Environmental Microbiology
September 2014, Volume 16, Issue 9, pages 2723–2738
http://dx.doi.org/10.1111/1462-2920.12406

The gill chamber epibiosis of deep-sea shrimp *Rimicaris exoculata*: an in-depth metagenomic investigation and discovery of *Zetaproteobacteria*

Cyrielle Jan[1], Jillian M. Petersen[2,†], Johannes Werner[2,3,†], Hanno Teeling[2,†], Sixing Huang[2],
Frank Oliver Glöckner[2,3], Olga V. Golyshina[4], Nicole Dubilier[2], Peter N. Golyshin[4], Mohamed Jebbar[1,5]
and Marie-Anne Cambon-Bonavita[6,*]

[1] UMR 6197-Laboratoire de Microbiologie des Environnements Extrêmes (LM2E), Institut Universitaire Européen de la Mer (IUEM), Université de Bretagne Occidentale, Plouzané, France
[2] Max Planck Institute for Marine Microbiology, Bremen, Germany
[3] Jacobs University Bremen gGmbH, Bremen, Germany
[4] School of Biological Sciences, Bangor University, Bangor, Gwynedd, Wales, UK
[5] CNRS, UMR6197-Laboratoire de Microbiologie des Environnements Extrêmes (LM2E), Institut Universitaire Européen de la Mer (IUEM), Place Nicolas Copernic, Plouzané, France
[6] Ifremer, Centre de Brest, REM/EEP/LM2E, Plouzané, France

[†] These authors contributed equally to this work.

*: Corresponding author : Marie-Anne Cambon-Bonavita, tel. (+33) 2 98 224756 ; Fax (+33) 2 98 224757 ; email address : Marie.Anne.Cambon@ifremer.fr

---

Abstract:

The gill chamber of deep-sea hydrothermal vent shrimp *Rimicaris exoculata* hosts a dense community of epibiotic bacteria dominated by filamentous *Epsilonproteobacteria* and *Gammaproteobacteria*. Using metagenomics on shrimp from the Rainbow hydrothermal vent field, we showed that both epibiont groups have the potential to grow autotrophically and oxidize reduced sulfur compounds or hydrogen with oxygen or nitrate. For carbon fixation, the *Epsilonproteobacteria* use the reductive tricarboxylic acid cycle, whereas the *Gammaproteobacteria* use the Calvin–Benson–Bassham cycle. Only the epsilonproteobacterial epibionts had the genes necessary for producing ammonium. This ability likely minimizes direct competition between epibionts and also broadens the spectrum of environmental conditions that the shrimp may successfully inhabit. We identified genes likely to be involved in shrimp–epibiont interactions, as well as genes for nutritional and detoxification processes that might benefit the host. Shrimp epibionts at Rainbow are often coated with iron oxyhydroxides, whose origin is intensely debated. We identified 16S rRNA sequences and functional genes affiliated with iron-oxidizing *Zetaproteobacteria*, which indicates that biological iron oxidation might play a role in forming these deposits. Fluorescence *in situ* hybridizations confirmed the presence of active *Zetaproteobacteria* in the *R. exoculata* gill chamber, thus providing the first evidence for a *Zetaproteobacteria*–invertebrate association.

1. Introduction

The shrimp *Rimicaris exoculata* (Williams and Rona, 1986) dominates the macrofauna at many hydrothermal vent sites along the Mid-Atlantic Ridge (MAR). *R. exoculata* swim along chimney walls in the gradient between hydrothermal fluids and cold oxygenated ambient seawater, in a temperature range between 3 and 25 °C (Schmidt *et al.*, 2008). Adult *R. exoculata* have a hypertrophied cephalothorax that housts a dense epibiotic bacterial community. These bacteria colonize the mouthparts and inner faces of the shrimp carapace (Zbinden *et al.*, 2004; Zbinden *et al.*, 2008 and references therein) and are morphologically diverse, including at least six filament types and bacilli (Zbinden *et al.*, 2004; Corbari *et al.*, 2008a; Zbinden *et al.*, 2008). Molecular studies have revealed two dominant phylotypes belonging to the *Epsilon*- and *Gammaproteobacteria*, and lower abundances of *Alpha-, Delta-* and *Betaproteobacteria* (Zbinden *et al.*, 2008; Petersen *et al.*, 2010; Hügler *et al.*, 2011; Guri *et al.*, 2012). Based on functional gene surveys of the epibiont community, it was suggested that epibiont chemolithoautotrophy could be powered by a variety of electron sources, including reduced sulfur compounds, hydrogen, and methane (Zbinden *et al.*, 2008; Hügler *et al.*, 2011; Petersen *et al.*, 2011; Guri *et al.*, 2012). Furthermore, based on PCR-amplification of two genes of the reductive tricarboxylic acid (rTCA) cycle and one of the Calvin-Benson-Bassham (CBB) cycle, it was hypothesized that the dominant epsilon- and gammaproteobacterial epibionts could be autotrophic and use different carbon fixation pathways (Hügler *et al.*, 2011). In support of this, CO_2 fixation activity has been demonstrated by both *in vitro* and *in vivo* experiments (Polz *et al.*, 1998; Ponsard *et al.*, 2013), although, only functional gene fragments have been amplified and sequenced from these complex epibiont communities so far, and these fragments could only be putatively assigned to specific epibionts. Hence, the sources of energy and CO_2 fixation pathways of the dominant symbionts remain to be resolved.

Shrimps from the Rainbow hydrothermal vent field of the MAR have a conspicuous rusty color due to iron oxyhydroxide particle accumulation in their gill chambers. Close association of epibionts with these chemically inert iron oxyhydroxides led to the suggestion that some epibionts might oxidize iron (Zbinden *et al.*, 2004; Corbari *et al.*, 2008b; Schmidt *et al.*, 2008; Zbinden *et al.*, 2008). However, no study has yet revealed the presence of iron-oxidizing bacteria in the gill chamber or provided evidence that previously identified epibionts can oxidize iron. Since the dominant epibionts have not yet been successfully cultivated, we undertook a metagenomic approach to (i) study the microbial diversity of the *R. exoculata* gill chamber community (while avoiding some of the pitfalls of traditional methods based on PCR amplification of 16S rRNA or functional genes) (ii) determine which energy sources could potentially power this symbiosis, (iii) pinpoint the differences between the epsilon- and gammaproteobacterial epibionts.

2. Results and discussion

2.1. Metagenome characteristics and microbial diversity

Taxonomic classification of the metagenome at the level of class and above revealed that a large number of sequences could be affiliated with *Bacteria* (84%), but also with *Eukarya* (14%, probably *R. exoculata*) and *Archaea* (2%) (Fig. 1a). Most of the bacterial sequences were affiliated with *Proteobacteria* (78%), including *Epsilonproteobacteria* (39%), *Gammaproteobacteria* (30%), *Alphaproteobacteria* (4%), *Zetaproteobacteria* (2.2%), *Betaproteobacteria* (1.6%) and *Deltaproteobacteria* (1.0%) (Fig. 1b). The metagenome harbored 1,029 partial small subunit (SSU) rRNA gene sequences of sufficient quality for taxonomic assignment (Fig. 1c). These SSU sequences affiliated with

Eukarya/*Eumalacostraca* (44%), *Epsilonproteobacteria*/*Sulfurovum* (38%), *Gammaproteobacteria*/*Leucothrix* (5%) and *Zetaproteobacteria*/*Mariprofundus* (3%). The difference in *R. exoculata* sequence abundance between the 14% taxonomically classified metagenome sequences and 44% taxonomically classified partial SSU sequence can likely be attributed to the high rRNA gene copy numbers of eukaryotic cells, which can number in thousands (Zhu *et al.*, 2005), whereas ten is considered a high number for prokaryotes.

Like other techniques, metagenomics does not necessarily reflect microbial diversity accurately, since biases affect DNA extraction and library preparation before sequencing. Such effects might explain the lack of 16S rRNA sequences affiliating with the methanotrophic gammaproteobacterial epibionts as described previously (Zbinden *et al.*, 2008; Guri *et al.*, 2012).

For functional analysis, we focused on taxonomically coherent bins of sequences (hereafter referred to as 'taxobins') affiliating with the dominating *Epsilon*- and *Gammaproteobacteria*, and with the *Zetaproteobacteria* aspotential sources of the iron oxyhydroxides. Previous studies indicated that the main epibiotic community is comprised of several closely related epsilon- and gammaproteobacterial phylotypes (Zbinden *et al.*, 2004; Zbinden *et al.*, 2008; Petersen *et al.*, 2010; Hügler *et al.*, 2011; Guri *et al.*, 2012). This was also evident in phylogenetic tree of the epsilonprteobacterial phylotypes that we reconstructed from metagenomic partial 16S rRNA sequences (Fig. S1), although this analysis might overestimate the strain diversity for a number of reasons: Firstly, many partial sequences were too short for it to be possible to confidently determine their phylogeny within the epsilonproteobacterial group. Secondly, many of the distinct phylotypes shown in Figure S1 were only represented by single reads and may constitute sequencing errors. Finally, as bacterial genomes can harbor multiple copies of slightly different 16S rRNA genes, it is unclear how much of the detected 16S rRNA gene sequence microdiversity was due to inter- or intra-strain level variation. It was not possible to discriminate sequences from distinct phylotypes in the metagenome. Hence, the analyzed taxobins must be considered as sub-metagenomes, and all inferred metabolic reconstructions may represent the combined metabolic potentials of a number of phylotypes, albeit of phylotypes that are so closely related that they are unlikely to show any major differences.

The epsilonproteobacterial taxobin consisted of 2,103 contigs (2.89 Mbp) (Table 1). Members of the *Epsilonproteobacteria* often constitute the dominant bacterial group in hydrothermal habitats (Campbell *et al.*, 2006, Dubilier *et al.*, 2008). Several related epsilonproteobacterial isolates have recently been characterized, such as the deep-sea strains *Sulfurovum* sp. NBC37-1 and *Nitratiruptor* sp. SB155-2 (Nakagawa *et al.*, 2007), and the coastal strain *Sulfurimonas denitrificans* DSM1251 (Sievert *et al.*, 2008a).

The gammaproteobacterial taxobin consisted of 2,441 contigs (3.43 Mbp) (Table 1). The filamentous chemolithoheterotrophic sulfur oxidizer *Leucothrix mucor* (Grabovich *et al.*, 1999) represents the closest cultured gammaproteobacterial epibiont relative (90.6% 16S rRNA sequence identity).Taxonomic classification attributed 0.47 Mbp (2.2%) of the metagenome to the class *Zetaproteobacteria* and thus allowed identification of zetaproteobacterial genes (Table S2). The class *Zetaproteobacteria* so far contains just a single cultivated species, the neutrophilic iron oxidizer *Mariprofundus ferrooxydans* (Emerson *et al.*, 2007; Singer *et al.*, 2011). FISH experiments on cephalothorax sections of three shrimps revealed *Zetaproteobacteria* on each specimen (Fig. 2). The *Zetaproteobacteria* were low in abundance and located on the surface of the branchiostegite membrane (inner face of the gill chamber) closely associated with the iron oxyhydroxide layer that accumulates between the shrimp's molts (Corbari *et al.*, 2008b; Corbari *et al.*, 2008a). This distinct association of close relatives of cultivated neutrophilic iron oxidizers and iron oxyhydroxides deposits in the gill chamber supports the hypothesis that *R. exoculata* associates with iron-oxidizing epibionts (Zbinden *et al.*, 2004; Corbari *et al.*, 2008b; Zbinden *et al.*, 2008; Guri *et al.*, 2012).

Zetaproteobacteria were found on all sampled *R. exoculata* individuals, indicating that *Zetaproteobacteria* might be regular members of the *R. exoculata* gill chamber community at the Rainbow hydrothermal vent site.

2.2. Metabolism

2.2.1. Carbon fixation

Based on functional gene surveys, it has been postulated that chemoautotrophy at deep-sea hydrothermal vents is mediated by at least two pathways, the rTCA and CBB cycles (Campbell and Cary, 2004; Hügler *et al.*, 2011). Our results could corroborate these hypotheses. The epsilonproteobacterial taxobin harbored a complete set of genes for the rTCA cycle (Fig. 3, sequences: HG799104 to HG799148), which is widespread in microaerophilic bacteria. In contrast, the gammaproteobacterial epibiont taxobin harbored genes for the complete CBB cycle (Fig. 4, sequences: HG799149 to HG799155), which is quantitatively the most important mechanism of autotrophic CO_2 fixation on Earth. The co-existence of both CBB and rTCA cycles was reported for the endosymbiont of the deep-sea tube worm *Riftia pachyptila* (Markert *et al.*, 2007), switching between these cycles according to environmental conditions. This endosymbiont preferentially uses the CBB cycle in high-energy high-sulfide conditions, and the rTCA in low-energy low-sulfide conditions (Markert *et al.*, 2007). Furthermore, the reversibility of the rTCA cycle may provide the *Riftia* endosymbiont with additional metabolic flexibility. Such metabolic adaptation strategies provide a selective advantage under varying environmental conditions (Berg, 2011). In the case of the *Rimicaris* symbiosis, the same advantage may be achieved by coexistence of epibionts with distinct carbon fixation pathways.

The zetaproteobacterial taxobin also contained genes for the CBB cycle, including the *cbbM* gene encoding a form II RuBisCO (Table S2, sequences: HG799156 to HG799158, and HG799163). This indicates that the zetaproteobacterial epibionts have the potential for autotrophic growth. The RuBisCO gene that we found in the gammaproteobacterial taxobin also encodes form II RuBisCO. This form is adapted to low oxygen and medium to high carbon dioxide concentrations (Badger and Bek, 2008) – conditions that likely prevail at hydrothermal vent sites, and particularly in the gill chamber of *R. exoculata* (Zbinden *et al.*, 2004). Both RuBisCO form II and the rTCA cycle are often found in autotrophic bacteria living under micro-oxic conditions (Campbell and Cary, 2004) (see Supplementary text).

2.2.2. Sulfur metabolism

Most free-living vent *Epsilonproteobacteria* are sulfur oxidizers (Inagaki *et al.*, 2003; Nakagawa *et al.*, 2005; Campbell *et al.*, 2006; Takai *et al.*, 2006; Sievert *et al.*, 2008b). So far, the only evidence that the epsilonproteobacterial *R. exoculata* epibionts can oxidize sulfur has come from the amplification and sequencing of *soxB* genes from the epibiont community, which are similar to those of cultivated *Epsilonproteobacteria* (Hügler *et al.*, 2011). We identified numerous *sox* genes in the epsilonproteobacterial taxobin (Fig. 3) in eight different contigs (Table S2, sequences: HG799164 to HG799177). Two clusters contained *soxYZA*, one cluster contained *soxYD* and another contained *soxAB. SoxD* and *soxX* also occurred as isolated genes. In facultative chemoautotrophs such as *Paracoccus pantotrophus* GB17 (Friedrich *et al.*, 2001), the *sox* genes are organized in a single gene cluster. In contrast, in obligate chemoautotrophs closely related to the *R. exoculata* epsilonproteobacterial epibiont, such as *Sulfurovum* sp. NBC37-1 (Yamamoto *et al.*, 2010) and *S. denitrificans* DSM1251 (Sievert *et al.*, 2008a), *sox* genes are found in separate clusters spread throughout the genome. As in *S. denitrificans* DSM1251, four copies of the sulfide-quinone reductase (SQR) gene were found (sequences: HG799183 to HG799185), all of which closely affiliated with SQR-encoding genes from *Sulfurovum* sp. NBC37-1. SQR could catalyze the oxidation of sulfide to extracellular elemental sulfur and thereby contribute to the formation of filamentous sulfur (Nakagawa *et al.*, 2007). It was hypothesized that sulfur-oxidizing bacteria may use SQR to switch from the complete oxidation of sulfide to sulfate to incomplete oxidation to sulfur, in order to prevent toxic acidification of their environment (Ferrer *et al.*, 2011). Similarly, SQR could play a role in preventing harmful acidification inside the gill chamber, but also could constitute a way to store biochemical energy in the form of sulfur when sulfide concentrations are high.

