The ISME Journal (2014) 8, 40–51

www.nature.com/ismej

ORIGINAL ARTICLE

Allying with armored snails: the complete genome of gammaproteobacterial endosymbiont

Satoshi Nakagawa[1,2], Shigeru Shimamura[2], Yoshihiro Takaki[3], Yohey Suzuki[4], Shun-ichi Murakami[1], Tamaki Watanabe[1], So Fujiyoshi[1], Sayaka Mino[1], Tomoo Sawabe[1], Takahiro Maeda[5,6], Hiroko Makita[2], Suguru Nemoto[7], Shin-Ichiro Nishimura[5,8], Hiromi Watanabe[9], Tomo-o Watsuji[2] and Ken Takai[2]

[1]*Laboratory of Microbiology, Faculty of Fisheries Sciences, Hokkaido University, Hakodate, Japan;* [2]*Subsurface Geobiology Advanced Research (SUGAR) Project, Institute of Biogeosciences, Japan Agency for Marine-Earth Science and Technology (JAMSTEC), Yokosuka, Japan;* [3]*Cooperative Research Project for Extremophiles, Institute of Biogeosciences, JAMSTEC, Yokosuka, Japan;* [4]*Department of Earth and Planetary Science, Graduate School of Science, University of Tokyo, Tokyo, Japan;* [5]*Laboratory of Advanced Chemical Biology, Graduate School of Life Science, Hokkaido University, Sapporo, Japan;* [6]*Amino Up Chemical Company, Sapporo, Japan;* [7]*Enoshima Aquarium, Fujisawa, Japan;* [8]*Medicinal Chemistry Pharmaceuticals, Co. Ltd., Sapporo, Japan and* [9]*Marine Biodiversity Research Program, JAMSTEC, Yokosuka, Japan*

Deep-sea vents harbor dense populations of various animals that have their specific symbiotic bacteria. Scaly-foot gastropods, which are snails with mineralized scales covering the sides of its foot, have a gammaproteobacterial endosymbiont in their enlarged esophageal glands and diverse epibionts on the surface of their scales. In this study, we report the complete genome sequencing of gammaproteobacterial endosymbiont. The endosymbiont genome displays features consistent with ongoing genome reduction such as large proportions of pseudogenes and insertion elements. The genome encodes functions commonly found in deep-sea vent chemoautotrophs such as sulfur oxidation and carbon fixation. Stable carbon isotope (^{13}C)-labeling experiments confirmed the endosymbiont chemoautotrophy. The genome also includes an intact hydrogenase gene cluster that potentially has been horizontally transferred from phylogenetically distant bacteria. Notable findings include the presence and transcription of genes for flagellar assembly, through which proteins are potentially exported from bacterium to the host. Symbionts of snail individuals exhibited extreme genetic homogeneity, showing only two synonymous changes in 19 different genes (13 810 positions in total) determined for 32 individual gastropods collected from a single colony at one time. The extremely low genetic individuality in endosymbionts probably reflects that the stringent symbiont selection by host prevents the random genetic drift in the small population of horizontally transmitted symbiont. This study is the first complete genome analysis of gastropod endosymbiont and offers an opportunity to study genome evolution in a recently evolved endosymbiont.

The ISME Journal (2014) **8,** 40–51; doi:10.1038/ismej.2013.131; published online 8 August 2013
Subject Category: Microbe-microbe and microbe-host interactions
Keywords: chemoautotrophy; deep-sea hydrothermal vent; symbiosis; scaly-foot gastropod

Introduction

Scaly-foot gastropods, tentatively named '*Crysomallon squamiferum*', were first described in 2001 (Van Dover *et al.*, 2001). These snails live a sedentary life at the base of a black smoker chimney, in the Kairei hydrothermal field in the Indian Ocean. Dark-colored sclerites cover the sides of the snails' foot in a roof-tile fashion (Warén *et al.*, 2003; Goffredi *et al.*, 2004). The dominant crystalline mineral phase is pyrite, whereas greigite, the sulfide analog of magnetite, is present in lower proportions and is responsible for the ferrimagnetism (Warén *et al.*, 2003; Suzuki *et al.*, 2006). In addition, the shell of scaly-foot gastropod is fortified by a unique multilayer structure (Yao *et al.*, 2010). These defenses are probably acquired for the protection against environmental threats including grazing (Suzuki *et al.*, 2006). The scaly-foot gastropods,

Correspondence: S Nakagawa, Laboratory of Microbiology, Faculty of Fisheries Sciences, Hokkaido University, 3-1-1 Minato-cho, Hakodate 041-8611, Japan.
E-mail: nakagawa@fish.hokudai.ac.jp
Received 5 March 2013; revised 26 June 2013; accepted 1 July 2013; published online 8 August 2013

however, share the habitat around the base of a black smoker chimney with a different kind of gastropod, *Alviniconcha* sp., which have a bare foot and very thin soft shell.

Many deep-sea vent endemic mollusks harbor specific symbionts, although *Alviniconcha* gastropods exhibit marked diversity in their symbiont ribotypes within and among populations (Suzuki *et al.*, 2005; Beinart *et al.*, 2012). The scaly-foot gastropod houses its gammaproteobacterial symbionts within the cells of its enlarged esophageal gland (Goffredi *et al.*, 2004). The endosymbiont population is highly clonal within and among snail individuals at the 16S rRNA gene sequence level (Goffredi *et al.*, 2004). In addition, *Deltaproteobacteria* and *Epsilonproteobacteria* densely colonize the surface of scales (Goffredi *et al.*, 2004). The gastropod nutritionally depends on the intracellular gammaproteobacterium that synthesize organics from CO_2 potentially by oxidizing reduced sulfur compounds (Goffredi *et al.*, 2004). Until recently, ecophysiological characteristics of symbiotic *Gammaproteobacteria* in deep-sea vents had remained mostly unknown because of their resistance to cultivation outside of specialized host tissues. However, genome sequences have been determined for endosymbionts of *Riftia pachyptila* (tubeworm) (Markert *et al.*, 2007) and *Calyptogena* species (bivalves endemic in deep-sea vents) (Kuwahara *et al.*, 2007; Newton *et al.*, 2007). Complete genome sequence has also been determined for the sulfur-oxidizing endosymbiont of *Olavius algarvensis* (shallow oligochaete) (Woyke *et al.*, 2006). These genome sequences have provided new insights into symbiont evolution, physiological functions and interactions with their host animals. For example, the marine endosymbionts, including heritable ones, have significantly larger genomes with higher GC contents compared with insect endosymbionts, suggesting their recent association with hosts (Woyke *et al.*, 2006; Kuwahara *et al.*, 2007, Markert *et al.*, 2007; Newton *et al.*, 2007). As expected from the large genome sizes, the marine gammaproteobacterial endosymbionts retain metabolically complete genomes, and even have multiple pathways both for energy metabolism (Woyke *et al.*, 2006; Kuwahara *et al.*, 2007; Newton *et al.*, 2007) and for carbon fixation (Markert *et al.*, 2007). These endosymbiont genomes therefore provide a rare opportunity to study the early stages of genome reduction during symbiont evolution.

In deep-sea vent ecosystems, chemoautotrophic *Epsilonproteobacteria* are dominant and metabolically versatile, suggesting that deep-sea vent chemoautotrophs, even if they are epi- or endosymbionts, experience broad fluctuations in the availabilities (both kind and concentration) of electron donors, electron acceptors and carbon sources (Nakagawa *et al.*, 2005, 2007). Specific ecophysiological differences between gammaproteobacterial and epsilonproteobacterial deep-sea vent chemoautotrophs have not been clarified yet. Several genetic and biochemical studies demonstrated that the chemoautotrophic *Gammaproteobacteria* and *Epsilonproteobacteria* fix CO_2 via Calvin-Benson-Basham cycle and reductive tricarboxylic acid cycle, respectively (Hügler *et al.*, 2005; Takai *et al.*, 2005; Nakagawa and Takai, 2008). In addition, repertoire of the sulfur-oxidation pathways differs between deep-sea vent *Gammaproteobacteria* and *Epsilonproteobacteria* (Nakagawa and Takai, 2008; Yamamoto and Takai, 2011). However, little is known about how these differences influence their ecological performance.

