

Brief Report

A hydrogen-oxidizing bacterium enriched from the open ocean resembling a symbiont

Katharina Sass,^{1,2} Simon Güllert,²
Wolfgang R. Streit² and Mirjam Perner^{1,†*}

¹Molecular Biology of Microbial Consortia, Universität Hamburg, Hamburg, Germany.

²Microbiology and Biotechnology, Universität Hamburg, Hamburg, Germany.

Summary

A new autotrophic hydrogen-oxidizing Chromatiaceae bacterium, namely bacterium CTD079, was enriched from a water column sample at 1500 m water depth in the southern Pacific Ocean. Based on the phylogeny of 16S rRNA genes, it was closely related to a scallop endosymbiont (99.2% DNA sequence identity) whose host so far is only known to colonize hydrothermal vents along the Indian ridge. The average nucleotide identity between the genomes of CTD079 and the scallop endosymbiont was 91%. The observed differences likely reflect adaptations to their specific habitats. For example, CTD079 encodes additional enzymes like the formate dehydrogenase increasing the organism's spectrum of energy generation pathways. Other additional physiological features of CTD079 included the increase of viral defence strategies, secretion systems and specific transporters for essential elements. These important genome characteristics suggest an adaptation to life in the open ocean.

Introduction

Microorganisms living in symbiosis with invertebrate hosts such as mussels, snails or tubeworms are particularly widespread at hydrothermal vent habitats (Cavanaugh *et al.*, 1981; 1987; Krueger *et al.*, 1996). These endosymbionts use a broad range of inorganic compounds, such as H₂, H₂S and so on, to gain energy, which can then be

used to fuel autotrophic CO₂ fixation (Takai *et al.*, 2006; Nakagawa and Takai, 2008; Petersen *et al.*, 2011). The synthesized organic compounds are made available to the host organism (Takai *et al.*, 2006; Jorgensen and Boetius, 2007; Nakagawa and Takai, 2008). Depending on the symbiont's metabolism, the host also benefits from the symbiont's ability to detoxify the environment, as is the case in sulfidic habitats (Kusel *et al.*, 2006). In return, the host provides shelter and access to essential chemical compounds to the symbiotic prokaryotes (Brissac *et al.*, 2009; Tian *et al.*, 2017).

Phylogenetic trees show that many lineages of chemosynthetic symbionts are interspersed with free-living bacteria, indicating that symbioses evolved multiple times (Dubilier *et al.*, 2008; Petersen *et al.*, 2012). Besides the symbiotic stage, a few symbionts also have a free-living stage; these organisms are coined facultative symbionts (Gros *et al.*, 2003; Aida *et al.*, 2008; Harmer *et al.*, 2008). Facultative endosymbionts are not exclusively dependent on their host, since they are taken up from the environment with each host generation and can live some time without the host (Nussbaumer *et al.*, 2006).

Especially obligate endosymbionts continuously colonizing a sheltered, stable host environment can afford to lose more genes than microorganisms with a free-living lifestyle exposed to the irregular influences of the ambient surrounding (McCutcheon and Moran, 2011). For free-living marine bacteria, the acquisition of ecologically relevant genes is well known. Recombination with distant relatives is a crucial mechanism for the adaptation to different environments and the emergence of diverse marine bacteria. In similar genomic core regions of various marine bacterial lineages, very contrastive recombination patterns have been detected, presumably due to recombination with distantly related organisms (Sun and Luo, 2018). These identified gene sections encode for functions that are beneficial for the bacterium's survival in a given habitat (e.g., antibiotic synthesis or cold adaptation) (Sun and Luo, 2018). Another important mechanism for horizontal gene transfer in the marine environment that can affect bacterial community diversity is bacteriophage-mediated horizontal transduction (Jiang and Paul, 1998).

*For correspondence. E-mail mperner@geomar.de; Tel. +49 431 600 2837; Fax +49 431 600 2941. [†]Present address: Geomicrobiology, GEOMAR Helmholtz Centre for Ocean Research Kiel, Kiel, Germany.

In this study, we present the complete genome of a newly cultured Chromatiaceae enriched from the water column of the southern Pacific Ocean. We compared its genome with that of the closest relative available in the database, namely a scaly snail endosymbiont found near hydrothermal vents along the Indian Ridge (Goffredi *et al.*, 2004). Special focus of this study lays on gene acquisition through horizontal gene transfer as an adaptation to conditions in the water column. We analysed for which cellular functions the newly acquired gene segments encode and to what extent the organism can benefit from the acquired phenotype. We compared the potential traits of our bacterium enriched from the water column with those from its closest known relative – an endosymbiont – and discuss differences in relation to metabolism, genome evolution and ecological functions.

Results and discussion

We identified two types of bacteria by 16S rRNA gene sequencing (Supporting Information Figs S1 and S2) and fluorescence *in situ* hybridization (Supporting Information Fig. S3) in a culture enriched with water from the open ocean (79 CTD V14_06; water depth 1500 m; $-35^{\circ}15.5010'S$ $178^{\circ}46.6270'W$). No hydrothermal influence was detected in the water sample according to hydrogen and methane measurements (Dr. Alexander Diehl, pers. comm., Universität Bremen). The enrichment cultivation was dominated by an uncultivated Gammaproteobacterium (90% of all 4',6-diamidino-2-phenylindole (DAPI) stained cells), named bacterium CTD079. The remaining 10% of all DAPI stained cells in the culture were similar to *Alteromonas* sp. based on 16S rRNA sequence analysis (99% DNA identity). Although we managed to isolate the *Alteromonas* species (by dilution series and with the PALM MicroTweezer), none of the tested isolation methods resulted in a pure culture of bacterium CTD079 (for details, see experimental procedure in the Supporting Information).

General genomic characteristics

For sequencing, the total DNA extracted from the enrichment culture was used, followed by hybrid genome assembly based reconstruction of bacterium CTD079's genome. The complete genome of bacterium CTD079 consisted of one circular chromosome: it is 2,745,378 bp in length with an average G + C content of 64.0 mol% (Fig. 1 and Table 1). In total, 2577 protein coding sequences (CDS), 44 tRNAs and 3 rRNAs were predicted, resulting in a coding density of 90.5%. The genome sequence and annotation were deposited under the NCBI accession number CP037918.

