The ISME Journal (2012) 6, 766–776

www.nature.com/ismej

ORIGINAL ARTICLE

Physiological homogeneity among the endosymbionts of *Riftia pachyptila* and *Tevnia jerichonana* revealed by proteogenomics

Antje Gardebrecht[1,2], Stephanie Markert[2,3], Stefan M Sievert[4], Horst Felbeck[5], Andrea Thürmer[6], Dirk Albrecht[2,7], Antje Wollherr[6], Johannes Kabisch[2,3], Nadine Le Bris[8,9], Rüdiger Lehmann[6], Rolf Daniel[6], Heiko Liesegang[6], Michael Hecker[2,3,7] and Thomas Schweder[1,2,3]

[1]*Department of Pharmaceutical Biotechnology, Ernst-Moritz-Arndt-University, Greifswald, Germany;* [2]*ZIK FunGene, Ernst-Moritz-Arndt-University, Greifswald, Germany;* [3]*Institute of Marine Biotechnology, Greifswald, Germany;* [4]*Biology Department, Woods Hole Oceanographic Institution, Woods Hole, MA, USA;* [5]*Marine Biology Research Division, Scripps Institution of Oceanography, San Diego, CA, USA;* [6]*Göttingen Genomics Laboratory, Georg-August-University, Göttingen, Germany;* [7]*Institute for Microbiology, Ernst-Moritz-Arndt-University, Greifswald, Germany;* [8]*Benthic Ecogeochemistry Laboratory, Université Pierre et Marie Curie, Banyuls-sur-Mer, France and* [9]*Laboratoire Environnement Profond, Ifremer, Brest, France*

The two closely related deep-sea tubeworms *Riftia pachyptila* and *Tevnia jerichonana* both rely exclusively on a single species of sulfide-oxidizing endosymbiotic bacteria for their nutrition. They do, however, thrive in markedly different geochemical conditions. A detailed proteogenomic comparison of the endosymbionts coupled with an *in situ* characterization of the geochemical environment was performed to investigate their roles and expression profiles in the two respective hosts. The metagenomes indicated that the endosymbionts are genotypically highly homogeneous. Gene sequences coding for enzymes of selected key metabolic functions were found to be 99.9% identical. On the proteomic level, the symbionts showed very consistent metabolic profiles, despite distinctly different geochemical conditions at the plume level of the respective hosts. Only a few minor variations were observed in the expression of symbiont enzymes involved in sulfur metabolism, carbon fixation and in the response to oxidative stress. Although these changes correspond to the prevailing environmental situation experienced by each host, our data strongly suggest that the two tubeworm species are able to effectively attenuate differences in habitat conditions, and thus to provide their symbionts with similar micro-environments.

The ISME Journal (2012) **6,** 766–776; doi:10.1038/ismej.2011.137; published online 20 October 2011
Subject Category: microbe–microbe and microbe–host interactions
Keywords: chemoautotrophy; endosymbiosis; hydrothermal vent; metagenome; proteomics; tubeworms

Introduction

Deep-sea hydrothermal vents are ephemeral features that are frequently perturbed by volcanic eruptions. These eruptions, such as the ones recorded on the East Pacific Rise near 9°50′N in 1991 and again in 2005–2006 (Haymon *et al.*, 1993; Tolstoy *et al.*, 2006; Cowen *et al.*, 2007), have a tremendous impact on the ecosystems present at these sites. Faunal assemblages are being destroyed and nascent diffuse flows develop (Von Damm, 2000), allowing the recolonization of the area by invertebrates, such as tubeworms and bivalves (Hessler *et al.*, 1988; Shank *et al.*, 1998). In this sequential process, the tubeworm *Tevnia jerichonana* (Jericho worm; Jones, 1985) stands out as an early colonizer of post-eruptive basalt-hosted vents (Mullineaux *et al.*, 2000). The pioneer species colonizes areas of vigorous diffuse hydrothermal flow characterized by microaerobic conditions, and relatively high levels of sulfide and, potentially, heavy metals (Von Damm *et al.*, 1995; Shank *et al.*, 2006; Moore *et al.*, 2009; Nees *et al.*, 2009). As the flux decreases and physicochemical properties of the fluid change over time, individuals of *Riftia pachyptila* (Giant tubeworm; Jones, 1981a) begin to colonize the same vent sites (Shank *et al.*, 1998). *R. pachyptila* preferentially thrives in zones characterized by higher

Correspondence: S Markert, Institute of Marine Biotechnology, Walther-Rathenau-Straße 49A, 17489 Greifswald, Germany.
E-mail: stephanie.markert@uni-greifswald.de
or T Schweder, Ernst-Moritz-Arndt-University of Greifswald, Pharmaceutical Biotechnology, Friedrich-Ludwig-Jahn-Straße 17, 17487 Greifswald, Germany.
E-mail: schweder@uni-greifswald.de
Received 7 February 2011; revised 23 August 2011; accepted 25 August 2011; published online 20 October 2011

oxygen and lower sulfide concentrations, eventually replacing *T. jerichonana* at these vents (Nees *et al.*, 2009). However, in contrast to their different colonization behavior, the two closely related siboglinid tubeworms (Pleijel *et al.*, 2009) were shown to harbor conspecific chemoautotrophic endosymbionts (Edwards and Nelson, 1991; Cary *et al.*, 1993; Di Meo *et al.*, 2000). Moreover, very similar 16S ribosomal RNA (rRNA) phylotypes could also be detected in the vent-dwelling siboglinids *Oasisia alvinae* and *Ridgeia piscesae* (Feldman *et al.*, 1997; Thornhill *et al.*, 2008), which implies that the same or closely related chemosynthetic symbionts are able to establish a symbiosis with different tubeworm genera. While relatively little is known about *T. jerichonana* and its symbionts, the *Riftia* symbiosis has been studied extensively in the past: adult worms lack a mouth and digestive tract (Jones, 1981b) and are therefore completely dependent on the symbionts, which were taxonomically classified as Gammaproteobacteria (Distel *et al.*, 1988). Substrates for bacterial chemosynthesis, that is, oxygen, sulfide and carbon dioxide, are taken up from the diffuse flow fluids through the worm's plume and delivered to the symbionts, which are enclosed in specialized host cells (bacteriocytes) in an organ termed trophosome (Hand, 1987; Bright and Sorgo, 2003). Bacterial sulfide oxidation provides the energy that fuels CO_2 fixation into organic matter (Cavanaugh *et al.*, 1981; Felbeck, 1981; Felbeck and Somero, 1982; Fisher and Childress, 1984; Childress *et al.*, 1991; Girguis and Childress, 2006). With the publication of its metagenome (Robidart *et al.*, 2008), the name *Candidatus* Endoriftia persephone was proposed for the *Riftia* symbiont. The metagenome sequence facilitated a first proteomic analysis (Markert *et al.*, 2007), which provided a detailed description of the physiological status of the *Riftia* symbiont.

Here, a cultivation-independent proteogenomic approach was pursued to elucidate and compare the physiological conditions of the presumably conspecific symbionts in the two different tubeworm species *R. pachyptila* and *T. jerichonana*. For this purpose, enriched symbiont fractions were used to sequence the metagenome of the *Tevnia* symbiont as well as to obtain a new, improved *Riftia* symbiont metagenome sequence. The metagenomes and metaproteomes of the endosymbionts of both hosts were found to be highly similar. Only marginal differences were detectable in the expression patterns of the symbionts, suggesting that both hosts are providing a comparable internal environment to the symbionts, despite the markedly different physicochemical conditions measured in the immediate environment at the plume level of each host species.