At present, no genome sequence is available for any gastropod endosymbionts. In this study, we describe the complete genome sequence analysis of gammaproteobacterial endosymbiont found in scaly-foot gastropod. To assess the physiology of gastropod, we also performed the ^{13}C tracer experiments and hemolymph glycan analysis. We sequenced the symbiont genome specifically as an effort to address (1) genomic differences with other chemoautotrophs, (2) genome evolution in a recently derived symbiont and (3) endosymbiont physiology. This study represents a step toward a better understanding of genome dynamics and evolution of symbiotic bacteria in deep-sea vents, and will lead to more specifically focused biochemical investigations.

Materials and methods

Sampling and DNA/RNA preparation

Samples used in this study were obtained by means of DSV *Shinkai 6500* at the Kairei hydrothermal field (Van Dover *et al.*, 2001) in the Central Indian Ridge (25° 19.23′S, 70° 02.42′E). All samples were collected from a single colony during dive no. 1169 on 6 November 2009, unless noted otherwise. Although *Alviniconcha* sp. represents the gastropods visually observed from the submersible's window in this hydrothermal field, we discovered a large and dense population of scaly-foot gastropods covered by *Rimicaris* shrimps (water depth = 2420 m) (Supplementary Figure S1). Once onboard the ship, the esophageal gland was dissected from a single snail (Supplementary Figure S1c) and homogenized with glass Teflon homogenizer powered by a standard electric drill in 5 ml of phosphate-buffered saline at pH 7.0 on ice. Debris was pelleted at $800 \times g$ for 5 min, and supernatant was filtered through a 10-μm nylon mesh filter (Millipore, Billerica, MA, USA). The filtrate was further centrifuged at 14 000 r.p.m. for 10 min to recover endosymbiont cells. The cell pellet was washed three times in phosphate-buffered saline and stored at −80 °C until taken to the laboratory. Genomic DNA (~ 500 μg) was extracted manually using the phenol/chloroform method. In addition, cell pellet was washed in RNAlater (Ambion, Austin, TX, USA) and stored at −80 °C.

Genome sequencing and assembly
The genome was sequenced using 454 Titanium (Roche, Branford, CT, USA) with the paired-end library of 3 kb. Library preparation and sequencing were performed at TaKaRa Bio Inc. (Otsu, Japan). The sequence reads (a total of 188 Mb in 492 255 reads) were assembled using the GS De Novo Assembler (Newbler) version 2.3 (Roche/454 Life Sciences, Branford, CT, USA). Low-sequence quality regions and potential sequence errors were searched and confirmed by Sanger sequencing. Together, all reads provided about $\times 65$ coverage of the genome.

Sequence analysis and annotation
Genes were identified using Glimmer (Salzberg *et al.*, 1998) and GeneMarkS (Besemer and Borodovsky, 2005), followed by manual screening. Additional open reading frames that were homologous to genes identified in other organisms were searched from intergenic region. The open reading frames were compared with the nonredundant database of the NCBI, UniProt and Kyoto Encyclopedia of Genes and Genomes GENES database, using BLASTP (Altschul *et al.*, 1997). Protein domains were searched against the PFAM database (Bateman *et al.*, 1999). The annotated functions were classified into clusters of orthologous group categories (Tatusov *et al.*, 2000). The metabolic pathway was reconstructed using the Kyoto Encyclopedia of Genes and Genomes and MetaCyc (Caspi *et al.*, 2008) databases. Signal peptides and membrane-spanning domains were predicted using the SIGNALP (Nielsen *et al.*, 1997), and SOSUI (Hirokawa *et al.*, 1998), respectively. Transfer RNAs were predicted by tRNAscan-SE (Lowe and Eddy, 1997). We defined orthologs using Inparanoid/MultiParanoid (Remm *et al.*, 2001; Alexeyenko *et al.*, 2006), in which one open reading frame is the closet relative of the other and *vice versa* with the BLAST score >40 and over 70% coverage. Pseudogenes were defined as genes with frameshift mutations, insertion sequence elements, other insertions or deletions, premature stop codons or truncations that nevertheless showed high homology to the functional genes.

Genetic individuality analyses
A total of 19 genes including a pseudogene were sequenced for each symbiont extracted from 32 different individuals. Genes were selected based on PubMLST database (http://pubmlst.org/). Primers used in this study were listed in Supplementary Table S1. Genes were PCR amplified with LA Taq polymerase (TaKaRa Bio Inc.) using initial denaturation at 96 °C for 1 min, 30 cycles of 96 °C for 20 s, 60 °C for 45 s and 72 °C for 2 min, and a final extension at 72 °C for 10 min. The PCR product was cleaned with exonuclease I (Affymetrix, Santa Clara, CA, USA) and shrimp alkaline phosphatase (GE Healthcare, Piscataway, NJ, USA), and subjected to direct sequencing with an ABI 3130xl sequencer (Applied Biosystems, Foster City, CA, USA) according to the manufacturer's instructions. Likewise, six different gene sequences were determined for host mitochondria (Supplementary Table S1).

Reverse transcription PCR
Total RNA was extracted from cells stored in RNAlater using High Pure RNA Isolation Kit (Roche), according to the manufacturer's instructions. cDNA was synthesized using transcriptor high-fidelity cDNA synthesis kit and random hexamers (Roche). Primers for nine flagellar genes were constructed (Supplementary Table S1). Amplification protocol was as described above. In the negative control, the reverse transcription was omitted. Sequences were confirmed as described above.

Incubation with ^{13}C-bicarbonate
Scaly-foot gastropods can be kept alive in a tank for over 3 weeks. Immediately after the collection, gastropods were transferred in 2.5 l plastic jars (BBL Microbiology Systems, Cockeysville, MD, USA) at 4 °C (Supplementary Figure S2a). Each jar was filled with one individual gastropod and 1.5 l of surface seawater containing 3 mM $NaH^{13}CO_3$. Potential energy substrates were added at final concentrations of 100 μM (sulfide) or 20% H_2 (v/v). Gastropods killed by freezing (1–2 h at -80 °C) were also incubated as negative control in the presence of sulfide or H_2. After 64.5 h incubation in dark, the gastropods were dissected into esophageal glands and feet and kept at -80 °C. At the laboratory, samples were lyophilized, powdered and then acid-fumed for over 10 h. The carbon isotopic compositions were analyzed by a Thermo Electron (Bremen, Germany) DELTA^{Plus} Advantage mass spectrometer connected to an elemental analyzer (EA1112) through a ConFlo III interface. When the alanine standard (δ^{13}C, -19.4‰) was measured for 10 times, the s.d. was 0.3‰. Samples of this tracer experiments were prepared and analyzed in 2007.

Glycan analysis
The hemolymph was collected from snails incubated for 18 h with sulfide, H_2 (as described above) and thiosulfate (50 μM) (Supplementary Figure S2b), centrifuged for a few minutes and then stored at -80 °C. *N*-glycans were released and purified from 20 μl of hemolymph as described previously (Amano *et al.*, 2010) (detailed in Supplementary Methods). The aoWR (aminooxy-functionalized peptide compound)-tagged glycans were crystallized with 2,5-dihydrozybenzoic acid and analyzed by MALDI-TOF MS (4700 Proteomics Analyzer; Applied Biosystems) in reflector, positive ion mode. Data were analyzed by GlycoMod Tool (Cooper *et al.*, 2001)

and glyfinTMS software (http://www.jamstec.go.jp/software/glyfintms/e/). Glycans were quantified by comparative analyses between the areas of the mass signals and the internal standard (20 pmol maltohexaose). Oligosaccharides with masses greater than ~1000 Da exhibited similar signal strengths, irrespective of structure (Naven and Harvey, 1996). Two-way analysis of variance was performed with Prism software (GraphPad Software, La Jolla, CA, USA). Less than 1 μl of hemolymph was required for this analysis.

Nucleotide sequence accession number
Genome sequences of gammaproteobacterial endosymbiont and host mitochondrion have been deposited in DDBJ/EMBL/GenBank databases under the project accession numbers AP012978 and AP013032, respectively.