Previous molecular studies on sulfur oxidation genes in the epibiotic community identified *aprA* genes that were related to *aprA* genes of *Gammaproteobacteria* (Zbinden et al. 2008, Hügler et al. 2011). In the gammaproteobacterial taxobin, in addition to the *aprA* gene, we found six *sox* genes in one cluster (*soxAZYX*) and we found other *sox* genes as single genes on separate contigs (*soxB*, *soxW*) (sequences: HG799187 to HG799193). A single SQR gene was found, but no *sox*CD (Fig. 4). It was unclear if these genes were missed due to the incompleteness of the metagenome or whether they are indeed absent from these gammaproteobacterial epibionts. Sulfur-oxidizing bacteria that lack *soxCD,* such as the anoxygenic phototrophic purple sulfur bacterium *Allochromatium vinosum,* produce elemental sulfur as an intermediate in sulfide oxidation (Hensen *et al.*, 2006). S*oxCD* genes were required for complete oxidation to sulfate (Friedrich *et al.*, 2001). The Sox system without SoxCD can oxidize sulfite and the sulfone group (SO_3^-) in thiosulfate, but cannot oxidize sulfide or elemental sulfur to thiosulfate (Friedrich *et al.*, 2001). Thus, the gammaproteobacterial epibiont might oxidize sulfide to elemental sulfur for storage using flavocytochrome *c* sulfide dehydrogenase. Genes were found for both the small subunit A (*fccA*) and the large subunit B (*fccB*) (sequences: HG799212 to HG799214). The best-studied enzyme system for the oxidation of stored elemental sulfur is the reversely-operating dissimilatory sulfate reductase (*dsr*) of *A. vinosum* encoded by *dsrABEFHCMKLJOPNRS* genes (Grimm *et al.*, 2010). We found *dsr* genes in the gammaproteobacterial taxobin in several clusters (*dsrAB*, *dsrFHCMK, dsrLJOP*) and as isolated genes (*dsrA*, *dsrE*, *dsrS*) (sequences: HG799196 to HG799211). Intracellular sulfur granules have been identified in the thin filamentous gammaproteobacterial epibionts (Zbinden *et al.*, 2008; Petersen *et al.*, 2010). *Dsr* might in this case be used to tap stores of intracellular elemental sulfur to compensate for the lower environmental sulfide concentrations that the epibionts encounter when their shrimp hosts swim outside of the hydrothermal stream. Deep-sea hydrothermal vent *Gammaproteobacteria* seem to possess two different sulfur-oxidization pathways: (i) the reverse sulfate reduction using Dsr and Apr and (ii) the Sox system without SoxCD (Yamamoto and Takai, 2011). In case of sulfur-oxidizing symbionts at hydrothermal vents, the supply of reduced sulfur compounds and oxygen for these pathways depends on the movements of the invertebrate host in the mixing zones of deep-sea hydrothermal vent fluids (Yamamoto and Takai, 2011).

In addition to the oxidation of reduced sulfur compounds, the epsilonproteobacterial epibionts, and possibly also the gammaproteobacterial epibionts, may have the potential to use oxidized sulfur compounds as electron acceptors. This hypothesis was supported by the presence of three genes that encode a polysulfide reductase NrfD-related membrane anchor (4Fe-4S ferredoxin iron-sulfur binding), Fe/S proteins (4Fe-4S ferredoxin) and molybdopterin oxidoreductase (sequences: HG799186, HG799215, HG799403). These genes were arranged in one cluster in the epsilonproteobacterial taxobin and on several contigs in the gammaproteobacterial taxobin. These three proteins form a periplasmic-facing, membrane-bound complex presumably involved in the reduction of sulfur compounds including elemental sulfur, thiosulfate and polysulfide in *Sulfurovum* sp. NBC37-1. *Sulfurovum* sp. NBC37-1 has been shown to grow using hydrogen as an electron donor and elemental sulfur as an electron acceptor (Yamamoto *et al.*, 2010). The epsilonproteobacterial epibionts therefore might also be capable of anaerobic growth on these substrates, which are both present at the Rainbow site (Charlou *et al.*, 2002; Charlou *et al.*, 2010).

2.2.3. Hydrogenases

Hydrogen use by chemosynthetic symbionts was recently discovered in the sulfur-oxidizing symbionts of hydrothermal *Bathymodiolus* mussels (Petersen *et al.*, 2011). These authors also amplified the key gene for hydrogen oxidation, a [Ni-Fe] hydrogenase, from the epibionts of *R. exoculata*, as did Hügler et al. (2011). Because these genes were amplified from a mixed community, H_2-oxidizing potential could not be assigned to a particular epibiont, but phylogenetic analysis suggested that the hydrogenases originated from *Epsilon*- and *Deltaproteobacteria* (Hügler et al. 2011). Hydrogenases catalyze the reversible oxidation of hydrogen, and can be classified into four groups depending on their physiological function. We found hydrogenases belonging to H_2-uptake and H_2-sensing hydrogenases, in the epsilonproteobacterial taxobin (sequences: HG799450, HG799456, HG799457). Surprisingly, we also found hydrogenases in the gammaproteobacterial taxobin (sequences: HG799425 to HG799433, HG799436, HG799437, HG799441 and HG799444), suggesting that both dominant epibionts have the potential to use hydrogen as an energy source. In addition, H_2-evolving hydrogenases were found in the epsilonproteobacterial taxobin (sequence HG799459). We found 17 hits for *Epsilonproteobacteria* (sequences: HG799449, HG799451 to HG799455, HG799458, HG799460) and 14 hits for *Gammaproteobacteria* (sequences: HG799422 to HG799424, HG799434, HG799435, HG799438, HG799440, HG799442, HG799443, HG799445 to HG799448) to hydrogenase expression, formation, and maturation proteins, which are required for functional hydrogenase expression (Vignais, 2008). Two distinct genes encoding H_2-uptake hydrogenases were identified in each epibiont taxobin, indicating the potential importance of H_2 as an electron donor for the symbiosis (Vignais and Colbeau, 2004). This is consistent with the high concentrations of hydrogen measured in Rainbow hydrothermal fluids (Charlou et al., 2002). H_2-sensing hydrogenase genes were found in both epibionts. These cytoplasmic enzymes regulate respiratory [Ni-Fe] hydrogenase expression depending on the environmental H_2 concentration (Vignais and Colbeau, 2004). For *Epsilonproteobacteria*, it has also been hypothesized that cytoplasmic hydrogenases could supply the rTCA cycle with low-potential electrons, increasing growth efficiency (Sievert *et al.*, 2008a). Three sequences for H_2-evolving hydrogenases were identified in the epsilonproteobacterial taxobin. Genes for hydrogenases belonging to multiple hydrogenase functional groups have been reported in the genomes of other deep-sea hydrothermal vent *Epsilonproteobacteria*. For example, *Sulfurovum* sp. NBC37-1 has four different hydrogenases (two H_2-uptake type, one H_2-sensing type and one H_2-evolving type) and *Nitratiruptor* sp. SB155-2 has three hydrogenases (one of each of these types) (Nakagawa *et al.*, 2007). The capacity to use hydrogen is known in many free-living deep-sea *Epsilonproteobacteria* such as *Hydrogenimonas thermophila* (Takai *et al.*, 2004), *Sulfurovum lithotrophicum* (Inagaki *et al.*, 2004), or *Sulfurimonas paralvinellae* (Takai *et al.*, 2005; Takai *et al.*, 2006). This knowledge, together with our data and other recent findings (Petersen et al., 2011), strongly suggest that hydrogen is commonly used as energy source in free-living and host-associated microorganisms in hydrothermal ecosystems.

2.2.4. *Nitrogen assimilation and nitrate reduction*

Two pathways for nitrate reduction were identified in the epsilonproteobacterial taxobin: denitrification with two types of nitrate reductase (Nap and Nas) and dissimilatory nitrate reduction to ammonium via the Nrf system (Fig. 3, sequences: HG799223 to HG799252). The presence of both cytoplasmic (*nas*) and periplasmic (*nap*) nitrate reductase genes indicate the putative ability to adjust the mode of nitrate reduction to fluctuating nitrate concentrations. The presence of periplasmic dissimilatory nitrate reductase (*nap*), cytoplasmic assimilatory nitrate reductase (*nas*) and ammonia-forming nitrite reductase (*nrf*) genes indicate that the epsilonproteobacterial epibionts have the potential to produce ammonium. Consequently, they would not depend on environmental ammonium uptake, unlike the free-living *S. denitrificans*, which lacks *nrf* genes (Sievert *et al.*, 2008a). It is intriguing that the host-associated *Epsilonproteobacteria*, including the *R. exoculata* epibiont and bacteria of the genera *Campylobacter* and *Helicobacter* have the potential to produce ammonium for assimilation, whereas free-living *Epsilonproteobacteria*, such as *S. denitrificans*, rely on ammonium uptake from the environment. Most aquatic animals excrete ammonium, which should be available to their associated microorganisms (Wright, 1995). Accordingly, ammonium oxidation was recently detected in biofilms of marine invertebrates (Heisterkamp *et al.*, 2012). For the *R. exoculata* epsilonproteobacterial epibionts, the ability to produce ammonium as a nitrogen source might provide a competitive advantage during its free-living stage. Furthermore, in the gill chamber, this would prevent competition with co-occurring epibionts for ammonium. In the gammaproteobacterial taxobin, one Nap/Nir denitrification system was found, which would enable dissimilatory nitrate reduction to N_2O (Fig. 4, sequences: HG799257 to HG799272).

A more detailed description of nitrogen assimilation and nitrate reduction pathways in the epsilonproteobacterial taxobin is presented in the Supplementary text.

2.2.5. *Oxygen respiration*

The epsilonproteobacterial taxobin contained genes for a proton pumping cbb3-type cytochrome c oxidase (*cooN, cooO, cooP*) (sequences: HG799280 to HG799292), whereas genes for the subunit CooQ and the CooGHIS complex were not retrieved. Even though the absence of certain genes from a metagenome might simply be due to the fact that they were not sequenced, it is noteworthy that the *cooGHIS* genes are also lacking in the complete genome of the closest cultured relative of the *R. exoculata* epsilonproteobacterial epibionts, the marine epsilonproteobacterium *S. denitrificans* DSM1251. The CooGHIS complex is involved in assembly and maturation of the cbb3-type cytochrome c oxidase (Pitcher and Watmough, 2004), and the consequences of its absence in *S. denitrificans* DSM1251 are unclear (Sievert *et al.*, 2008a). Cbb3-type cytochrome c oxidase genes were also identified in the gammaproteobacterial taxobin (*cooP, cooN* and *cooG*) (sequences: HG799309 to HG799317). In *Proteobacteria*, cbb3-type cytochrome c oxidases are involved in micro-aerobic respiration (Kulajta *et al.*, 2006). The presence of these genes, especially *cooG*, indicates the ability to use oxygen under micro-oxic conditions, which is consistent with modeling studies that predict micro-oxic conditions in the gill chamber (Schmidt *et al.*, 2009). This is also supported by measurements from within shrimp aggregates, where levels of oxygen measured were half of those in the surrounding seawater (Zbinden *et al.*, 2004; Schmidt *et al.*, 2009). Moreover, the epibionts are located in the upper part of the gill chamber where the water flow is outgoing and thus, depleted in oxygen and enriched in carbon oxide due to shrimp respiration (Zbinden et al., 2004). The *Sulfurimonas gotlandica* GD-1 genome (Grote *et al.*, 2012) also features cbb3-type cytochrome c oxidase genes, but growth experiments showed that this strain could not grow with oxygen as the sole electron acceptor. It has been suggested that the cbb3-type cytochrome c oxidase genes might be used for protection from oxidative stress rather than for respiration (Sievert *et al.*, 2008a; Grote *et al.*, 2012).

Other genes encoding oxygen-reducing terminal oxidases such as cytochrome d ubiquinol oxidases were retrieved in the epsilon- (2 hits) and gammaproteobacterial (4 hits) taxobins. These cytochrome d ubiquinol oxidases are typically found in organisms living in low-oxygen environment which is probably the case for the epibionts.

2.2.6. Insights into the zetaproteobacterial epibiont metabolism

All the characterized *Zetaproteobacteria* are microaerophilic iron oxidizers. In the zetaproteobacterial taxobin, a molybdopterin oxidoreductase Fe_4S_4 region and ferredoxin encoding genes (5 hits) were identified (sequences: HG799404 to HG799409). Homologs of these genes are expressed during iron oxidation by *A. ferrooxidans* (Yarzábal *et al.*, 2002). One of the ferredoxin genes was located in a cluster with a polysulfide reductase, as identified in the iron oxidizers *M. ferrooxydans*, *Sideroxydans lithotrophicus* and *Gallionella capsiferriformans* (Singer *et al.*, 2011). It is therefore likely that the *Zetaproteobacteria* in the *R. exoculata* gill chamber are also iron oxidizers.

The zetaproteobacterial taxobin also contained a number of genes for the biosynthesis of energy storage compounds, including glycogen and polyphosphate (polyP) (9 hits with 5 polyP kinase sequences, sequences: HG799414 to HG799417). Such genes were also found in the gammaproteobacterial taxobin: polyP kinase (2 hits, sequences HG799419 and HG799420), dinucleoside polyP hydrolase (1 hit, sequence HG799418) and exopolyphosphatase (1 hit, sequence HG799421) (Fig. 4). Thus, both the zeta- and gammaproteobacterial epibionts might produce polyphosphate. Previous TEM observations and X-ray analyses corroborate these results by showing phosphate-rich granules in filamentous *R. exoculata* epibionts and in *M. ferrooxydans* cells (Zbinden *et al.*, 2008; Singer *et al.*, 2011). The polyphosphate granules serve as phosphorus and energy storage and may also have a detoxifying function by chelation of harmful heavy metal ions (Kornberg, 1995).

2.3. Symbiont-host interactions

The provision of organic compounds for nutrition is one of the potential benefits that animal host's derive from symbiotic associations with chemoautotrophic bacteria. The symbionts in such associations typically oxidize reduced sulfur compounds, methane or hydrogen with inorganic electron acceptors and couple this process to the synthesis of organic molecules from single-carbon compounds such as carbon dioxide or methane (Cavanaugh *et al.*, 2006; Dubilier *et al.*, 2008). Additional benefits for the host include detoxification of potent metabolic inhibitors such as iron and sulfide, to less toxic compounds such as elemental sulfur via oxidation (Fig. 5).

2.3.1. Role of the epibionts

Recently, it was demonstrated that *R. exoculata* shrimp can take up dissolved organic molecules from their gill epibionts, proving that the shrimp gain a nutritional benefit from their epibionts (Ponsard *et al.*, 2013). Amino acids, sugars and a variety of vitamins (such as vitamin B6 or pyridoxal 5'-phosphate) synthesized by the gamma- and epsilonproteobacterial epibionts could be suitable compounds for transfer to the *R. exoculata* host (Fig. 5). Amino acids are likely candidates, since lysine uptake has been demonstrated (Ponsard *et al.*, 2013), and various amino acid exporters were identified in the epibiont taxobins (Table S2).

An additional likely benefit for the *R. exoculata* shrimp is the contribution of its epibionts to the detoxification of abundant harmful compounds at Rainbow, such as heavy metals, sulfide, nitrite and Fe(II) (Charlou *et al.*, 2002; Charlou *et al.*, 2010). The epsilon- and gammaproteobacterial epibionts can convert H_2S to elemental sulfur, the epsilonproteobacterial epibionts can convert nitrite to dinitrogen, and the zetaproteobacterial epibionts can convert Fe(II) to insoluble iron oxides (Fig. 3). These processes are not only the key energy-generating pathways of the epibionts, but also have the beneficial side-effects that their products are less toxic for the shrimp than their reactants. For example, H_2S and sulfide have toxic effects on the aerobic respiration of hydrothermal fauna (Powell and Somero, 1986). Moreover, *Gamma*- and *Zetaproteobacteria* have the potential to form polyphosphate granules, which — as discussed above — may chelate toxic heavy metal ions.