Materials and methods

Sampling

Tubeworm specimens were collected from diffuse flow vents on the East Pacific Rise using DSV Alvin during two expeditions with R/V Atlantis in December 2007/January 2008 and in October 2008 (for details, see Supplementary Table S1). The bacterial symbionts were isolated from the host trophosome by density gradient centrifugation (Distel and Felbeck, 1988) following a protocol specifically adapted to the purification of *Riftia* symbionts (Felbeck and Jarchow, 1998). The symbiont-containing trophosome tissue was homogenized in imidazole-buffered saline (0.49 M NaCl, 0.03 M $MgSO_4$, 0.011 M $CaCl_2$, 0.003 M KCl, 0.05 M imidazole, pH 7.1, supplemented with 30% *R. pachyptila* blood) and subjected to Percoll gradient centrifugation. Owing to their elemental sulfur inclusions, the bacterial cells accumulate as a highly pure symbiont fraction in the bottom region of the gradient, while host cell debris stays at the top. To circumvent oxidative stress caused by atmospheric air, the tissue disruption and symbiont isolation were performed anoxically. The isolated *Riftia* and *Tevnia* symbionts were washed in TE buffer (0.1 M Tris pH 7.5 and 1 mM EDTA pH 8.0) or imidazole-buffered saline and then stored at $-80\,^{\circ}C$ until further processing. The protein reference maps of this study originated from bacterial samples of three representatives of each tubeworm species, which were collected during the same Alvin dive at a single sampling site. On average, the trophosome tissue of the *Tevnia* specimens harvested during this dive was slightly lighter in color (indicating a slightly higher sulfur content) than that of the *Riftia* specimens (see Supplementary Table S1).

In situ *habitat chemistry*

Free sulfide and pH were measured at different levels within the mixed *Tevnia* and *Riftia* clump to be sampled during two dives several days before collection (Supplementary Table S1). Measurements were performed in four different locations: (1) among the tubes and plumes of *Tevnia*, (2) at the level of some *Riftia* plumes, (3) at the base of tubes and (4) among tubes, close to *Tevnia* plumes. The probe was positioned using the manipulator arm of the submersible and maintained for about 3–5 min at each position. The sensor is composed of a miniaturized glass electrode sensitive to pH (diameter: 1.5 mm) and an Ag/Ag_2S electrode sensitive to S^{2-} (diameter: 0.8 mm), attached to autonomous data loggers as described by Le Bris *et al.* (2001). Calibrations were performed in the laboratory, before and after recovery, allowing the conversion of electrode potentials into pH and log $[HS^-]$. Free sulfide concentrations (that is, $[H_2S]+[HS^-]$) were calculated from these data using the thermodynamic constants given in Rickard and Luther (2007). The recording rate was set to one measurement every 5 s, allowing an average value and s.d. to be determined for each location.

DNA extraction and genome sequencing

Bacterial cells used for DNA extraction were solely derived from light green (that is, sulfur-rich) trophosomal host tissue. The total DNA was

prepared with the MasterPure DNA Purification Kit (Epicentre, Madison, WI, USA) according to the manufacturer's instructions. In all, 454 libraries were constructed using the GS FLX General Shotgun DNA Library Preparation Method Manual (Roche, Mannheim, Germany). As the previously published *Riftia* symbiont metagenome (Robidart *et al.*, 2008) is highly fragmented and thus not suitable for comparative analyses, a new metagenome sequence was obtained for the *Riftia* symbiont, in addition to the sequencing of the *Tevnia* symbiont metagenome. To verify if potential genomic differences between the *Riftia* symbiont ('*Riftia* 1 symbiont') and the *Tevnia* symbiont were due to the symbionts' origin from different host species or rather sequencing artifacts, an additional *Riftia* symbiont sample was sequenced and included in the genomic comparisons ('*Riftia* 2 symbiont'). The *Riftia* 1 symbiont library was sequenced with the Genome Sequencer FLX (Roche) using the standard chemistry, whereas Titanium chemistry was used for the sequencing of the *Riftia* 2 symbiont and of the *Tevnia* symbiont libraries. Two large lanes were used for the sequencing of the *Riftia* 1 symbiont library. The *Riftia* 2 symbiont and the *Tevnia* symbiont were each sequenced on one medium lane. Shot gun reads (average read lengths for the *Riftia* 1 symbiont: 185 bp, for the *Riftia* 2 symbiont: 288 bp and for the *Tevnia* symbiont: 242 bp) were assembled *de novo* using the Roche Newbler assembly software (Margulies *et al.*, 2005). The resulting draft genomes served as database for proteome identifications and for all DNA comparisons.

Gene prediction, annotation and genome accession
Automated detection of open-reading frames and annotation of all draft genomes were automatically performed by means of the RAST service pipeline (Aziz *et al.*, 2008). Annotation results for genes of those proteins identified on two-dimensional (2D) gels were manually verified using the NCBI Protein BLAST tool (blastp, Altschul *et al.*, 1997). The metagenome sequences of both symbionts are publicly available via the GenBank database under accession no. 60887 (*Tevnia* symbiont) and no. 60889 (*Riftia* 1 symbiont), whereas the additional metagenome of the *Riftia* 2 symbiont is accessible at ftp://www.g2l.bio.uni-goettingen.de.

Multiple genome comparisons
Global nucleotide sequence alignments of the internal transcribed spacer (ITS) and 16S rRNA gene region and of selected key genes with relevant metabolic functions were performed using the Geneious Pro tool (Drummond *et al.*, 2010). For whole genome computations of homologies, the BiBaG software (Bidirectional BLAST for identification of bacterial pan and core genomes, Göttingen Genomics Laboratory, Germany) was applied. The pairwise BiBlast alignment is focused on substrings (sequence regions) of genomes, while the global, dynamic Needleman–Wunsch algorithm compares every residue of the DNA and thereby considers the length of sequences.

Protein extraction, 2D electrophoresis and reference maps
Symbiont cells were disrupted by sonication in lysis buffer (10 mM Tris pH 7.5, 10 mM EDTA pH 8.0) containing protease inhibitors (1.7 mM phenylmethanesulfonylfluoride, Roche Complete Inhibitor) and the cell debris was removed by centrifugation. To ensure an adequate sample quality, extracts of intracellular soluble proteins were subjected to precipitation using the 2D Clean-Up Kit (GE Healthcare, München, Germany). The symbionts' protein extracts were separated in 2D sodium dodecyl sulfate-polyacrylamide gels according to their isoelectric point (pI) in the pI range 4–7 and molecular weight. The protein spots were stained with the fluorescent Krypton Protein Stain (Pierce, Bonn, Germany). Using the software Delta2D (Decodon, Greifswald, Germany), spots were detected and quantitatively analyzed. The colored protein maps of the *Riftia* and *Tevnia* symbionts were overlaid and compared as dual-channel images. Delta2D calculates protein spot volumes as percent of all proteins on the same gel (% vol). Spot ratios ('fold change') are the factor by which spot volumes differ when comparing two gels. Spots with an at least 1.5-fold increased volume (ratio >1.5 or <-1.5, respectively) compared with the respective protein spot of the other symbiont fraction were labeled on the dual-channel image and considered in this study.

Mass spectrometry and protein identification
All visible protein spots were automatically excised from the 2D gels (Ettan Spot Picker, GE Healthcare), digested with trypsin and spotted on a matrix-assisted laser desorption/ionization (MALDI)-target (Ettan Spot Handling Workstation, GE Healthcare). High-throughput MALDI- time-of-flight (TOF) measurements combined with MALDI-TOF-TOF tandem mass spectrometry were carried out on the 4800 MALDI TOF/TOF Analyzer (Applied Biosystems, Darmstadt, Germany). During the MALDI-TOF analysis, spectra were recorded in a mass range from 900 to 3700 Da with a focus mass of 2000 Da. The mass spectrometry (MS)/MS screen measured the three strongest peaks from the TOF-spectra. After calibration, peak lists were created based on the script of the GPS Explorer Software (Applied Biosystems) Version 3.6. Peptide assignment was realized by comparison with the corresponding symbiont sequence database in FASTA format using the Mascot search engine Version 2.1.04 (Matrix Science Ltd, Boston, MA, USA). Proteins that met the following requirements were regarded as identified on the reference maps of both symbionts: sequence coverage of at least 30%, minimum

peptide count of 2 and probability-based protein Mowse score value of 75 or higher.

