.33	0.00	0.39	0.00
Cytochrome c oxidase subunit *CcoN*	1.74	3.64	4.13	0.59		
Cytochrome c oxidase subunit *CcoO*	2.05	1.69	3.67	0.00		
Cytochrome c oxidase subunit *CcoQ*	0.00	2.95	0.00	0.00		
Cytochrome c oxidase subunit *CcoP*	0.00	0.56	0.96	0.96		
Cytochrome c oxidase polypeptide I *coxA*	2.26	4.08	3.65	5.31	3.88	3.70
Cytochrome c oxidase polypeptide II *coxB*	0.00	0.64	1.19	3.27	2.40	1.15
Cytochrome c oxidase polypeptide III *coxC*	0.00	0.62	2.60	4.47	0.42	0.00
SoxA	0.00	1.23	0.21	1.02	0.00	0.00

SoxB	0.00	0.00	0.69	0.00	0.00	0.00
SoxC	0.00	0.00	0.27	0.00	0.04	0.00
SoxD	0.00	0.54	0.56	0.00	0.06	0.00
SoxX	0.00	1.46	1.98	0.00	0.00	3.66
SoxY	16.53	2.23	3.00	2.82	0.13	2.82
SoxZ	4.16	1.69	1.71	2.84	0.20	12.84

b)

	BHT1	*BHT2*	*BHT3*	*BHT4*	*ST1*	*ST2*
pHMO A-subunit (Group Z)	12.48	35.52	252.82	34.79	56.11	50.06
pHMO B-subunit (Group Z)	10.29	12.12	135.09	37.30	73.56	54.94
pHMO C-subunit (Group Z)	9.09	3.52	63.77	33.05	22.74	25.04
pHMO A-subunit (Group X-like)	3.58	6.29	73.95	5.25		
pHMO B-subunit (Group X-like)	2.13	5.57	53.82	6.09		
pHMO C-subunit (Group X-like)	10.75	5.19	171.68	19.29		
Serine hydroxymethyltransferase	0.32	1.67	0.95	0.48	0.39	0.96
Quino(hemo)protein alcohol dehydrogenase PQQ-dependent	1.57	3.58	8.82	7.53	9.46	11.42
Tungsten-containing aldehyde:ferredoxin oxidoreductase	4.55	3.20	23.59	14.57	13.50	20.67
Acetyl-coenzyme A synthetase	1.14	2.24	12.14	4.17	1.68	1.24
Isocitrate lyase	6.69	5.24	9.91	7.59	0.59	0.00
Malate synthase	0.00	0.20	1.55	1.79	0.35	0.00
Citrate synthase	0.00	0.62	2.82	0.63	0.13	0.93
Aconitate hydratase	0.62	3.23	7.05	1.63	1.05	1.16
Isocitrate dehydrogenase	0.00	0.58	1.64	0.55	0.31	1.63
Dihydrolipoamide succinyltransferase component (E2)	0.16	4.81	6.23	1.79	1.37	1.45
Dihydrolipoamide dehydrogenase	0.28	1.01	6.82	0.85	0.34	0.00
2-Oxoglutarate dehydrogenase E1 component	0.29	1.20	4.91	0.50	0.34	1.92
Succinyl-CoA ligase [ADP-forming] beta chain	0.17	0.28	7.68	1.21	0.90	4.13
Succinyl-CoA ligase [ADP-forming] alpha chain	0.00	0.37	5.14	0.92	0.26	0.00
Succinate dehydrogenase iron-sulfur protein	0.00	0.00	1.55	1.03	0.37	0.00
Succinate dehydrogenase flavoprotein subunit	0.34	2.24	8.14	1.69	0.38	0.00
Succinate dehydrogenase cytochrome b-556 subunit	0.00	0.00	0.00	0.51	0.41	1.52
Succinate dehydrogenase hydrophobic	0.00	0.00	0.00	0.53	0.00	4.73

membrane anchor protein						
Fumarate hydratase	0.26	0.11	1.32	0.00	0.17	0.00
Malate dehydrogenase	0.41	5.44	3.32	0.63	1.47	0.00
Propioniyl-coenzyme A synthetase *PrpE*	1.49	3.23	29.09	7.13	2.97	1.90
2-Methylcitrate synthase *PrpC*	2.69	10.71	24.14	6.10	1.97	1.06
2-Methylcitrate dehydratase FeS dependent *PrpD*	1.55	6.03	29.00	2.42	2.18	4.86
2-Methylaconitate isomerase *prpF acnB*	0.00	2.87	1.73	0.68	0.25	0.00
Methylisocitrate lyase *PrpB*	3.18	22.15	81.91	10.02	1.54	8.80
Butyryl-CoA synthetase	0.62	1.79	3.77	3.25	2.26	1.11
Trans-2-enoyl-CoA reductase (NAD+)	0.67	0.13	1.68	0.00	0.22	2.00
Butyryl-CoA dehydrogenase	0.00	0.66	0.00	0.00	0.03	0.00
Enoyl-CoA hydratase/isomerase	0.00	0.00	0.00	0.00	0.43	0.78
3-Hydroxybutyryl-CoA dehydratase	1.77	1.22	8.18	0.51	0.71	1.45
Polyhydroxyalkanoic acid (PHB) synthase	0.47	1.59	33.95	2.23	5.53	0.70
Polyhydroxyalkanoic acid (PHB) synthase copy 2	0.00	0.46	0.36	0.00	0.03	0.00
3-Hydroxybutyryl-CoA dehydrogenase	0.71	0.76	1.45	0.00	0.00	0.00
3-Hydroxyacyl-CoA dehydrogenase	0.35	0.35	0.55	0.43	0.08	0.00
Acetoacetyl-CoA reductase	0.00	2.59	17.41	0.00	4.89	9.66
Acetoacetyl-CoA reductase copy 2	0.00	1.31	3.86	2.19	3.05	7.32
Acetyl-CoA acetyltransferase	3.40	5.43	13.32	1.03	4.78	2.53
ATP synthase alpha chain	1.31	10.76	36.09	2.10	1.10	1.17
ATP synthase beta chain	0.15	7.34	26.45	4.24	0.99	1.30
ATP synthase gamma chain	1.64	3.35	23.05	1.16	2.87	4.86
ATP synthase delta chain	0.00	4.23	13.68	4.17	1.04	1.13
ATP synthase epsilon chain	0.00	2.32	1.00	2.42	0.16	1.45
ATP synthase F0 sector subunit a	0.00	0.00	0.00	0.00	0.40	0.00
ATP synthase F0 sector subunit b	0.00	2.69	33.86	2.11	2.11	1.25
ATP synthase F0 sector subunit c	0.00	1.36	24.00	0.00	1.22	2.51
Ubiquinol-cytochrome C reductase iron-sulfur subunit	0.00	0.55	2.09	0.00	0.11	2.04
Ubiquinol-cytochrome c reductase cytochrome B subunit	0.00	0.66	3.32	0.00	0.05	0.49
Ubiquinol cytochrome C oxidoreductase cytochrome C1	0.27	0.00	5.50	0.00	0.22	0.00
Cytochrome c oxidase subunit *CcoN*	0.28	1.15	9.77	0.28		
Cytochrome c oxidase subunit *CcoO*	0.33	0.53	8.68	0.00		
Cytochrome c oxidase subunit *CcoQ*	0.00	0.93	0.00	0.00		
Cytochrome c oxidase subunit *CcoP*	0.00	0.18	2.27	0.45		
Cytochrome c oxidase polypeptide I *coxA*	0.36	1.28	8.64	2.50	2.15	3.43
Cytochrome c oxidase polypeptide II *coxB*	0.00	0.20	2.82	1.54	1.33	1.06

Cytochrome c oxidase polypeptide III						
coxC	0.00	0.19	6.14	2.11	0.23	0.00
SoxA	0.00	0.39	0.50	0.48	0.00	0.00
SoxB	0.00	0.00	1.64	0.00	0.00	0.00
SoxC	0.00	0.00	0.64	0.00	0.02	0.00
SoxD	0.00	0.17	1.32	0.00	0.03	0.00
SoxX	0.00	0.46	4.68	0.00	0.00	3.40
SoxY	2.64	0.70	7.09	1.33	0.07	2.61
SoxZ	0.66	0.53	4.05	1.34	0.11	11.91

c)