2.3.2. Responses to oxidative and thermal stress

R. exoculata live in steep gradients between oxygenated seawater and anoxic hydrothermal fluids. Consequently, epibionts of the *R. exoculata* gill chamber microbes are continuously subjected to oxidative stress and fluctuating physicochemical conditions as the shrimp move in and out of the vent fluid-seawater mixing zone. Genes related to oxidative stress responses were also found here in epsilonproteobacterial taxobin (see Supplementary text for details). Hydrothermal fluids can reach up to 350°C and the temperatures in the mixing zones can reach up to 35°C (Ravaux *et al.*, 2003). Heat-inducible stress proteins (HSP70 family) were identified in *R. exoculata* (Ravaux *et al.*, 2003; Ravaux *et al.*, 2007). Genes encoding heat shock proteins were also identified in the metagenome for epsilon-, gamma- and zetaproteobacterial epibionts (Table S2, sequences: HG799339 to HG799381, and HG799461), indicating their ability to cope with the fluctuating conditions of the hydrothermal habitat.

2.3.3. Virulence gene homologues

Many *Epsilonproteobacteria*, such as *Campylobacter jejuni* or *Helicobacter pylori*, are well-known pathogens with characteristic virulence genes. Even non-pathogenic species, such as the deep-sea vent *Epsilonproteobacteria Sulfurovum* sp. and *Nitratiruptor* sp., share many virulence genes with pathogenic species (Nakagawa *et al.*, 2007). Homologs of these virulence factors were found in the *R. exoculata* epibiont metagenome (Table 2, sequencecs: HG799382 to HG799402). N-linked glycosylation (NLG) long believed to occur only in eukaryotes, was then discovered in the epsilonproteobacterium *Campylobacter jejuni* (Szymanski *et al.*, 1999) and now is recognized as a feature that occurs in members of all domains of life (Dell *et al.*, 2010). In eukaryotes, glycosylated proteins are involved in cell-cell and cell-matrix recognition. These proteins can also play a role in host-bacteria interactions, for example allowing *C. jejuni* to evade the human immune system when causing enteric infections (Szymanski and Wren, 2005). The NLG in deep-sea *Epsilonproteobacteria* might have emerged during the evolutionary transition from free-living to symbiotic association with hydrothermal vent invertebrates (Nakagawa *et al.*, 2007). It is possible that the NLG in these epsilonproteobacterial epibionts is involved in selective recognition between the epibionts and the shrimp. No such virulence genes were detected in the gammaproteobacterial taxobin.

2.3.4. Surface attachment

The specific colonization of animals by bacterial symbionts is mediated by recognition at the molecular level. Different mechanisms are well understood for some of the model associations, such as the squid *Euprymna scolopes* and its bioluminescent symbiont *Vibrio fisheri* (Troll *et al.*, 2009), but are still unknown for the majority of marine symbioses. The *R. exoculata* epibiont metagenome contained several genes coding for well-known cell-surface structures and extracellular attachment components, including type IV pili, mannose-sensitive haemagglutinin pili and O-antigens (capsular polysaccharides), which were found here in the epsilon- or gammaproteobacterial taxobins (Table S2). These taxobins also contained genes that may be involved in extracellular biofilm matrix formation, such as type II and III secretion systems, MscS family mechanosensitive channels, von Willebrand factor type A domain proteins and genes for exopolysaccharide biosynthesis (one of which had a best BLASTp hit to the epibiont of the deep-sea worm *Alvinella pompejana*) (Danese *et al.*, 2000; Whittaker and Hynes, 2002; Durand *et al.*, 2003; Thomas *et al.*, 2008). These data indicate that the epibionts might produce various extracellular structures as well as matrix polysaccharides allowing surface adhesion. The latter has been observed using TEM (Corbari et al., 2008, Zbinden et al., 2008).

3. Conclusions

Two distinct sulfur oxidizers dominate the *R. exoculata* gill chamber epibiosis. On all the microscopic observations, the *Epsilon*- and *Gammaproteobacteria* occupy the same zones and are present on the whole surface of the branchiostegite. These co-occurring epibionts carry out similar processes such as sulfur and hydrogen oxidation and carbon dioxide fixation. This seeming functional redundancy raises the question of why the epibiont with the greater fitness does not prevail in a Darwinian extinction by the process of competition. We propose that even subtle differences between co-occurring epibionts, such as the use of different CO_2 fixation pathways or the possible independence of the *Epsilonproteobacteria* from ammonia provide sufficient niche differentiation between the epibionts, which prevent direct competition and allow their stable co-occurrence. Similar associations with multiple distinct sulfur-oxidizing symbionts have also been described for bivalves (Duperron *et al.*, 2007; Duperron *et al.*, 2008; Raggi *et al.*, 2013).

Furthermore, Prosser (2012) proposed that even if two strains appear to have almost identical genetic characteristics, they will probably respond differently to environmental changes. Here, environmental drivers of niche differentiation might include oxygen or sulfide availability, as has been demonstrated for closely associated sulfur-oxidizing *Gamma*- and *Epsilonproteobacteria* in a terrestrial cave (Macalady *et al.*, 2008).

Thus, the metabolic plasticity of the epibiotic community associated with *R. exoculata* probably confers an adaptive advantage for the shrimp in the highly dynamic hydrothermal mixing zone they colonize, and could help to explain their large success in colonizing many geochemically and physically contrasting hydrothermal vents along the Mid-Atlantic Ridge.

4. Experimental procedures

4.1. Shrimp collection and DNA extraction

Rimicaris exoculata specimens were sampled during the 2007 MoMARDREAM-Naut oceanographic cruise at the Rainbow hydrothermal site (36°14’N – 33°54’W, 2,320 m depth) on the MAR. Shrimps were collected using the Ifremer research vessel *Pourquoi Pas?* and the submersible *Nautile*. Once aboard, some *R. exoculata* specimens were immediately frozen at -80 °C. In the laboratory, three of these specimens were dissected to separate the mouthparts and the inner face of the gill chamber (branchiostegite). These parts were immediately used to extract DNA using a phenol, chloroform and isoamyl alcohol extraction (Zbinden and Cambon-Bonavita, 2003). The quality and size of the extracted genomic DNA was assessed by 0.6% agarose gel electrophoresis. Aliquots of extracted genomic DNA, fosmid control DNA (40 kb size, Epicentre Biotechnologies, Madison, WI, USA), and BAC-Tracker™ Supercoiled DNA Ladder (Epicentre Biotechnologies) were loaded in parallel for size estimation.

4.2. DNA sequencing and assembly

DNA sequencing was performed with a Roche 454 GS FLX Ti sequencer (454 Life Sciences, Branford, CT, USA) at the Centre for Genomic Research of the University of Liverpool with 1.5 picotiter plates (1,011,151 reads amounting to 318 Mbp). Assembly was carried out with a Roche 454 Newbler v.2.3 assembler using default parameters, resulting in 24,529 contigs (831 contigs ≥ 2.5 kbp) of 15.3 Mbp in total (11.1 Mbp ≥ 500 bp). For further details see Table S1.

4.3. Taxonomic classification and annotation

A consensus from five individual taxonomic prediction tools was used to infer the taxonomic affiliation of the metagenome sequences as previously described (Ferrer *et al.*, 2012), with modifications as follows. Competitive read recruitment was run with 339 genomes of marine microbes plus three bacterial genomes that, based on 16S rRNA analyses, are currently closest to the main epibionts and zetaproteobacterial epibiont: the *gammaproteobacterium Leucothrix mucor* (unpublished), the *epsilonproteobacterium Sulfurovum* sp. NBC37-1 (NCBI RefSeq NC_009663.1), and the *zetaproteobacterium Mariprofundus ferrooxydans* PV-1 (NCBI RefSeq NZ_AATS0100000000.1). A rank-based approach was used to consolidate taxonomic predictions of the five tools into a consensus using a weighted assessment on all 27 existing ranks of the NCBI taxonomy from superkingdom to species. As a result, a substantial fraction of the sequences, from the *R. exoculata* gill microbiota metagenome could be classified [superkingdom: 9.9 Mbp (64%); phylum: 9.6 Mbp (63%); class: 9.0 Mbp (59%); genus: 6.6 Mbp (44%)].

Annotation was done using a modified GenDB v2.2.1 annotation system (Meyer *et al.*, 2003) as described previously (Ferrer *et al.*, 2012). The JCoast software (Richter *et al.*, 2008) was used for manual annotation and data mining. Annotations were compared with annotations obtained by RAST (Aziz *et al.*, 2008) for epsilon-, gamma- and zetaproteobacterial taxonomic sequence bins (taxobins) individually. The sequence data reported in this study are available at the European Bioinformatics Institute (study number ERP001477; http://www.ebi.ac.uk/ena/data/view/ERP001477). Sequence data have been submitted to the EMBL database under accession numbers HG799104 to HG799461.

4.4. Fluorescence in situ hybridization (FISH)

FISH was performed on 0.6 µm transverse sections of *R. exoculata* branchiostegite (Guri et al., 2012) as previously described (Durand *et al.*, 2010), except for two main modifications: the hybridization and washing temperatures were 35°C and 37°C, respectively, and the reaction mix contained 55% formamide hybridization buffer. Branchiostegite sections were hybridized using probes Eub338 (Amann *et al.*, 1990) and ZETA123 (Kato *et al.*, 2009). Observations and imaging were performed using an Apotome Axio Imager Z2 with a COLIBRI system (Carl Zeiss Microimaging GmbH, Göttingen, Germany).

5. Acknowledgements

This research was supported by the European Community project MAMBA (FP7-KBBE-2008-226977) and by the International Doctoral College of UEB (European University of Brittany). We thank Joël Querellou for coordination between the MAMBA project and Ifremer, and him and Marie Roumagnac for their contributions. We thank the chief scientist of the MOMARDREAM-Naut cruise, the captain and crew of *Pourquoi Pas?*, and the *Nautile* team. We are grateful to Helen McCombie for helpful language improvement. Finally, we thank M. R. Belgacem for his help with the graphics.

6. References

Amann R.I., Binder B.J., Olson R.J., Chisholm S.W., Devereux R. and Stahl D.A. (1990). Combination of 16S ribosomal-RNA-targeted oligonucleotide probes with flow-cytometry for analyzing mixed microbial-populations. *Appl Environ Microbiol,* **56:** 1919-1925.

Aziz R., Bartels D., Best A., DeJongh M., Disz T., Edwards R. *et al.* (2008). The RAST server: Rapid Annotations using Subsystems Technology. *BMC Genomics,* **9:** 75.

Badger M.R. and Bek E.J. (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.

Bassler B.L. (2002). Small talk. Cell-to-cell communication in bacteria. *Cell,* **109:** 421-424.

Berg I.A. (2011). Ecological spects of the distribution of different autotrophic CO_2 fixation pathways. *App Environ Microbiol,* **77:** 1925-1936.

Campbell B.J. and Cary S.C. (2004). Abundance of reverse tricarboxylic acid cycle genes in free-living microorganisms at deep-sea hydrothermal vents. *Appl Environ Microbiol,* **70:** 6282-6289.

Campbell B.J., Engel A.S., Porter M.L. and Takai K. (2006). The versatile *Epsilon-proteobacteria*: key players in sulphidic habitats. *Nat Rev Micro,* **4:** 458-468.

Cavanaugh C.M., McKiness Z.P., Newton I.L.G. and Stewart F.J. (2006). The Prokaryotes : a handbook on the biology of bacteria., Springer: New York.

Charlou J.L., Donval J.P., Fouquet Y., Jean-Baptiste P. and Holm N. (2002). Geochemistry of high H_2 and CH_4 vent fluids issuing from ultramafic rocks at the Rainbow hydrothermal field (36°14 ' N, MAR). *Chem Geol,* **191:** 345-359.

Charlou J.L., Donval J.P., Konn C., Ondreas H., Fouquet Y., Jean-Baptiste P. *et al.* (2010). High production of H_2 and CH_4 and abiotic hydrocarbons in ultramafic-hosted hydrothermal systems on the Mid-Atlantic Ridge. *Geoch Cosmochim Acta,* **74:** A163-A163.

Corbari L., Zbinden M., Cambon-Bonavita M.A., Gaill F. and Compere P. (2008a). Bacterial symbionts and mineral deposits in the branchial chamber of the hydrothermal vent shrimp *Rimicaris exoculata*: relationship to moult cycle. *Aqua Biol,* **1:** 225-238.

Corbari L., Cambon-Bonavita M.A., Long G.J., Grandjean F., Zbinden M., Gaill F. *et al.* (2008b). Iron oxide deposits associated with the ectosymbiotic bacteria in the hydrothermal vent shrimp *Rimicaris exoculata*. *Biogeosciences,* **5:** 1295-1310.

Danese P.N., Pratt L.A. and Kolter R. (2000). Exopolysaccharide production is required for development of Escherichia coli K-12 biofilm architecture. *J Bacteriol,* **182:** 3593-3596.

Dubilier N., Bergin C. and Lott C. (2008). Symbiotic diversity in marine animals: the art of harnessing chemosynthesis. *Nat Rev Micro,* **6:** 725-740.

Duperron S., Sibuet M., MacGregor B.J., Kuypers M.M.M., Fisher C.R. and Dubilier N. (2007). Diversity, relative abundance and metabolic potential of bacterial endosymbionts in three *Bathymodiolus* mussel species from cold seeps in the Gulf of Mexico. *Environ Microbiol,* **9:** 1423-1438.

Duperron S., Halary S., Lorion J., Sibuet M. and Gaill F. (2008). Unexpected co-occurrence of six bacterial symbionts in the gills of the cold seep mussel *Idas sp.* (*Bivalvia*: *Mytilidae*). *Environ Microbiol,* **10:** 433-445.

Durand É., Bernadac A., Ball G., Lazdunski A., Sturgis J.N. and Filloux A. (2003). Type II protein secretion in *Pseudomonas aeruginosa*: the pseudopilus is a multifibrillar and adhesive structure. *J Bacteriol,* **185:** 2749-2758.

Durand L., Zbinden M., Cueff-Gauchard V., Duperron S., Roussel E.G., Shillito B. *et al.* (2010). Microbial diversity associated with the hydrothermal shrimp *Rimicaris exoculata* gut and occurrence of a resident microbial community. *FEMS Microbiol Ecol,* **71** 291–303.

Emerson D., Rentz J.A., Lilburn T.G., Davis R.E., Aldrich H., Chan C. *et al.* (2007). A novel lineage of *Proteobacteria* involved in formation of marine Fe-oxidizing microbial mat communities. *PLOS One,* **2:** e667.

Federle M.J. and Bassler B.L. (2003). Interspecies communication in bacteria. *J Clin Invest,* **112:** 1291-1299.

Ferrer M., Guazzaroni M.-E., Richter M., García-Salamanca A., Yarza P., Suárez-Suárez A. *et al.* (2011). Taxonomic and functional metagenomic profiling of the microbial community in the anoxic sediment of a sub-saline shallow lake (Laguna de Carrizo, Central Spain). *Micro Ecol,* **62:** 824-837.

Ferrer M., Werner J., Chernikova T.N., Bargiela R., Fernández L., La Cono V. *et al.* (2012). Unveiling microbial life in the new deep-sea hypersaline lake *Thetis*. Part II: a metagenomic study. *Environ Microbiol,* **14:** 268-281.

Filloux A. and Vallet I. (2003). Biofilm: set-up and organization of a bacterial community. *Med Sci (Paris),* **19:** 77-83.