Examination of the 16S rRNA gene of bacterium CTD079 against the SILVA 132 SSU database using the

SINA aligner (Pruesse *et al.*, 2012) indicated an affiliation with the family Chromatiaceae of the Gammaproteobacteria and a very high sequence identity of 99.2% to an endosymbiont of scaly snails (Nakagawa *et al.*, 2014) (Supporting Information Fig. S2). Bacteria of the family Chromatiaceae are mainly found in sediments of marine shorelines, coastal waters, lagoons, fjords (Imhoff, 2014) or as symbionts of marine sponges (Tian *et al.*, 2014) and are less common in the marine water column. Based on phylogenetic analysis of 16S rRNA genes, bacterium CTD079 clustered with other endosymbionts and uncultured bacteria from a microbial mat at Lo'ihi Seamount, Hawaii (Supporting Information Fig. S4). These findings prompted us to compare the complete genome of the new bacterium CTD079 with that of the scaly snail endosymbiont. We found a high level of synteny between the two bacterial genomes. A syntenic dotplot showed an x-like alignment pattern which is typical for genomic inversion events (Supporting Information Fig. S5). The comparison of the genome sequences of both organisms showed an average nucleotide identity (ANI) of 91.0% and an average amino acid identity (AAI) of 92.6% indicating that they belong to two different species, but to the same genus (thresholds suggested for species definition are 95% [ANI] and > 95% [AAI]) (Goris *et al.*, 2007). Based on these results, we set out to study differences and similarities between the bacterium CTD079 enriched from the open ocean and the hydrothermal vent snail endosymbiont.

First, we analysed the two genomes using IslandViewer 4 (Bertelli *et al.*, 2017) to identify genomic islands in both genomes. In total, 11 islands were found in bacterium CTD079 (Fig. 1) and 17 in the snail's endosymbiont (Supporting Information Fig. S6). Genomic islands are regions of the genome indicating horizontal gene transfer, that is, uptake of genetic material from the environment/other organisms. The annotation by IslandViewer is based on the combination of the programs Island-Path-DIMOB and Islander. The former seeks dinucleotide biases in a region of at least eight consecutive genes, including a mobility gene (Hsiao *et al.*, 2005), the latter looks for regions flanked by a tRNA gene or a tRNA gene fragment which contains an integrase gene (Mantri and Williams, 2004). Thus, genomic regions with an abnormal sequence composition suggest that these may have been acquired by horizontal gene transfer (for detailed discussion, see below: distinctions from the symbiont).

Similarities to the symbiont

The genomic information from bacterium CTD079 suggests a hydrogen/sulfur-oxidizing chemolithoautotrophic lifestyle (Fig. 2, Supporting Information Table S1) as can be typically found in hydrothermally influenced environments. This is in line with what has been predicted for

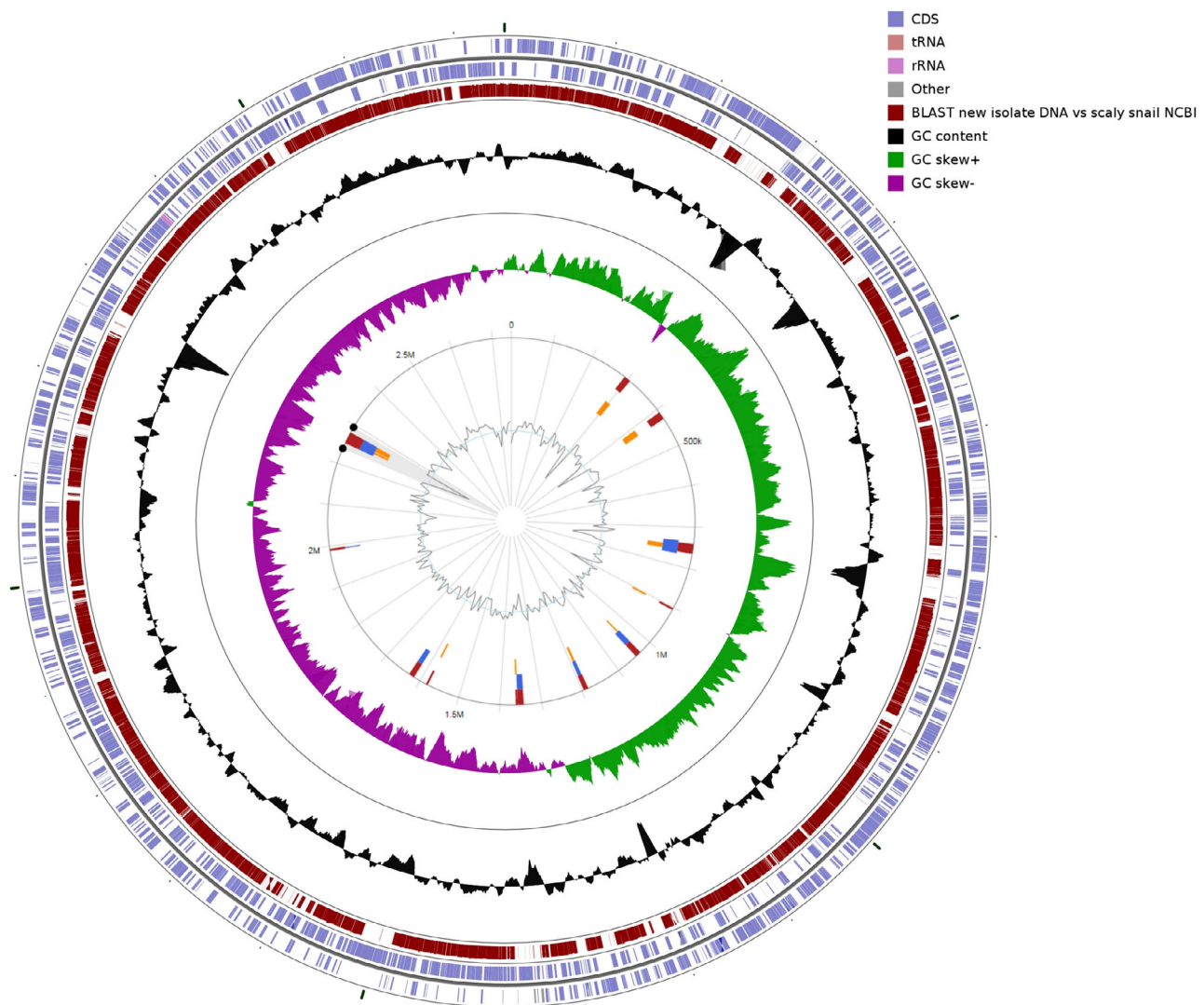


Fig 1. Circular genome representation of bacterium CTD079 using the GCView Comparison Tool (Stothard *et al.*, 2019). The two blue outer circles depict predicted open reading frames on the plus and minus strand. The third red circle illustrates homologous regions between the genome of CTD079 and the genome of the scaly snail endosymbiont. The next tracks indicate the change in GC (guanine-cytosine) content and GC skew. The inner circle shows genomic islands in the genome of bacterium CTD079 as predicted by the IslandViewer 4 software (Bertelli *et al.*, 2017). The differently coloured blocks indicate the island prediction method (Red = predicted by at least one method, blue = IslandPath-DIMOB, orange = SIGI-HMM).

its closest relative the endosymbiont of a scaly snail (Nakagawa *et al.*, 2014). Nakagawa and colleagues already reported in detail on the genes from the symbiont potentially involved in sulfur cycling and hydrogen oxidation (Nakagawa *et al.*, 2014). In short, the genome of the endosymbiont has the *sox* gene cluster with the exception of *soxCD*. As commonly found for chemolithoautotrophic Gammaproteobacteria, the genes *sorAB* are also missing (Kappler *et al.*, 2000). Additionally, in the endosymbiont's and CTD079's (Supporting Information Table S1) genome hydrogenases of the groups 1 and 2 are present.