Results and Discussion

General genomic characteristics
Pyrosequence reads were assembled into three large scaffolds, which were assigned to the genome fragments from gastropod (15 611 bp, 2 contigs), mitochondria (7764 bp, 1 contig) and symbiotic bacteria (2 591 967 bp, 53 contigs). The gaps between contigs were closed by primer walking. We could assemble the genome of gammaproteobacterial endosymbiont into a complete circular genome, which is 2 597 759 bp in length, with an average G+C content of 65.1 mol% (Figure 1 and Table 1). The genome had a 16S rRNA gene sequence identical to that reported in previous clone analysis (Goffredi *et al.*, 2004) (Supplementary Figure S3a). No additional large contig was obtained. The successful assemblage indicated high genetic clonality of endosymbiont within a host gastropod, contrasting to the co-occurrence of divergent endosymbionts in deep-sea vent bivalves (Distel *et al.*, 1995). Nucleotide position 1 of chromosome was assigned to the predicted replication origin (*ori*), flanked on one side by *dnaA, dnaN* and *recF* genes and on the other by *rpmH* gene. Cumulative GC skew analysis also suggested that *ori* should have been localized to the same region (Figure 1).

The endosymbiont genome contained a total of 2249 protein-coding DNA sequences, resulting in a coding density of 82.6%. Of the predicted coding DNA sequences, 1677 (74.6%) have been assigned to a putative function using BLASTP searches with an *E*-value cutoff of $\geqslant 10^{-5}$ (Table 1). Approximately one-third of the predicted proteins revealed top BLAST hits to tubeworm symbionts (492 proteins to *Tevnia* symbiont and 240 proteins to *Riftia* symbiont). In addition, 586 coding DNA sequences had their top hits to one of the purple sulfur bacteria (*Chromatiaceae*/*Ectothiorhodospiraceae* group), including *Thioalkalivibrio* (112 hits), *Thiocapsa* (82 hits) and *Allochromatium* (73 hits) (Table 1). Compared with genomes of these free-living relatives (size, 2.7–5.7 Mb; G+C content, 46.2–68.0 mol%; coding density, 86.8–91.7%), the symbiont genome was relatively smaller in size, gene-sparse, but not apparently AT-rich (Table 1).

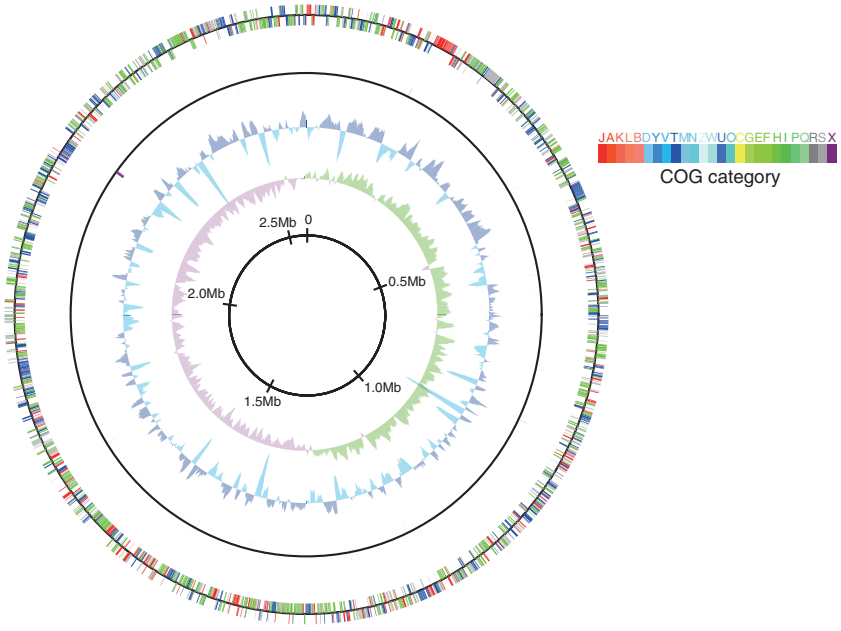

Figure 1 Schematic circular diagram of the scaly-foot endosymbiont genome. Outer circle, predicted coding regions on the plus and minus strands, color-coded by functional categories; second circle, RNA genes on the plus and minus strands; third circle, GC content showing deviation from average (65.1%); and fourth circle, GC skew. The GC content, and GC skew were calculated using a sliding window of 5 kb in step of 2.5 kb.

Table 1 General genomic features of the scaly-foot gastropod endosymbiont and free-living purple sulfur bacteria

Characteristic	*Scaly-foot gastropod endosymbiont*	*Allochromatium vinosum*	*Thioalkalivibrio sulfidophilus*	*Thioalkalivibrio* sp. K90mix
Size (bp)	2 597 759	3 669 074	3 464 554	2 985 056
G + C content (%)	65.1	64.2	65.1	65.5
Number of protein-coding gene	2249	3302	3319	2942
Number of pseudogenes	152	82	36	33
Coding density (%)	82.6	90.6	87.5	89.6
Number of identified RNA genes				
rRNA operons	1	3	3	3
tRNA	40	51	44	48
Number of chemotaxis genes				
che	13	34	8	9
mcp	12	28	14	6

We also assembled the mitochondrial genome into a complete circular genome that was 15 388 bp in length (Supplementary Figure S4).

Pseudogenes and genomic islands

We discovered 152 endosymbiont sequences that appeared to be pseudogenes (Table 1). In previously reported genomes of free-living purple sulfur bacteria, number of pseudogenes varied ranging from 33 (*Thioalkalivibrio* sp. K90mix) to 82 (*Allochromatium vinosum*). Recently evolved pathogens/symbionts generally have many pseudogenes and low coding density, reflecting that gene inactivation outpaces DNA deletion (Burke and Moran, 2011). Pseudogenized genes in the snail symbiont included putative ABC transporters of organics and putative sugar phosphotransferase systems, suggesting the endosymbiont had an ability to grow heterotrophically until recently (Figure 2). Tubeworm symbiont was suggested to have an ability to grow heterotrophically in the free-living state (Markert *et al.*, 2007). At least 39 regions were identified as insertion sequence elements. None of the insertion sequence elements exhibited significant similarities to those of the phylogenetic relatives or gammaproteobacterial symbionts.

Flagellum

Flagellar motility is essential for the infection of some symbionts, including both facultative and obligate ones (Dulla *et al.*, 2012; Rio *et al.*, 2012). At least 20 genes have been identified to be necessary for bacterial flagellar motility (Pallen and Matzke, 2006). The endosymbiont of scaly-foot gastropod had all but one flagellin gene, and these genes were mostly organized in three regions. Although some hook-associated genes have been pseudogenized, the flagellar genes in the endosymbiont genome mostly remained intact, which included core proteins of the type III secretion system (Supplementary Table S2). The reverse transcription PCR showed that genes encoding the type III secretion system were all expressed (Supplementary Figure S5), potentially suggesting the flagellar genes have undergone functional divergence. In many endosymbiotic bacteria of insects, for example, flagellar proteins (basal body with or without hook) have been regarded to serve as a protein transporter to maintain the symbiotic system (Maezawa *et al.*, 2006; Toft and Fares, 2008). The actual function of flagellar genes should be confirmed directly by experimental works in the future.

Chemotaxis proteins have important roles in bacterial adaptation to deep-sea vents (Scott *et al.*, 2006; Nakagawa *et al.*, 2007; Takaki *et al.*, 2010; Xie *et al.*, 2011). Even endosymbionts are no exception when in the free-living state. The endosymbiont of scaly-foot gastropod, compared with free-living relatives, has a similar number of chemotaxis proteins (Table 1). The genome included 13 chemotaxis (*che*) genes and 12 methyl-accepting chemotaxis proteins. In addition, the symbiont has a wide array of transport machineries including detoxification systems of heavy metals such as arsenate and copper (Figure 2). Even though the host animal must provide the endosymbiont with a relatively stable habitat, these functions probably reflect the endosymbiont's ability to sense and respond to the surrounding environment in its free-living state.