	BHT1	*BHT2*	*BHT3*	*BHT4*	*ST1*	*ST2*
pHMO A-subunit (Group Z)	64.57	32.63	35.88	88.24	34.65	10.90
pHMO B-subunit (Group Z)	53.26	11.13	19.17	94.61	45.43	11.96
pHMO C-subunit (Group Z)	47.04	3.24	9.05	83.83	14.04	5.45
pHMO A-subunit (Group X-*like*)	18.52	5.78	10.50	13.32		
pHMO B-subunit (Group X-*like*)	11.00	5.12	7.64	15.44		
pHMO C-subunit (Group X-*like*)	55.61	4.77	24.37	48.93		
Serine hydroxymethyltransferase	1.65	1.53	0.14	1.22	0.24	0.21
Quino(hemo)protein alcohol dehydrogenase PQQ-dependent	8.13	3.29	1.25	19.10	5.84	2.49
Tungsten-containing aldehyde:ferredoxin oxidoreductase	23.57	2.94	3.35	36.95	8.34	4.50
Acetyl-coenzyme A synthetase	5.91	2.06	1.72	10.59	1.04	0.27
Isocitrate lyase	34.61	4.81	1.41	19.24	0.36	0.00
Malate synthase	0.00	0.19	0.22	4.54	0.22	0.00
Citrate synthase	0.00	0.57	0.40	1.59	0.08	0.20
Aconitate hydratase	3.22	2.96	1.00	4.15	0.65	0.25
Isocitrate dehydrogenase	0.00	0.53	0.23	1.39	0.19	0.35
Dihydrolipoamide succinyltransferase component (E2)	0.83	4.41	0.88	4.54	0.85	0.32
Dihydrolipoamide dehydrogenase	1.43	0.93	0.97	2.15	0.21	0.00
2-Oxoglutarate dehydrogenase E1 component	1.52	1.10	0.70	1.27	0.21	0.42
Succinyl-CoA ligase [ADP-forming] beta chain	0.87	0.26	1.09	3.07	0.56	0.90
Succinyl-CoA ligase [ADP-forming] alpha chain	0.00	0.34	0.73	2.34	0.16	0.00
Succinate dehydrogenase iron-sulfur protein	0.00	0.00	0.22	2.61	0.23	0.00
Succinate dehydrogenase flavoprotein subunit	1.78	2.06	1.15	4.29	0.24	0.00

Succinate dehydrogenase cytochrome b-556 subunit	0.00	0.00	0.00	1.29	0.25	0.33
Succinate dehydrogenase hydrophobic membrane anchor protein	0.00	0.00	0.00	1.34	0.00	1.03
Fumarate hydratase	1.35	0.10	0.19	0.00	0.11	0.00
Malate dehydrogenase	2.13	4.99	0.47	1.59	0.91	0.00
Propioniyl-coenzyme A synthetase *PrpE*	7.70	2.96	4.13	18.10	1.83	0.41
2-Methylcitrate synthase *PrpC*	13.91	9.84	3.43	15.46	1.22	0.23
2-Methylcitrate dehydratase FeS dependent *PrpD*	8.04	5.54	4.12	6.15	1.35	1.06
2-Methylaconitate isomerase *prpF acnB*	0.00	2.64	0.25	1.73	0.15	0.00
Methylisocitrate lyase *PrpB*	16.43	20.34	11.63	25.41	0.95	1.92
Butyryl-CoA synthetase	3.22	1.64	0.54	8.24	1.39	0.24
Trans-2-enoyl-CoA reductase (NAD+)	3.48	0.12	0.24	0.00	0.14	0.44
Butyryl-CoA dehydrogenase	0.00	0.61	0.00	0.00	0.02	0.00
Enoyl-CoA hydratase/isomerase	0.00	0.00	0.00	0.00	0.26	0.17
3-Hydroxybutyryl-CoA dehydratase	9.17	1.12	1.16	1.29	0.44	0.32
Polyhydroxyalkanoic acid (PHB) synthase	2.43	1.46	4.82	5.66	3.41	0.15
Polyhydroxyalkanoic acid (PHB) synthase copy 2	0.00	0.42	0.05	0.00	0.02	0.00
3-Hydroxybutyryl-CoA dehydrogenase	3.70	0.70	0.21	0.00	0.00	0.00
3-Hydroxyacyl-CoA dehydrogenase	1.83	0.32	0.08	1.10	0.05	0.00
Acetoacetyl-CoA reductase	0.00	2.38	2.47	0.00	3.02	2.10
Acetoacetyl-CoA reductase copy 2	0.00	1.21	0.55	5.56	1.88	1.59
Acetyl-CoA acetyltransferase	17.61	4.99	1.89	2.61	2.95	0.55
ATP synthase alpha chain	6.78	9.88	5.12	5.32	0.68	0.25
ATP synthase beta chain	0.78	6.74	3.75	10.76	0.61	0.28
ATP synthase gamma chain	8.48	3.08	3.27	2.95	1.77	1.06
ATP synthase delta chain	0.00	3.88	1.94	10.59	0.64	0.24
ATP synthase epsilon chain	0.00	2.13	0.14	6.15	0.10	0.32
ATP synthase F0 sector subunit a	0.00	0.00	0.00	0.00	0.25	0.00
ATP synthase F0 sector subunit b	0.00	2.47	4.81	5.34	1.30	0.27
ATP synthase F0 sector subunit c	0.00	1.25	3.41	0.00	0.75	0.55
Ubiquinol-cytochrome C reductase iron-sulfur subunit	0.00	0.50	0.30	0.00	0.07	0.44
Ubiquinol-cytochrome c reductase cytochrome B subunit	0.00	0.61	0.47	0.00	0.03	0.11
Ubiquinol cytochrome C oxidoreductase cytochrome C1	1.39	0.00	0.78	0.00	0.14	0.00
Cytochrome c oxidase subunit *CcoN*	1.43	1.05	1.39	0.71		
Cytochrome c oxidase subunit *CcoO*	1.70	0.49	1.23	0.00		
Cytochrome c oxidase subunit *CcoQ*	0.00	0.85	0.00	0.00		
Cytochrome c oxidase subunit *CcoP*	0.00	0.16	0.32	1.15		

Cytochrome c oxidase polypeptide I *coxA*	1.87	1.18	1.23	6.34	1.33	0.75
Cytochrome c oxidase polypeptide II *coxB*	0.00	0.19	0.40	3.90	0.82	0.23
Cytochrome c oxidase polypeptide III *coxC*	0.00	0.18	0.87	5.34	0.14	0.00
SoxA	0.00	0.36	0.07	1.22	0.00	0.00
SoxB	0.00	0.00	0.23	0.00	0.00	0.00
SoxC	0.00	0.00	0.09	0.00	0.01	0.00
SoxD	0.00	0.16	0.19	0.00	0.02	0.00
SoxX	0.00	0.42	0.66	0.00	0.00	0.74
SoxY	13.65	0.64	1.01	3.37	0.04	0.57
SoxZ	3.43	0.49	0.57	3.39	0.07	2.59

Supplementary Table 8 | Relative abundance (%) of all *Cycloclasticus* proteins detected by proteomic analysis of three *B. heckerae*. Locus tags correspond to the genome submitted to NCBI in BioProject PRJNA318571. Proteins are ordered by relative abundance. Key proteins involved in the oxidation of short-chain alkanes are highlighted in bold.