As hydrogen is one of the most abundant elements and plays a key role in most bacterial metabolic pathways (Dilling and Cypionka, 1990; Bothe *et al.*, 2010; Tang

et al., 2011), we added hydrogen as the sole electron donor to the medium. Furthermore, it is an ideal electron transporter and the ability to use hydrogen as an energy source is widespread among chemolithoautotrophic organisms. Our enrichment culture containing bacterium CTD079 and *Alteromonas* sp. consumed hydrogen ($0.6 \text{ fM cell}^{-1} \text{ h}^{-1}$) under autotrophic growth conditions (Fig. 3). The amount of consumed hydrogen is in the lower range of what has been reported for other hydrogen-oxidizing Gammaproteobacteria in *Hydrogenovibrio* (e.g., *Hydrogenovibrio* sp. MA-3 $0.63 \text{ fM cell}^{-1} \text{ h}^{-1}$, *Hydrogenovibrio* sp. SP41 $1.47 \text{ fM cell}^{-1} \text{ h}^{-1}$) (Hansen and Perner, 2015; 2016). Hydrogenase transcripts of CTD079 and *Alteromonas* species were also found in the

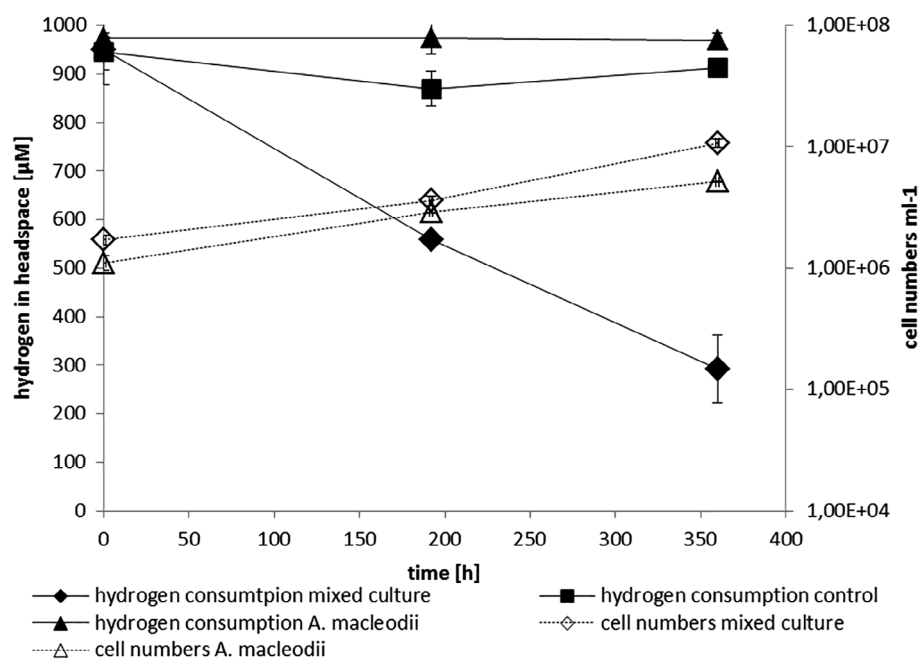


Fig 3. *In vivo* hydrogen consumption and growth of the mixed culture containing bacterium CTD079 and *Alteromonas* sp. and of the *Alteromonas* sp. isolate. Hydrogen consumption of the enrichment is shown with diamonds, of *Alteromonas* sp. isolate with triangles and controls with squares. Growth of the enrichment is shown with open diamonds and broken lines with dots. Growth of *Alteromonas* sp. is shown with open triangles and a dotted line. Hydrogen concentration was measured using gas chromatography and cell numbers were determined using an Olympus BX41 (Olympus, K. K., Tokyo, Japan) microscope as described before (Hansen and Perner, 2015).

modABCD operon (Scott and Amy, 1989). When comparing the molybdenum operon found in the bacterium CTD079's genome against the sequences deposited in the NCBI non-redundant database, the highest similarity was found to an unspecified Gammaproteobacterium (Accession PKM03231.1, 73% AA identity). The next described relative according to blastx searches was *Methylobacter* sp. (PPD34451.1, 56% AA identity), a marine methanotrophic gammaproteobacterium (Boden *et al.*, 2011).

Furthermore, we identified two additional two-component signal transduction systems, the EnvZ/OmpR and the RegB/RegA system in CTD079 (Supporting Information Table S1). These systems are used by bacteria to respond to changes in environmental conditions (Stock *et al.*, 2000). In *Escherichia coli*, for example, the EnvZ/OmpR system is described to respond to variations in environmental osmolarity (Forst and Roberts, 1994; Pratt and Silhavy, 1995; Egger *et al.*, 1997). For the RegB/RegA two-component system, it is known that it constitutes a conserved regulatory system for redox control of many energy-producing and energy-consuming processes (Madigan and Gest, 1979; Joshi and Tabita, 1996; Qian and Tabita, 1996). The use of the RegB/RegA two-component system initiates changes of different metabolic pathways, for example, carbon fixation, nitrogen fixation, hydrogen oxidation or denitrification. Under the provided cultivation conditions, CTD079 appears to be associated with *Alteromonas* sp. Thus, CTD079 may use organic matter released from *Alteromonas* sp. In the natural environment from which CTD079 was enriched, the deep ambient water, it might attach to sinking particulate organic matter (POM). This macro organic matter drives the lower aquatic food web by providing energy in

form of carbohydrates, sugars and other polymers that can be degraded by, for example, Gammaproteobacteria (Cole *et al.*, 2011; Taipale *et al.*, 2016; Gomez-Consmau *et al.*, 2019). Depending on the composition of POM, the corresponding metabolic pathway can then be activated using the RegB/RegA system. Those bacterial relationships with sedimenting particles have been described many times (Iturriaga, 1979; Ducklow *et al.*, 1982; Karl and Knauer, 1984). In general, free-living bacteria are often exposed to environmental changes, for example, temperature, pH, nutrients and so on, while obligate symbionts encounter comparatively stable surroundings. Thus, both the EnvZ/OmpR and the RegB/RegA system may provide a benefit for organisms living outside of a host – temporarily, as is the case for facultative symbionts, or permanently as is the case for free-living organisms.