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

Galperin M.Y. (2006). Structural classification of bacterial response regulators: diversity of output domains and domain combinations. *J Bacteriol,* **188:** 4169-4182.

Grabovich M.Y., Muntyan M.S., Lebedeva V.Y., Ustiyan V.S. and Dubinina G.A. (1999). Lithoheterotrophic growth and electron transfer chain components of the filamentous gliding bacterium *Leucothrix mucor* DSM 2157 during oxidation of sulfur compounds. *FEMS Microbiol Lett,* **178:** 155-161.

Grimm F., Dobler N. and Dahl C. (2010). Regulation of dsr genes encoding proteins responsible for the oxidation of stored sulfur in *Allochromatium vinosum*. *Microbiol,* **156:** 764-773.

Grote J., Schott T., Bruckner C.G., Glöckner F.O., Jost G., Teeling H. *et al.* (2012). Genome and physiology of a model *epsilonproteobacterium* responsible for sulfide detoxification in marine oxygen depletion zones. *Proc Natl Acad Sci USA,* **109:** 506-510.

Guri M., Durand L., Cueff-Gauchard V., Zbinden M., Crassous P., Shillito B. *et al.* (2012). Acquisition of epibiotic bacteria along the life cycle of the hydrothermal shrimp *Rimicaris exoculata*. *ISME J,* **6:** 597-609.

Heisterkamp I.M., Schramm A., Larsen L.H., Svenningsen N.B., Lavik G., de Beer D. *et al.* (2012). Shell biofilm-associated nitrous oxide production in marine molluscs: processes, precursors and relative importance. *Environ Microbiol,* **15:** 1943-1955.

Hensen D., Sperling D., Trüper H.G., Brune D.C. and Dahl C. (2006). Thiosulphate oxidation in the phototrophic sulphur bacterium *Allochromatium vinosum*. *Mol Microbiol,* **62:** 794-810.

Hughes D.T. and Sperandio V. (2008). Inter-kingdom signalling: communication between bacteria and their hosts. *Nat Rev Micro,* **6:** 111-120.

Hügler M., Petersen J.M., Dubilier N., Imhoff J.F. and Sievert S.M. (2011). Pathways of carbon and energy metabolism of the epibiotic community associated with the deep-sea hydrothermal vent shrimp *Rimicaris exoculata PLOS ONE,* **6:** e16018.

Inagaki F., Takai K., Kobayashi H., Nealson K.H. and Horikoshi K. (2003). *Sulfurimonas autotrophica* gen. nov., sp. nov., a novel sulfur-oxidizing ε-proteobacterium isolated from hydrothermal sediments in the Mid-Okinawa Trough. *Int J Syst Evol Microbiol* **53:** 1801-1805.

Inagaki F., Takai K., Nealson K.H. and Horikoshi K. (2004). *Sulfurovum lithotrophicum* gen. nov., sp. nov., a novel sulfur-oxidizing chemolithoautotroph within the ε-*Proteobacteria* isolated from Okinawa Trough hydrothermal sediments. *Int J Syst Evol Microbiol,* **54:** 1477-1482.

Kato S., Yanagawa K., Sunamura M., Takano Y., Ishibashi J.-i., Kakegawa T. *et al.* (2009). Abundance of *Zetaproteobacteria* within crustal fluids in back-arc hydrothermal fields of the Southern Mariana Trough. *EnvironMicrobiol,* **11:** 3210-3222.

Kornberg A. (1995). Inorganic polyphosphate: toward making a forgotten polymer unforgettable. *J Bacteriol,* **177:** 491-496.

Kulajta C., Thumfart J.O., Haid S., Daldal F. and Koch H.-G. (2006). Multi-step assembly pathway of the cbb3-type cytochrome c oxidase complex. *J Mol Biol* **355:** 989-1004.

Macalady J.L., Dattagupta S., Schaperdoth I., Jones D.S., Druschel G.K. and Eastman D. (2008). Niche differentiation among sulfur-oxidizing bacterial populations in cave waters. *ISME J,* **2:** 590-601.

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

Meyer F., Goesmann A., McHardy A.C., Bartels D., Bekel T., Clausen J. *et al.* (2003). GenDB - an open source genome annotation system for prokaryote genomes. *Nucleic Acids Res* **31:** 2187-2195.

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

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

Petersen J.M., Ramette A., Lott C., Cambon-Bonavita M.A., Zbinden M. and Dubilier N. (2010). Dual symbiosis of the vent shrimp *Rimicaris exoculata* with filamentous *Gamma*- and

Epsilonproteobacteria at four Mid-Atlantic Ridge hydrothermal vent fields. *Environ Microbiol,* **12:** 2204–2218.

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

Pitcher R.S. and Watmough N.J. (2004). The bacterial cytochrome cbb3 oxidases. *BBA - Bioenergetics,* **1655:** 388-399.

Polz M.F., Robinson J.J., Cavanaugh C.M. and Van Dover C.L. (1998). Trophic ecology of massive shrimp aggregations at a Mid-Atlantic Ridge hydrothermal vent site. *Limnol Oceanograph,* **43:** 1631-1638.

Ponsard J., Cambon-Bonavita M.-A., Zbinden M., Lepoint G., Joassin A., Corbari L. *et al.* (2013). Inorganic carbon fixation by chemosynthetic ectosymbionts and nutritional transfers to the hydrothermal vent host-shrimp *Rimicaris exoculata*. *ISME J,* **7:** 96–109.

Raggi L., Schubotz F., Hinrichs K.U., Dubilier N. and Petersen J.M. (2013). Bacterial symbionts of *Bathymodiolus* mussels and *Escarpia* tubeworms from Chapopote, an asphalt seep in the southern Gulf of Mexico. *Environ Microbiol,* **15:** 1969-1987.

Ravaux J., Gaill F., Le Bris N., Sarradin P.M., Jollivet D. and Shillito B. (2003). Heat-shock response and temperature resistance in the deep-sea vent shrimp *Rimicaris exoculata*. *J Exp Biol,* **206:** 2345-2354.

Ravaux J., Toullec J.Y., Leger N., Lopez P., Gaill F. and Shillito B. (2007). First hsp70 from two hydrothermal vent shrimps, *Mirocaris fortunata* and *Rimicaris exoculata*: characterization and sequence analysis. *Gene,* **386:** 162-172.

Richter M., Lombardot T., Kostadinov I., Kottmann R., Duhaime M., Peplies J. *et al.* (2008). JCoast - A biologist-centric software tool for data mining and comparison of prokaryotic (meta)genomes. *BMC Bioinformatics,* **9:** 177.

Ryjenkov D.A., Tarutina M., Moskvin O.V. and Gomelsky M. (2005). Cyclic diguanylate is a ubiquitous signaling molecule in bacteria: insights into biochemistry of the GGDEF protein domain. *J Bacteriol,* **187:** 1792-1798.

Schmidt C., Le Bris N. and Gaill F. (2008). Interactions of deep-sea vent invertebrates with their environment: the case of *Rimicaris exoculata*. *J Shellfish Res,* **27:** 79-90.

Schmidt C., Corbari L., Gaill F. and Le Bris N. (2009). Biotic and abiotic controls on iron oxyhydroxide formation in the gill chamber of the hydrothermal vent shrimp *Rimicaris exoculata*. *Geobiology,* **7:** 454-464.

Sievert S.M., Scott K.M., Klotz M.G., Chain P.S.G., Hauser L.J., Hemp J. *et al.* (2008a). Genome of the epsilonproteobacterial chemolithoautotroph *Sulfurimonas denitrificans*. *Appl Environ Microbiol,* **74:** 1145-1156.

Sievert S.M., Hügler M., Taylor C.D. and Wirsen C.O. (2008b). Sulfur oxidation at deep-sea hydrothermal vents. *Microbial sulfur metabolism*. Springer Berlin Heidelberg. pp 238-258.

Singer E., Emerson D., Webb E.A., Barco R.A., Kuenen J.G., Nelson W.C. *et al.* (2011). *Mariprofundus ferrooxydans* PV-1 the first genome of a marine Fe(II) oxidizing *Zetaproteobacterium*. *PLOS One,* **6:** e25386.

Szymanski C.M., Yao R., Ewing C.P., Trust T.J. and Guerry P. (1999). Evidence for a system of general protein glycosylation in *Campylobacter jejuni*. *Mol Microbiol,* **32:** 1022-1030.

Szymanski C.M. and Wren B.W. (2005). Protein glycosylation in bacterial mucosal pathogens. *Nat Rev Micro,* **3:** 225-237.

Takai K., Nealson K.H. and Horikoshi K. (2004). *Hydrogenimonas thermophila* gen. nov., sp. nov., a novel thermophilic, hydrogen-oxidizing chemolithoautotroph within the *ε-Proteobacteria*, isolated from a black smoker in a Central Indian Ridge hydrothermal field. *Int J Syst Evol Microbiol,* **54:** 25-32.

Takai K., Campbell B.J., Cary S.C., 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., Suzuki M., Nakagawa S., Miyazaki M., Suzuki Y., Inagaki F. *et al.* (2006). *Sulfurimonas paralvinellae* sp. nov., a novel mesophilic, hydrogen- and sulfur-oxidizing chemolithoautotroph within the *Epsilonproteobacteria* isolated from a deep-sea hydrothermal vent polychaete nest, reclassification of *Thiomicrospira denitrificans* as *Sulfurimonas denitrificans* comb. nov. and emended description of the genus *Sulfurimonas*. *Int J Syst Evol Microbiol,* **56:** 1725-1733.

Thomas T., Evans F.F., Schleheck D., Mai-Prochnow A., Burke C., Penesyan A. *et al.* (2008). Analysis of the *Pseudoalteromonas tunicata* genome reveals properties of a surface-associated life style in the marine environment. *PLOS ONE,* **3:** e3252.

Troll J.V., Adin D.M., Wier A.M., Paquette N., Silverman N., Goldman W.E. *et al.* (2009). Peptidoglycan induces loss of a nuclear peptidoglycan recognition protein during host tissue development in a beneficial animal-bacterial symbiosis. *Cell Microbiol,* **11:** 1114-1127.

Vignais P.M. and Colbeau A. (2004). Molecular biology of microbial hydrogenases. *Cur Iss Mol Biol* **4:** 159-188.

Vignais P.M. (2008). Regulation of hydrogenase gene expression. In: Govindjee (ed). *The purple phototrophic bacteria*. Springer Netherlands. pp 743-757.

Wahl M., Goecke F., Labes A., Dobretsov S. and Weinberger F. (2012). The second skin: ecological role of epibiotic biofilms on marine organisms. *Front Microbiol,* **3:** 292.

Whittaker C.A. and Hynes R.O. (2002). Distribution and evolution of von Willebrand/Integrin A domains: widely dispersed domains with roles in cell adhesion and elsewhere. *Mol Biol Cell,* **13:** 3369-3387.

Williams A.B. and Rona P.A. (1986). Two new caridean shrimps (*Bresiliidae*) from a hydrothermal field on the mid-atlantic ridge, **vol. 6**. Crustacean Society: San Antonio, TX, USA.

Wright P.A. (1995). Nitrogen excretion: three end products, many physiological roles. *J Exp Biol,* **198:** 273-281.

Yamamoto M., Nakagawa S., Shimamura S., Takai K. and Horikoshi K. (2010). Molecular characterization of inorganic sulfur-compound metabolism in the deep-sea *epsilonproteobacterium Sulfurovum* sp. NBC37-1. *Environ Microbiol,* **12:** 1144-1153.

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

Yarzábal A., Brasseur G., Ratouchniak J., Lund K., Lemesle-Meunier D., DeMoss J.A. *et al.* (2002). The high-molecular-weight cytochrome c Cyc2 of *Acidithiobacillus ferrooxidans* is an outer membrane protein. *J Bacteriol,* **184:** 313-317.

Zbinden M. and Cambon-Bonavita M.A. (2003). Occurrence of Deferribacterales and Entomoplasmatales in the deep-sea Alvinocarid shrimp *Rimicaris exoculata* gut. *FEMS Microbiol Ecol,* **46:** 23-30.

Zbinden M., Bris N.L., Gaill F. and Compère P. (2004). Distribution of bacteria and associated minerals in the gill chamber of the vent shrimp *Rimicaris exoculata* and related biogeochemical processes. *Mar Ecol Progr Series,* **284:** 237-251.

Zbinden M., Shillito B., Le Bris N., de Villardi de Montlaur C., Roussel E., Guyot F. *et al.* (2008). New insights on the metabolic diversity among the epibiotic microbial communitiy of the hydrothermal shrimp *Rimicaris exoculata*. *J Experim Mar Biol Ecol,* **359:** 131-140.

Zhu F., Massana R., Not F., Marie D. and Vaulot D. (2005). Mapping of picoeucaryotes in marine ecosystems with quantitative PCR of the 18S rRNA gene. *FEMS Microbiol Ecol,* **52:** 79-92.

Tables

Table 1: General features of the epsilonproteobacterial, gammaproteobacterial and zetaproteobacterial sequences of the binning metagenomic dataset from the symbiotic community of the gill chamber of *R. exoculata*. The reads with the zetaproteobacterial 16S fragments could not be assembled and/or classified, explaining the absence of 16S rRNA sequences within the zetaproteobacterial taxobin. *The numbers between brackets correspond to the number of 16S rRNAs fragments identified in the whole metagenome, including taxobins and the unclassified reads.

	Gill chamber metagenome - *Epsilonproteobacteria*	**Gill chamber metagenome - *Gammaproteobacteria***	**Gill chamber metagenome - *Zetaproteobacteria***
Total contigs	2,106	2,441	558
Assembly size (bp)	2,887,614	3,427,427	468,699
Min size (bp)	100	100	109
Max size (bp)	31,453	14,111	6,412
GC content	34.3%	38.8%	44.5%
Coding density	87%	84%	88%
Number of tRNAs	71	32	0
Number of rRNAs *	7 (373)	1 (71)	0 (25)
Total ORFs	9,728	11,094	1,809
Hypothetical proteins	1,167	2,054	269
Conserved hypothetical proteins	855	689	106

Table 2: Comparison of genes from the epsilonproteobacterial taxobin with two related reference genomes.

Response to environment		**Gill chamber *Epsilonproteobacteria* 2.9 Mbp**	***Sulfurovum* sp. NBC37-1 2.6 Mbp**	***Nitratiruptor* sp. SB155-2 1.9 Mbp**
Sensory systems	Sensor histidine protein kinases (HPK) genes	26	16	26
	Transcriptional factor genes	58	34	26
	Chemotaxis proteins	2	0	12
	Methyl-accepting chemotaxis genes	5	0	6
	Proteins with PAS/GGDEF domains genes	17	23	36
Heavy metal transport		20 (Ni, Fe, Zn, Cd, Co, Mg, pb, Mo)	22 (Ni, Fe, Zn, Mn, Co, Cd, Cu, As, Mo, Hg, Ag)	17 (Ni, Fe, Zn, Mn, Co, Cd, Cu, As, Mo, Hg, Ag)
Oxidative stress	Catalase	-	+	-
	Superoxide dismutase	+	-	-
	Cytochrome c peroxidase	+	+	+
	Thiol peroxidase	+	+	+
Virulence factors	Virulence factor MviN	+	+	+
	Hemolysin	+	+	+
	Fibronectin- and fibrinogen-binding protein	+	+	+
	N-linked glycosylation pathway	+	+	+
	Invasion antigen CiaB	+	+	+
	Lytic murein transglycosylase	+	+	+

Figures

Figure 1: Distribution of (a) superkingdom and (b) class affiliations of the assembled metagenomic sequences, and (c) distribution of domain, class and genus affiliations of metagenomic 16S and 18S rRNA reads.