Locus Tag	*Description*	*BHP1*	*BHP2*	*BHP3*
A6F70_01965	F0F1 ATP synthase subunit alpha	14.38	10.22	8.46
A6F70_06240	**particulate hydrocarbon monooxygenase subunit A (group Z)**	13.10	11.36	6.01
A6F70_01955	F0F1 ATP synthase subunit beta	13.09	8.40	8.24
A6F70_06740	alkyl hydroperoxide reductase	11.65	7.52	5.77
A6F70_08785	hypothetical transmembrane protein associated with Amt	4.05	10.34	4.78
A6F70_07635	hypothetical protein	0.00	16.63	0.00
A6F70_08210	elongation factor G	6.16	6.27	3.09
A6F70_08025	DNA-binding protein HU	5.75	3.16	3.97
A6F70_01220	50S ribosomal protein L5	5.01	4.27	3.04
A6F70_01800	30S ribosomal protein S6	3.73	3.22	2.37
A6F70_01275	30S ribosomal protein S4	3.73	2.74	1.45
A6F70_06680	thiosulfate oxidation carrier protein SoxY	5.11	1.87	0.18
A6F70_06340	cold-shock protein	0.00	1.23	3.75
A6F70_10670	hypothetical protein	1.65	1.93	0.90
A6F70_03735	hypothetical protein	2.43	1.29	0.48
A6F70_06200	hypothetical protein	1.40	1.32	0.45
A6F70_01655	molecular chaperone	1.43	0.96	0.39
A6F70_00545	succinate--CoA ligase subunit alpha	0.89	0.00	1.52
A6F70_07700	metal-dependent hydrolase	0.00	1.43	0.97
A6F70_06235	**particulate hydrocarbon monooxygenase subunit C (group Z)**	0.00	0.89	1.34
A6F70_01960	F0F1 ATP synthase subunit gamma	0.45	0.42	1.05
A6F70_02540	hypothetical protein	0.61	0.48	0.84
A6F70_09600	aconitate hydratase B	0.52	0.94	0.38
A6F70_03045	nitrate/sulfonate/bicarbonate ABC transporter ATP-binding protein	0.00	0.00	1.68
A6F70_01880	selenium-binding protein	1.59	0.00	0.00
A6F70_01265	30S ribosomal protein S13	0.00	0.34	1.16
A6F70_01980	ATP F0F1 synthase subunit C	0.00	0.00	1.41
A6F70_03040	hypothetical protein	0.00	0.00	1.37
A6F70_01200	50S ribosomal protein L29	0.00	0.00	1.27
A6F70_01180	30S ribosomal protein S19	0.00	0.00	1.22
A6F70_00515	ATP-dependent chaperone ClpB	0.45	0.28	0.38
A6F70_06540	nucleic acid-binding protein	0.00	0.00	0.97

A6F70_00315	malic enzyme	0.60	0.32	0.00
A6F70_07705	phosphomethylpyrimidine synthase ThiC	0.52	0.20	0.18
A6F70_02575	23S rRNA (guanine(2445)-N(2))/(guanine(2069)-N(7))-methyltransferase	0.00	0.00	0.86
A6F70_04610	threonine synthase	0.55	0.00	0.31
A6F70_01820	replicative DNA helicase	0.00	0.00	0.83
A6F70_09085	acetyl-CoA acetyltransferase	0.00	0.00	0.77
A6F70_07170	hypothetical protein	0.00	0.00	0.75
A6F70_00685	haloacid dehalogenase	0.26	0.16	0.33
A6F70_06590	arsenate reductase	0.00	0.00	0.72
A6F70_00100	30S ribosomal protein S20	0.00	0.00	0.62
A6F70_01485	apolipoprotein acyltransferase	0.00	0.00	0.60
A6F70_09720	hypothetical protein	0.00	0.55	0.05
A6F70_03270	hypothetical protein	0.00	0.00	0.56
A6F70_07415	class II glutamine amidotransferase	0.00	0.00	0.53
A6F70_03565	nitrate ABC transporter substrate-binding protein	0.00	0.00	0.52
A6F70_01975	F0F1 ATP synthase subunit B 21507:21986	0.00	0.00	0.52
A6F70_05930	oxaloacetate decarboxylase 49133:50923	0.00	0.00	0.51
A6F70_01245	30S ribosomal protein S5 51500:52009	0.00	0.00	0.49
A6F70_09290	phenylalanine--tRNA ligase subunit alpha	0.00	0.00	0.49
A6F70_01240	50S ribosomal protein L18	0.00	0.00	0.47
A6F70_01235	50S ribosomal protein L6	0.00	0.00	0.47
A6F70_01160	50S ribosomal protein L3	0.00	0.00	0.39
A6F70_06115	hypothetical protein	0.00	0.00	0.39
A6F70_04275	hypothetical protein	0.00	0.00	0.39
A6F70_02550	pressure-regulated protein	0.00	0.00	0.38
A6F70_06260	tRNA-specific adenosine deaminase	0.00	0.00	0.37
A6F70_02760	phosphoglucomutase	0.33	0.00	0.03
A6F70_03065	(Fe-S)-binding protein	0.00	0.00	0.35
A6F70_05035	single-stranded DNA-binding protein	0.00	0.00	0.35
A6F70_09370	30S ribosomal protein S7	0.00	0.00	0.35
A6F70_08030	endopeptidase La	0.24	0.00	0.10
A6F70_07520	pyruvate kinase	0.00	0.00	0.34
A6F70_00460	phosphate transport regulator	0.00	0.34	0.00
A6F70_05495	2-isopropylmalate synthase	0.00	0.24	0.11
A6F70_03060	cyclic pyranopterin phosphate synthase	0.00	0.00	0.33
A6F70_07225	hypothetical protein	0.00	0.00	0.32
A6F70_02230	DNA-(apurinic or apyrimidinic site) lyase	0.00	0.00	0.32

A6F70_06110	hypothetical protein	0.00	0.00	0.31
A6F70_02145	PAS domain-containing sensor histidine kinase	0.00	0.00	0.31
A6F70_00940	lipopolysaccharide kinase	0.00	0.00	0.31
A6F70_00065	N-acetyl-anhydromuranmyl-L-alanine amidase	0.00	0.00	0.30
A6F70_06500	4-amino-4-deoxychorismate lyase	0.00	0.00	0.29
A6F70_04010	co-chaperone GroES	0.00	0.00	0.29
A6F70_07090	hypoxanthine-guanine phosphoribosyltransferase	0.00	0.00	0.29
A6F70_04580	hypothetical protein	0.00	0.00	0.28
A6F70_00700	acyl-CoA dehydrogenase	0.00	0.00	0.28
A6F70_06615	deoxycytidine triphosphate deaminase	0.00	0.00	0.27
A6F70_02600	primosomal protein N'	0.00	0.00	0.26
A6F70_10285	hypothetical protein	0.00	0.00	0.26
A6F70_06180	hypothetical protein	0.00	0.00	0.26
A6F70_08985	**particulate hydrocarbon monooxygenase subunit B (group X-like)**	0.00	0.00	0.25
A6F70_01280	DNA-directed RNA polymerase subunit alpha	0.00	0.00	0.25
A6F70_01435	DNA polymerase III subunit delta	0.00	0.00	0.24
A6F70_05265	chorismate synthase	0.00	0.00	0.23
A6F70_00680	haloacid dehalogenase	0.00	0.00	0.22
A6F70_07350	sigma-54-dependent Fis family transcriptional regulator	0.00	0.22	0.00
A6F70_04420	GTPase Era	0.22	0.00	0.00
A6F70_01270	30S ribosomal protein S11	0.00	0.00	0.21
A6F70_00465	hypothetical protein	0.00	0.00	0.21
A6F70_04780	30S ribosomal protein S2	0.00	0.00	0.21
A6F70_01950	F0F1 ATP synthase subunit epsilon	0.00	0.00	0.20
A6F70_03070	**aldehyde ferredoxin oxidoreductase (AOR)**	0.00	0.07	0.13
A6F70_09560	DNA polymerase III subunit chi	0.00	0.00	0.20
A6F70_00850	50S ribosomal protein L13	0.00	0.00	0.19
A6F70_07420	metal-dependent phosphohydrolase	0.00	0.00	0.19
A6F70_00215	thiol reductase thioredoxin	0.00	0.00	0.19
A6F70_00710	**PQQ-dependent alcohol dehydrogenase (PQQ-ADH)**	0.00	0.00	0.18
A6F70_04475	hypothetical protein	0.00	0.00	0.18
A6F70_04565	SUF system Fe-S cluster assembly regulator	0.00	0.00	0.18
A6F70_05100	aconitate hydratase	0.00	0.00	0.17
A6F70_00490	phosphate ABC transporter substrate-binding protein PstS	0.00	0.00	0.17