Physiology. Genomic sequence data suggested that bacterium CTD079 appears to be able to use two additional ways to generate energy: carbon monoxide (CO) and formate metabolism (Supporting Information Table S1). In the CO metabolism, the enzyme CO dehydrogenase (CODH) mediates the oxidation of CO to CO₂ (Hausinger, 1993). The formate dehydrogenase (FDH) catalyses the oxidation of formate to CO₂ (Ferry, 1990). Both, the CODH of aerobic organisms as well as the FDH contain molybdenum in the active centre (Ragsdale, 2004), which may explain CTD079's acquisition of a molybdenum transporter (see discussion above).

We found the complete CODH operon (aerobic CODH) located on one of the predicted genomic islands, encoding the genes for the three subunits *coxL* (molybdenum protein), *coxM* (flavoprotein) and *coxS* (iron-sulfur protein) (Hugendieck and

Meyer, 1992; Gremer *et al.*, 2000). It is known that some aerobic carboxydrotrophs like *Oligotropha carboxidovorans* (Meyer and Schlegel, 1983) can use the energy from CO oxidation to fix CO₂ in the Calvin-Benson-Bassham (CBB) cycle (Ragsdale, 2004). Since all genes for the CBB fixation pathway are present on our newly sequenced genome, it is likely that bacterium CTD079 also uses CO oxidation to generate energy – possibly for CO₂ fixation. The comparison of the sequences of the CODH operon with sequences from the NCBI non-redundant database gave the highest similarity to genes from an uncharacterized Gammaproteobacterium (Accession PKM13574.1, 81% AA identity). Phylogenetic analysis based on the large CODH subunit showed that bacterium CTD079 clustered within a group of gammaproteobacterial free-living marine prokaryotes and a bacterium associated with a sand snail (*Cocleimona flava*) (Supporting Information Fig. S8). Since the sequences of free-living bacteria are basal to the monophyletic group into which CTD079 clusters, it may be that the CODH gene cluster originated from a free-living ancestor and the scaly snail symbiont lost it. Alternatively, the CODH may not have been in the common ancestor and CTD079 may have acquired it later.

The genes encoding all subunits of the bacterial FDH are most similar to the ones of the gammaproteobacterial symbiont of *Stewartia floridana* (Accession RLW52734.1, 82% AA identity). The FDH potentially enables our cultured organism to utilize an additional energy source.

Another difference between the expected physiology of CTD079 and the symbiont is the presence of a gene segment encoding for the nitrous-oxide reductase on the genome of CTD079 (Supporting Information Table S1). This enzyme catalyses the conversion of nitrous oxide to dinitrogen, which is the last step of the denitrification process and is an autonomous form of respiration (Zumft, 2005). The ability to use a dissimilatory denitrification pathway provides a new complementary source of energy for CTD079. Although the *nosZ* gene was described as a pseudogene (Nakagawa *et al.*, 2014), we identified a gene cluster for the nitrous-oxide reductase in the genome of CTD079. This part of the genome shares similarity to genes from *Thiolapillus brandeum* (Accession WP_041065396.1, 89% AA identity), a sulfur oxidizing Gammaproteobacterium, which is phylogenetically placed within a group of symbionts (Nunoura *et al.*, 2014).

Secretion system. In the genomes of both – the endosymbiont and bacterium CTD079 – we could identify the type I secretion system, the *sec*-system and the *tat*-system (Supporting Information Table S1). In bacterium CTD079's genome, we additionally discovered all genes necessary for the type II secretion system (Supporting Information Table S1). These listed systems all mediate protein secretion, which is an essential cell function in all prokaryotes. They enable the transport of proteins from the cytoplasm to other cell compartments, the environment or other microorganisms (Gerlach and Hensel, 2007).

The type II secretion system transports folded proteins like hydrolases. Furthermore, substances that mediate biofilm formation (Duncan *et al.*, 2011; Sikora, 2013) or symbiosis (Maltz and Graf, 2011) are released to the environment by this secretion system. To better understand the function of the type II secretion system, we compared the coding sequences

with the sequences deposited in the NCBI non-redundant database. They resembled those from *Sedimenticola selenatireducens* (68% AA identity; Accession PLX60149.1), a strictly anaerobic, selenite-respiring Gammaproteobacterium (Narasimgarao and Haggblom, 2006). Studies on the type II secretion system of the marine organism *Vibrio cholerae* showed that this type of secretion system supports biofilm formation on biotic and abiotic surfaces (Sikora, 2013), which enhances the growth of planktonic cells (Stauder *et al.*, 2012) and in this way helps to increase the cell density. In the case of CTD079, the acquired genes may enable CTD079 to protect itself from external influences via biofilm formation.

Defence mechanisms. Bacteria and archaea are exposed to the constant threat of viruses, which are especially prevalent in the marine environment (Suttle, 2005). The antiviral defence mechanisms consist of special Cas proteins and a Clustered Regularly Interspaced Short Palindromic Repeats (CRISPR) array (Barrangou *et al.*, 2007), which consists of short repeated sequences separated by spacers derived from nucleic acids of viruses and plasmids (Bolotin *et al.*, 2005; Mojica *et al.*, 2005; Pourcel *et al.*, 2005). The genomes of the symbiont and of bacterium CTD079 differ considerably with respect to their special defence mechanisms (Supporting Information Table S1). A much higher number of repeat sequences with viral DNA spacers (55 in total with 12 associated *cas* genes) was identified in the genome of CTD079 than in the endosymbiont's genome (4 in total). This implies that CTD079 is using the CRISPR-Cas defence system more frequently than the scaly snail endosymbiont. As hydrothermal vent environments have been shown to hold higher numbers of viruses than the open ocean (Cochlan *et al.*, 1993; Wommack *et al.*, 2004; Ortmann and Suttle, 2005), one would expect that a facultative endosymbiont has an imprint reflecting the free-living stages on its genome. Conclusively, the very low number of repeat sequences with viral DNA spacers detected on the scaly snail's symbiont's genome would be common for an obligate symbiont. Investigations of obligate symbiont genomes have shown that in some cases the CRISPR-Cas system is completely absent (Burstein *et al.*, 2016), while others have acquired a very high number of viral DNA spacers (Anbutsu and Fukatsu, 2011; He *et al.*, 2018). Thus, the number of spacers cannot serve as an indicator for facultative or obligate symbiosis. However, it has been observed that at high virus concentrations, like at hydrothermal vents, the defence takes place via constitutive mechanisms, such as recognition by surface proteins (Westra *et al.*, 2015) – a mechanism that becomes effective very quickly. Since CRISPR-Cas is an induced defence system, too much time passes before it becomes active. With a low virus density, like in the water column, the fitness costs for a constitutive, that is, permanent virus defence are too high for the organism. Therefore, the defence mainly takes place via the induced CRISPR-Cas system (Westra *et al.*, 2015).