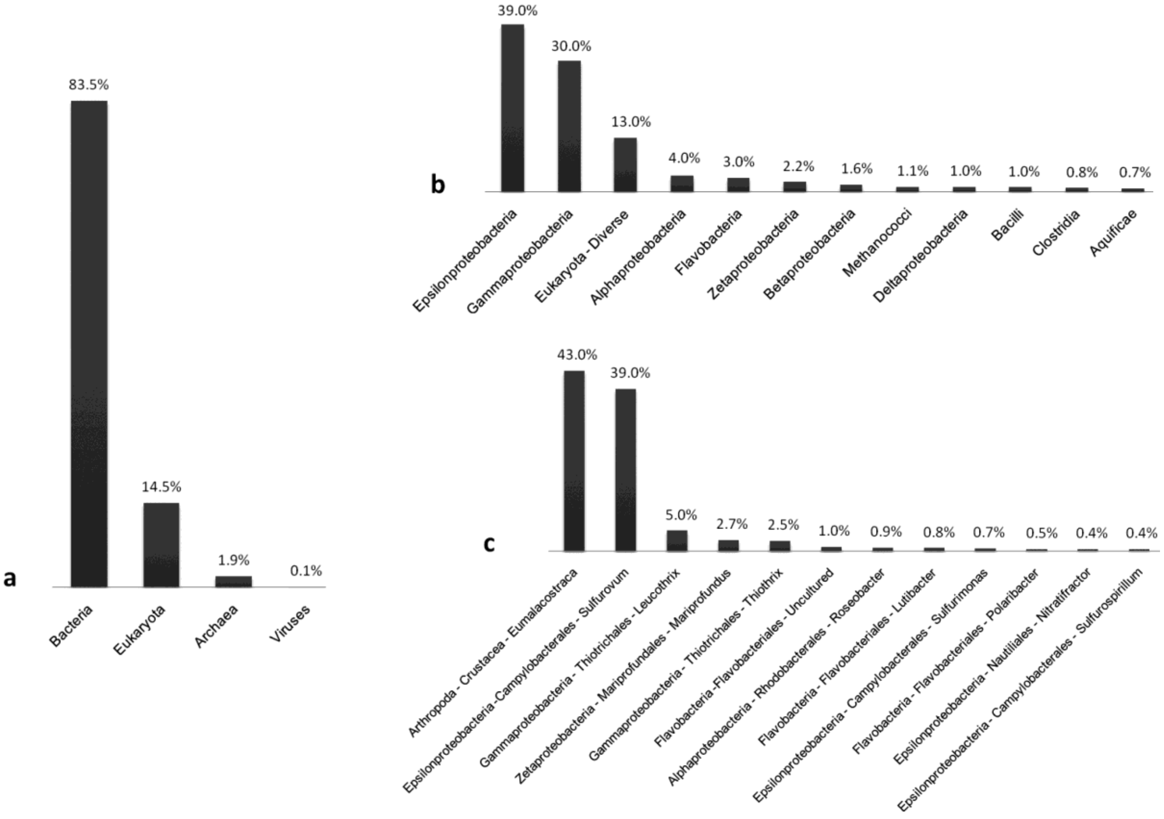


Figure 2: FISH observation of Rainbow *R. exoculata* specimen gill branchiostegite epibionts. Transverse view of a branchiostegite from a Rainbow vent shrimp with epibionts. *Zetaproteobacteria* (red) were hybridized with the ZETA123 probe. All cells were labeled with DAPI (blue). N: eukaryotic nuclei. Z*etaproteobacteria* were located along the shrimp membrane. Z*etaproteobacteria* were located inside a thick layer of iron oxides (IL).

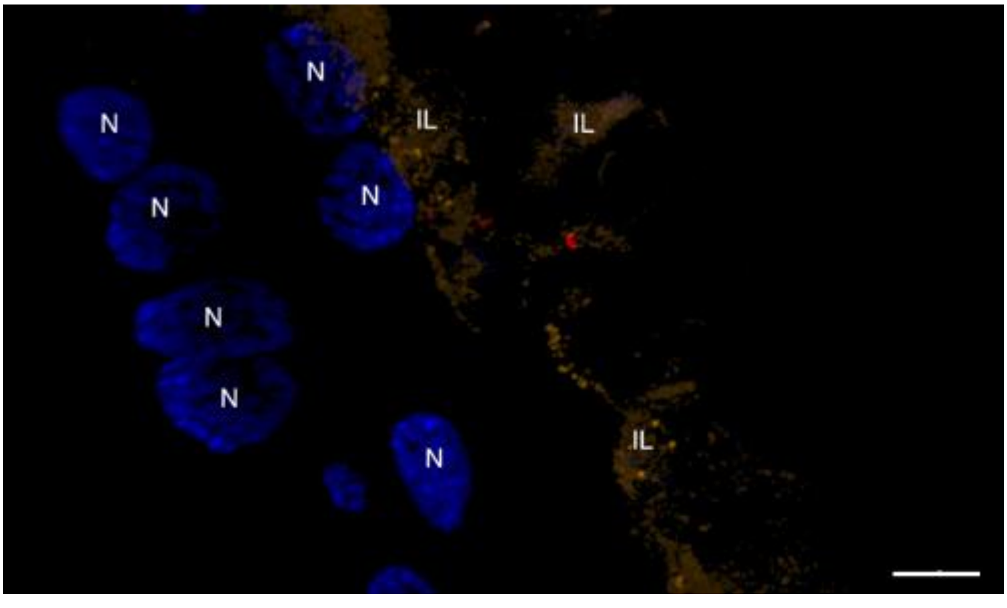


Figure 3: Central metabolism and solute transport in the epsilonproteobacterial epibionts. Cyt: cytochrome; H_2ase: hydrogenase; Sqr: sulfide-quinone oxidoreductase; Nap: periplasmic nitrate reductase; Nas: cytoplasmic nitrate reductase; Anr: ammonia-forming siroheme nitrite reductase; Nir: cytochrome *cd*$_1$-dependent nitrite reductase; Nrf: cytochrome c nitrite reductase; cNor: nitric oxide reductase; cNos: nitrous oxide reductase. Amino acids are indicated in red on a pink background. Blue striped ovals correspond to diffusing CO_2 particles that are used for the rTCA cycle. Yellow striped stars correspond to final products likely excretion products. Dotted arrows correspond to indirect synthesis (for amino acids), enzymatic action (for cytochrome c) or diffusion (for H_2 and H^+).

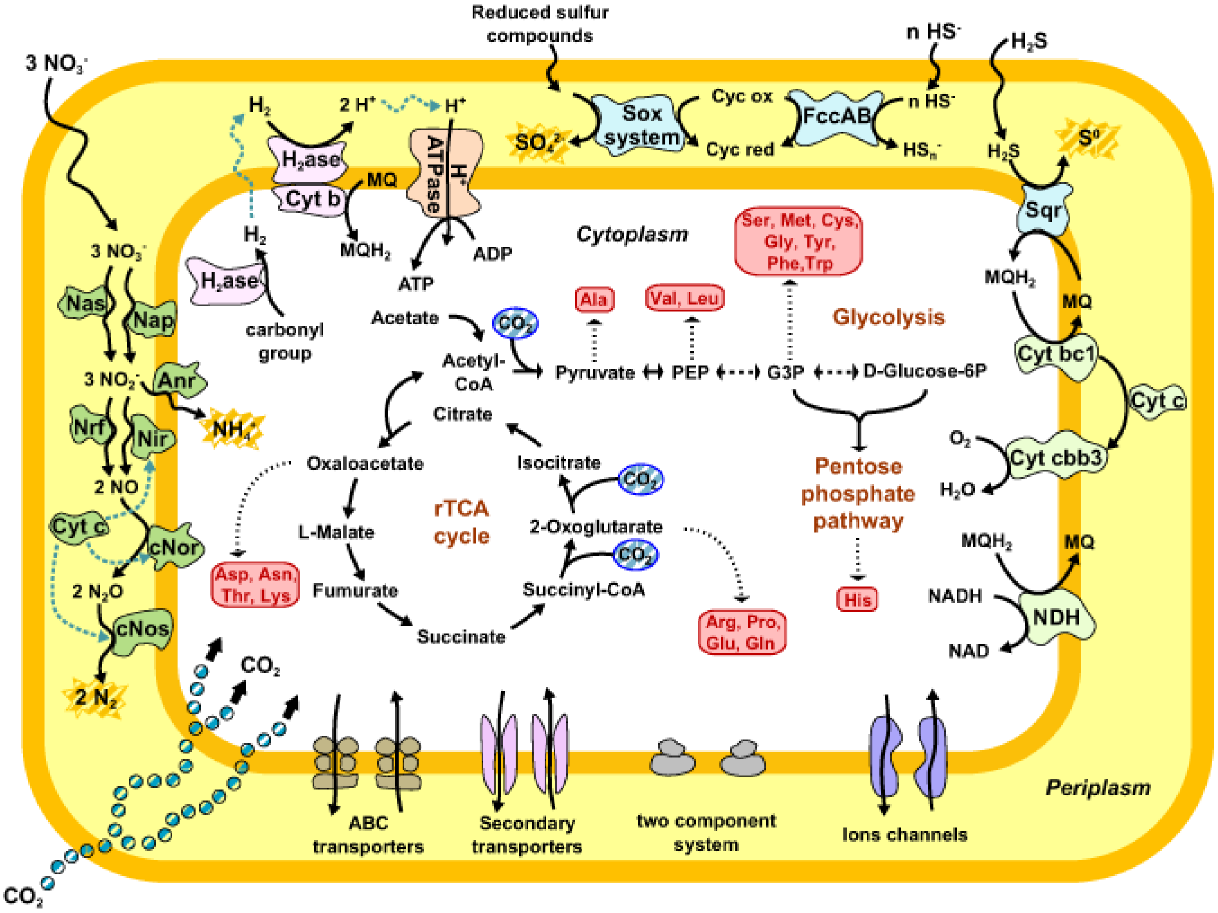


Figure 4: Central metabolism and solute transport in the gammaproteobacterial epibionts. Cyt: cytochrome; H_2ase: hydrogenase; Sqr: sulfide-quinone oxidoreductase; Nap: periplasmic nitrate reductase; Nir: cytochrome *cd*$_1$-dependent nitrite reductase; cNor: nitric oxide reductase; PolyP: polyphosphates. Amino acids are shown in red on a pink background. Blue striped ovals correspond to diffusing CO_2 particles used for the rTCA cycle. Yellow striped stars correspond to final products likely excretion products. Dotted arrows correspond to indirect synthesis (for amino acids), enzymatic action (for cytochrome c) or diffusion (for H_2 and H^+).

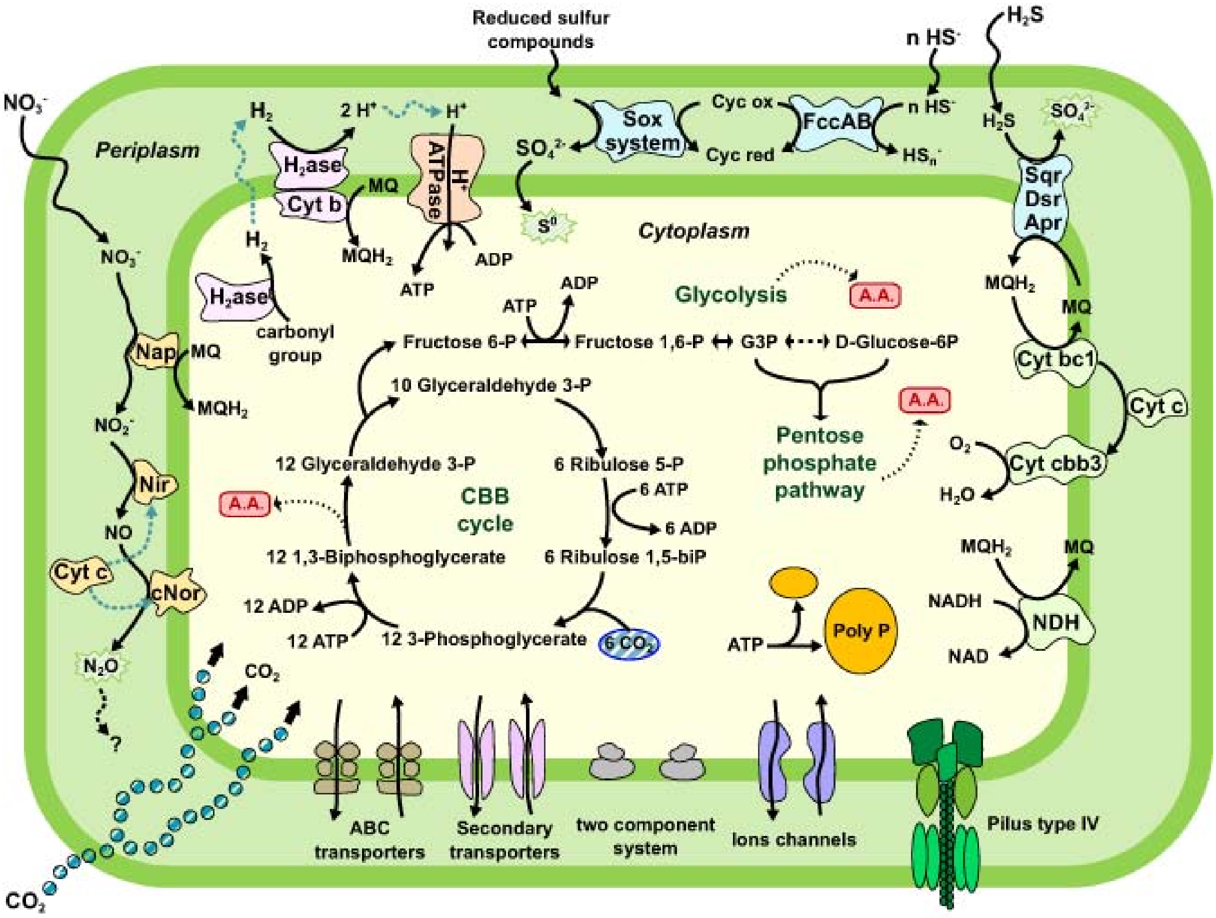
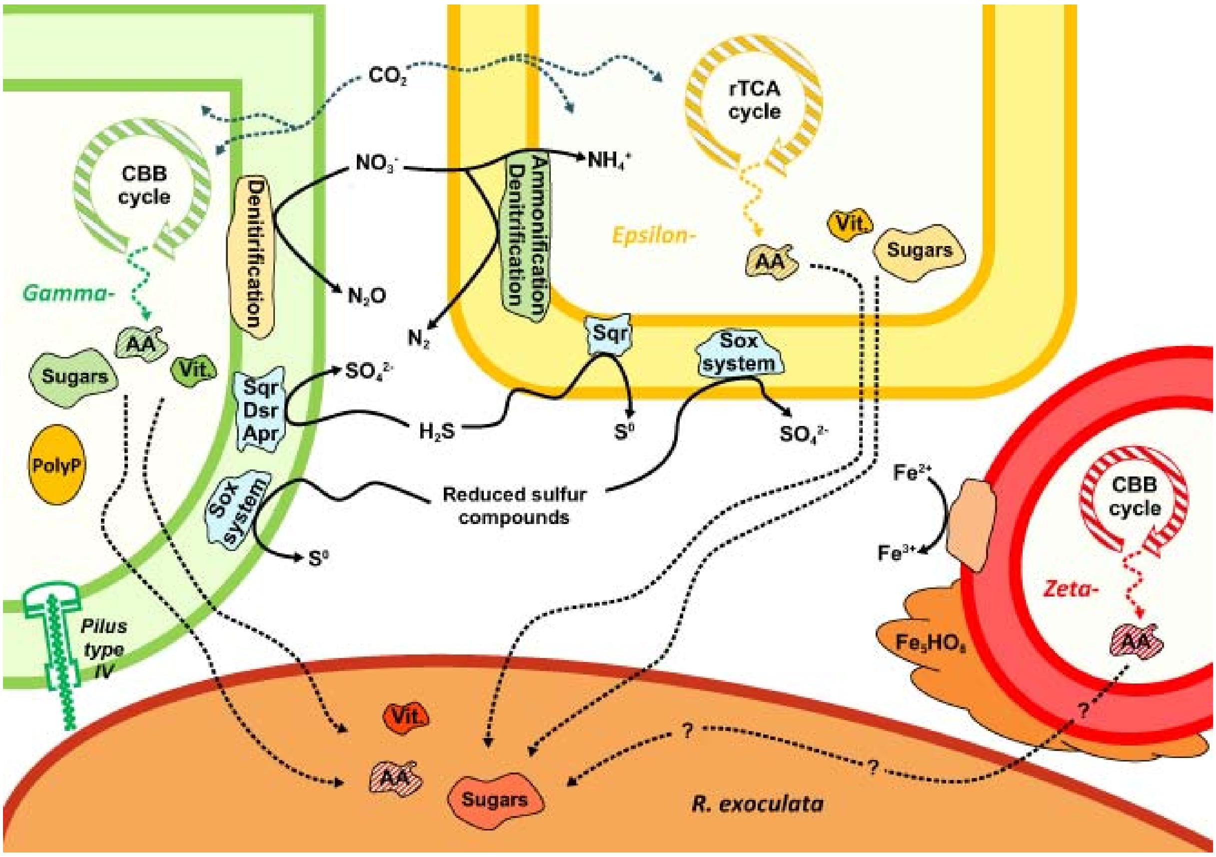


Figure 5: Scheme of the proposed functioning of the *R. exoculata* gill chamber symbiosis. AA: amino acids; Sqr: sulfide-quinone oxidoreductase; Vit.: vitamins; PolyP: polyphosphates. Dotted arrows indicate indirect synthesis (for amino acids), diffusion (for CO_2) or transfer (for amino acids, sugars and vitamins).