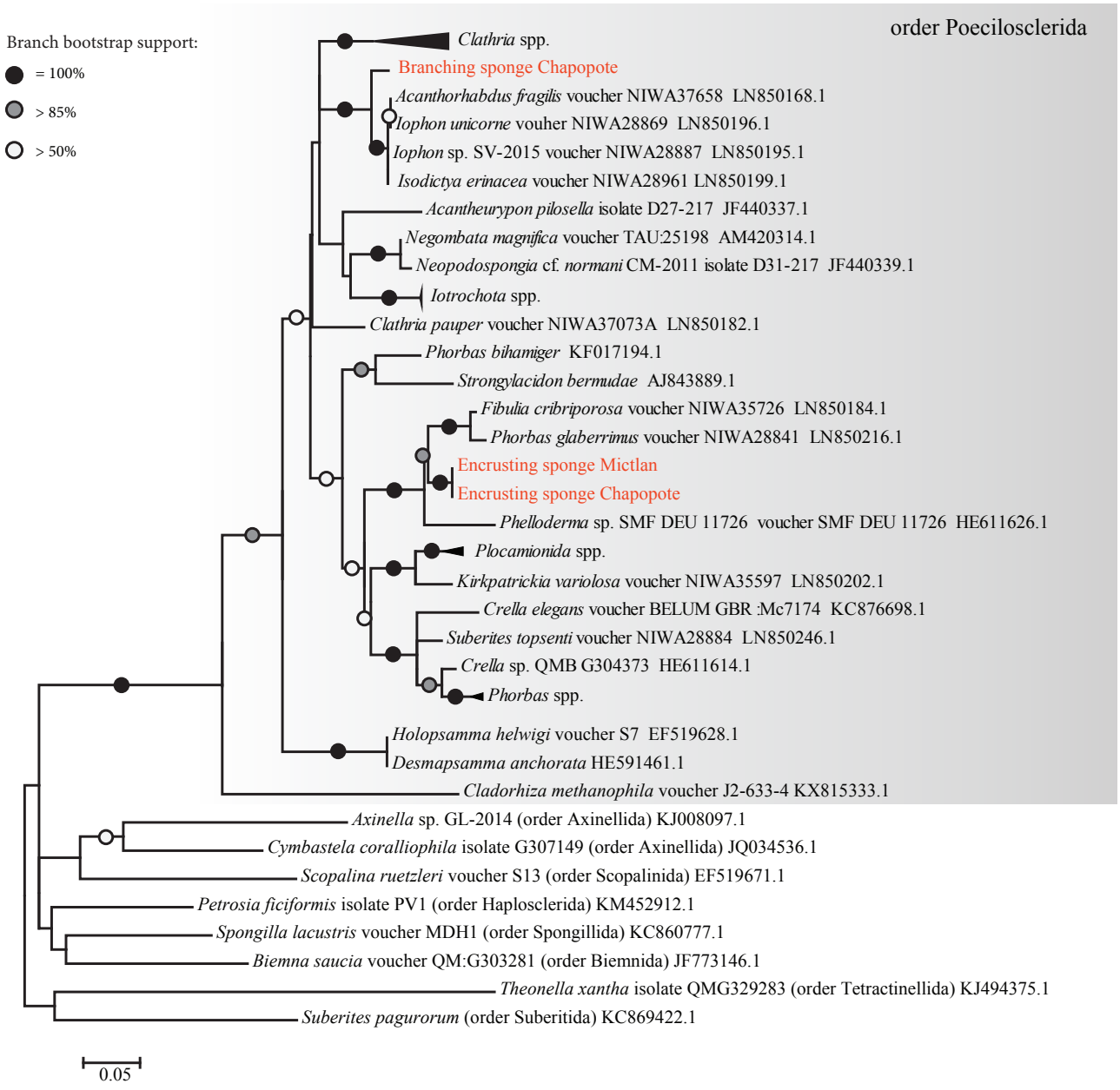
A6F70_01865	hypothetical protein	0.00	0.00	0.16
A6F70_02385	GTPase	0.10	0.06	0.00
A6F70_04895	ketol-acid reductoisomerase	0.00	0.00	0.16
A6F70_05040	cytochrome-c peroxidase	0.00	0.00	0.16
A6F70_09330	transcription termination/antitermination protein NusG	0.00	0.00	0.16
A6F70_02285	3-dehydroquinate synthase	0.00	0.00	0.15
A6F70_08535	cobaltochelatase subunit CobN	0.00	0.00	0.15
A6F70_07395	GTP cyclohydrolase I FolE	0.00	0.00	0.15
A6F70_06730	nitrate reductase catalytic subunit	0.00	0.00	0.15
A6F70_03425	tRNA threonylcarbamoyladenosine biosynthesis protein RimN	0.00	0.00	0.15
A6F70_08795	ATP-dependent protease	0.00	0.00	0.15
A6F70_09075	cobalamin adenosyltransferase	0.00	0.00	0.15
A6F70_07620	hypothetical protein	0.00	0.00	0.14
A6F70_00955	carbamoyltransferase	0.00	0.00	0.14
A6F70_07840	electron transport complex subunit	0.00	0.00	0.14
A6F70_01700	3-phosphoglycerate dehydrogenase	0.00	0.00	0.14
A6F70_05440	RNA helicase	0.00	0.00	0.14
A6F70_00330	beta-ketoacyl-[acyl-carrier-protein] synthase I	0.00	0.00	0.14
A6F70_04995	chemotaxis protein CheD	0.00	0.00	0.13
A6F70_08410	D-alanine--D-alanine ligase	0.00	0.13	0.00
A6F70_00800	Xaa-Pro aminopeptidase	0.00	0.00	0.13
A6F70_06610	hypothetical protein	0.00	0.00	0.13
A6F70_04380	bifunctional folylpolyglutamate synthase/ dihydrofolate synthase	0.00	0.00	0.13
A6F70_05550	pyridine nucleotide-disulfide oxidoreductase	0.00	0.00	0.13
A6F70_06650	adenylate kinase	0.00	0.00	0.13
A6F70_07040	adenylosuccinate synthase	0.00	0.00	0.13
A6F70_04040	hypothetical protein	0.00	0.00	0.13
A6F70_09670	transketolase	0.00	0.00	0.12
A6F70_00765	glycine dehydrogenase	0.00	0.00	0.12
A6F70_08085	ethanolamine ammonia-lyase	0.00	0.00	0.12
A6F70_02615	hypothetical protein	0.00	0.00	0.12
A6F70_04015	molybdopterin-synthase adenylyltransferase MoeB	0.00	0.00	0.12
A6F70_05660	acylneuraminate cytidylyltransferase	0.00	0.00	0.11
A6F70_03505	GTP-binding protein TypA	0.00	0.07	0.05
A6F70_00140	beta-ketoacyl-ACP reductase	0.00	0.00	0.11
A6F70_07200	tRNA (guanosine(37)-N1)-methyltransferase TrmD	0.00	0.00	0.11

A6F70_08705	anthranilate synthase	0.00	0.00	0.11
A6F70_05120	hypothetical protein	0.00	0.00	0.11
A6F70_08850	enoyl-CoA hydratase	0.00	0.00	0.11
A6F70_04295	ribonuclease E	0.00	0.00	0.11
A6F70_00975	16S rRNA (adenine(1518)-N(6)/adenine(1519)-N(6))-dimethyltransferase	0.00	0.00	0.11
A6F70_00230	ATPase	0.00	0.00	0.10
A6F70_07245	hypothetical protein	0.00	0.00	0.10
A6F70_07525	segregation and condensation protein A	0.00	0.00	0.10
A6F70_05900	DNA-formamidopyrimidine glycosylase	0.00	0.00	0.10
A6F70_01780	short chain dehydrogenase	0.00	0.00	0.10
A6F70_03195	ubiquinone biosynthesis regulatory protein kinase UbiB	0.00	0.00	0.10
A6F70_09035	pantoate--beta-alanine ligase	0.00	0.00	0.10
A6F70_05455	Vi polysaccharide biosynthesis protein VipA/TviB	0.00	0.10	0.00
A6F70_04885	acetolactate synthase 3 catalytic subunit	0.00	0.00	0.10
A6F70_04775	elongation factor Ts	0.00	0.00	0.09
A6F70_05985	hypothetical protein	0.00	0.00	0.09
A6F70_06210	hypothetical protein	0.00	0.00	0.09
A6F70_03590	cation acetate symporter	0.00	0.00	0.09
A6F70_00530	RNA pseudouridine synthase	0.00	0.00	0.08
A6F70_01035	lytic murein transglycosylase B	0.00	0.00	0.08
A6F70_02025	oxidoreductase	0.00	0.00	0.08
A6F70_07315	DNA recombination/repair protein RecA	0.00	0.00	0.08
A6F70_02255	pilus assembly protein PilM	0.00	0.00	0.08
A6F70_07800	hypothetical protein	0.00	0.00	0.08
A6F70_04055	glycine oxidase	0.00	0.00	0.08
A6F70_07050	methylmalonyl-CoA mutase	0.00	0.00	0.08
A6F70_06165	response regulator receiver protein	0.00	0.00	0.08
A6F70_09550	LPS export ABC transporter permease LptF	0.00	0.00	0.07
A6F70_08830	peroxidase	0.00	0.00	0.07
A6F70_03445	DNA processing protein DprA	0.00	0.00	0.07
A6F70_04175	cell division protein FtsK	0.00	0.00	0.07
A6F70_02450	hypothetical protein	0.00	0.00	0.07
A6F70_09750	hypothetical protein	0.00	0.00	0.07
A6F70_01095	glutamyl-tRNA reductase	0.00	0.00	0.07
A6F70_03960	translocation protein TolB	0.00	0.00	0.06
A6F70_08265	type I secretion protein TolC	0.00	0.00	0.06
A6F70_07490	GTPase HflX	0.00	0.00	0.06