Results and Discussion

In situ *habitat chemistry measurements*

Our habitat chemistry measurements revealed that the plumes of the *Tevnia* specimens at the sampling site were exposed to comparatively high sulfide concentrations (~2 mM), moderate temperature (20 °C) and quite low pH (5.4, see Table 1). As previously described (Le Bris *et al.*, 2006), this low pH suggests a very low seawater contribution, implying low oxygen conditions. In contrast to that, low temperature (2.6 °C), high pH (7.4) and a low sulfide concentration (<0.01 mM) were detected at *Riftia* plume level, indicating a high seawater contribution (maximum oxygen concentration in seawater: 110 μM (Johnson, 1988)). Relatively high temperatures (66 °C), low pH (4.9) and high sulfide concentrations (1.7 mM) indicative of a proximal fluid source were detected at the base of the tubes. These values directly reflect the relative positions of *Riftia* and *Tevnia* specimens at the sampling site. The *Riftia* plumes were more distant from the basalt rock and thus from the hydrothermal diffuse flow at the base of the bush, while *Tevnia* tubes appeared more curved and their plumes were located closer to the hydrothermal outflow (Supplementary Figure S1). The gradient observed at plume level between the tube worm species in the sampled bush is consistent with the more general results of Nees *et al.* (2009) obtained from several tubeworm aggregations, who also showed that *Tevnia* plumes were exposed to both higher sulfide and lower oxygen concentrations compared with the *Riftia* plumes when both tube worm species were co-existing at 9°N East Pacific Rise in 2007.

General genome characteristics

Resequencing and assembly of the *Riftia* symbiont's metagenome resulted in a considerable reduction of individual DNA fragments (316 large contigs, that is, >500 bp, in the *Riftia* 1 symbiont sequence, 414 large contigs in the *Riftia* 2 symbiont sequence; Table 2) in comparison with the previously published Sanger sequence (2472 contigs, Robidart *et al.*, 2008). While Robidart *et al.* reported a genome size of 3.2 Mb for Cand. *E. persephone*, the draft

Table 1 Physico-chemical conditions at different locations in the tubeworm clump from which individual tubeworms were subsequently collected for the proteomic work

Location of probe	*(1) Among* Tevnia *tubes and plumes*	*(2) At* Riftia *plume*	*(3) At base of* Riftia *and* Tevnia *tubes*	*(4) Close to* Tevnia *plumes*
Date	2007/12/30	2008/01/01	2008/01/01	2008/01/01
Temperature (°C)	19.8	2.6	65.9	19.4
±	1.1	0.5	1.5	0.8
pH	5.5	7.4	4.9	5.4
±	<0.1	0.1	0.2	<0.1
Free sulfide (mM)	0.7	<0.01	1.7	2.1
±	0.06	<0.01	0.8	0.5
Total no. of measurements (n)	31	48	29	49

Average values and s.d. obtained for 3–5 min scans combining temperature, pH and free sulfide ($[H_2S] + [HS^-]$) are given. These data sets represent a few discrete positions and illustrate the strong chemical and thermal gradients experienced at the scale of the tubeworm aggregation. Owing to the position of the worm plumes in this gradient, systematic differences in the habitat conditions of the two species can be hypothesized. These observations are consistent with the more complete measurement series previously presented by Nees *et al.* (2009). A rather large s.d. for sulfide reflects the limited precision of sulfide potentiometric measurements at low pH and high sulfide.

Table 2 Overview of the symbiont metagenomes

	Riftia *1 symbiont*	Riftia *2 symbiont*	Tevnia *symbiont*
Draft genome size (large contigs, Mbp)	3.68	3.71	3.64
GC content (%) of large contigs	58.6	58.6	58.2
Coverage	25 ×	13 ×	15 ×
Number of shotgun reads total	467 070	205 880	212 833
Number of large contigs	316	414	184
Average size (kb) of large contigs	11.6	9.0	19.8
N50 size (kb) of large contigs	28.4	24.6	92.7
Size (kb) of largest contig	95.4	76.8	248.7
Q40 value (quality of reads, %)	99.4	98.8	99.3
Protein-coding genes (CDS)	3209	3515	3230
RNA-coding genes	45	51	47

General features of the 454 genome sequences obtained from endosymbionts of the deep-sea tubeworms *Riftia pachyptila* and *Tevnia jerichonana*. For the genomic comparison, *Riftia* symbiont sequences were obtained for bacteria from two different *Riftia* tubeworms. Large contigs exhibit a minimum size of 500 kb. Short genomic fragments (<500 bp) were excluded from the sequences and were not used for annotation.

genome size for the two *Riftia* symbiont fractions sequenced in this study averages 3.7 Mb. In all, 454 pyrosequencing of the *Tevnia* symbiont fraction yielded fewer and larger genome segments (184 large contigs) and a genome size of 3.64. The sequencing of the symbiont metagenomes yielded a 25-fold coverage for the *Riftia* 1 symbiont, a 13-fold coverage for the *Riftia* 2 symbiont and a 20-fold coverage for the *Tevnia* symbiont. Highest read quality was achieved for the *Tevnia* (Q40: 99.3%) and the *Riftia* 1 symbiont sequencing (Q40: 99.4%). Subsequent automated open-reading frame identification and function assignment using the RAST pipeline (Aziz *et al.*, 2008) predicted 3230 protein coding sequences (CDS) for the *Tevnia* symbiont metagenome, whereas 3209 CDS were determined for the *Riftia* 1 symbiont metagenome and 3515 CDS, respectively, for the *Riftia* 2 symbiont sequence. Generally, short genomic fragments (<500 bp), which resulted from repetitive genomic regions and host DNA contaminations, were excluded from the sequences, as they hampered the search for overlaps during assembly and reconstruction of complete genomes and did not have a significant probability to contain untruncated annotatable genes. Contigs smaller than 500 bp were therefore also excluded from the annotation.

Metagenome comparison

As the metagenomes could not be closed, it is not possible to unambiguously ascertain whether genomic differences were caused artificially by the remaining sequence gaps or are in fact of biological nature, that is, because of the symbionts' origin from different host species. To circumvent this problem, we included the symbiont metagenome from an additional *Riftia* specimen (*Riftia* 2 symbiont) in our evaluation. Single gene comparisons of *Riftia* 1 symbiont, *Riftia* 2 symbiont and the *Tevnia* symbiont showed only minimal nucleotide sequence differences (average heterogeneity 0.1%, Supplementary Table S2 and Supplementary Figure S2) among key genes involved in main metabolic pathways like sulfide oxidation (*aprA/aprB, dsrA/dsrB, sopT*) and carbon fixation (*cbbM, pgk, acnA, icd, sdhA1*), and between genes related to the response to oxidative stress (*ahpC, sodB*). An alignment of the 16S rRNA and internal transcribed spacer regions determined an identity of 100% between the bacterial communities from the three siboglinid tubeworm samples (Supplementary Table S2, Supplementary Figure S3) and thus supports the assumption that *R. pachyptila* and *T. jerichonana* share the same endosymbiotic bacterial species (Edwards and Nelson, 1991; Di Meo *et al.*, 2000). This hypothesis is furthermore confirmed by comparative alignments of the above mentioned metabolic key genes with their respective homologs from the published *Riftia* symbiont metagenome (Robidart *et al.*, 2008), which yielded an average homogeneity of 99.6% (Supplementary Figure S2, Supplementary Table S2). The comparison with previously published *Riftia* and *Tevnia* symbiont internal transcribed spacer and 16S rRNA sequences (Di Meo *et al.*, 2000; Harmer *et al.*, 2008; Robidart *et al.*, 2008) revealed 99.2% (internal transcribed spacer) and 99.3% (16S rRNA) homogeneity to the sequences presented in this study (see Supplementary Figure S3). On the level of whole genome comparisons the degree of DNA variations between *Riftia* 1 symbiont, *Riftia* 2 symbiont and the *Tevnia* symbiont was determined by a combined BiBlast search and Needleman–Wunsch alignment (Needleman and Wunsch, 1970) using the BiBaG software. The bidirectional Blast screenings computed about 2760 matching protein coding genes (CDS), which corresponds to 78–85% of the total metagenomes (depending on the varying number of genes in each metagenome; *E*-value cut-off of 10^{-5}). The subsequent global alignment (Figure 1) determined that about 70% (2531–2551) of all genes could be assigned to the core genome (homology score >50%). Within this triple comparison, the metagenome of the *Riftia* 2 symbiont exhibited the lowest similarity, probably due to the smaller contig sizes (9 kb in average) and the resulting increased gene truncation (414 contigs). Overall, the *Tevnia* symbiont metagenome showed more resemblance to both *Riftia* symbiont