Supplementary information

Supplementary text: Description of metabolic pathways – carbon assimilation and nitrogen assimilation – metabolic response to environment – monitoring proteins and oxidative stress for epsilonproteobacterial epibionts, CRISPRs and transposases.

Supplementary figure S1: Phylogenetic tree showing the affiliation of partial 16S rRNA sequences from the *R. exoculata* gill chamber metagenome that fell into the epsilonproteobacterial epibiont group. The basis tree (only full-length 16S rRNA sequences, > 1,300 nt) was calculated with PhyML implemented in ARB, using a positional variability filter for Bacteria. After trimming the first 10 bases and quality filtering to Q28, the 16S rRNA reads were very short (most shorter than 250 bp). We therefore only included sequences that aligned to highly variable regions in the 16S rRNA (these were identified in the alignment by eye). These sequences, together with the 16S rRNA gene fragment sequences from the Epsilon taxobin were added to the basis tree using the ARB 'quick add' tool. Bootstrap values are from 100 replicates. The outgroup (not shown) was *Flavobacterium psychrolimnae* (AB455260, AJ585428). Sequences in purple are 16S tags from the metagenome, sequences in green are from the epsilonproteobacterial taxobin, and sequences in red are isolates. It was not possible to do this analysis with the gammaproteobacterial 16S reads, as previously identified strains of gammaproteobacterial epibionts are more closely related that those from the Epsilonproteobacteria, and such short reads (<250 bp) did not provide enough resolution for strain analysis in this group.

Supplementary table S1: General features of the metagenomic dataset from the symbiotic community of the gill chamber of *R. exoculata*.

Supplementary table S2: Annotation of genes relevant to the principal metabolic pathways and specific genes in epsilon-, gamma- and zetaproteobacterial epibionts.

Supplementary Text

Metabolism

Carbon fixation

Members of the *Epsilonproteobacteria* commonly use the reductive tricarboxylic acid (rTCA) cycle as the main autotrophic carbon dioxide (CO_2) fixation pathway (Peretó et al., 1999, Atomi, 2002, Berg, 2011) (Figure 3), which is confirmed by the thus far sequenced epsilonproteobacterial genes or genomes (Takai et al., 2005, Nakagawa et al., 2007, Sievert et al., 2008). As in *S. denitrificans* DSM1251 (Sievert et al., 2008), the epsilonproteobacterial *R. exoculata* gill epibionts could convert acetyl-CoA produced by the rTCA cycle using anapleurotic reactions via pyruvate oxidoreductase and phosphoenolpyruvate (PEP) carboxykinase to oxaloacetate which allows the funneling of TCA cycle intermediates into the central metabolism, or they could directly carboxylate acetyl-CoA (acetyl-CoA carboxylase) to malonyl-CoA for fatty acid biosynthesis. Two possible systems can accomplish this reaction: PEP synthase together with PEP carboxylase or pyruvate carboxylase, an enzyme of the citrate cycle. Pyruvate carboxylase was found in the epsilonproteobacterial taxonomic bin (taxobin). Genes encoding all other required enzymes for autotrophic growth within the rTCA cycle were identified in the epsilonproteobacterial taxobin.

Epsilonproteobacteria are known for their metabolic versatility using a variety of electron donors and acceptors, generally under microaerophilic conditions (Takai et al., 2003) consistent with using the rTCA cycle. The overwhelming abundance and phylogenetic diversity of rTCA genes in vent samples indicates that the principal CO_2 fixation pathway used by hydrothermal vent microbial epsilonproteobacterial communities is the rTCA cycle (Campbell and Cary, 2004).

Further carbon assimilation could be provided by the assimilation of acetate, since the key enzyme acetyl-CoA synthetase was present in our epsilonproteobacterial taxobin as well as in the reference genomes. The assimilation of acetate in epsilonproteobacterial epibionts was recently demonstrated, confirming the expression and the activity of this enzyme (Ponsard et al., 2013). All genes necessary for the gluconeogenesis were found (Figure 3), which are necessary to convert acetyl-CoA to sugars monomers.

The CBB (Calvin-Benson-Bassham) cycle, using the key enzyme ribulose 1,5-bisphosphate carboxylase/oxygenase (RubisCO), is the most significant carbon fixation pathway on Earth (Bassham et al., 1954, Tabita et al., 2008, Raven, 2009). Many symbionts at hydrothermal vents belong to the *Gammaproteobacteria* and use the CBB cycle for carbon fixation (Dubilier et al., 2008). Genes encoding for key CBB cycle enzymes were retrieved in our study, including ribulose 1,5-bisphosphate carboxylase/oxygenase form II (RubisCO II), phosphoribulokinase, phosphoglycerate kinase, fructose 1,6 bispohosphate aldolase and transketolase (Table S2). CBB cycle genes were also found in the zetaproteobacterial taxobin, corroborating the description of carbon assimilation in *M. ferrooxydans* PV-1 (Singer et al., 2011).

The use of the rTCA cycle additionally to CBB cycle by epibionts provides a possible explanation for the heavy stable carbon isotopic composition of *Rimicaris exoculata* (Pond et al., 2000, Hügler et al., 2011), similar to what has been found for gastropods (Suzuki et al., 2005) and the tubeworm *Riftia pachyptila* (Markert et al., 2007).

Nitrogen assimilation and nitrate reduction

Epsilonproteobacterial taxobin

Genes required for the complete reduction of nitrate to N_2 were identified (Figure 3, Table S2). Similar to other nitrate-reducing host-associated bacteria, a periplasmic nitrate reductase (Nap) and a cytoplasmic assimilatory nitrate reductase (Nas) were present but no cytoplasmic membrane-bound nitrate reductase (Nar) was found (Richardson and Watmough, 1999, Van Spanning et al., 2005). The *nap* genes (Kern and Simon, 2009) were divided in two clusters: *napAGHB* and *napFLD*. These clusters differ in their arrangement compared to the *nap* gene cluster (*napABHGFLD*) identified in *S. denitrificans* DSM1251 and in the pathogenic epsilonproteobacterium *Wolinella succinogenes*. However, all three lack the gene encoding the NapC subunit (Simon et al., 2003).

For nitrite reduction to nitric oxide, genes coding for a cytochrome *cd*$_1$-dependant nitrite reductase (*nirF* and *nirS*) were identified. Except in the epsilonproteobacterial taxobin, the Nir system was only present in the genomes of *Epsilonproteobacteria S. denitrificans* DSM1251, *Sulfurovum* sp. NBC37-1 and *Nitratiruptor* sp. SB155-2 (Kern and Simon, 2009). The presence of both cytoplasmic (Nas) and periplasmic (Nap) nitrate reductase in the epibionts may indicate the ability to adapt their metabolism to fluctuating nitrate concentrations in the *R. exoculata* gill chamber. The systems cNor and cNos were found, which complete the denitrification pathway. Similar to *Nitratiruptor* sp. SB155-2, a cytochome c: nitric oxide reductase (NorCB complex, cNor) was identified but no quinol-dependent nitric oxide reductase of the heme copper oxidase (gNor) or quinol:nitric oxide reductase (qNor) were found. A partial cytochrome c nitrous oxide reductase (cNos) with only the genes *nosZ* and *nosL* was also found. The epsilonproteobacterial *nos* gene cluster was similar to that of *W. succinogenes* and contained a unique assembly found exclusively in *Epsilonproteobacteria* (Kern and Simon, 2009).

Genes coding for a periplasmic cytochrome c nitrite reductase (Nrf) were identified, but they are absent in the deep-sea *Epsilonproteobacteria* reference genomes. The Nrf system (*nrfHA*) was retrieved in bacteria affiliated to the genera *Campylobacter* or *Helicobacter*, also host-associated species (Kern and Simon, 2009). This enzyme permits the dissimilatory nitrate reduction to ammonium (DNRA). The presence of this second system indicates that epsilonproteobacterial symbionts could use different pathways for nitrate reduction, one corresponding to the denitrification pathway and the second one for DNRA. Genes encoding for a NADPH-dependent ammonia-forming siroheme nitrite reductase (ferredoxin) and a cytochrome c552 complete the nitrate ammonification pathway. The presence of these elements and *nrfHA* indicated that epsilonproteobacterial epibionts could produce ammonia and would therefore not depend on ammonia uptake from the environment. The epsilonproteobacterial taxobin also contained genes for the regulation of nitrogen assimilation (Leigh and Dodsworth, 2007).

Gammaproteobacterial taxobin

Nitrate/nitrite transporters, a nitrate/nitrite sensor protein, a periplasmic nitrate reductase (Nap), an incomplete cytochrome-containing nitrite reductase (Nir) and a cytochome-containing nitric oxide reductase (NorCB complex, cNor) were identified (Figure 4, Table S2). The *nap* genes were retrieved in one cluster: *napDAGHB*. Nap is activated in the presence of nitrate and nitrite by the sensor kinase NarX (2 hits) (Table S2). No nitrous oxide reductase for the reduction of N_2O to N_2 was found. Genes for the assimilation of nitrate and nitrite are found in the *A. thiooxidans* DSM 17138 genome, grouped in one genomic region (Levican et al., 2008). Nevertheless, the nitrate reductase in *A. thiooxidans* DSM 17138 is a cytoplasmic enzyme (Nas) and not periplasmic as in the gammaproteobacterial taxobin. The presence of genes encoding periplasmic Nap and cytoplasmic Nir could indicate the presence of both systems (Nas/Nir and Nap/Nrf) as in epsilonproteobacterial epibionts, albeit Nas and Nrf genes were not found.

Because nitrate is the dominant nitrogen form at vents (Johnson et al., 1988), it is not surprising that both main *R. exoculata* epibionts can metabolize it through periplasmic and/or cytoplasmic systems. Nitrate can be used by bacteria as an alternative electron acceptor as well as a potential nitrogen source. The epibionts may also assimilate ammonia via recycling of the host's amino acid waste.

The ammonia assimilation enzymes glutamine synthetase (GS) and glutamate dehydrogenase (GDH) were detected in epsilon- and zetaproteobacterial epibionts (Table S2), indicating that ammonia resulting from nitrate reduction or from environmental uptake could be incorporated into amino acids. In other marine chemoautotrophic symbioses with an invertebrate, GS and GDH are both present in host (Lee et al., 1999).

Response to environment

Monitoring proteins in epsilonproteobacterial epibionts

Microorganisms have various mechanisms to sense and respond to their environment, such as two-component signal transcriptional regulators, transduction factors, PAS/GGDEF systems, sensory histidine kinases or methyl-accepting chemotaxis proteins. The number of such systems in the epsilonproteobacterial taxobin (Table 2) was similar to what has been observed in the free-living chemolithoautotrophs *S. denitrificans* DSM1251 and *Nitrosococcus oceani* ATCC19707 (Klotz et al., 2006, Sievert et al., 2008). As in *S. denitrificans* DSM1251, the large number of histidine kinases and transcriptional regulators might reflect the ability to switch between the use of different electron donors (e.g. H_2, reduced sulfur compounds) and electron acceptors (e.g. O_2, NO_x), permitting adaptation to the spatially and temporally heterogeneous environment at hydrothermal vents.

Cell-cell communication

The presence of GGDEF/EAL type response regulators constitutes an indicator of cell communication and environmental response. Based on a PFAM model research, we identified 27 sequences with GGDEF and EAL domains in the metagenome. Proteins with GGDEF and EAL domains have been implicated particularly in formation of biofilms, regulation of gene expression, regulation of enzymatic activities, or in protein-protein interactions in response to environmental variations (Ryjenkov *et al.*, 2005; Galperin, 2006), Cell-cell communication between bacteria termed quorum sensing (QS) plays a crucial role in some symbioses. QS mediated by extracellular autoinducer molecules allows bacteria to 'communicate' (Bassler, 2002; Hughes and Sperandio, 2008). In the metagenome, ten sequences were identified with possible links to QS: two *luxS* sequences, three *luxR* sequences, three genes encoding homoserine lactone efflux proteins, a homoserine dehydrogenase and an O-acetylhomoserine sulfhydrylase sequence. These sequences affiliated with the gammaproteobacterial and epsilonproteobacterial epibionts but also with *Betaproteobacteria* (Table S2). QS plays a major role in gene regulation in many environments, and previous data suggest that many pathogens rely on QS to promote infection (Federle and Bassler, 2003). In a symbiotic community, QS could help to sustain the biofilm on the host via secreted autoinducers and acyl-homoserine lactone (acyl-HSL) (Filloux and Vallet, 2003). In addition, sequences for antimicrobial peptide transport systems were identified in the gamma- and epsilonproteobacterial taxobins (Table S2). Antimicrobial peptides could also play a role in host-symbiont communication. For instance, they could prevent other potentially harmful bacteria from colonizing the shrimp. Similar processes have been described in other symbiosis (Wahl *et al.*, 2012).

Oxidative stress in epsilonproteobacterial epibionts

Known members of the subgroup F, including the epsilonproteobacterial epibionts, and of the subgroup G of the *Epsilonproteobacteria* (Corre et al., 2001) are sensitive to oxygen (Inagaki et al., 2004, Nakagawa et al., 2005b). Oxygen concentrations vary around hydrothermal chimneys and decrease from the surrounding seawater to the anoxic hydrothermal vent fluids. *R. exoculata* specimens are continuously swimming in the area where vent fluid and seawater mix and oxygen is present. Hence, their gill-colonizing microbes are subjected to oxidative stress. The epsilonproteobacterial taxobin revealed some of their coping mechanisms (Table S2). A nickel superoxide dismutase was present that was phylogenetically distant to other superoxide dismutases and different from the iron/manganese superoxide dismutase of *S. denitrificans* (Sievert et al., 2008). Likewise, a cytochrome c peroxidase, a superoxide dismutase and a peroxide stress regulator were identified, but no catalase. A nitric oxide dioxygenase was found, which may convert nitric oxide to nitrate. Similar to S*. denitrificans*, thioredoxins and thioredoxin reductases were abundant, but no glutathione synthetases nor gamma-glutamate-cysteine ligases. These enzymes have been speculated to funnel electrons to oxidative stress proteins such as glutathione (Sievert et al., 2008). Genes of oxidative DNA damage repair were found as well, such as, endonucleases III and IV and *mutS*.