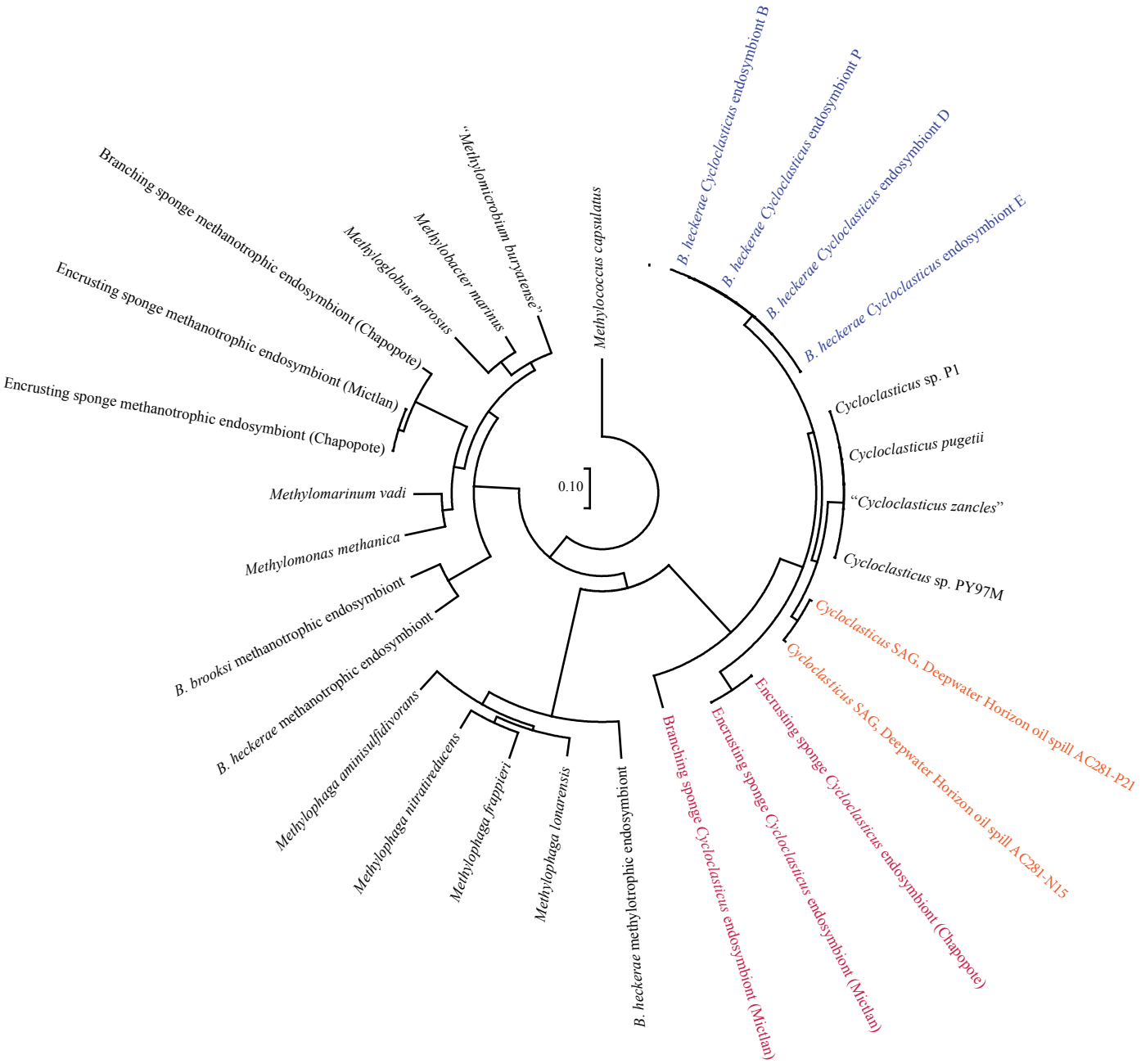
A6F70_04820	DNA helicase	0.00	0.00	0.06
A6F70_00050	dimethylmenaquinone methyltransferase	0.00	0.00	0.06
A6F70_09985	integrase	0.00	0.00	0.06
A6F70_04705	DNA polymerase III subunit alpha	0.00	0.03	0.02
A6F70_03730	hypothetical protein	0.00	0.00	0.06
A6F70_02065	membrane protein insertase YidC	0.00	0.00	0.05
A6F70_09830	hypothetical protein	0.00	0.00	0.05
A6F70_07610	glutamine--tRNA ligase	0.00	0.00	0.05
A6F70_02445	hypothetical protein	0.00	0.00	0.05
A6F70_02785	dihydrolipoamide dehydrogenase	0.00	0.00	0.05
A6F70_06700	thiosulfohydrolase SoxB	0.00	0.00	0.05
A6F70_00930	lipid ABC transporter permease/ATP-binding protein	0.00	0.00	0.05
A6F70_01940	glutamine--fructose-6-phosphate aminotransferase	0.00	0.00	0.05
A6F70_08255	signal protein	0.00	0.00	0.04
A6F70_04865	ResB-like family protein	0.00	0.00	0.04
A6F70_02865	sodium-translocating pyrophosphatase	0.00	0.00	0.04
A6F70_09240	hypothetical protein	0.00	0.00	0.04
A6F70_08000	peptide ABC transporter substrate-binding protein	0.00	0.00	0.04
A6F70_01480	hypothetical protein	0.00	0.00	0.02

Supplementary Table 9 | Relative abundance of proteins (%) from *B. heckerae* hosts, symbiotic *Cycloclasticus* and other symbionts.