Sulfur-compound oxidation

Figure 2 shows the reconstruction of gastropod endosymbiont metabolism. The endosymbiont is a hydrogen/sulfur-oxidizing chemoautotroph, sharing the pathways involved in converting inorganic carbon into organics with other gammaproteobacterial chemoautotrophs. Similar with other sulfur-oxidizing *Gammaproteobacteria*, the endosymbiont has *sqr* (EBS_1809, EBS_1831), *dsrAB* (EBS_1783-1784), *aprAB* (EBS_2148-2149), *sat* (EBS1086) and

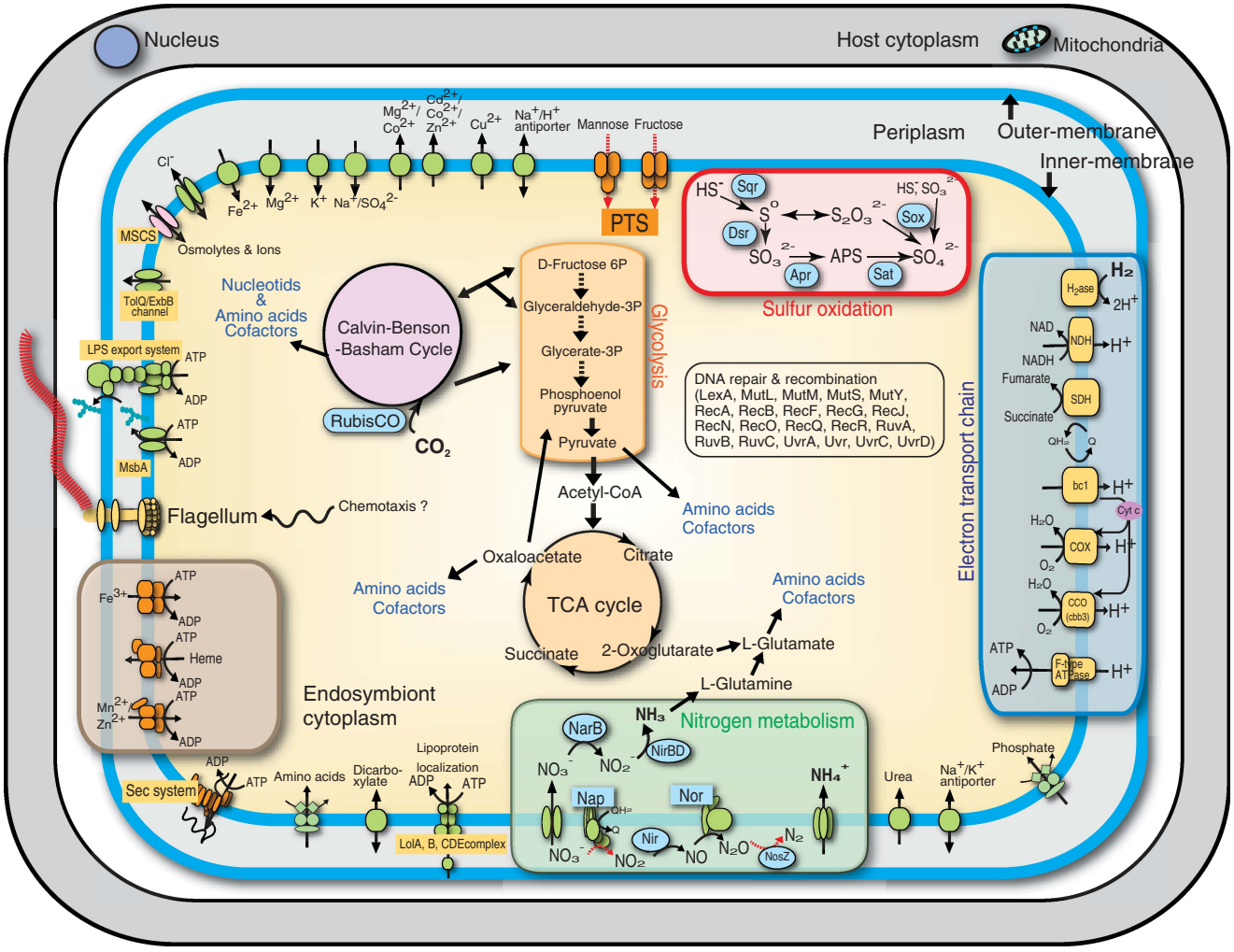

Figure 2 The major metabolic pathways and solute transport in the scaly-foot gastropod endosymbiont. The Kyoto Encyclopedia of Genes and Genomes database was used for the reconstruction of the metabolic pathways. Red dotted lines are inoperative due to pseudogenes.

soxXYZA (EBS_1635-1638)/*soxB* (EBS_0922) genes for sulfur oxidation (Figure 2). The order and composition of putative *dsr* gene cluster was almost identical with those of *A. vinosum* (Dahl *et al.*, 2005) and *Riftia* symbiont (Gardebrecht *et al.*, 2012) (Figure 3a). The isolated location of *soxB* relative to other *sox* genes was unusual but previously reported in gammaproteobacterium *Thiomicrospira crunogena* (Scott *et al.*, 2006). SoxB of the scaly-foot symbiont and *T. crunogena* share 60% amino acid sequence identity (Supplementary Figure S3b). Incongruence between tree topologies (Supplementary Figures S3a and b) could be explained by horizontal gene transfer (Petersen *et al.*, 2012). For alphaproteobacterial *Paracoccus pantotrophus,* a total of 15 genes comprising a single *sox* gene cluster, have been investigated in detail (Friedrich *et al.*, 2005). Of the *sox* genes, seven genes, *soxXYZABCD*, code for periplasmic proteins responsible for the oxidation of sulfur compounds (sulfide, elemental sulfur, thiosulfate and sulfite) (Friedrich *et al.*, 2005). The genome of scaly-foot gastropod endosymbiont lacks *soxCD* genes, which encode the sulfur dehydrogenase and mediate a unique oxidative six-electron transfer. To our knowledge, only sulfur-oxidizing *Alphaproteobacteria* and *Epsilonproteobacteria* have the *soxCD* genes (Friedrich *et al.*, 2005; Nakagawa *et al.*, 2007). In addition, chemoautotrophic *Gammaproteobacteria* including scaly-foot gastropod endosymbiont generally lack SorAB, which catalyze the oxidation of sulfite to sulfate and is usually present in mesophilic chemoautotrophic *Epsilonproteobacteria* and *Alphaproteobacteria* (Kappler *et al.*, 2000). In contrast, deep-sea vent chemoautotrophic *Epsilonproteobacteria* lack Dsr/Apr/Sat (Nakagawa *et al.*, 2007). Although some patterns have been emerging in the repertoire of the sulfur-oxidation pathways in chemoautotrophs, how it affects the ecological performance in deep-sea vents remains to be studied.