We detected the type II system (class 2 CRISPR-Cas system) in both genomes. The type I-C system (class 1 CRISPR-Cas system) was detected in the symbiont only, while the type III-B system (class 1 CRISPR-Cas system) was merely identified in the bacterium CTD079. By cleaving RNA and also DNA, the type III-B system destroys the viral genome and the

corresponding transcripts (Hale *et al.*, 2009). In contrast, the type I-C CRISPR-Cas system only cleaves DNA (Makarova *et al.*, 2011; Sinkunas *et al.*, 2011). We were able to assign the operon for the type III-B CRISPR-Cas system on the genome of CTD079 to one of the predicted genomic islands suggesting an uptake of the genes via horizontal gene transfer. In addition, when compared with sequences in the NCBI non-redundant database, the highest similarity (76–94%, AA identity) was found to an unspecified Gammaproteobacterium from a hydrothermal vent (Accession RTZ73208.1), indicating that CTD079's ancestor may have had contact with vent environments/organisms.

Symbiotic versus water column habitat. We identified all important metabolic pathways in CTD079's genome that should enable it to be viable on its own. While under specific conditions CTD079 may be able to survive in a pure culture, we could not isolate it under the provided culture conditions and thus cannot rule out that it may be in a syntrophic relationship with *Alteromonas* sp. A similar relationship was suggested for a *Beggiatoa* strain that depends on *Pseudovibrio*, which appears to supply vitamins or detoxifies metabolic intermediates (Bondarev *et al.*, 2013).

Based on the phylogenetic analysis of the 16S rRNA genes, it is shown that bacterium CTD079 forms a monophyletic group with symbionts of two snail species (*Alviniconcha* sp.) and uncultured bacteria from microbial mats at Lo'ihi Seamount, Hawaii (Supporting Information Fig. S4). CTD079 and the snail symbionts share some potential phenotypes (e.g., hydrogen or sulfur oxidation, autotrophic CO₂ fixation). Hence, a common ancestor may have colonized a habitat where such metabolisms are favourable, for example, hydrothermal vent systems. Genes only identified on the CTD079 genome (e.g., encoding carbon monoxide dehydrogenase, FDH, nitrous oxide reductase, molybdate ABC-transporter, RegB/RegA and EnvZ/OmpR two-component system, type II secretion system, type III-B CRISPR-Cas system) exhibited the highest similarities to those from free-living but also symbiotic bacteria (Supporting Information Table S1). Thus, either (i) these traits were acquired multiple times by a common ancestor and the endosymbiont lost them or (ii) they were acquired horizontally multiple times by bacterium CTD079. However, the two available genomes are not sufficient to make a clear statement on whether the discussed metabolic features were lost by one or acquired by the other organism.

The most likely mode of transmission for endosymbionts takes place vertically, that is, from one generation to the next maternally. Such strictly vertically transmitted symbionts are described as obligate symbionts and have co-evolved with their hosts (Bright and Bulgheresi, 2010). The endosymbiont of the scaly snail was described to be located in the enlarged oesophageal glands (Goffredi *et al.*, 2004), but nothing is stated about its occurrence in the gonads. It remains to be clarified whether the symbiont is transmitted vertically and thus obligatory symbiotic or not. For other symbionts, a mix of vertical and horizontal (picked up from the environment) transmission has been proposed – thus an obligate and a facultative state (Dmytrenko *et al.*, 2014; Zimmermann *et al.*, 2016). For the scaly snail endosymbiont, flagellar genes and those encoding chemotaxis proteins were recognized. Since these

genes encode functions that allow the bacterium's survival outside the host, it appears highly likely that the endosymbiont also has a free-living stage and thus can be considered a facultative symbiont. Other examples exist: for the intracellular bacterium of the coastal bivalve *Solemya velum* genes that support a facultative endosymbiosis (e.g., genes for membrane-associated functions) have been described (Dmytrenko *et al.*, 2014). For some endosymbionts like the lucinid clam (*Codakia orbicularis*), the giant tubeworm (*Riftia pachyptila*) and the beard worm (*Oligobranchia mashikoi*) also free-living facultative symbiotic forms exist, although they have only been found in the vicinity of their hosts (Gros *et al.*, 2003; Aida *et al.*, 2008; Harmer *et al.*, 2008). In these cases, it appears that facultative symbionts are only able to survive for a short time outside their host, because of essential missing functions like important two-component systems for sensing the environment (Dmytrenko *et al.*, 2014).

In summary, given that we identified horizontally acquired gene clusters (from free-living, symbiotic, hydrothermal vent and other marine habitats), CTD079 appears well adapted to life in the water column. Based on some of the potential chemosynthetic phenotypes and the relatedness of many genes to those of hydrothermal vent organisms, it appears that the common ancestor colonized a hydrothermal vent comparable habitat.

Conclusion

We here successfully enriched a very close free-living relative of a sulfur-oxidizing endosymbiont. Its genome analysis provides information on the adaptation to a life in the open ocean. By comparing the genome of bacterium CTD079 with the available genome of the nearest living relative, a snail endosymbiont, we were able to show that the bacterium from the deep ambient water has several additional gene segments apparently obtained by horizontal gene transfer. The acquisition of genes and respective enzymes for using alternative energy sources, as well as the specific systems for secretion and defence against viruses, clearly show an adaptation to life in the water column. Further comparative studies of the symbiotic and the cultivated free-living organism could provide valuable insights into the evolution and the establishment of symbiosis.

Acknowledgements

We thank the captain and crew of the RV *Sonne* as well as Maren Walter and Sharon Walker for helping us to obtain the Kermadec samples with the CTD. Furthermore, we thank Dominik Danso for helpful discussions. Funding for this cruise was granted by the Federal Ministry of Education and Research (BMBF, project 03G0253), Germany. Permission to work in the Kermadec Arc by the New Zealand authorities is gratefully acknowledged.

Conflict of interest

The authors declare no conflict of interest.