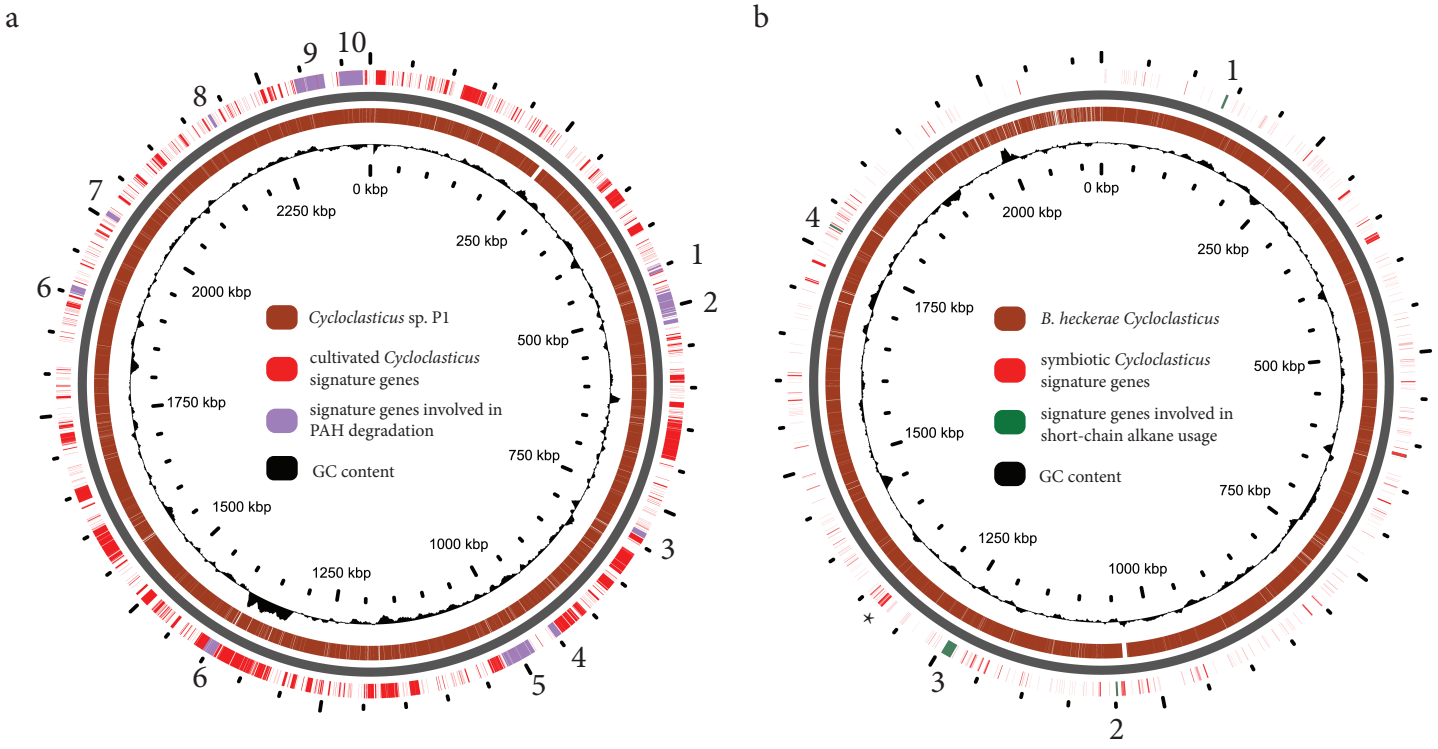
	BHP1	*BHP2*	*BHP3*
Cycloclasticus	0.23	0.34	0.42
B. heckerae	88.73	84.76	90.94
Other symbionts	11.04	14.90	8.64



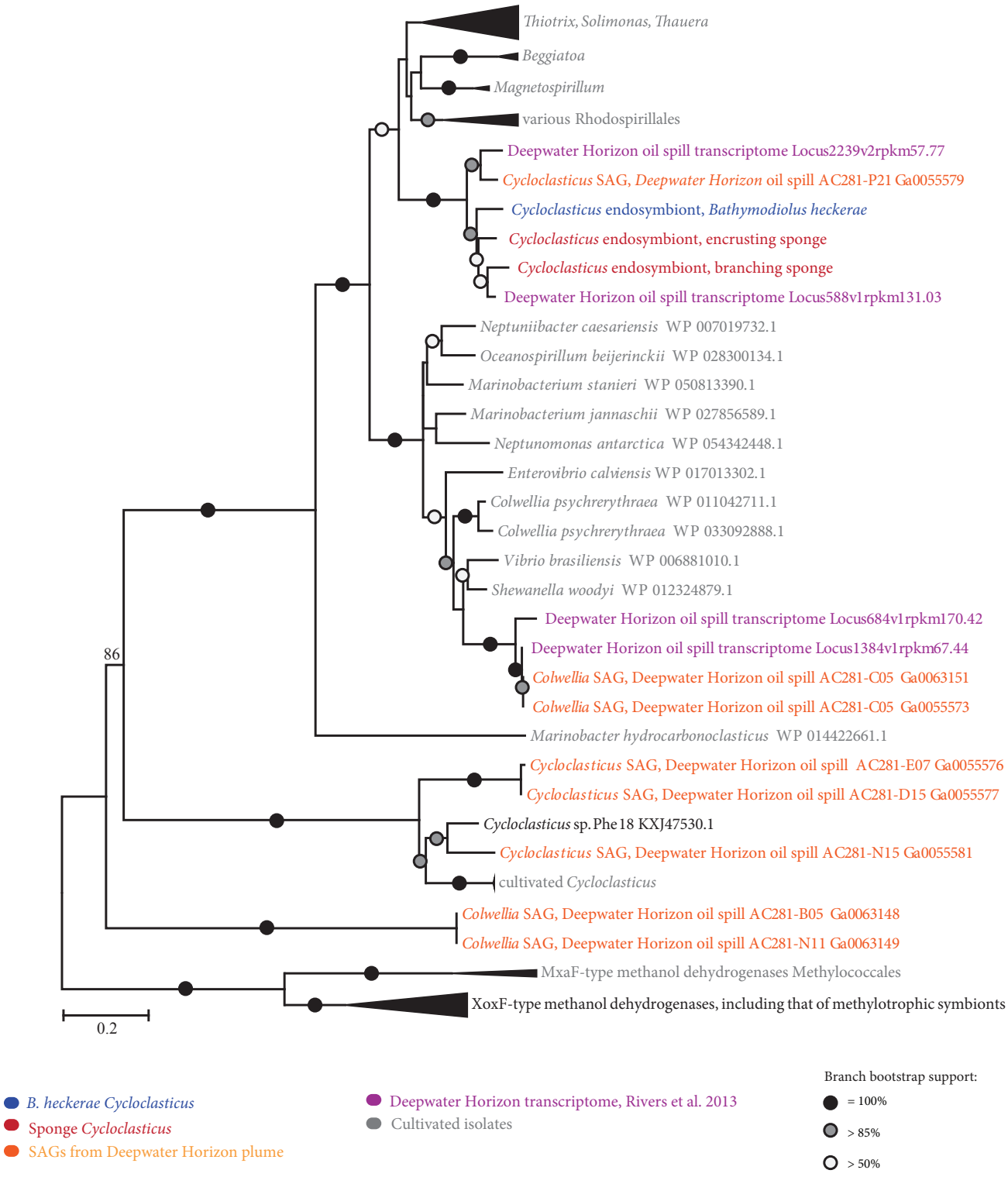
Supplementary Figure 1 | Sponge phylogeny based on mitochondrial COI gene sequences. The evolutionary history was inferred using the Maximum Likelihood method based on the Hasegawa-Kishino-Yano model. The tree is drawn to scale, with branch lengths measured as the number of substitutions per site. The analysis included 44 nucleotide sequences. There were a total of 549 positions in the final dataset. A comparison of the COI sequences from the Campeche Knoll sponges to those of the poecilosclerid sponge *Myxilla methanophila*, which also harbors *Cycloclasticus*, was not possible, because no COI sequences are available for *M. methanophila.*



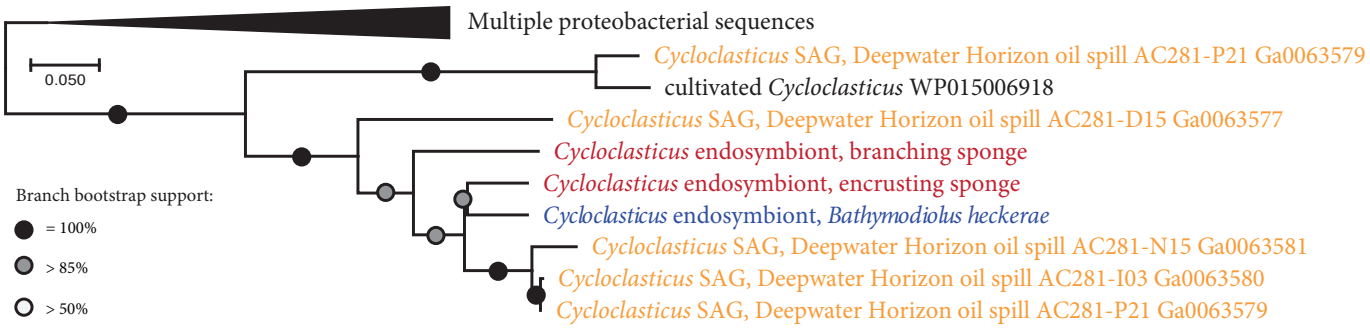
Supplementary Figure 2 | Phylogenomic tree showing the relationships between cultivated, free-living and symbiotic *Cycloclasticus*, as well as methylotrophs in the *Piscirickettsiaceae* and *Methylococcaceae* families of the Gammaproteobacteria. Eleven single-copy markers as defined in the AMPHORA2 core bacterial phylogenetic marker database were used in the analysis. All node support values in the tree were ≥ 89 and therefore not shown. Symbiotic *Cycloclasticus* from *B. heckerae* and sponge hosts are highlighted with blue and red colors, respectively. Genomic library name is mentioned in parentheses to distinguish between *Cycloclasticus* symbionts from different host individuals.