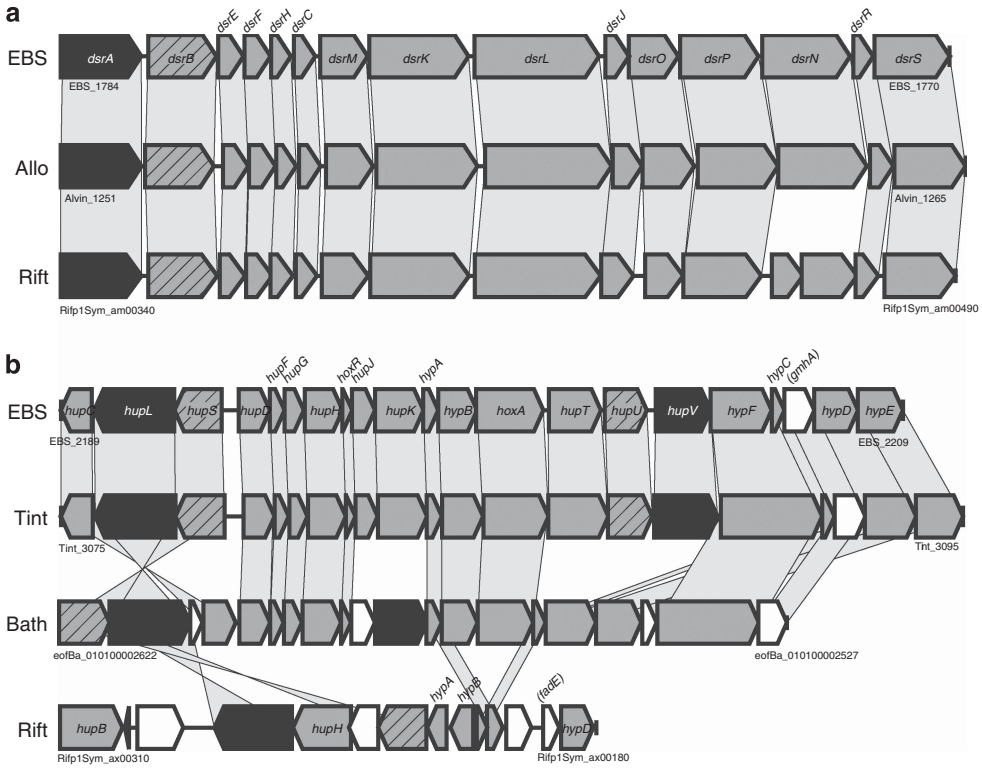

Figure 3 Comparative organization of *dsr* (**a**) and hydrogenase gene (**b**) clusters. EBS, scaly-foot gastropod endosymbiont; Allo, *A. vinosum*; Bath, *Bathymodiolus* sp. endosymbiont; Rift, *R. pachyptila* endosymbiont; Tint, *Thiomonas intermedia* K12. Genes are represented by arrows of length proportional to the gene length. Structural genes are represented by hatched arrows (small subunits) or black arrows (large subunits). White arrows without gene names are hypothetical protein genes. Gray bars indicate orthologous genes. Sequence information was retrieved from GenBank.

Hydrogen oxidation

Hydrogen gas could be the alternative energy source to sulfur compounds for deep-sea vent-dominating chemoautotrophs (Nakagawa *et al.*, 2005). Hydrogenases catalyze the reversible oxidation of H_2 to protons and electrons. The hydrogenases are classified into four distinct groups based on cellular functions and amino acid sequences: group 1, membrane-bound H_2-uptake hydrogenase; group 2, H_2-sensing or cyanobacterial hydrogenase; group 3, F_{420}-reducing, bifunctional hyperthermophilic hydrogenase, methylviologen-reducing hydrogenase and bidirectional NAD-linked hydrogenase; and group 4, membrane-bound H_2-evolving hydrogenase (Vignais *et al.*, 2001). The hydrogenase genes usually occur as tightly clustered functional units or operons, which consist of hydrogenase subunit genes (structural genes) and accessory or maturation genes encoding proteins necessary for the assembly of functional holoenzyme (for review, see Vignais *et al.*, 2001). The compositions of hydrogenase gene clusters vary among *Gammaproteobacteria* (Vignais *et al.*, 2001). Although heritable endosymbiont of *Calyptogena* clams have no hydrogenase (Kuwahara *et al.*, 2007; Newton *et al.*, 2007), other previously characterized deep-sea vent gammaproteobacterial chemoautotrophs appear to have group 1 hydrogenases (Scott *et al.*, 2006; Markert *et al.*, 2007; Anantharaman *et al.*, 2013). In the snail endosymbiont genome, we found one of the group 1 hydrogenase (EBS_2190-2191) and one of the group 2 hydrogenase (EBS_2203-2204) (Figure 3b and Supplementary Figure S6). Recent studies suggested that not only sulfur compounds but also H_2 was potentially utilized as an energy source by endosymbionts of mussel *Bathymodiolus*, tubeworm *R. pachyptila* and shrimp *Rimicaris exoculata* (Petersen *et al.*, 2011). Sequences and gene order of hydrogenase genes of scaly-foot gastropod endosymbiont showed the highest similarities, not to those of phylogenetic relatives, but to those of *Thiomonas intermedia* K12 (betaproteobacterium) (Arsène-Ploetze *et al.*, 2010) (Figure. 3b). Only 10.3% (235 hits) of the predicted proteins of endosymbiont showed top BLAST hits to members of *Betaproteobacteria*. This may suggest that the hydrogenase gene cluster has been horizontally transferred, although members of the class *Betaproteobacteria* are rare in deep-sea vents (Takai *et al.*, 2006). Recent studies suggested that acquisition of hydrogenase genes via horizontal gene transfer might have a profound role in the metabolic evolution of various deep-sea vent *Gammaproteobacteria* (Kleiner *et al.*, 2012). Deep-sea vent-dominating epsilonproteobacterial chemoautotrophs commonly have the group 4 hydrogenase as

well as groups 1 and 2 hydrogenases (Nakagawa *et al.*, 2007), which is suggestive of the efficient energy metabolism similar to the 'intracellular H_2-cycling mechanism' of sulfate-reducing bacteria (Odom and Peck Jr, 1981).

Aerobic respiration

The symbiont genome encodes all the necessary machinery for aerobic respiration, including NADH-quinone oxidoreductase, succinate dehydrogenase and cytochrome bc_1-type ubiquinol oxidoreductase (Figure 2). The composition of terminal oxidases varies among gammaproteobacterial chemoautotrophs; the scaly-foot symbiont has two cytochrome cbb_3-type oxidases and one *bd*-type cytochrome oxidase as *Calyptogena* symbionts do (Kuwahara *et al.*, 2007; Newton *et al.*, 2007) (Figure 2). The *Escherichia coli bd*-type cytochrome oxidase has a high affinity for O_2, and it is expressed under microaerobic conditions (D'mello *et al.*, 1996). The presence of both cbb_3-type and *bd*-type cytochromes may allow the symbionts to survive under a wide range of redox conditions.

Nitrogen metabolism

Among deep-sea vent-dominating chemoautotrophs, most members of *Epsilonproteobacteria* utilize nitrate both as an electron acceptor and as a nitrogen source (Nakagawa *et al.*, 2005; Nakagawa and Takai, 2008). In contrast, lack of a dissimilatory denitrification pathway (sequential reduction of nitrate to nitrogen gas) is common in previously sequenced deep-sea vent gammaproteobacterial chemoautotrophs (Scott *et al.*, 2006; Markert *et al.*, 2007). Among the genes necessary for the denitrification, *napA* (periplasmic nitrate reductase gene), *napF* (gene necessary for the assembly of the iron-sulfur center of NapA) and *nosZ* (nitrous-oxide reductase gene) have become pseudogenized in the scaly-foot gastropod symbiont (Figure 2). We identified intact genes associated with assimilatory ammonification (sequential reduction of nitrate to ammonium): ferredoxin-nitrate reductase (NarB; EBS_0633, EBS_0643) and nitrite reductase (NAD(P)H) (NirBD; EBS_0631, EBS_0632).

Central metabolism and carbon fixation

The gastropod symbiont genome contained nearly all genes needed to reconstruct the complete central pathways, that is, TCA cycle, Embden–Meyerhof–Parnas and pentose phosphate pathways (Figure 2). The complete pathways for the formation of the all amino acids, nucleotides, fatty acids and phospholipids from intermediates of the central metabolism could also be reconstructed (Figure 2). Previous research has suggested that chemoautotrophic deep-sea vent *Gammaproteobacteria* fix CO_2 by way of the Calvin–Benson–Bassham (CBB) cycle (Robinson *et al.*, 1998; Scott *et al.*, 2006). Ribulose 1,5-bisphosphate carboxylase/oxygenase (RubisCO) is a key enzyme of the CBB cycle and is classified into four different groups, that is forms I to IV. The scaly-foot gastropod symbiont genome encodes genes of both a form IAq (EBS_0695-0696) and a form II (EBS_0699) RubisCO for CO_2 fixation. These genes are organized in an operon with two posttranscriptional activators (*cbbOQ;* EBS_0693-0694, EBS_0697-0698) and a transcriptional regulator (*lysR*; EBS_0700). No additional form of RubisCO was found. The repertoire of RubisCOs in the endosymbiont genome is unique because previously sequenced chemoautotrophic *Gammaproteobacteria* with both form IAq and form II RubisCO include form IAc (Badger and Bek, 2008). Form I RubisCO, which has a higher CO_2/O_2 specificity, is probably more highly expressed under oxic conditions, whereas form II is expressed under reducing conditions (Badger and Bek, 2008). Although the co-occurrence of both CBB and reductive TCA cycles was suggested in the *Riftia* symbiont (Markert *et al.*, 2007), there is no sequence evidence for the alternative carbon-fixation pathways in the genome of scaly-foot gastropod symbiont.