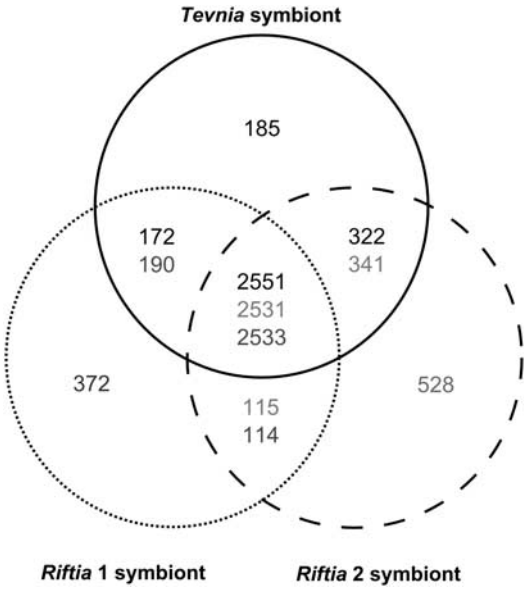

Figure 1 Metagenome comparison. The Venn diagram depicts the number of genes shared (global alignment score >50%) or specific for the three symbionts. All bidirectional BLAST hits were globally screened by means of the Needleman–Wunsch algorithm. 2531–2551 open-reading frames (ORFs) are highly conserved among the three symbiont metagenomes and can be considered as orthologs (in the center). Variations in core genome size result from gene content redundancy in the metagenomes caused by paralogs and domain duplications. The number of 'unique' genes is primarily attributed to the sequence fragmentation caused by repetitive elements and unavoidable host DNA contaminations. Black numbers, solid line: *Tevnia* symbiont; dark gray numbers, dotted line: *Riftia* 1 symbiont; light gray numbers, dashed line: *Riftia* 2 symbiont.

sequences (2723 CDS of the *Tevnia* symbiont metagenome were shared with *Riftia* 1 symbiont, 2873 CDS with *Riftia* 2 symbiont) than the *Riftia* symbiont metagenomes among each other (2647 homologous CDS between *Riftia* 1 and *Riftia* 2 symbiont). The results suggest that genomic differences, such as the 'unique' genes (that is, genes that were exclusively identified in one of the three metagenomes: 6% for the *Tevnia* symbiont, 12% for the *Riftia* 1 symbiont and 15% for the *Riftia* 2 symbiont), were predominantly caused by the incompleteness of the 454-derived metagenome sequences, possibly resulting from differences in the quality of the symbiotic DNA used for the analyses.

Gene functions

The following paragraphs are intended to provide an overview of the symbionts' metabolic properties as encoded in the *Riftia* and *Tevnia* symbiont metagenomes. Supplementary Table S3 summarizes the key genes of major metabolic pathways as well as several genes, which are putatively involved in symbiont–host interactions and protein secretion.

Autotrophy and energy generation. Unlike other organisms, the symbionts are able to use two alternative pathways, the Calvin–Benson–Bassham cycle (Felbeck, 1981; see Proteomics section below) and the reductive tricarboxylic acid (rTCA) cycle (Markert *et al.*, 2007) for the synthesis of organic matter from CO_2. Potentially, most enzymatic reactions of the TCA cycle can be catalyzed by the same enzymes in the oxidative or the reductive direction. However, our metagenome data indicate the presence of separate sets of enzymes for either the oxidative or the reductive direction of the pathway. A gene encoding a citrate synthase, the key enzyme of the oxidative TCA cycle, for example, is clustered with genes coding for a malate dehydrogenase and a four subunit succinate dehydrogenase, which differ from the ones supposedly involved in the reductive TCA cycle. None of the genes in this cluster were found to be expressed (this study; Markert *et al.*, 2011), which indicates that the examined symbionts used the TCA cycle to fix CO_2, rather than to oxidize organic carbon. In contrast, the rTCA cycle key enzyme ATP citrate lyase (*aclAB*) genomically clusters with genes encoding an isocitrate dehydrogenase (*icd*), an aconitase (*acnA*) and a malate dehydrogenase (*mdh1*), suggesting that these enzymes—all of which were detected with high spot volumes on the protein level—function in the rTCA cycle. The clustered organization of these genes might facilitate the regulation of the pathways depending on environmental conditions. Details of the energy-generating sulfide oxidation pathway in the thioautotrophic symbionts are given in the proteomics section below and in Markert *et al.* (2011).

Nitrogen metabolism. With *in situ* concentrations ranging from 18.3 to 37.5 μM (Lee and Childress, 1994) nitrate is the most easily accessible form of inorganic nitrogen at hydrothermal vent sites. Besides bacterial nitrate assimilation, which provides the host with nitrogen required for biosynthesis (Girguis *et al.*, 2000), the *Riftia* symbionts have been proposed to be capable of using nitrate as an alternative electron acceptor under anaerobic conditions. The detection of nitrate respiration in purified *Riftia* symbionts as reported by Hentschel and Felbeck (1993) was confirmed by Pospesel *et al.* (1998), who measured nitrate concentrations of up to 1.5 mM in *Riftia* blood, suggesting that nitrate—in addition to oxygen—may play an important role for the symbionts' energy generation. However, a gene encoding a nitrous oxide reductase appeared to be absent in the previously published *Riftia* symbiont metagenome (Robidart *et al.*, 2008), which led to the conclusion that the symbionts lack the ability to perform canonical denitrification, that is, the reduction of nitrate to dinitrogen gas (N_2). Interestingly, in the present study all three symbiont metagenomes were found to encode all the enzymes needed to respire nitrate completely to N_2: the membrane-bound respiratory nitrate reductase NarGHJI, a cytochrome cd1/nitrite reductase (*nirNJHGLDFCTS*), a nitric-oxide reductase (*norBC*) and a nitrous-oxide reductase (*nosZ*). Expression of the NarGH subunits was detected in both the *Tevnia* and the *Riftia* symbiont proteome (similar % vol values of 0.3–0.4) in this study. Moreover, the expression of the repressor NarL, the transporter NarK, NirHFTS and NosZ were reported for the *Riftia* symbiont in an accompanying proteome analysis (Markert *et al.*, 2011), which was based on the *Riftia* 1 genome sequence established in this study. These results strongly imply that the *Riftia* and *Tevnia* symbionts indeed use nitrate for respiratory purposes under the given conditions. In addition to *narGHJI*, the *Riftia* and *Tevnia* symbiont metagenomes also include the *napFDAGHBC* operon, encoding the periplasmic dissimilatory nitrate reductase NapABC and associated electron carriers. The symbiotic Nap system shows similarity to the anaerobically induced *nap* operon of *Escherichia coli* (*E. coli*) that might compensate the low efficiency of the NarGHI enzyme under limiting nitrate concentrations (Stewart *et al.*, 2002). Clearly, further research is warranted to determine if in fact the symbionts are able to gain energy from denitrification, and to elucidate the details of this process.