The rTCA cycle has two oxygen-sensitive enzymes, and hence is restricted to anaerobic or microaerophilic bacteria (Hügler and Sievert, 2011). Nonetheless, the closest epsilonproteobacterial relative, *Sulfurovum* sp., can tolerate relatively high amounts of oxygen (Nakagawa et al., 2005a), *S. gotlandica* GC-1 can also deal with lateral intrusions of oxygenated seawater into its otherwise anoxic habitat (Grote et al., 2012). Epsilonproteobacterial symbionts thus could be microaerophilic, which is clearly in accordance with the host-associated lifestyle.

CRISPRs

To resist invading genetic elements such as viruses and plasmids, microorganisms have a defense system that is characterized by clustered regularly interspaced short palindromic repeats (CRISPRs) (Barrangou et al., 2007, Brouns et al., 2008, Horvath and Barrangou, 2010, Makarova et al., 2011). CRISPRs in the *R. exoculata* gill chamber metagenome were identified with CRISPR-finder (Grissa et al., 2007). 118 CRISPRs were identified in the complete metagenome. Three CRISPRs were clearly identified belonging to the *Epsilonproteobacteria* and one to *the Gammaproteobacteria* (Table S2). 14 epsilonproteobacterial and 16 gammaproteobacterial sequences coded for CRISPR-associated (CAS) proteins (Haft et al., 2005). An effective protection against viral infection might appear essential to survive in non-mobile communities, avoiding huge impact on the bacterial population size (Sorek et al., 2008, Tyson and Banfield, 2008).

Transposases

Genes coding for transposases were retrieved in the epsilonproteobacterial (26 hits) and the gammaproteobacterial taxobins (51 hits), with the detection of multiple copies of the insertion sequences (IS) related to the IS4 family (Rezsöhazy et al., 1993), which play an important role in the transposition process. A gene coding for a transposase was affiliated to eukaryotes, possibly *R. exoculata*. Transposases could enable DNA exchange between differing species when the type is shared among community members, which is indicated by our data.

References

Atomi H. (2002). Microbial enzymes involved in carbon dioxide fixation. J Biosciences & Bioengineering 94**:** 497-505.

Barrangou R., Fremaux C., Deveau H.l.n., Richards M., Boyaval P., Moineau S. et al. (2007). CRISPR Provides Acquired Resistance Against Viruses in Prokaryotes. Science 315**:** 1709-1712.

Bassham J.A., Benson A.A., Kay L.D., Harris A.Z., Wilson A.T. and Calvin M. (1954). The Path of Carbon in Photosynthesis. XXI. The Cyclic Regeneration of Carbon Dioxide Acceptor1. J Am Chem Soc 76**:** 1760-1770.

Berg I.A. (2011). Ecological Aspects of the Distribution of Different Autotrophic CO_2 Fixation Pathways. App Environ Microbiol 77**:** 1925-1936.

Brouns S.J.J., Jore M.M., Lundgren M., Westra E.R., Slijkhuis R.J.H., Snijders A.P.L. et al. (2008). Small CRISPR RNAs Guide Antiviral Defense in Prokaryotes. Science 321**:** 960-964.

Campbell B.J. and Cary S.C. (2004). Abundance of Reverse Tricarboxylic Acid Cycle Genes in Free-Living Microorganisms at Deep-Sea Hydrothermal Vents. Appl Environ Microbiol 70**:** 6282-6289.

Corre E., Reysenbach A.L. and Prieur D. (2001). *Epsilon*-proteobacterial diversity from a deep-sea hydrothermal vent on the Mid-Atlantic Ridge. FEMS Microbiol Lett 205**:** 329-335.

Dubilier N., Bergin C. and Lott C. (2008). Symbiotic diversity in marine animals: the art of harnessing chemosynthesis. Nat Rev Micro 6**:** 725-740.

Grissa I., Vergnaud G. and Pourcel C. (2007). CRISPRFinder: a web tool to identify clustered regularly interspaced short palindromic repeats. Nucleic Acids Res 35**:** W52-W57.

Grote J., Schott T., Bruckner C.G., Glockner F.O., Jost G., Teeling H. et al. (2012). Genome and physiology of a model Epsilonproteobacterium responsible for sulfide detoxification in marine oxygen depletion zones. Proc Natl Acad Sci USA 109**:** 506-510.

Haft D.H., Selengut J., Mongodin E.F. and Nelson K.E. (2005). A Guild of 45 CRISPR-Associated (Cas) Protein Families and Multiple CRISPR/Cas Subtypes Exist in Prokaryotic Genomes. PLoS Comput Biol 1**:** e60.

Horvath P. and Barrangou R. (2010). CRISPR/Cas, the Immune System of Bacteria and Archaea. Science 327**:** 167-170.

Hügler M., Petersen J.M., Dubilier N., Imhoff J.F. and Sievert S.M. (2011). Pathways of Carbon and Energy Metabolism of the Epibiotic Community Associated with the Deep-Sea Hydrothermal Vent Shrimp *Rimicaris exoculata* PLoS ONE 6**:** e16018.

Hügler M. and Sievert S.M. (2011). Beyond the Calvin Cycle: Autotrophic Carbon Fixation in the Ocean. An Rev Mar Sci 3**:** 261-289.

Inagaki F., Takai K., Nealson K.H. and Horikoshi K. (2004). *Sulfurovum lithotrophicum* gen. nov., sp. nov., a novel sulfur-oxidizing chemolithoautotroph within the ε-*Proteobacteria* isolated from Okinawa Trough hydrothermal sediments. Int J Syst Evol Microbiol 54**:** 1477-1482.

Johnson K.S., Childress J.J. and Beehler C.L. (1988). Short-term temperature variability in the Rose Garden hydrothermal vent field: an unstable deep-sea environment. Deep Sea Research Part A Oceanograph Res Papers 35**:** 1711-1721.

Kern M. and Simon J. (2009). Electron transport chains and bioenergetics of respiratory nitrogen metabolism in *Wolinella succinogenes* and other *Epsilonproteobacteria*. Biochimica et Biophysica Acta (BBA) - Bioenergetics 1787**:** 646-656.

Klotz M.G., Arp D.J., Chain P.S.G., El-Sheikh A.F., Hauser L.J., Hommes N.G. et al. (2006). Complete Genome Sequence of the Marine, Chemolithoautotrophic, Ammonia-Oxidizing Bacterium *Nitrosococcus oceani* ATCC 19707. Appl Environ Microbiol 72**:** 6299-6315.

Lee R.W., Robinson J.J. and Cavanaugh C.M. (1999). Pathways of inorganic nitrogen assimilation in chemoautotrophic bacteria-marine invertebrate symbioses: expression of host and symbiont glutamine synthetase. J Experim Biol 202**:** 289-300.

Leigh J.A. and Dodsworth J.A. (2007). Nitrogen Regulation in *Bacteria* and *Archaea*. An Revi Microbiol 61**:** 349-377.

Levican G., Ugalde J., Ehrenfeld N., Maass A. and Parada P. (2008). Comparative genomic analysis of carbon and nitrogen assimilation mechanisms in three indigenous bioleaching bacteria: predictions and validations. Bmc Genomics 9**:** 581.

Makarova K.S., Haft D.H., Barrangou R., Brouns S.J.J., Charpentier E., Horvath P. et al. (2011). Evolution and classification of the CRISPR/Cas systems. Nat Rev Micro 9**:** 467-477.

Markert S., Arndt C., Felbeck H., Becher D., Sievert S.M., Hügler M. et al. (2007). Physiological Proteomics of the Uncultured Endosymbiont of *Riftia pachyptila*. Science 315**:** 247-250.

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

Nakagawa S., Takai K., Inagaki F., Horikoshi K. and Sako Y. (2005b). *Nitratiruptor tergarcus* gen. nov., sp. nov. and *Nitratifractor salsuginis* gen. nov., sp. nov., nitrate-reducing chemolithoautotrophs of the ε-*Proteobacteria* isolated from a deep-sea hydrothermal system in the Mid-Okinawa Trough. Int J Syst Evol Microbiol 55**:** 925-933.

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

Peretó J.G., Velasco A.M., Becerra A. and Lazcano A. (1999). Comparative biochemistry of CO_2 fixation and the evolution of autotrophy. Int Microbiol 2(1)**:** 3-10.

Pond D.W., Gebruk A., Southward E.C., Southward A.J., Fallick A.E., Bell M.V. et al. (2000). Unusual fatty acid composition of storage lipids in the bresilioid shrimp *Rimicaris exoculata* couples the photic zone with MAR hydrothermal vent sites. Mar Ecol Progr Series 198**:** 171-179.

Ponsard J., Cambon-Bonavita M.-A., Zbinden M., Lepoint G., Joassin A., Corbari L. et al. (2013). Inorganic carbon fixation by chemosynthetic ectosymbionts and nutritional transfers to the hydrothermal vent host-shrimp *Rimicaris exoculata*. ISME J 7**:** 96–109.

Raven J.A. (2009). Contributions of anoxygenic and oxygenic phototrophy and chemolithotrophy to carbon and oxygen fluxes in aquatic environments. Aqua Micro Ecol 56**:** 177-192.

Rezsöhazy R., Hallet B., Delcour J. and Mahillon J. (1993). The IS4 family of insertion sequences: evidence for a conserved transposase motif. Mol Microbiol 9**:** 1283-1295.

Richardson D.J. and Watmough N.J. (1999). Inorganic nitrogen metabolism in bacteria. Cur Opin Chem Biol 3**:** 207-219.

Sievert S.M., Scott K.M., Klotz M.G., Chain P.S.G., Hauser L.J., Hemp J. et al. (2008). Genome of the Epsilonproteobacterial Chemolithoautotroph *Sulfurimonas denitrificans*. Appl Environ Microbiol 74**:** 1145-1156.

Simon J., Sänger M., Schuster S.C. and Gross R. (2003). Electron transport to periplasmic nitrate reductase (NapA) of *Wolinella succinogenes* is independent of a NapC protein. Mol Microbiol 49**:** 69-79.

Singer E., Emerson D., Webb E.A., Barco R.A., Kuenen J.G., Nelson W.C. et al. (2011). *Mariprofundus ferrooxydans* PV-1 the first genome of a marine Fe(II) oxidizing *Zetaproteobacterium*. Plos One 6**:** e25386.

Sorek R., Kunin V. and Hugenholtz P. (2008). CRISPR - a widespread system that provides acquired resistance against phages in *bacteria* and *archaea*. Nat Rev Micro 6**:** 181-186.

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.

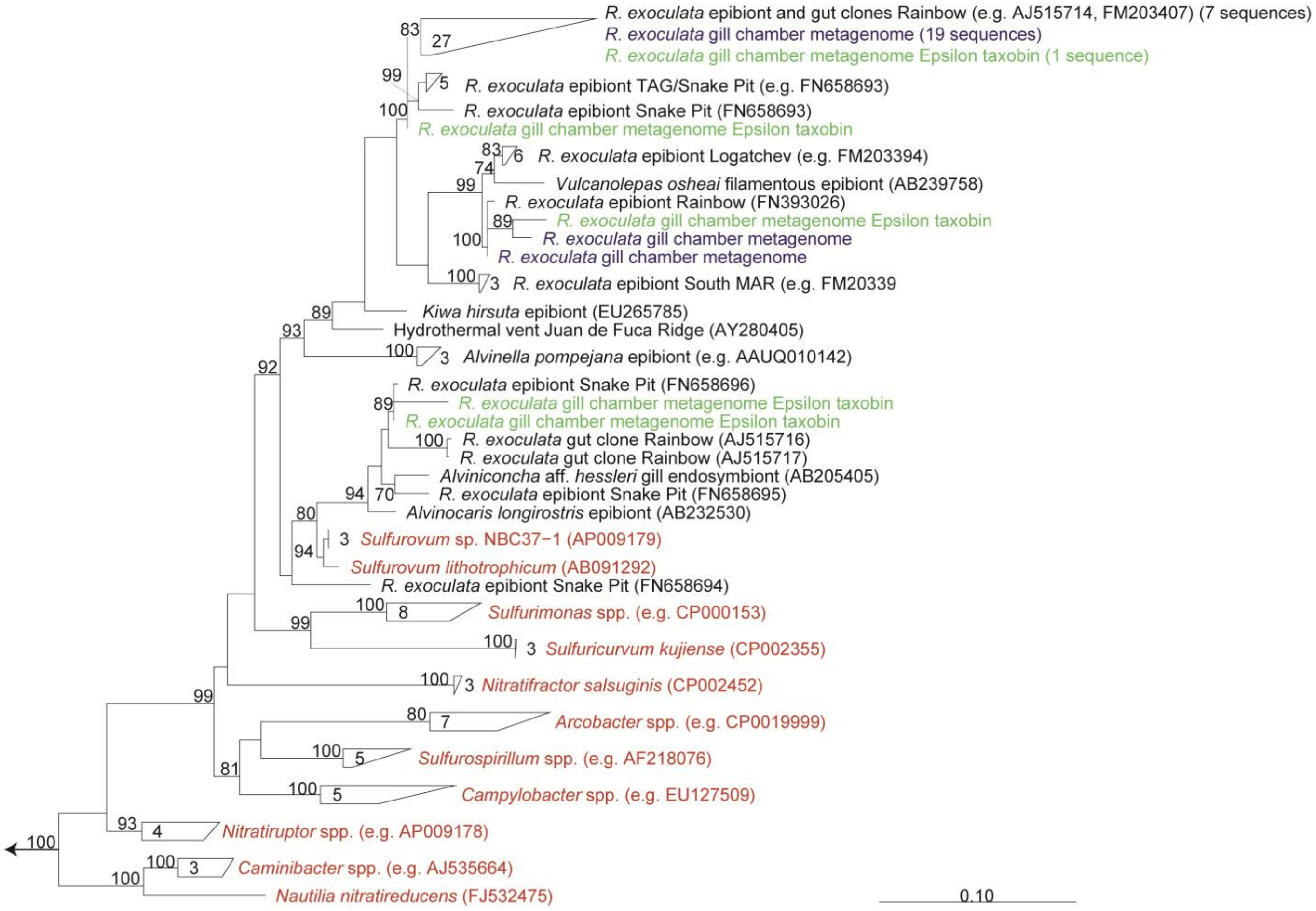
Tabita F.R., Satagopan S., Hanson T.E., Kreel N.E. and Scott S.S. (2008). Distinct form I, II, III, and IV Rubisco proteins from the three kingdoms of life provide clues about Rubisco evolution and structure/function relationships. J Experim Botan 59**:** 1515-1524.

Takai K., Inagaki F., Nakagawa S., Hirayama H., Nunoura T., Sako Y. et al. (2003). Isolation and phylogenetic diversity of members of previously uncultivated *Epsilon-Proteobacteria* in deep-sea hydrothermal fields. FEMS Microbiology Letters 218**:** 167-174.

Takai K., Campbell B.J., Cary S.C., 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.

Tyson G.W. and Banfield J.F. (2008). Rapidly evolving CRISPRs implicated in acquired resistance of microorganisms to viruses. Environ Microbiol 10**:** 200-207.

Van Spanning R., Delgado M. and Richardson D. (2005). The Nitrogen Cycle: Denitrification and its Relationship to N_2 Fixation. In: Werner D, Newton W (eds). Nitrogen Fixation in Agriculture, Forestry, Ecology, and the Environment. Springer Netherlands. pp 277-342.