^{13}C incorporation

We incubated gastropods onboard the ship with ^{13}C-bicarbonate under hydrogen/sulfur-oxidizing conditions. The ^{13}C incorporation occurred even in the absence of H_2 or sulfide (Table 2), suggesting the presence of stored energy source such as intracellular S° observed in marine sulfur-oxidizing *Gammaproteobacteria* (Brock and Schulz-Vogt, 2011). The above-mentioned repertoire of sulfur-oxidation genes in the endosymbiont, that is, the presence of *dsr* genes and the absence of *soxCD* genes, suggests that the bacterium forms sulfur globules (Friedrich *et al.*, 2005). The δ^{13}C values fluctuated highly, probably due to individual differences in metabolic status, and thus their availability/preference of energy source provided was unclear from this experiment (Table 2). The δ^{13}C values, however, do indicate that all tissue samples, except for those from the killed control are enriched in 13C, demonstrating symbiont chemoautotrophy (Table 2). Compared with similar tracer experiments (Watsuji *et al.*, 2010, 2012), the degree of ^{13}C enrichment was apparently modest. Future tracer experiments should be performed at higher temperatures and for longer durations. In addition, the effect of hydrostatic pressure should also be addressed.

Host hemolymph glycan analysis

Glycans are highly sensitive biomarkers mirroring the cell/organism status (Varki, 1993; Gagneux and Varki, 1999). We therefore analyzed hemolymph *N*-glycans of snails incubated in the presence of H_2 and/or sulfur compounds to assess their response to

Table 2 Stable isotope composition of scaly-foot gastropods incubated with ^{13}C-bicarbonate

Incubation conditons[a,b]	*δ^{13}C of tissue (Average ± s.d.)*		*Number of samples*	*Reference*
	Esophageal gland	*Foot*		
SW	+6.8	+12.9	1	This study
SW + 100 μM sulfide	+7.4	+6.0	1	This study
SW + 20% H_2 (v/v)	−11.3	+17.6	1	This study
SW + 100 μM sulfide + 20 % (v/v) H_2	+12.4 ± 9.5	−1.5 ± 1.3	2	This study
Control				
SW + 100 μM sulfide (dead)	−20.1	−17.6	1	This study
SW + 20% H_2 (v/v) (dead)	−20.5	−17.9	1	This study
Not incubated	−20.7 ± 0.9	−18.2 ± 0.6	4	Goffredi *et al.*, 2004

[a]SW, surface seawater containing 3mM ^{13}C-bicarbonate.
[b]Dead, snails killed by freezing were incubated.

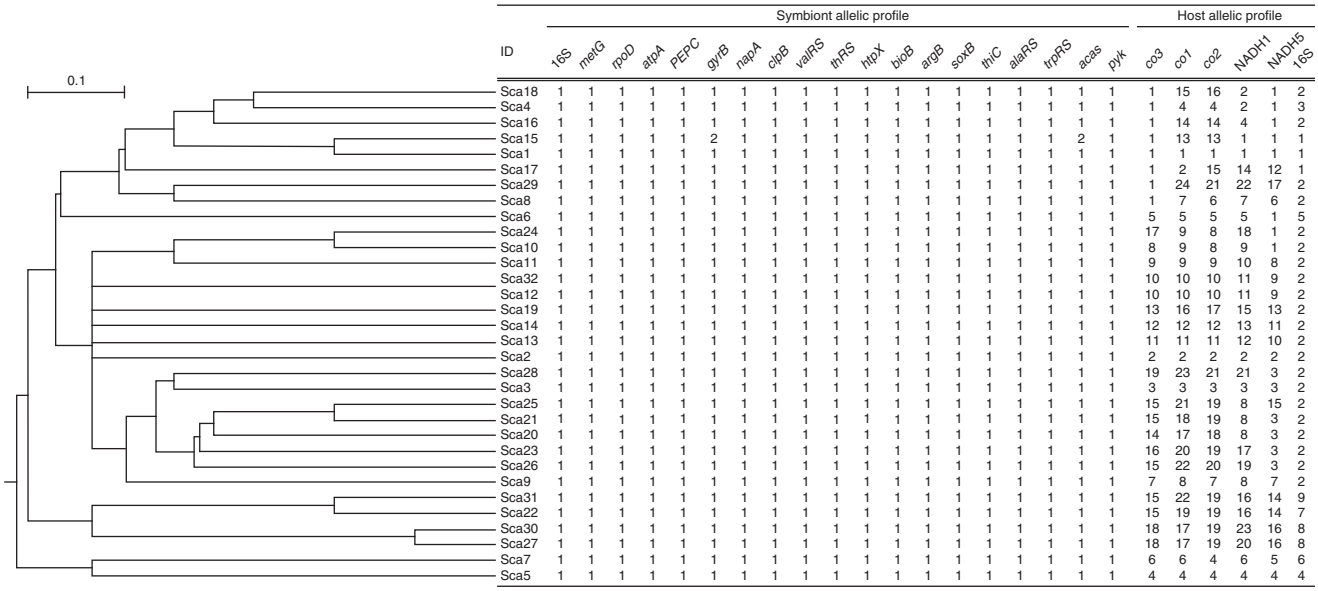

Figure 4 A unweighted pair group method with arithmetic mean dendrogram drawn on the basis of the alleles present at each of the 19 loci shown together with the symbiont and host allelic profiles.

the potential energy sources for symbionts. All of the prominent glycans were not significantly changed during the incubation without potential energy sources (Supplementary Figure S7). In addition, the amounts of two glycans, that is, IDs, 1826 and 2160, remained constant under any incubation conditions. In contrast, other hemolymph *N*-glycan profile changed significantly when snails were incubated with sulfur compounds and H_2 (Supplementary Figure S7). In particular, the amount of 1664 glycan changed significantly almost in all culture conditions, including the H_2-oxidizing condition. Although the glycome data indicated the gastropod cells quickly responded to both H_2 and sulfur compounds, much remains to be studied, including whether the substrates affect directly or indirectly (symbionts first and then host), and how proteins in hemolymph are changed.

Genetic polymorphism

Previous studies showed that increased genomic plasticity might confer a competitive advantage enabling deep-sea vent chemoautotrophs to thrive in ever-changing steep physical–chemical gradients (Nakagawa *et al.*, 2005, 2007). We sequenced a total of 19 genes including one pseudogene (*napA*) for each symbiont extracted from 32 different individuals. Unexpectedly, the endosymbiont exhibited extreme genetic homogeneity, showing only two synonymous changes in 19 different genes (13 810 positions in total) determined for 32 individual gastropods (Figure 4). In contrast, host mitochondria genes were highly diverged, and 31 haplotypes were identified in 32 individuals (Figure 4). Silent mutations of *gyrB* and *acas* of symbiont were exclusively detected in a single snail (ID, Sca15); however, its host haplotype was not unique,

suggesting the horizontal transmission of endosymbiont (Figure 4). The fate of genetic variations depends both on the fitness effect (selection coefficient) and on the effective population size. Effective population sizes should be small in vertically transmitted symbionts, because the symbionts experience a population bottleneck during transmission from one host generation to the next (Moran, 1996). This host-mediated reduction in population size must have consequences for their evolution, including high levels of random genetic drift (stochastic changes in gene frequency) and a faster rate of substitution at nearly neutral sites (Ohta, 1987, Burke and Moran, 2011). Therefore, extreme homogeneity of gastropod symbiont population again rejected their vertical transmission. Horizontally transmitted symbionts are specifically selected and taken up from the environment anew by each host generation. We consider that the effective symbiont selection by host gastropod overcomes the effect of random genetic drift in the symbiont population. Previous studies also reported slow rates of nucleotide substitution in the 16S rRNA gene of various environmentally acquired symbionts, reflecting efficient purifying selection (Peek *et al.*, 1998). Likewise, highly clonal population has also been reported for the environmentally transmitted, symbiotic sulfur-oxidizing gammaproteobacterium of ciliate *Zoothamnium* (Rinke *et al.*, 2009).