References

- Aida, M., Kanemori, M., Kubota, N., Matada, M., Sasayama, Y., and Fukumori, Y. (2008) Distribution and population of free-living cells related to endosymbiont a harbored in *Oligobranchia mashikoi* (a siboglinid polychaete) inhabiting Tsukumo Bay. *Microbes Environ* **23**: 81–88.
- Anbutsu, H., and Fukatsu, T. (2011) *Spiroplasma* as a model insect endosymbiont. *Environ Microbiol Rep* **3**: 144–153.
- Barrangou, R., Fremaux, C., Deveau, H., Richards, M., Boyaval, P., Moineau, S., et al. (2007) CRISPR provides acquired resistance against viruses in prokaryotes. *Science* **315**: 1709–1712.
- Bertelli, C., Laird, M.R., Williams, K.P., Simon Fraser University Research Computing Group, Lau, B.Y., Hoad, G., et al. (2017) IslandViewer 4: expanded prediction of genomic islands for larger-scale datasets. *Nucleic Acids Res* **45**: W30–W35.
- Boden, R., Cunliffe, M., Scanlan, J., Moussard, H., Kits, K. D., Klotz, M.G., et al. (2011) Complete genome sequence of the aerobic marine methanotroph *Methylomonas methanica* MC09. *J Bacteriol* **193**: 7001–7002.
- Bolotin, A., Quinquis, B., Sorokin, A., and Ehrlich, S.D. (2005) Clustered regularly interspaced short palindrome repeats (CRISPRs) have spacers of extrachromosomal origin. *Microbiology* **151**: 2551–2561.
- Bondarev, V., Richter, M., Romano, S., Piel, J., Schwedt, A., and Schulz-Vogt, H.N. (2013) The genus *Pseudovibrio* contains metabolically versatile bacteria adapted for symbiosis. *Environ Microbiol* **15**: 2095–2113.
- Bothe, H., Schmitz, O., Yates, M.G., and Newton, W.E. (2010) Nitrogen fixation and hydrogen metabolism in cyanobacteria. *Microbiol Mol Biol Rev* **74**: 529–551.
- Bright, M., and Bulgheresi, S. (2010) A complex journey: transmission of microbial symbionts. *Nat Rev Microbiol* **8**: 218–230.
- Brissac, T., Gros, O., and Mercot, H. (2009) Lack of endosymbiont release by two *Lucinidae* (*Bivalvia*) of the genus *Codakia*: consequences for symbiotic relationships. *FEMS Microbiol Ecol* **67**: 261–267.
- Burstein, D., Sun, C.L., Brown, C.T., Sharon, I., Anantharaman, K., Probst, A.J., et al. (2016) Major bacterial lineages are essentially devoid of CRISPR-Cas viral defence systems. *Nat Commun* **7**: 10613.
- Cavanaugh, C.M., Gardiner, S.L., Jones, M.L., Jannasch, H. W., and Waterbury, J.B. (1981) Prokaryotic cells in the hydrothermal vent tube worm *Riftia-Pachyptila* Jones - possible chemoautotrophic symbionts. *Science* **213**: 340–342.
- Cavanaugh, C.M., Levering, P.R., Maki, J.S., Mitchell, R., and Lidstrom, M.E. (1987) Symbiosis of methylotrophic bacteria and deep-sea mussels. *Nature* **325**: 346–348.
- Cochlan, W.P., Wikner, J., Steward, G.F., Smith, D.C., and Azam, F. (1993) Spatial-distribution of viruses, bacteria and chlorophyll-a in neritic, oceanic and estuarine environments. *Mar Ecol Prog Ser* **92**: 77–87.
- Cole, J.J., Carpenter, S.R., Kitchell, J., Pace, M.L., Solomon, C.T., and Weidel, B. (2011) Strong evidence for terrestrial support of zooplankton in small lakes based on stable isotopes of carbon, nitrogen, and hydrogen. *Proc Natl Acad Sci U S A* **108**: 1975–1980.
- Dilling, W., and Cypionka, H. (1990) Aerobic respiration in sulfate-reducing bacteria. *FEMS Microbiol Lett* **71**: 123–127.
- Dmytrenko, O., Russell, S.L., Loo, W.T., Fontanez, K.M., Liao, L., Roeselers, G., et al. (2014) The genome of the intracellular bacterium of the coastal bivalve, *Solemya velum*: a blueprint for thriving in and out of symbiosis. *BMC Genomics* **15**: 924.
- Dubilier, N., Bergin, C., and Lott, C. (2008) Symbiotic diversity in marine animals: the art of harnessing chemosynthesis. *Nat Rev Microbiol* **6**: 725–740.
- Ducklow, H.W., Kirchman, D.L., and Rowe, G.T. (1982) Production and vertical flux of attached bacteria in the Hudson River Plume of the New-York Bight as studied with floating sediment traps. *Appl Environ Microbiol* **43**: 769–776.
- Duncan, C., Prashar, A., So, J., Tang, P., Low, D.E., Terebiznik, M., and Guyard, C. (2011) Lcl of *Legionella pneumophila* is an immunogenic GAG binding adhesin that promotes interactions with lung epithelial cells and plays a crucial role in biofilm formation. *Infect Immun* **79**: 2168–2181.
- Egger, L.A., Park, H., and Inouye, M. (1997) Signal transduction via the histidyl-aspartyl phosphorelay. *Genes Cells* **2**: 167–184.
- Ferry, J.G. (1990) Formate dehydrogenase. *FEMS Microbiol Rev* **7**: 377–382.
- Forst, S.A., and Roberts, D.L. (1994) Signal transduction by the EnvZ-OmpR phosphotransfer system in bacteria. *Res Microbiol* **145**: 363–373.
- Gerlach, R.G., and Hensel, M. (2007) Protein secretion systems and adhesins: the molecular armory of Gram-negative pathogens. *Int J Med Microbiol* **297**: 401–415.
- Goffredi, S.K., Warren, A., Orphan, V.J., Van Dover, C.L., and Vrijenhoek, R.C. (2004) Novel forms of structural integration between microbes and a hydrothermal vent gastropod from the Indian Ocean. *Appl Environ Microbiol* **70**: 3082–3090.
- Gomez-Consarnau, L., Needham, D.M., Weber, P.K., Fuhrman, J.A., and Mayali, X. (2019) Influence of light on particulate organic matter utilization by attached and free-living marine bacteria. *Front Microbiol* **10**: 1204.
- Goris, J., Konstantinidis, K.T., Klappenbach, J.A., Coenye, T., Vandamme, P., and Tiedje, J.M. (2007) DNA-DNA hybridization values and their relationship to whole-genome sequence similarities. *Int J Syst Evol Microbiol* **57**: 81–91.
- Gremer, L., Kellner, S., Dobbek, H., Huber, R., and Meyer, O. (2000) Binding of flavin adenine dinucleotide to molybdenum-containing carbon monoxide dehydrogenase from *Oligotropha carboxidovorans*. Structural and functional analysis of a carbon monoxide dehydrogenase species in which the native flavoprotein has been replaced by its recombinant counterpart produced in *Escherichia coli*. *J Biol Chem* **275**: 1864–1872.
- Gros, O., Liberge, M., Heddi, A., Khatchadourian, C., and Felbeck, H. (2003) Detection of the free-living forms of sulfide-oxidizing gill endosymbionts in the lucinid habitat