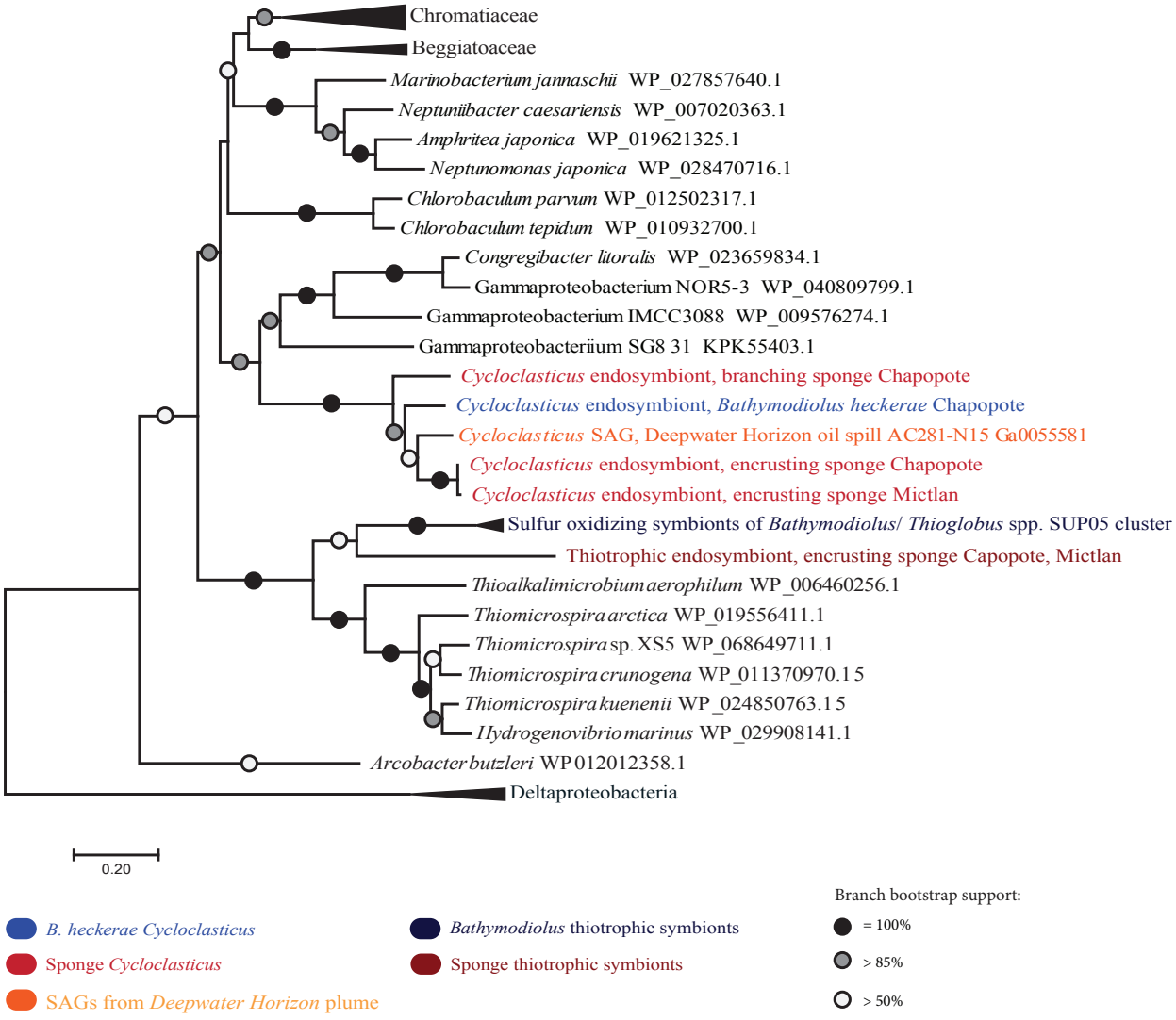
Supplementary Figure 3 | Signature genes in cultivated PAH degrading *Cycloclasticus* (a) and symbiotic *Cycloclasticus* (b). Signature genes for a group of organisms are those that are present within this group, but lack homologs in other closely related organisms. The cutoff used to determine homology was 40% at amino acid level. The two groups compared are: Symbiotic *Cycloclasticus* from *B. heckerae* and two sponges vs. cultivated *C. pugetii* P-1, *C. zancles* and *Cycloclasticus* PY97M. In panel a, signature genes of cultivated PAH degrading *Cycloclasticus* are mapped on the genome of *C. pugetii* P-1. In panel b, signature genes of symbiotic *Cycloclasticus* are mapped on the genome of *B. heckerae Cycloclasticus* (library P). Annotations are: a) 1 - 4-hydroxybenzoate transporter, 2 - benzoate, cresol, catechol usage proteins, 3 - ring-hydroxylating dioxygenase naphthalene 1,2-dioxygenase, 4 - ring-hydroxylating dioxygenase naphthalene 1,2-dioxygenase and 2,3-dihydroxybiphenyl 1,2 dioxygenase, 5 - multiple genes involved in PAH degradation, 6 - ring hydroxylating dioxygenase, 7 - biphenyl monooxygenase and related proteins, 8 - multiple genes involved in PAH degradation and alkane-1 monooxygenase, 9 - ring-hydroxylating dioxygenase naphthalene 1,2-dioxygenase, 9,10 - multiple genes involved in PAH degradation. In addition to PAH degradation genes, cultivated *Cycloclasticus* signature genes included genes encoding hydrogenases and related proteins, urea transport and urease proteins, ferric siderophore transport system proteins, heavy metal transport systems, flagellar biosynthesis proteins, etc. b) 1 - PQQ-dependent alcohol dehydrogenase (related to ethanol and propanol oxidizing dehydrogenase), 2 -cytochrome c550 and other proteins associated with PQQ-dependent alcohol dehydrogenase, 3 - particulate hydrocarbon monooxygenase (3 subunits, group Z), 4- 3-hydroxybutyryl-CoA dehydratase and dehydrogenase (butanoate metabolism). * - sulfur oxidation proteins (sox operon, see Supplement for details).