Storage compounds. The detailed genomic examination led to the discovery of a gene annotated as cyanophycin synthetase (*cphA*) in all three symbiont genomes. This functionally interesting enzyme generates cyanophycin (multi-L-arginyl-poly-L-aspartate), a storage biopolymer containing both nitrogen and carbon (Ziegler *et al.*, 1998; Krehenbrink *et al.*, 2002). So far, glycogen and sulfur were the only storage compounds observed

in the deep-sea symbionts (Sorgo *et al.*, 2002; Pflugfelder *et al.*, 2005).

Oxidative stress response. In accordance with the early description of bacterial superoxide dismutase (FeSOD) activity and peroxidase activity in *Riftia* trophosome tissue (Blum and Fridovich, 1984), genes encoding the superoxide dismutase SodB and the alkylhydroperoxide reductase AhpC were identified in all three symbiont genomes in this study and were found to be expressed in both *Tevnia* and *Riftia* symbionts (see below). As predicted by Blum and Fridovich (1984), no catalase gene was detected in the symbiont metagenomes.

Horizontal gene transfer. Comparative analyses of the genomes revealed the existence of parts of an integrated F (fertility) plasmid in the *Riftia* 1 symbiont, which encodes several genes for horizontal gene transfer (especially the *tra* locus) that enable bacteria to generate pili necessary for conjugation. Horizontal gene transfer is suggested to be an important contribution to symbiont–host adaptation (Toft and Anderson, 2010). As already observed by Robidart *et al.* (2008), several transposases were detected in all three metagenome sequences. However, no reliable insertion sequence (IS) elements of known IS classes could be detected by means of the IS-Finder tool (http://www-is.biotoul.fr/) in any of the three metagenomes.

Host infection. The metagenome data presented in this study suggest that bacterial attachment to the worm surface during the infection process might be mediated by fimbriae, pili or even the flagellum (as shown for *Burkholderia pseudomallei* by Inglis *et al.*, 2003), and may involve a putative adhesin-like protein, hyalin or a fibronectin type III domain protein, all of which are encoded in the metagenomes (Supplementary Table S3). A cell wall-associated biofilm protein may possibly facilitate the colonization of the mucus layer that coats the tubeworm larvae (Nussbaumer *et al.*, 2006), while a putative hemolysin III-homolog (*hlyIII*, a virulence factor of *Staphylococcus aureus*; Bhakdi and Tranum-Jensen, 1991), a lysine 2,3-aminomutase (*kamA*, produces the cationic detergent beta lysine), and the colicin V production protein (*cvpA*, originally discovered in invasive *E. coli* strains; Waters and Crosa, 1991) may enable the bacteria to permeabilize the cuticle and epidermis cells of their host, allowing them to inter- and intracellularly migrate to the newly developed trophosome tissue (Bright and Bulgheresi, 2010).

Comparative proteomics

Protein reference maps provided detailed insights into the physiological conditions for the endosymbionts in both tubeworm species. The expression patterns were analyzed by generating dual-channel images from overlaid master gels of the intracellular *Riftia* and *Tevnia* symbiont metaproteomes (Figure 2). Repeated metaproteome analyses revealed a surprisingly reproducible, highly congruent spot distribution of the intracellular soluble protein fractions on 2D gels of the *Tevnia* and the *Riftia* symbiont. Only minor differences were detected in the relative abundance of key enzymes involved in sulfide oxidation, carbon fixation and oxidative stress response (Figure 3). To identify even small variations between the metabolic footprints of the *Tevnia* and the *Riftia* endosymbionts in their respective tubeworms, a very low ratio threshold of 1.5 (that is, an at least 1.5-fold decrease or increase in spot volume) was chosen. Only larger protein spots, which were detected with relative volumes of at least 0.1% vol on the average-fused images ($n = 3$) were included in the analysis (Figure2 and Supplementary Table S4).

Sulfide oxidation. The stepwise oxidation of sulfide to sulfate is a fundamental process in the *Riftia* symbiont, fueling the production of organic matter through chemosynthesis. Oxidation of reduced sulfur compounds is mediated via the central cytoplasmic enzymes DsrA/DsrB (dissimilatory sulfite reductase), AprA/AprB (APS reductase) and SopT (ATP sulfurylase) (Markert *et al.*, 2007; reviewed in Stewart and Cavanaugh, 2006), whose spot volumes contributed to 14% vol of all proteins in both intracellular metaproteomes on the 2D gels in the pH range 4–7 in this study. The expression pattern of sulfur oxidation-related proteins was very similar among both symbionts. However, the symbiont proteins DsrC, DsrE and DsrF, which are putatively involved in the conversion of stored elemental sulfur to sulfide (as suggested by their homology to products of the *dsr* gene cluster of the purple sulfur bacterium *Allochromatium vinosum*; Dahl *et al.*, 2008), showed significantly higher spot volumes in the *Riftia* symbionts compared with the *Tevnia* symbionts (DsrC: ratio 1.9, DsrE 3.1, DsrF 2.3). This is in line with the observation that the *Riftia* trophosomes used for the proteomic analysis were on average slightly darker than the *Tevnia* trophosomes, which might reflect lower sulfide availability and thus an increased necessity to activate sulfur reserves in case of *Riftia*. In contrast, the *Tevnia* symbiont proteome exhibited a slightly larger spot volume for the APS oxidation enzyme SopT (−1.7), possibly implying greater sulfide availability for the *Tevnia* symbionts. This assumption is corroborated by the results of the physico–chemical measurements performed at the sampling site (Table 1): While free sulfide concentrations were in the micromolar range around the *Riftia* plumes, sulfide concentrations in the vent