- (*Thalassia testudinum* environment). *Appl Environ Microbiol* **69**: 6264–6267.
- Hale, C.R., Zhao, P., Olson, S., Duff, M.O., Graveley, B.R., Wells, L., *et al.* (2009) RNA-guided RNA cleavage by a CRISPR RNA-Cas protein complex. *Cell* **139**: 945–956.
- Hansen, M., and Perner, M. (2015) A novel hydrogen oxidizer amidst the sulfur-oxidizing *Thiomicrospira* lineage. *ISME J* **9**: 696–707.
- Hansen, M., and Perner, M. (2016) Hydrogenase gene distribution and H₂ consumption ability within the *Thiomicrospira* lineage. *Front Microbiol* **7**: 99.
- Harmer, T.L., Rotjan, R.D., Nussbaumer, A.D., Bright, M., Ng, A.W., DeChaine, E.G., and Cavanaugh, C.M. (2008) Free-living tube worm endosymbionts found at deep-sea vents. *Appl Environ Microbiol* **74**: 3895–3898.
- Hausinger, R.P. (1993) *Carbon Monoxide Dehydrogenase*. New York, NY, USA: Springer.
- He, L.S., Zhang, P.W., Huang, J.M., Zhu, F.C., Danchin, A., and Wang, Y. (2018) The enigmatic genome of an obligate ancient *Spiroplasma* symbiont in a hadal holothurian. *Appl Environ Microbiol* **84**(1), e01965–17.
- Hsiao, WWL, Ung, K, Aeschliman, D, Bryan, J, Brett Finlay, B, & Brinkman, FSL (2005). Evidence of a large novel gene pool associated with prokaryotic genomic Islands. *PLoS Genet.*, **1**(5), 540–550.
- Hugendieck, I., and Meyer, O. (1992) The structural genes encoding CO dehydrogenase subunits (cox L, M and S) in *Pseudomonas carboxydovorans* OM5 reside on plasmid pHCG3 and are, with the exception of *Streptomyces thermoaerophilus*, conserved in carboxydophilic bacteria. *Arch Microbiol* **157**: 301–304.
- Imhoff, J.F. (2014) The family *Chromatiaceae*. In *The Prokaryotes*. Rosenberg, E., DeLong, E.F., Lory, S., Stackebrandt, E., and Thompson, F. (eds). Berlin, Germany: Springer.
- Iturriaga, R. (1979) Bacterial-activity related to sedimenting particulate matter. *Mar Biol* **55**: 157–169.
- Jiang, S.C., and Paul, J.H. (1998) Gene transfer by transduction in the marine environment. *Appl Environ Microbiol* **64**: 2780–2787.
- Jorgensen, B.B., and Boetius, A. (2007) Feast and famine—microbial life in the deep-sea bed. *Nat Rev Microbiol* **5**: 770–781.
- Joshi, H.M., and Tabita, F.R. (1996) A global two component signal transduction system that integrates the control of photosynthesis, carbon dioxide assimilation, and nitrogen fixation. *Proc Natl Acad Sci U S A* **93**: 14515–14520.
- Kappler, U., Bennett, B., Rethmeier, J., Schwarz, G., Deutzmann, R., McEwan, A.G., and Dahl, C. (2000) Sulfite:cytochrome c oxidoreductase from *Thiobacillus novellus*. Purification, characterization, and molecular biology of a heterodimeric member of the sulfite oxidase family. *J Biol Chem* **275**: 13202–13212.
- Karl, D.M., and Knauer, G.A. (1984) Vertical-distribution, transport, and exchange of carbon in the northeast Pacific-Ocean - evidence for multiple zones of biological-activity. *Deep Sea Res A Oceanogr Res Pap* **31**: 221–243.
- Krueger, D.M., Dubilier, N., and Cavanaugh, C.M. (1996) Chemoautotrophic symbiosis in the tropical clam *Solemya occidentalis* (Bivalvia: Protobranchia): ultrastructural and phylogenetic analysis. *Mar Biol* **126**: 55–64.
- Kusel, K., Trinkwalter, T., Drake, H.L., and Devereux, R. (2006) Comparative evaluation of anaerobic bacterial communities associated with roots of submerged macrophytes growing in marine or brackish water sediments. *J Exp Mar Biol Ecol* **337**: 49–58.
- McCutcheon, J.P., and Moran, N.A. (2011) Extreme genome reduction in symbiotic bacteria. *Nat Rev Microbiol* **10**: 13–26.
- Madigan, M.T., and Gest, H. (1979) Growth of the photosynthetic bacterium *Rhodospseudomonas capsulata* chemoautotrophically in darkness with H₂ as the energy source. *J Bacteriol* **137**: 524–530.
- Makarova, K.S., Haft, D.H., Barrangou, R., Brouns, S.J., Charpentier, E., Horvath, P., *et al.* (2011) Evolution and classification of the CRISPR-Cas systems. *Nat Rev Microbiol* **9**: 467–477.
- Maltz, M., and Graf, J. (2011) The type II secretion system is essential for erythrocyte lysis and gut colonization by the leech digestive tract symbiont *Aeromonas veronii*. *Appl Environ Microbiol* **77**: 597–603.
- Mantri, Y, & Williams, KP (2004). Islander: a database of integrative islands in prokaryotic genomes, the associated integrases and their DNA site specificities. *Nucleic Acids Res*, **32**, D55–D58.
- Meyer, O., and Schlegel, H.G. (1983) Biology of aerobic carbon monoxide-oxidizing bacteria. *Annu Rev Microbiol* **37**: 277–310.
- Mojica, F.J., Diez-Villasenor, C., Garcia-Martinez, J., and Soria, E. (2005) Intervening sequences of regularly spaced prokaryotic repeats derive from foreign genetic elements. *J Mol Evol* **60**: 174–182.
- Nakagawa, S., and Takai, K. (2008) Deep-sea vent chemoautotrophs: diversity, biochemistry and ecological significance. *FEMS Microbiol Ecol* **65**: 1–14.
- Nakagawa, S., Shimamura, S., Takaki, Y., Suzuki, Y., Murakami, S., Watanabe, T., *et al.* (2014) Allying with armored snails: the complete genome of gammaproteobacterial endosymbiont. *ISME J* **8**: 40–51.
- Narasimharao, P., and Haggblom, M.M. (2006) *Sedimenticola selenatireducens*, gen. nov., sp. nov., an anaerobic selenate-respiring bacterium isolated from estuarine sediment. *Syst Appl Microbiol* **29**: 382–388.
- Nunoura, T., Takaki, Y., Kazama, H., Kakuta, J., Shimamura, S., Makita, H., *et al.* (2014) Physiological and genomic features of a novel sulfur-oxidizing gammaproteobacterium belonging to a previously uncultivated symbiotic lineage isolated from a hydrothermal vent. *PLoS One* **9**: e104959.
- Nussbaumer, A.D., Fisher, C.R., and Bright, M. (2006) Horizontal endosymbiont transmission in hydrothermal vent tubeworms. *Nature* **441**: 345–348.
- Ortmann, A.C., and Suttle, C.A. (2005) High abundances of viruses in a deep-sea hydrothermal vent system indicates viral mediated microbial mortality. *Deep Sea Res I Oceanogr Res Pap* **52**: 1515–1527.
- Petersen, J.M., Zielinski, F.U., Pape, T., Seifert, R., Moraru, C., Amann, R., *et al.* (2011) Hydrogen is an energy source for hydrothermal vent symbioses. *Nature* **476**: 176–180.