Conclusions

Our genome analysis of gastropod endosymbiont gives new insights into symbiosis, as well as into an ability to survive in deep-sea vents. Many findings must be further elucidated by experimental studies. As the gastropod can be kept alive at least for several weeks, the availability of the genome sequence will allow the detailed assessment of their ecophysiology when used in combination with high-throughput techniques such as microarray, proteome and metabolome analyses. As additional genome sequences become available, it will be of particular interest to carry out more detailed analyses of the common and specific genes corresponding to the lifestyles, that is, (i) free-living vs symbiotic in deep-sea hydrothermal environments and (ii) gamma- vs epsilon-*Proteobacteria*. This would help to clarify the factors controlling host–symbiont combinations in deep-sea vents.

Conflict of Interest

The authors declare no conflict of interest.

Acknowledgements

We would like to thank the captain and crew of the R/V *Yokosuka* and *Shinkai 6500* for helping us to obtain samples. This work was partially supported by Grant-in-Aid for Science Research (No. 23687004, 20109005) from the Ministry of Education, Culture, Sports, Science and Technology of Japan.

References

Alexeyenko A, Tamas I, Liu G, Sonnhammer E. (2006). Automatic clustering of orthologs and inparalogs shared by multiple proteomes. *Bioinformatics* **22**: e9–e15.

Altschul S, Madden T, Schäffer A, Zhang J, Zhang Z, Miller W *et al.* (1997). Gapped BLAST and PSI-BLAST: a new generation of protein database search programs. *Nucleic Acids Res* **25**: 3389–3402.

Amano M, Yamaguchi M, Takegawa Y, Yamashita T, Terashima M, Furukawa J *et al.* (2010). Threshold in stage-specific embryonic glycotypes uncovered by a full portrait of dynamic *N*-glycan expression during cell differentiation. *Mol Cell Proteomics* **9**: 523–537.

Anantharaman K, Breier J, Sheik C, Dick G. (2013). Evidence for hydrogen oxidation and metabolic plasticity in widespread deep-sea sulfur-oxidizing bacteria. *Proc Natl Acad Sci USA* **110**: 330–335.

Arsène-Ploetze F, Koechler S, Marchal M, Coppée J, Chandler M, Bonnefoy V *et al.* (2010). Structure, function, and evolution of the *Thiomonas* spp. genome. *PLoS Genet* **6**: e1000859.

Badger M, Bek E. (2008). Multiple rubisco forms in proteobacteria: their functional significance in relation to CO_2 acquisition by the CBB cycle. *J Exp Bot* **59**: 1525–1541.

Bateman A, Birney E, Durbin R, Eddy S, Finn R, Sonnhammer E. (1999). Pfam 3.1: 1313 multiple alignments and profile HMMs match the majority of proteins. *Nucleic Acids Res* **27**: 260–262.

Beinart R, Sanders J, Faure B, Sylva S, Lee R, Becker E *et al.* (2012). Evidence for the role of endosymbionts in regional-scale habitat partitioning by hydrothermal vent symbioses. *Proc Natl Acad Sci USA* **109**: E3241–E3250.

Besemer J, Borodovsky M. (2005). GeneMark: web software for gene finding in prokaryotes, eukaryotes and viruses. *Nucleic Acids Res* **33**: W451–W454.

Brock J, Schulz-Vogt H. (2011). Sulfide induces phosphate release from polyphosphate in cultures of a marine *Beggiatoa* strain. *ISME J* **5**: 497–506.

Burke G, Moran N. (2011). Massive genomic decay in Serratia symbiotica, a recently evolved symbiont of aphids. *Genome Biol Evol* **3**: 195–208.

Caspi R, Foerster H, Fulcher C, Kaipa P, Krummenacker M, Latendresse M *et al.* (2008). The MetaCyc database of metabolic pathways and enzymes and the BioCyc collection of pathway/genome databases. *Nucleic Acids Res* **36**: D623–D631.

Cooper C, Gasteiger E, Packer N. (2001). GlycoMod—a software tool for determining glycosylation compositions from mass spectrometric data. *Proteomics* **1**: 340–349.

D'mello R, Hill S, Poole R. (1996). The cytochrome bd quinol oxidase in *Escherichia coli* has an extremely high oxygen affinity and two oxygen-binding haems: implications for regulation of activity in vivo by oxygen inhibition. *Microbiology* **142**: 755–763.

Dahl C, Engels S, Pott-Sperling A, Schulte A, Sander J, Lübbe Y *et al.* (2005). Novel genes of the dsr gene cluster and evidence for close interaction of Dsr proteins during sulfur oxidation in the phototrophic sulfur bacterium *Allochromatium vinosum*. *J Bacteriol* **187**: 1392–1404.

Distel D, Lee H, Cavanaugh C. (1995). Intracellular coexistence of methano- and thioautotrophic bacteria in a hydrothermal vent mussel. *Proc Natl Acad Sci USA* **92**: 9598–9602.

Dulla G, Go R, Stahl D, Davidson S. (2012). *Vermiphrobacter eiseniae* type IV pili and flagella are required to colonize earthworm nephridia. *ISME J* **6**: 1166–1175.

Friedrich C, Bardischewsky F, Rother D, Quentmeier A, Fischer J. (2005). Prokaryotic sulfur oxidation. *Curr Opin Microbiol* **8**: 253–259.

Gagneux P, Varki A. (1999). Evolutionary considerations in relating oligosaccharide diversity to biological function. *Glycobiology* **9**: 747–755.

Gardebrecht A, Markert S, Sievert S, Felbeck H, Thürmer A, Albrecht D *et al.* (2012). Physiological homogeneity among the endosymbionts of *Riftia pachyptila* and *Tevnia jerichonana* revealed by proteogenomics. *ISME J* **6**: 766–776.

Goffredi S, Warén A, Orphan V, Van Dover C, Vrijenhoek R. (2004). Novel forms of structural integration between microbes and a hydrothermal vent gastropod from the Indian Ocean. *Appl Environ Microbiol* **70**: 3082–3090.

Hirokawa T, Boon-Chieng S, Mitaku S.. (1998). SOSUI: classification and secondary structure prediction system for membrane proteins. *Bioinformatics* **14**: 378–379.

Hügler M, Wirsen C, Fuchs G, Taylor C, Sievert S. (2005). Evidence for autotrophic CO2 fixation via the reductive tricarboxylic acid cycle by members of the epsilon subdivision of proteobacteria. *J Bacteriol* **187**: 3020–3027.

Kappler U, Bennett B, Rethmeier J, Schwarz G, Deutzmann R, McEwan A *et al.* (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.

Kleiner M, Petersen J, Dubilier N. (2012). Convergent and divergent evolution of metabolism in sulfur-oxidizing symbionts and the role of horizontal gene transfer. *Curr Opin Microbiol* **15**: 621–631.

Kuwahara H, Yoshida T, Takaki Y, Shimamura S, Nishi S, Harada M *et al.* (2007). Reduced genome of the thioautotrophic intracellular symbiont in a deep-sea clam, *Calyptogena okutanii*. *Curr Biol* **17**: 881–886.

Lowe T, Eddy S. (1997). tRNAscan-SE: a program for improved detection of transfer RNA genes in genomic sequence. *Nucleic Acids Res* **25**: 955–964.

Maezawa K, Shigenobu S, Taniguchi H, Kubo T, Aizawa S, Morioka M. (2006). Hundreds of flagellar basal bodies cover the cell surface of the endosymbiotic bacterium *Buchnera aphidicola* sp. strain APS. *J Bacteriol* **188**: 6539–6543.

Markert S, Arndt C, Felbeck H, Becher D, Sievert S, Hügler M *et al.* (2007). Physiological proteomics of the uncultured endosymbiont of *Riftia pachyptila*. *Science* **315**: 247–250.

Moran N. (1996). Accelerated evolution and Muller's rachet in endosymbiotic bacteria. *Proc Natl Acad Sci USA* **93**: 2873–2878.

Nakagawa S, Takai K, Inagaki F, Hirayama H, Nunoura T, Horikoshi K *et al.* (2005). Distribution, phylogenetic diversity and physiological characteristics of epsilon-*Proteobacteria* in a deep-sea hydrothermal field. *Environ Microbiol* **7**: 1619–1632.