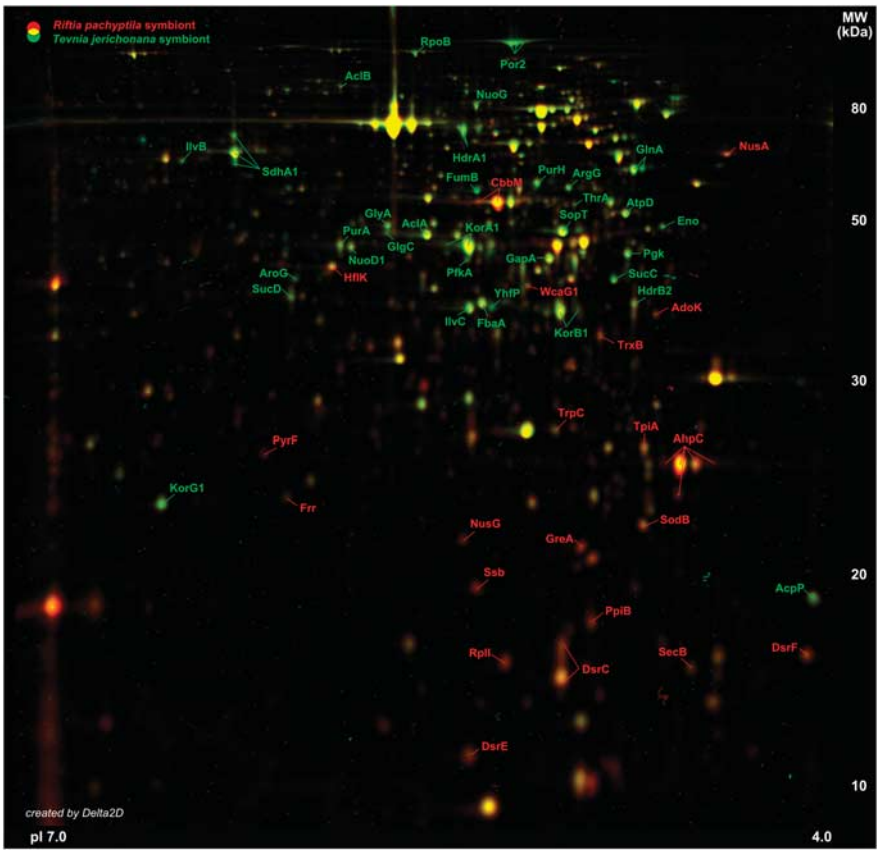

Figure 2 Differences in intracellular symbiont metaproteomes. Representative master gel images for the *Riftia* symbiont fraction and the *Tevnia* symbiont fraction, respectively, were created with Delta 2D by average-fusion of three individual symbiont gel images (pI range 4–7) for each tube worm genus. These master gel images were colored (*Riftia* symbiont: red spots, *Tevnia* symbiont: green spots) and overlaid to visualize differences in the soluble symbiont proteomes in a dual-channel gel image. Yellow spots indicate similar protein spot volume values in both symbiont fractions. Proteins with an at least 1.5-fold change in spot volume with regard to the reference organism are labeled accordingly and were considered in this study (see also Figure 3 and Supplementary Table S4). Variations could mainly be identified for the sulfur oxidation pathway, carbon fixation and the response to oxidative stress.

fluid surrounding *Tevnia* plumes were significantly higher (milimolar range).

Carbon fixation. Most of the enzymes putatively involved in the rTCA cycle were detected with high spot volumes in both metaproteomes, confirming the importance of this pathway for carbon fixation in the symbioses. Largest protein spot volumes were obtained for proteins catalyzing the incorporation of CO_2 into organic molecules, such as Icd (*Riftia* symbiont: 2.2% vol, *Tevnia* symbiont: 2.7% vol) and 2-oxoglutarate-ferredoxin oxidoreductase (subunit KorA1; *Riftia* symbiont: 1.3% vol, *Tevnia* symbiont: 2.1% vol). However, marginally enhanced spot volumes of most rTCA cycle enzymes were determined for the *Tevnia* symbiont (ratios: AclA/AclB −1.6, FumB −2.2, KorA1/KorB1/KorG1 −1.7/−1.5/−1.5, SdhA1 −1.5, SucC/SucD −1.9/−1.5) as compared with the respective *Riftia* symbiont spots. Considering that the oxygen levels in the vent fluid surrounding the plumes of the sampled *Riftia* specimens were supposedly higher than those around the *Tevnia* plumes (see above and Table 1), one might speculate that the highly oxygen-sensitive rTCA is either upregulated in the supposedly low-oxygen environment in *Tevnia* or inhibited under the more oxidized conditions presumed to exist in *Riftia*. Obviously, further analyses are required to corroborate these observations.

Stress response. Although oxygen is indispensible for the respiration of the heterotrophic host, increased quantities of oxygen can—despite the high oxygen buffering capacity of the worm's hemoglobin (Arp and Childress, 1983)—be passed on to the microaerophilic bacteria in the trophosome, potentially forcing them to cope with oxidative stress (Fisher *et al*., 1989). The enzymes alkyl hydroperoxide reductase (AhpC) and superoxide dismutase (SodB), which are involved in the resistance to oxidative stress, exhibited a substantial increase in spot volume in the *Riftia* symbiont, compared with the respective *Tevnia* symbiont protein spots (AhpC: ratio 2.1, SodB: ratio 1.8). Physiological *in situ* tests (Markert *et al*., 2007) confirmed that increased expression of the protective AhpC enzyme, as detected for the *Riftia* symbiont, is a decisive indicator for oxidative stress situations. In this context, enhanced expression of the thioredoxin

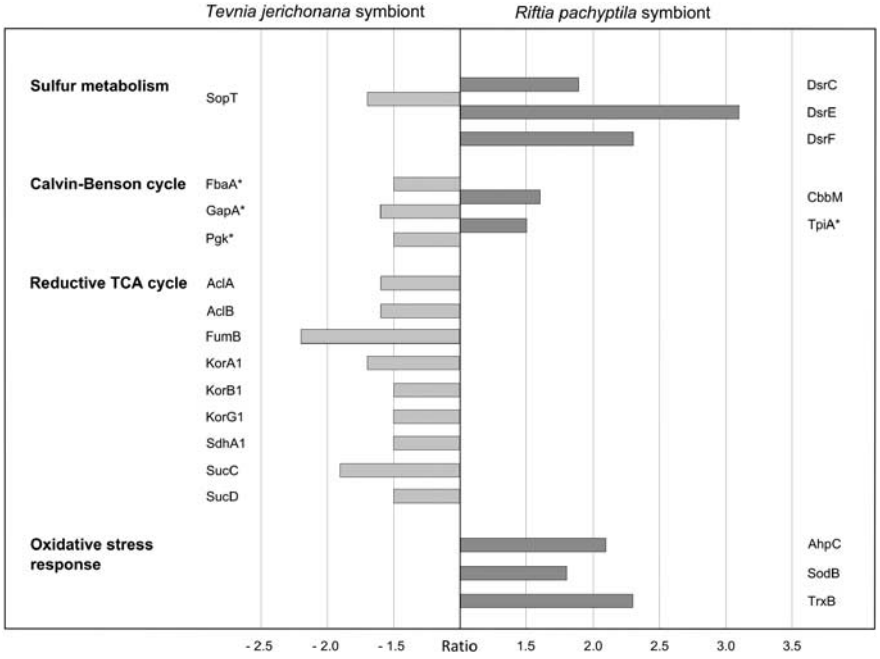

Figure 3 Protein spot ratios. Differences in spot volume as detected on the dual-channel image (Figure 2) when comparing the soluble fractions of the *Riftia* symbiont and the *Tevnia* symbiont metaproteomes are indicated for relevant proteins of the sulfur and carbon metabolism as well as for proteins involved in the response to oxidative stress. Spot volumes (see Supplementary Table S4) are calculated as percentages (% vol) of the total proteomes on the respective 2D reference gels. The ratio (fold change) reflects the quotient of two corresponding spot volumes. Ratio threshold: 1.5-fold change compared with the reference organism; minimum spot volume: 0.1% vol. Positive ratio values (dark gray bars) mark a relatively higher spot volume for proteins on the *Riftia* symbiont gel, compared with the *Tevnia* symbiont. Negative ratios (light gray bars) indicate higher spot volumes on the *Tevnia* symbiont gel. See Supplementary Table S4 for protein functions. *Protein is also involved in organic carbon metabolism.

reductase (TrxB, ratio 2.3), a component of the thiol/disulfide redox system, might also be part of the oxidative stress response in *Riftia* symbiont cells as observed for other bacteria (Helmann *et al.*, 2003; Ballal and Manna, 2010). These results suggest that the *in situ* oxygen concentrations in the *Riftia* trophosome were slightly higher than those in the *Tevnia* trophosome, which would be in line with the habitat chemistry measurements (see above and Table 1) that indicated a higher proportion of oxygen-rich sea water in the immediate surroundings of the *Riftia* plumes compared with the vent fluid surrounding the *Tevnia* plumes.

Conclusions

Our comparative proteogenomic approach revealed highly consistent physiological features in the symbionts *of R. pachyptila* and *T. jerichonana*. The results strongly support the hypothesis that the same symbiont species is able to establish a monospecific symbiosis with highly comparable physiological properties in both animals. On the metaproteomic level, only minor variations between both symbionts were detected, despite the drastically different geochemical conditions existing in the immediate outside environment of each host species (Table 1). Overall, the symbiotic association appears to be very well adapted to the dynamic environmental conditions in the hydrothermal vent habitat, with the hosts providing a similar and relatively stable internal environment for the symbionts. In addition, the symbionts seem to be able to fine-tune their physiological status in response to variations passed on from the host as attenuated reflections of the outside conditions. Furthermore, our data suggest that the differences in colonization behavior and competitive mechanisms, which have previously been observed between both siboglinid tubeworms and that contribute to the successional replacement of *T. jerichonana* by *R. pachyptila*, are not due to differences between the respective endosymbionts, but are more likely caused by variations specific to each tubeworm species and thus should be further investigated at the host level.