- Petersen, J.M., Wentrup, C., Verna, C., Knittel, K., and Dubilier, N. (2012) Origins and evolutionary flexibility of chemosynthetic symbionts from deep-sea animals. *Biol Bull* **223**: 123–137.
- Pourcel, C., Salvignol, G., and Vergnaud, G. (2005) CRISPR elements in *Yersinia pestis* acquire new repeats by preferential uptake of bacteriophage DNA, and provide additional tools for evolutionary studies. *Microbiology* **151**: 653–663.
- Pratt, L.A., and Silhavy, T.J. (1995) Identification of base pairs important for OmpR-DNA interaction. *Mol Microbiol* **17**: 565–573.
- Pruesse, E., Peplies, J., and Glockner, F.O. (2012) SINA: accurate high-throughput multiple sequence alignment of ribosomal RNA genes. *Bioinformatics* **28**: 1823–1829.
- Qian, Y., and Tabita, F.R. (1996) A global signal transduction system regulates aerobic and anaerobic CO₂ fixation in *Rhodobacter sphaeroides*. *J Bacteriol* **178**: 12–18.
- Ragsdale, S.W. (2004) Life with carbon monoxide. *Crit Rev Biochem Mol Biol* **39**: 165–195.
- Scott, D., and Amy, N.K. (1989) Molybdenum accumulation in *chlD* mutants of *Escherichia coli*. *J Bacteriol* **171**: 1284–1287.
- Sikora, A.E. (2013) Proteins secreted via the type II secretion system: smart strategies of *Vibrio cholerae* to maintain fitness in different ecological niches. *PLoS Pathog* **9**: e1003126.
- Sinkunas, T., Gasiunas, G., Fremaux, C., Barrangou, R., Horvath, P., and Siksnys, V. (2011) Cas3 is a single-stranded DNA nuclease and ATP-dependent helicase in the CRISPR/Cas immune system. *EMBO J* **30**: 1335–1342.
- Stauder, M., Huq, A., Pezzati, E., Grim, C.J., Ramoino, P., Pane, L., et al. (2012) Role of GbpA protein, an important virulence-related colonization factor, for *Vibrio cholerae*'s survival in the aquatic environment. *Environ Microbiol Rep* **4**: 439–445.
- Stock, A.M., Robinson, V.L., and Goudreau, P.N. (2000) Two-component signal transduction. *Annu Rev Biochem* **69**: 183–215.
- Stothard, P., Grant, J.R., and Van Domselaar, G. (2019) Visualizing and comparing circular genomes using the CGView family of tools. *Brief Bioinform* **20**(4): 1576–1582.
- Sun, Y., and Luo, H. (2018) Homologous recombination in core genomes facilitates marine bacterial adaptation. *Appl Environ Microbiol* **84**: e02545-17.
- Suttle, C.A. (2005) Viruses in the sea. *Nature* **437**: 356–361.
- Taipale, S.J., Galloway, A.W., Aalto, S.L., Kahilainen, K.K., Strandberg, U., and Kankaala, P. (2016) Terrestrial carbohydrates support freshwater zooplankton during phytoplankton deficiency. *Sci Rep* **6**: 30897.
- Takai, K., Nakagawa, S., Reysenbach, A.L., and Hoek, J. (2006) Microbial ecology of mid-ocean ridges and back-arc basins. In *Back-Arc Spreading Systems - Geological, Biological, Chemical, and Physical Interactions*. Christie, D.M., Fisher, C.R., Lee, S.M., and Givens, S. (eds). Washington, DC, USA: American Geophysical Union, pp. 185–213.
- Tang, K.H., Tang, Y.J., and Blankenship, R.E. (2011) Carbon metabolic pathways in phototrophic bacteria and their broader evolutionary implications. *Front Microbiol* **2**: 165.
- Tian, R.M., Wang, Y., Bougouffa, S., Gao, Z.M., Cal, L., Bajic, V., and Qian, P.Y. (2014) Genomic analysis reveals versatile heterotrophic capacity of a potentially symbiotic sulfur-oxidizing bacterium in sponge. *Environ Microbiol* **16**: 3548–3561.
- Tian, R.M., Zhang, W., Cai, L., Wong, Y.H., Ding, W., and Qian, P.Y. (2017) Genome reduction and microbe-host interactions drive adaptation of a sulfur-oxidizing bacterium associated with a cold seep sponge. *mSystems* **2**: e00184–16.
- Vargas, W.A., Weyman, P.D., Tong, Y., Smith, H.O., and Xu, Q. (2011) [NiFe] hydrogenase from *Alteromonas macleodii* with unusual stability in the presence of oxygen and high temperature. *Appl Environ Microbiol* **77**: 1990–1998.
- Westra, E.R., van Houte, S., Oyesiku-Blakemore, S., Makin, B., Broniewski, J.M., Best, A., et al. (2015) Parasite exposure drives selective evolution of constitutive versus inducible defense. *Curr Biol* **25**: 1043–1049.
- Wommack, K.E., Williamson, S.J., Sundbergh, A., Helton, R. R., Glazer, B.T., Portune, K., and Cary, S.C. (2004) An instrument for collecting discrete large-volume water samples suitable for ecological studies of microorganisms. *Deep Sea Res I Oceanogr Res Pap* **51**: 1781–1792.
- Zimmermann, J., Wentrup, C., Sadowski, M., Blazejak, A., Gruber-Vodicka, H.R., Kleiner, M., et al. (2016) Closely coupled evolutionary history of ecto- and endosymbionts from two distantly related animal phyla. *Mol Ecol* **25**: 3203–3223.
- Zumft, W.G. (2005) Nitric oxide reductases of prokaryotes with emphasis on the respiratory, heme-copper oxidase type. *J Inorg Biochem* **99**: 194–215.

Supporting Information

Additional Supporting Information may be found in the online version of this article at the publisher's web-site:

Appendix S1: Supporting Information