Nakagawa S, Takai K. (2008). Deep-sea vent chemoautotrophs: diversity, biochemistry and ecological significance. *FEMS Microbiol Ecol* **65**: 1–14.

Nakagawa S, Takaki Y, Shimamura S, Reysenbach A, Takai K, Horikoshi K. (2007). Deep-sea vent epsilon-proteobacterial genomes provide insights into emergence of pathogens. *Proc Natl Acad Sci USA* **104**: 12146–12150.

Naven T, Harvey D. (1996). Effect of structure on the signal strength of oligosaccharides in matrix-assisted laser desorption/ionization mass spectrometry on time-of-flight and magnetic sector instruments. *Rapid Commun Mass Spectrom* **10**: 1361–1366.

Newton I, Woyke T, Auchtung T, Dilly G, Dutton R, Fisher M *et al.* (2007). The *Calyptogena magnifica* chemoautotrophic symbiont genome. *Science* **315**: 998–1000.

Nielsen H, Engelbrecht J, Brunak S, von Heijne G. (1997). A neural network method for identification of prokaryotic and eukaryotic signal peptides and prediction of their cleavage sites. *Int J Neural Syst* **8**: 581–599.

Odom J, Peck Jr H. (1981). Hydrogen cycling as a general mechanism for energy coupling in the sulfate-reducing bacteria, *Desulfovibrio* sp. *FEMS Microbiol Lett* **12**: 47–50.

Ohta T. (1987). Very slightly deleterious mutations and the molecular clock. *J Mol Evol* **26**: 1–6.

Pallen M, Matzke N. (2006). From the origin of species to the origin of bacterial flagella. *Nat Rev Microbiol* **4**: 784–790.

Peek A, Vrijenhoek R, Gaut B. (1998). Accelerated evolutionary rate in sulfur-oxidizing endosymbiotic bacteria associated with the mode of symbiont transmission. *Mol Biol Evol* **15**: 1514–1523.

Petersen J, Wentrup C, Verna C, Knittel K, Dubilier N. (2012). Origins and evolutionary flexibility of chemosynthetic symbionts from deep-sea animals. *Biol Bull* **223**: 123–137.

Petersen J, Zielinski F, Pape T, Seifert R, Moraru C, Amann R *et al.* (2011). Hydrogen is an energy source for hydrothermal vent symbioses. *Nature* **476**: 176–180.

Remm M, Storm C, Sonnhammer E. (2001). Automatic clustering of orthologs and in-paralogs from pairwise species comparisons. *J Mol Biol* **314**: 1041–1052.

Rinke C, Schmitz-Esser S, Loy A, Horn M, Wagner M, Bright M. (2009). High genetic similarity between two geographically distinct strains of the sulfur-oxidizing symbiont *'Candidatus* Thiobios zoothamnicoli'. *FEMS Microbiol Ecol* **67**: 229–241.

Rio R, Symula R, Wang J, Lohs C, Wu Y, Snyder A *et al.* (2012). Insight into the transmission biology and species-specific functional capabilities of tsetse (Diptera: glossinidae) obligate symbiont Wigglesworthia. *MBio* **3**: e00240-11.

Robinson J, Polz M, Fiala-Medioni A, Cavanaugh C. (1998). Physiological and immunological evidence for two distinct C1-utilizing pathways in *Bathymodiolus puteoserpentis* (Bivalvia: Mytilidae), a dual endosymbiotic mussel from the Mid-Atlantic Ridge. *Mar Biol* **132**: 625–633.

Salzberg S, Delcher A, Kasif S, White O. (1998). Microbial gene identification using interpolated Markov models. *Nucleic Acids Res* **26**: 544–548.

Scott K, Sievert S, Abril F, Ball L, Barrett C, Blake R *et al.* (2006). The genome of deep-sea vent chemolithoautotroph *Thiomicrospira crunogena* XCL-2. *PLoS Biol* **4**: e383.

Suzuki Y, Kopp R, Kogure T, Suga A, Takai K, Tsuchida S *et al.* (2006). Sclerite formation in the hydrothermal-vent "scaly-foot" gastropod—possible control of iron sulfide biomineralization by the animal. *Earth Planet Sci Lett* **242**: 39–50.

Suzuki Y, Sasaki T, Suzuki M, Nogi Y, Miwa T, Takai K *et al.* (2005). Novel chemoautotrophic endosymbiosis between a member of the *Epsilonproteobacteria* and the hydrothermal-vent gastropod *Alviniconcha* aff. *hessleri* (Gastropoda: Provannidae) from the Indian Ocean. *Appl Environ Microbiol* **71**: 5440–5450.

Takai K, Campbell B, Cary S, Suzuki M, Oida H, Nunoura T *et al.* (2005). Enzymatic and genetic characterization of carbon and energy metabolisms by deep-sea hydrothermal chemolithoautotrophic isolates of *Epsilonproteobacteria*. *Appl Environ Microbiol* **71**: 7310–7320.

Takai K, Nakagawa S, Reysenbach A-L, Hoek J. (2006). Microbial ecology of mid-ocean ridges and back-arc basins. In: Christie DM, Fisher CR, Lee S-M, Givens S (eds) *Back-arc spreading systems—Geological, biological, chemical, and physical interactions. Geophysical Monograph Series* vol. 166. American Geophysical Union: Washington, DC, pp 185–213.

Takaki Y, Shimamura S, Nakagawa S, Fukuhara Y, Horikawa H, Ankai A *et al.* (2010). Bacterial lifestyle in a deep-sea hydrothermal vent chimney revealed by the genome sequence of the thermophilic bacterium *Deferribacter desulfuricans* SSM1. *DNA Res* **17**: 123–137.

Tatusov R, Galperin M, Natale D, Koonin E. (2000). The COG database: a tool for genome-scale analysis of protein functions and evolution. *Nucleic Acids Res* **28**: 33–36.

Toft C, Fares M. (2008). The evolution of the flagellar assembly pathway in endosymbiotic bacterial genomes. *Mol Biol Evol* **25**: 2069–2076.

Van Dover C, Humphris S, Fornari D, Cavanaugh C, Collier R, Goffredi S *et al.* (2001). Biogeography and ecological setting of Indian Ocean hydrothermal vents. *Science* **294**: 818–823.

Varki A. (1993). Biological roles of oligosaccharides: all of the theories are correct. *Glycobiology* **3**: 97–130.

Vignais P, Billoud B, Meyer J. (2001). Classification and phylogeny of hydrogenases. *FEMS Microbiol Rev* **25**: 455–501.

Warén A, Bengtson S, Goffredi S, Van Dover C. (2003). A hot-vent gastropod with iron sulfide dermal sclerites. *Science* **302**: 1007.

Watsuji T, Nakagawa S, Tsuchida S, Toki T, Hirota A, Tsunogai U *et al.* (2010). Diversity and function of epibiotic microbial communities on the galatheid crab, *Shinkaia crosnieri*. *Microbes Environ* **25**: 288–294.

Watsuji T, Nishizawa M, Morono Y, Hirayama H, Kawagucci S, Takahata N *et al.* (2012). Cell-specific thioautotrophic productivity of epsilon-proteobacterial epibionts associated with *Shinkaia crosnieri*. *PLoS One* **7**: e46282.

Woyke T, Teeling H, Ivanova N, Huntemann M, Richter M, Gloeckner F *et al.* (2006). Symbiosis insights through metagenomic analysis of a microbial consortium. *Nature* **443**: 950–955.

Xie W, Wang F, Guo L, Chen Z, Sievert S, Meng J *et al.* (2011). Comparative metagenomics of microbial communities inhabiting deep-sea hydrothermal vent chimneys with contrasting chemistries. *ISME J* **5**: 414–426.

Yamamoto M, Takai K. (2011). Sulfur metabolisms in epsilon- and gamma-proteobacteria in deep-sea hydrothermal fields. *Front Microbiol* **2**: 192.

Yao H, Dao M, Imholt T, Huang J, Wheeler K, Bonilla A *et al.* (2010). Protection mechanisms of the iron-plated armor of a deep-sea hydrothermal vent gastropod. *Proc Natl Acad Sci USA* **107**: 987–992.

Supplementary Information accompanies this paper on The ISME Journal website (http://www.nature.com/ismej)