Acknowledgements

We are grateful to the captains, crews and pilots of R/V *Atlantis* and DSV *Alvin* for their outstanding help in obtaining the samples. The cruises were funded through Grant OCE-0452333 of the US National Science Foundation (NSF) to SMS. Cruise participation of NLB was funded by Ifremer. Development of chemical sensors was supported by Ifremer and CNRS. This work was supported by Grant SCHW 595/3-3 of the Deutsche Forschungsgemeinschaft (DFG) to TS and by the German Federal Ministry of Education and Research (BMBF; reference 03F0480A/B). SMS was supported through NSF and a senior fellowship awarded by the Alfried Krupp Wissenschaftskolleg Greifswald (Germany).

References

Altschul SF, Madden TL, Schaffer AA, 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.

Arp AJ, Childress JJ. (1983). Sulfide binding by the blood of the hydrothermal vent tube worm *Riftia pachyptila*. *Science* **219**: 295–297.

Aziz RK, Bartels D, Best AA, DeJongh M, Disz T, Edwards RA *et al.* (2008). The RAST server: rapid annotations using subsystems technology. *BMC Genomics* **9**: 75.

Bhakdi S, Tranum-Jensen J. (1991). Alpha-toxin of *Staphylococcus aureus*. *Microbiol Rev* **55**: 733–751.

Ballal A, Manna AC. (2010). Control of thioredoxin reductase gene (*trxB*) transcription by SarA in *Staphylococcus aureus*. *J Bacteriol* **192**: 336–345.

Blum J, Fridovich I. (1984). Enzymatic defenses against oxygen toxicity in the hydrothermal vent animals *Riftia pachyptila* and *Calyptogena magnifica*. *Arch Biochem Biophys* **228**: 617–620.

Bright M, Bulgheresi S. (2010). A complex journey: transmission of microbial symbionts. *Nat Rev* **8**: 218–230.

Bright M, Sorgo A. (2003). Ultrastructural reinvestigation of the trophosome of adults of *Riftia pachyptila* (Annelida, Siboglinidae). *Invertebr Biol* **122**: 345–366.

Cary SC, Warren W, Anderson E, Giovannoni SJ. (1993). Identification and localization of bacterial endosymbionts in hydrothermal vent taxa with symbiont-specific polymerase chain reaction amplification and *in situ* hybridization techniques. *Mol Mar Biol Biotechnol* **2**: 51–62.

Cavanaugh CM, Gardiner SL, Jones ML, Jannasch HW, Waterbury JB. (1981). Prokaryotic cells in the hydrothermal vent tube worm *Riftia pachyptila* Jones: possible chemoautotrophic symbionts. *Science* **213**: 340–342.

Childress JJ, Fisher CR, Favuzzi JA, Kochevar RE, Sanders NK, Alayse AM. (1991). Sulfide-driven autotrophic balance in the bacterial symbiont-containing hydrothermal vent tubeworm, *Riftia pachyptila* Jones. *Biol Bull* **180**: 135–153.

Cowen JP, Fornari DJ, Shank TM, Love B, Glazer B, Treuch AH *et al.* (2007). Volcanic eruptions at East Pacific Rise near 9°50′N. *EOS Trans Am Geophys Un* **88**: 81–83.

Dahl C, Schulte A, Stockdreher Y, Hong C, Grimm F, Sander J *et al.* (2008). Structural and molecular genetic insight into a widespread sulfur oxidation pathway. *J Mol Biol* **384**: 1287–1300.

Di Meo CA, Wilbur AE, Holben WE, Feldman RA, Vrijenhoek RC, Cary SC. (2000). Genetic variation among endosymbionts of widely distributed vestimentiferan tubeworms. *Appl Environ Microbiol* **66**: 651–658.

Distel DL, Felbeck H. (1988). Pathways of inorganic carbon fixation in the endosymbiont-bearing lucinid clam *Lucinoma aequizonata*. *J Exp Zool* **247**: 1–10.

Distel DL, Lane DJ, Olsen GJ, Giovannoni SJ, Pace B, Pace NR *et al.* (1988). Sulfur-oxidizing bacterial endosymbionts: analysis of phylogeny and specificity by 16S rRNA sequences. *J Bacteriol* **170**: 2506–2510.

Drummond AJ, Ashton B, Buxton S, Cheung M, Cooper A, Duran C *et al.* (2010) Geneious v5.3, Available from http://www.geneious.com/.

Edwards DB, Nelson DC. (1991). DNA-DNA solution hybridization studies of the bacterial symbionts of hydrothermal vent tube worms (*Riftia pachyptila* and *Tevnia jerichonana*). *Appl Environ Microbiol* **57**: 1082–1088.

Felbeck H. (1981). Chemoautotrophic potential of the hydrothermal vent tube worm *Riftia pachyptila* Jones (Vestimentifera). *Science* **213**: 336–338.

Felbeck H, Jarchow J. (1998). Carbon release from purified chemoautotrophic bacterial symbionts of the hydrothermal vent tubeworm *Riftia pachyptila*. *Physiol Biochem Zool* **71**: 294–302.

Felbeck H, Somero GN. (1982). Primary production in deep-sea hydrothermal vent organisms: roles of sulfide-oxidizing bacteria. *Trends Biochemical Sci* **7**: 201–204.

Feldman RA, Black MB, Cary CS, Lutz RA, Vrijenhoek RC. (1997). Molecular phylogenetics of bacterial endosymbionts and their vestimentiferan hosts. *Mol Mar Biol Biotechnol* **6**: 268–277.

Fisher CR, Childress JJ. (1984). Substrate oxidation by trophosome tissue from *Riftia pachyptila* Jones (phylum pogonophora). *Mar Biol Lett* **5**: 171–183.

Fisher CR, Childress JJ, Minich E. (1989). Autotrophic carbon fixation by the chemoautotrophic symbionts of *Riftia pachyptila*. *Biol Bull* **177**: 372–385.

Girguis PR, Childress JJ. (2006). Metabolite uptake, stoichiometry and chemoautotrophic function of the hydrothermal vent tubeworm *Riftia pachyptila*: responses to environmental variations in substrate concentrations and temperature. *J Exp Biol* **209**: 3516–3528.

Girguis PR, Raymond WL, Desaulniers N, Childress JJ, Pospesel M *et al.* (2000). Fate of nitrate acquired by the tubeworm *Riftia pachyptila*. *Appl Environ Microbiol* **66**: 2783–2790.

Hand SC. (1987). Trophosome ultrastructure and the characterization of isolated bacteriocytes from invertebrate-sulfur bacteria symbioses. *Biol Bull* **173**: 260–276.

Harmer TL, Rotjan RD, Nussbaumer AD, Bright M, Ng AW, DeChaine EG *et al.* (2008). Free-living tube worm endosymbionts found at deep-sea vents. *Appl Environ Microbiol* **74**: 3895–3898.

Haymon RM, Fornari DJ, Von Damm KL, Lilley MD, Perfit MR, Edmond JM *et al.* (1993). Volcanic eruption of the mid-ocean ridge along the East Pacific Rise crest at 9°45-52′N: Direct submersible observations of seafloor phenomena associated with an eruption event in April, 1991. *Earth Planet Sc Lett* **119**: 85–101.

Helmann JD, Wu MF, Gaballa A, Kobel PA, Morshedi MM, Fawcett P *et al.* (2003). The global transcriptional response of *Bacillus subtilis* to peroxide stress is coordinated by three transcriptional factors. *J Bacteriol* **185**: 243–253.

Hentschel U, Felbeck H. (1993). Nitrate respiration in the hydrothermal vent worm *Riftia pachyptila*. *Nature* **366**: 338–340.