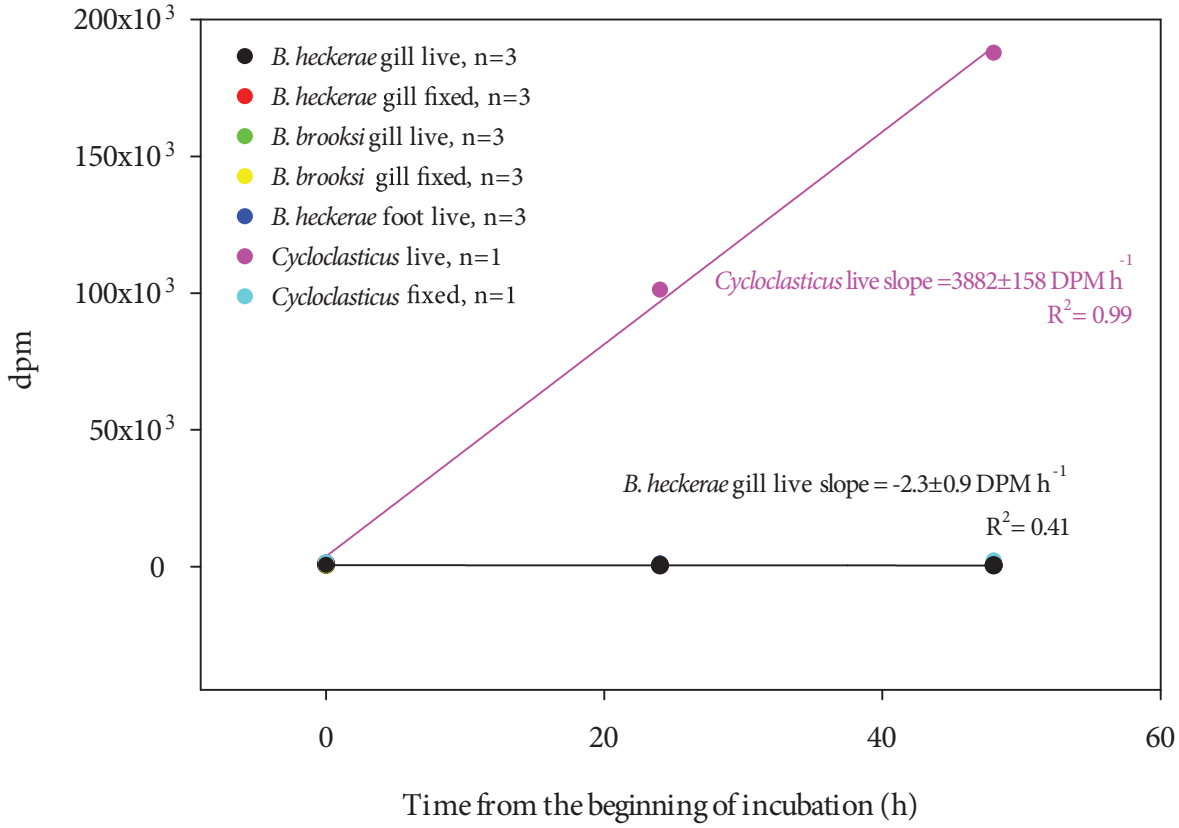
Supplementary Figure 4 | Phylogeny of PQQ-dependent alcohol dehydrogenase protein sequences. The tree is drawn to scale, with branch lengths representing the number of substitutions per site. The percentage of trees in which the associated taxa clustered together is shown next to the branches (only values above 50 %). The analysis included 69 amino acid sequences. There were a total of 413 positions in the final dataset.



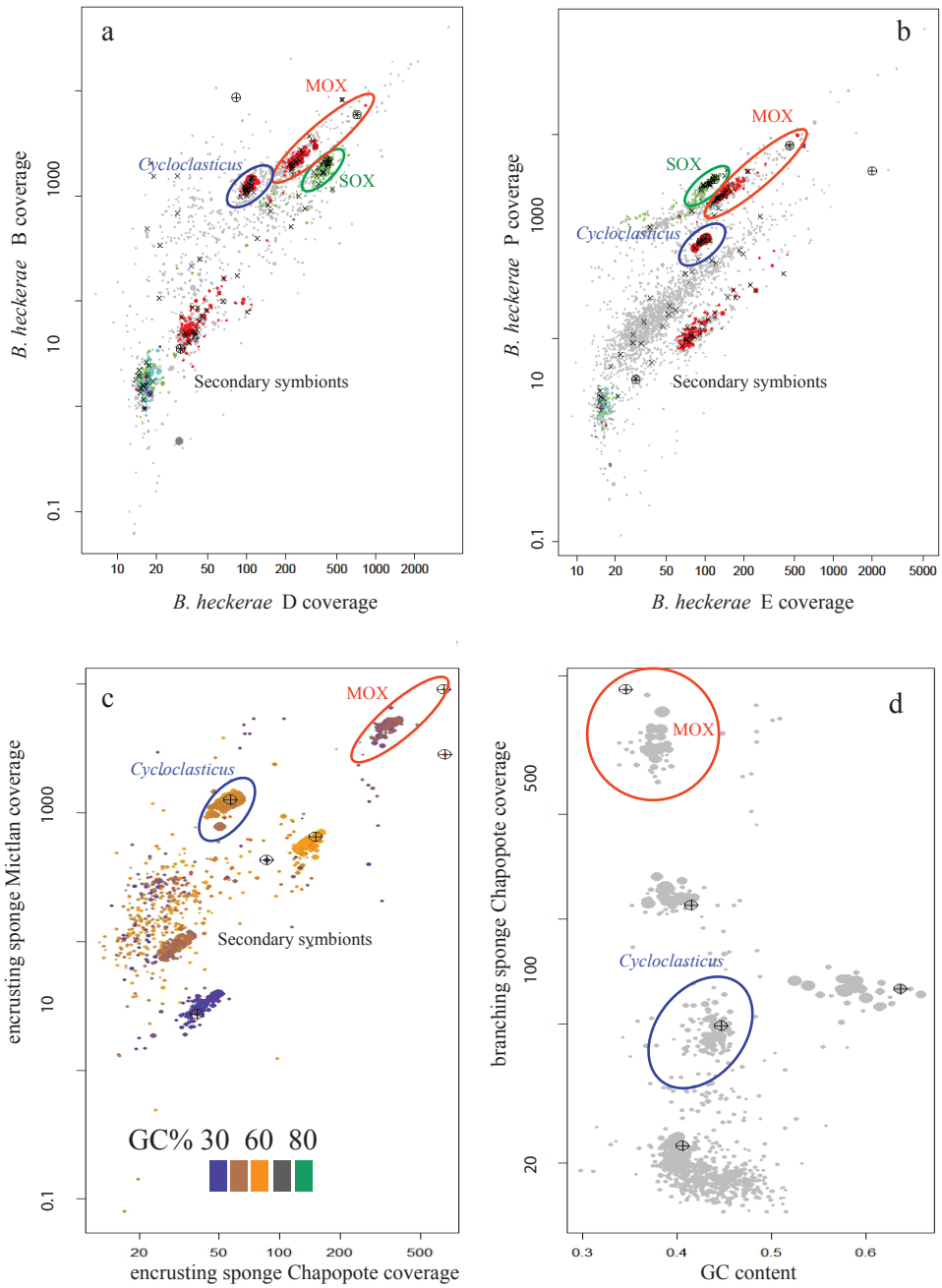
Supplementary Figure 5 | Phylogeny of tungsten-containing aldehyde ferredoxin oxidoreductase (AOR) protein sequences. The tree is drawn to scale, with branch lengths representing the number of substitutions per site. The percentage of trees in which the associated taxa clustered together is shown next to the branches (only values above 50 %). The analysis included 21 amino acid sequences. There were a total of 613 positions in the final dataset.



Supplementary Figure 6 | Phylogeny of thiosulfohydrolase SoxB protein sequences. The tree is drawn to scale, with branch lengths representing the number of substitutions per site. The percentage of trees in which the associated taxa clustered together is shown next to the branches (only values above 50 %). The analysis included 47 amino acid sequences. There were a total of 555 positions in the final dataset. SOX = sulfur-oxidizing bacteria.



Supplementary Figure 7 | Isotopically labeled naphthalene was mineralized by *Cycloclasticus pugetii* ATCC 51542 but not by symbiotic *B. heckerae Cycloclasticus*. This plot shows disintegrations per minute (DPM) counts for CO_2 in the headspace after addition of ^{14}C labeled naphthalene (400000 DPM). Naphthalene was mineralized by live *Cycloclasticus* strain ATCC 51542 (magenta circles), but mineralization was not detected in all other samples (all other colors, circles overlap so that only black circles are visible). The error bars are smaller than the symbol size. R^2 values are inferred from linear regression analysis. n represents the number of biological replicates.



Supplementary Figure 8 | Genome-binning-tools generated plots demonstrate binning methods used in this study. Panels a, b show the differential coverage plots for *B. heckerae* metagenomes, with the color produced by the phylotyping approach (red: *Methylococcaceae*-related, green: thiotroph-related). Only contigs larger than 3 kb are shown. Panel c shows the differential coverage plots for the encrusting sponge genomes, with the color representing % GC content. Only contigs larger than 5 kb are shown. Panel d shows read coverage as a function of GC content, which was used to bin the branching sponge metagenome. Only contigs larger than 8 kb are shown. Crossed circles mark contigs that contain 16S rRNA genes, x symbols mark the tRNA containing contigs. MOX and SOX are methane- and sulfur-oxidizing symbionts, respectively.