Hessler RR, Smithey WM, Boudrias MA, Keller CH, Lutz RA, Childress JJ. (1988). Temporal change in megafauna at the Rose Garden hydrothermal vent (Galapagos Rift; eastern tropical Pacific). *Deep Sea Res* **35**: 1681–1709.

Inglis TJJ, Robertson T, Woods DE, Dutton N, Chang BJ. (2003). Flagellum-mediated adhesion by *Burkholderia pseudomallei* precedes invasion of *Acanthamoeba astronyxis*. *Infect Immun* **71**: 2280–2282.

Johnson KS. (1988). Chemical and biological interactions in the Rose Garden hydrothermal field, Galapagos spreading center. *Deep Sea Res Pt A* **35**: 1723–1744.

Jones ML. (1981a). *Riftia pachyptila*, new genus, new species, the vestimentiferan worm from the Galápagos Rift geothermal vents (Pogonophora). *Proc Biol Soc Wash* **93**: 1295–1313.

Jones ML. (1981b). *Riftia pachyptila* Jones: observations on the vestimentiferan worm from the Galápagos Rift. *Science* **213**: 333–336.

Jones ML. (1985). On the Vestimentifera, new phylum: six new species and other taxa from hydrothermal vents and elsewhere. *Bull Biol Soc Wash* **6**: 117–158.

Krehenbrink M, Oppermann-Sanio FB, Steinbüchel A. (2002). Evaluation of non-cyanobacterial genome sequences for occurrence of genes encoding proteins homologous to cyanophycin synthetase and cloning of an active cyanophycin synthetase from *Acinetobacter* sp. strain DSM 587. *Arch Microbiol* **177**: 371–380.

Le Bris N, Govenar B, Le Gall C, Fisher CR. (2006). Variability of physico-chemical conditions in 9°N EPR diffuse flow vent habitat. *Mar Chem* **98**: 167–182.

Le Bris N, Sarradin PM, Pennec S. (2001). A new deep-sea probe for *in situ* pH measurement in the environment of hydrothermal vent biological communities. *Deep Sea Res Pt I* **48**: 1941–1951.

Lee RW, Childress JJ. (1994). Assimilation of inorganic nitrogen by marine invertebrates and their chemoautotrophic and methanotrophic symbionts. *Appl Environ Microbiol* **60**: 1852–1858.

Margulies M, Egholm M, Altman WE, Attiya S, Bader JS, Bemben LA *et al.* (2005). Genome sequencing in microfabricated high-density picolitre reactors. *Nature* **437**: 376–380.

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

Markert S, Gardebrecht A, Felbeck H, Sievert SM, Klose J, Becher D *et al.* (2011). *Status quo* in physiological proteomics of the uncultured *Riftia pachyptila* endosymbiont. *Proteomics* **11**: 3106–3117.

Moore TS, Shank TM, Nuzzio DB, Luther III GW. (2009). Time-series chemical and temperature habitat characterization of diffuse flow hydrothermal sites at 9°50′N East Pacific Rise. *Deep Sea Res Pt II: Topical Stud Oceanograph* **56**: 1616–1621.

Mullineaux LS, Fisher CR, Peterson CH, Schaeffer SW. (2000). Tubeworm succession at hydrothermal vents: use of biogenic cues to reduce habitat selection error? *Oecologia* **123**: 275–284.

Needleman SB, Wunsch CD. (1970). A general method applicable to the search for similarities in the amino acid sequence of two proteins. *J Mol Biol* **48**: 443–453.

Nees HA, Lutz RA, Shank TM, Luther GW. (2009). Pre- and post-eruption diffuse flow variability among tubeworm habitats at 9°50′ north on the East Pacific Rise. *Deep Sea Res II* **56**: 1607–1615.

Nussbaumer AD, Fisher CR, Bright M. (2006). Horizontal endosymbiont transmission in hydrothermal vent tubeworms. *Nature* **441**: 345–348.

Pflugfelder B, Fisher CR, Bright M. (2005). The color of the trophosome: elemental sulfur distribution in the endosymbionts of *Riftia pachyptila* (Vestimentifera; Siboglinidae). *Mar Biol* **146**: 895–901.

Pleijel F, Dahlgren TG, Rouse GW. (2009). Progress in systematics: from Siboglinidae to Pogonophora and Vestimentifera and back to Siboglinidae. *Comp Rend Biol* **332**: 140–148.

Pospesel MA, Hentschel U, Felbeck H. (1998). Determination of nitrate in the blood of the hydrothermal vent tubeworm *Riftia pachyptila* using a bacterial nitrate reduction assay. *Deep Sea Res Pt I Oceanograph Res Papers* **45**: 2189–2200.

Rickard D, Luther III GW. (2007). Chemistry of iron sulfides. *Chem Rev* **107**: 514–562.

Robidart JC, Bench SR, Feldman RA, Novoradovsky A, Podell SB, Gaasterland T *et al.* (2008). Metabolic versatility of the *Riftia pachyptila* endosymbiont revealed through metagenomics. *Environ Microbiol* **3**: 121–126.

Shank TM, Fornari DJ, Von Damm KL, Lilley MD, Haymon RM, Lutz RA. (1998). Temporal and spatial patterns of biological community development at nascent deep-sea hydrothermal vents (9°50′ N, East Pacific Rise). *Deep Sea Res II* **45**: 465–515.

Shank TM, Govenar B, Buckman K, Fornari DJ, Soule SA, Luther GW *et al.* (2006). Initial biological, chemical and geological observations after the 2005–6 volcanic eruption on the East Pacific Rise. *EOS Trans Am Geophys Un* **87**, Fall Meet. Suppl., Abstract V13C-04..

Sorgo A, Gaill F, Lechaire JP, Arndt C, Bright M. (2002). Glycogen storage in the *Riftia pachyptila* trophosome: contribution of host and symbionts. *Mar Ecol Prog Ser* **231**: 115–120.

Stewart FJ, Cavanaugh CM. (2006). Symbiosis of thioautotrophic bacteria with *Riftia pachyptila*. *Prog Mol Subcell Biol* **41**: 197–225.

Stewart V, Lu Y, Darwin AJ. (2002). Periplasmic nitrate reductase (NapABC enzyme) supports anaerobic respiration by *Escherichia coli* K-12. *J Bacteriol* **184**: 1314–1323.

Thornhill DJ, Fielmann KT, Santos SR, Halanych KM. (2008). Siboglinid-bacteria endosymbiosis: a model system for studying symbiotic mechanisms. *Commun Integr Biol* **1**: 163–166.

Toft C, Anderson SGE. (2010). Evolutionary microbial genomics: insights into bacterial host adaptation. *Nat Rev Genet* **11**: 465–475.

Tolstoy M, Cowen JP, Bakter ET, Fornari DJ, Rubin KH, Shank TM *et al.* (2006). A sea-floor spreading event captured by seismometers. *Science* **314**: 1920–1922.

Von Damm KL. (2000). Chemistry of hydrothermal vent fluids from 9°–10° N, East Pacific Rise. 'Time zero', the immediate posteruptive period. *J Geophys Res* **105**: 203–222.

Von Damm KL, Oosting SE, Kozlowski R, Buttermore LG, Colodner DC, Edmonds HN *et al.* (1995). Evolution of East Pacific Rise hydrothermal vent fluids following a volcanic eruption. *Nature* **357**: 47–50.

Waters VL, Crosa JH. (1991). Colicin V virulence plasmids. *Microbiol Rev* **55**: 437–450.

Ziegler K, Diener A, Herpin C, Richter R, Deutzmann R, Lockau W. (1998). Molecular characterization of cyanophycin synthetase, the enzyme catalyzing the biosynthesis of the cyanobacterial reserve material multi-L-arginyl-poly-L-aspartate (cyanophycin). *Eur J Biochem* **254**: 154–159.

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