Supplementary figure S1: Phylogenetic tree showing the affiliation of partial 16S rRNA sequences from the *R. exoculata* gill chamber metagenome that fell into the epsilonproteobacterial epibiont group. The basis tree (only full-length 16S rRNA sequences, > 1,300 nt) was calculated with PhyML implemented in ARB, using a positional variability filter for Bacteria. After trimming the first 10 bases and quality filtering to Q28, the 16S rRNA reads were very short (most shorter than 250 bp). We therefore only included sequences that aligned to highly variable regions in the 16S rRNA (these were identified in the alignment by eye). These sequences, together with the 16S rRNA gene fragment sequences from the Epsilon taxobin were added to the basis tree using the ARB 'quick add' tool. Bootstrap values are from 100 replicates. The outgroup (not shown) was *Flavobacterium psychrolimnae* (AB455260, AJ585428). Sequences in purple are 16S tags from the metagenome, sequences in green are from the epsilonproteobacterial taxobin, and sequences in red are isolates. It was not possible to do this analysis with the gammaproteobacterial 16S reads, as previously identified strains of gammaproteobacterial epibionts are more closely related that those from the Epsilonproteobacteria, and such short reads (<250 bp) did not provide enough resolution for strain analysis in this group.

Supplementary table S1: General features of the metagenomic dataset from the symbiotic community of the gill chamber of *R.exoculata*.

Read summary	
Before filtering:	
Total reads	1,011,151
Total bases	317,729,916
Assembled	32.7%
Singletons	39.8%
Contig summary	
All contigs:	
Total contigs	24,529
Assembly size (bp)	15,291,786
Mean size (bp)	623.42
Min size (bp)	100
Max size (bp)	31,453
Average GC content	37.33%
Total ORFs	65,699
Hypothetical proteins	14,478
Conserved hypothetical proteins	1,673
Large contigs (≥500 bp)	
Total	8,625
Assembly size (bp)	11,121,975
Mean size (bp)	1,289
Max size (bp)	31,453

Supplementary table S2: Annotation of genes relevant to the principal metabolic pathways and specific genes in epsilon-, gamma- and zetaproteobacterial epibionts.

1. Inorganic carbon fixation

Gene name (contig)	Gene product name	Organism (class)	Genus (best hit)	Strain (best hit)	Expected Value (E) and Identities (I)
14418_0	ATP-citrate lyase subunit 1 (rTCA)	*Bacteria – Proteobacteria – Epsilonproteobacteria*	*Sulfurimonas*	*Sulfurimonas denitrificans* DSM 1251	E=7e-28; I=85%
14418_1	ATP-citrate lyase subunit 1 (rTCA)	*Bacteria – Proteobacteria – Epsilonproteobacteria*	*Sulfurimonas*	*Sulfurimonas sp.* Go25-1	E=7e-43; I=88%
14418_2	ATP-citrate lyase subunit 1 (rTCA)	*Bacteria – Proteobacteria – Epsilonproteobacteria*	*Sulfurimonas*	*Sulfurimonas denitrificans* DSM 1251	E=9e-41; I=81%
10781_0	ATP-citrate lyase subunit 2 (rTCA)	*Bacteria – Proteobacteria – Epsilonproteobacteria*	Unclassified	*Alvinella pompejana* epibiont 6C5	E=1e-32; I=96%
16894_0	ATP-citrate lyase subunit 2 (rTCA)	*Bacteria – Proteobacteria – Epsilonproteobacteria*	Unclassified	*Alvinella pompejana* epibiont 6C5	E=7e-34; I=80%
28104_0	ATP-citrate lyase subunit 2 (rTCA)	*Bacteria – Proteobacteria – Epsilonproteobacteria*	*Sulfurimonas*	*Sulfurimonas denitrificans* DSM 1251	E=1e-24; I=79%
00259_3	Malate dehydrogenase (rTCA)	*Bacteria – Proteobacteria – Epsilonproteobacteria*	*Sulfurovum*	*Sulfurovum sp*. NBC37-1	E=1e-126; I=73%
00259_10	Malate dehydrogenase (rTCA)	*Bacteria – Proteobacteria – Epsilonproteobacteria*	*Sulfurovum*	*Sulfurovum sp*. NBC37-1	E=1e-87; I=93%
00816_0	Malate dehydrogenase (rTCA)	*Bacteria – Proteobacteria – Epsilonproteobacteria*	*Sulfurovum*	*Sulfurovum sp*. NBC37-1	E=1e-61; I=87%
02897_0	Malate dehydrogenase (rTCA)	*Bacteria – Proteobacteria – Epsilonproteobacteria*	*Sulfurovum*	*Sulfurovum sp*. NBC37-1	E=2e-100; I=92%
14136_0	Malate dehydrogenase (rTCA)	*Bacteria – Proteobacteria – Epsilonproteobacteria*	*Sulfurovum*	*Sulfurovum sp*. NBC37-1	E=1e-84; I=76%
02815_0	Fumarate hydratase (rTCA)	*Bacteria – Proteobacteria – Epsilonproteobacteria*	*Sulfurovum*	*Sulfurovum sp*. NBC37-1	E=3e-61; I=82%
01121_0	Fumarate reductase flavoprotein subunit (rTCA)	*Bacteria – Proteobacteria – Epsilonproteobacteria*	*Sulfurovum*	*Sulfurovum sp*. NBC37-1	E=5e-78; I=88%
02440_0	Fumarate reductase flavoprotein subunit (rTCA)	*Bacteria – Proteobacteria – Epsilonproteobacteria*	*Sulfurovum*	*Sulfurovum sp*. NBC37-1	E=8e-65; I=80%
03275_0	Fumarate reductase flavoprotein subunit (rTCA)	*Bacteria – Proteobacteria – Epsilonproteobacteria*	*Sulfurovum*	*Sulfurovum sp*. NBC37-1	E=1e-76; I=75%
18054_3	Fumarate reductase flavoprotein subunit (rTCA)	*Bacteria – Proteobacteria – Epsilonproteobacteria*	*Sulfurovum*	*Sulfurovum sp*. NBC37-1	E=0.0; I=87%
20886_5	Fumarate reductase flavoprotein subunit (rTCA)	*Bacteria – Proteobacteria – Epsilonproteobacteria*	*Sulfurovum*	*Sulfurovum sp*. NBC37-1	E=0.0; I=61%
10937_1	Fumarate reductase Fe-S subunit (rTCA)	*Bacteria – Proteobacteria – Epsilonproteobacteria*	*Sulfurovum*	*Sulfurovum sp*. NBC37-1	E=2e-60; I=80%
10938_0	Fumarate reductase Fe-S subunit (rTCA)	*Bacteria – Proteobacteria – Epsilonproteobacteria*	*Sulfurovum*	*Sulfurovum sp*. NBC37-1	E=6e-78; I=84%
17378_0	Fumarate reductase Fe-S subunit (rTCA)	*Bacteria – Proteobacteria – Epsilonproteobacteria*	*Sulfurovum*	*Sulfurovum sp*. NBC37-1	E=7e-65; I=77%
18054_2	Fumarate reductase Fe-S subunit (rTCA)	*Bacteria – Proteobacteria – Epsilonproteobacteria*	*Sulfurovum*	*Sulfurovum sp*. NBC37-1	E=2e-151; I=88%
20886_6	Fumarate reductase Fe-S subunit (rTCA)	*Bacteria – Proteobacteria – Epsilonproteobacteria*	*Sulfurovum*	*Sulfurovum sp*. NBC37-1	E=3e-80; I=62%
02695_0	Succinyl-CoA synthetase subunit A (rTCA)	*Bacteria – Proteobacteria – Epsilonproteobacteria*	*Sulfurovum*	*Sulfurovum sp*. NBC37-1	E=9e-88; I=89%
14100_1	Succinyl-CoA synthetase subunit A (rTCA)	*Bacteria – Proteobacteria – Epsilonproteobacteria*	*Sulfurovum*	*Sulfurovum sp*. NBC37-1	E=8e-131; I=86%
14100_0	Succinyl-CoA synthetase subunit B (rTCA)	*Bacteria – Proteobacteria – Epsilonproteobacteria*	*Sulfurovum*	*Sulfurovum sp*. NBC37-1	E=1e-69; I=89%
14101_0	Succinyl-CoA synthetase subunit B (rTCA)	*Bacteria – Proteobacteria – Epsilonproteobacteria*	*Sulfurimonas*	*Sulfurimonas denitrificans* DSM 1251	E=5e-111; I=85%
13375_4	2-Oxoglutarate oxidoreductase subunit A (rTCA)	*Bacteria – Proteobacteria – Epsilonproteobacteria*	*Sulfurovum*	*Sulfurovum sp*. NBC37-1	E=8e-89; I=84%
13375_2	2-Oxoglutarate oxidoreductase subunit B (rTCA)	*Bacteria – Proteobacteria – Epsilonproteobacteria*	*Sulfurovum*	*Sulfurovum sp*. NBC37-1	E=1e-133; I=75%
13375_1	2-Oxoglutarate oxidoreductase subunit C (rTCA)	*Bacteria – Proteobacteria – Epsilonproteobacteria*	*Sulfurovum*	*Sulfurovum sp*. NBC37-1	E=8e-78; I=77%
05861_0	2-Oxoglutarate oxidoreductase subunit D (rTCA)	*Bacteria – Proteobacteria – Epsilonproteobacteria*	*Sulfurovum*	*Sulfurovum sp*. NBC37-1	E=7e-47; I=80%
09678_0	Isocitrate dehydrogenase (rTCA)	*Bacteria – Proteobacteria – Epsilonproteobacteria*	*Sulfurimonas*	*Sulfurimonas denitrificans* DSM 1251	E=2e-19; I=72%
23083_0	Isocitrate dehydrogenase (rTCA)	*Bacteria – Proteobacteria – Epsilonproteobacteria*	*Sulfurimonas*	*Sulfurimonas denitrificans* DSM 1251	E=5e-19; I=70%
00183_0	Isocitrate dehydrogenase (rTCA)	*Bacteria – Proteobacteria – Epsilonproteobacteria*	*Sulfurovum*	*Sulfurovum sp*. NBC37-1	E=5e-49; I=77%

00202_1	Isocitrate dehydrogenase (rTCA)	*Bacteria – Proteobacteria – Epsilonproteobacteria*	*Sulfurovum*	*Sulfurovum sp*. NBC37-1	E=8e-18; I=85%
00203_0	Isocitrate dehydrogenase (rTCA)	*Bacteria – Proteobacteria – Epsilonproteobacteria*	*Sulfurovum*	*Sulfurovum sp*. NBC37-1	E=0.0; I=78%
00314_0	Aconitate hydratase (rTCA)	*Bacteria – Proteobacteria – Epsilonproteobacteria*	*Sulfurovum*	*Sulfurovum sp*. NBC37-1	E=0.0; I=87%
12146_0	Aconitate hydratase (rTCA)	*Bacteria – Proteobacteria – Epsilonproteobacteria*	*Sulfurovum*	*Sulfurovum sp*. NBC37-1	E=7e-135; I=74%
10901_1	Acetyl-CoA synthetase (rTCA)	*Bacteria – Proteobacteria – Epsilonproteobacteria*	*Sulfurovum*	*Sulfurovum sp*. NBC37-1	E=0.0; I=75%
11357_0	Acetyl-CoA synthetase (rTCA)	*Bacteria – Proteobacteria – Epsilonproteobacteria*	*Sulfurovum*	*Sulfurovum sp*. NBC37-1	E=1e-67; I=83%
11337_12	Pyruvate: ferredoxin oxidoreductase subunit A (rTCA)	*Bacteria – Proteobacteria – Epsilonproteobacteria*	*Sulfurovum*	*Sulfurovum sp*. NBC37-1	E=0.0; I=84%
16366_0	Pyruvate: ferredoxin oxidoreductase subunit A (rTCA)	*Bacteria – Proteobacteria – Epsilonproteobacteria*	*Sulfurovum*	*Sulfurovum sp*. NBC37-1	E=1e-83; I=87%
11337_13	Pyruvate: ferredoxin oxidoreductase subunit B (rTCA)	*Bacteria – Proteobacteria – Epsilonproteobacteria*	*Sulfurovum*	*Sulfurovum sp*. NBC37-1	E=6e-173; I=89%
06451_2	Pyruvate: ferredoxin oxidoreductase subunit C (rTCA)	*Bacteria – Proteobacteria – Epsilonproteobacteria*	*Sulfurovum*	*Sulfurovum sp*. NBC37-1	E=2e-94; I=87%
11337_10	Pyruvate: ferredoxin oxidoreductase subunit C (rTCA)	*Bacteria – Proteobacteria – Epsilonproteobacteria*	*Sulfurovum*	*Sulfurovum sp*. NBC37-1	E=4e-95; I=88%
14968_0	Pyruvate: ferredoxin oxidoreductase subunit C (rTCA)	*Bacteria – Proteobacteria – Epsilonproteobacteria*	*Sulfurovum*	*Sulfurovum sp*. NBC37-1	E=1e-41; I=87%
06451_3	Pyruvate: ferredoxin oxidoreductase subunit D (rTCA)	*Bacteria – Proteobacteria – Epsilonproteobacteria*	*Sulfurovum*	*Sulfurovum sp*. NBC37-1	E=9e-12; I=60%
11337_11	Pyruvate: ferredoxin oxidoreductase subunit D (rTCA)	*Bacteria – Proteobacteria – Epsilonproteobacteria*	*Sulfurovum*	*Sulfurovum sp*. NBC37-1	E=1e-61; I=78%
02459_1	Ribulose-1,5-bisphosphate carboxylase/oxygenase form II (CBB)	*Bacteria – Proteobacteria – Gammaproteobacteria*	*Acidithiobacillus*	*Acidithiobacillus ferrooxidans* ATCC 53993	E=0.0; I=83%
00465_0	Phosphoribulokinase (CBB)	*Bacteria – Proteobacteria – Gammaproteobacteria*	*Halothiobacillus*	*Halothiobacillus neapolitanus* c2	E=2e-125; I=74%
09760_3	Phosphoglycerate kinase (CBB)	*Bacteria – Proteobacteria – Gammaproteobacteria*	*Thioalkalivibrio*	*Thioalkalivibrio sulfidophilus* HL-EbGr7	E=3e-119; I=75%
15041_0	Phosphoglycerate kinase (CBB)	*Bacteria – Proteobacteria – Gammaproteobacteria*	*Thiomicrospira*	*Thiomicrospira crunogena* XCL-2	E=6e-93; I=83%
09760_5	Fructose-1,6-bisphosphate aldolase (CBB)	*Bacteria – Proteobacteria – Gammaproteobacteria*	Unclassified	uncultured SUP05 cluster *bacterium*	E=0.0; I=89%
01783_0	Transketolase (CBB)	*Bacteria – Proteobacteria – Gammaproteobacteria*	*Allochromatium*	*Allochromatium vinosum* DSM 180	E=0.0; I=68%
02550_0/1/2	Carbonic anhydrase	*Bacteria – Proteobacteria – Gammaproteobacteria*	*Idiomarina*	*Idiomarina loihiensis* L2TR	E=6e-13; I=53%
13062_0	Ribulose-1,5-bisphosphate carboxylase/oxygenase form II (CBB)	*Bacteria – Proteobacteria – Zetaproteobacteria*	*Mariprofundus*	*Mariprofundus ferrooxydans* PV-1	E=0.0; I=90%
03565_1	Phosphoribulokinase (CBB)	*Bacteria – Proteobacteria – Zetaproteobacteria*	*Mariprofundus*	*Mariprofundus ferrooxydans* PV-1	E=1e-35; I=77%
04371_1/2	Fructose-1,6-bisphosphate aldolase (CBB)	*Bacteria – Proteobacteria – Zetaproteobacteria*	*Mariprofundus*	*Mariprofundus ferrooxydans* PV-1	E=5e-16; I=84%
02437_0	Fructose-1,6-bisphosphate aldolase (CBB)	*Bacteria – Proteobacteria – Zetaproteobacteria*	*Mariprofundus*	*Mariprofundus ferrooxydans* PV-1	E=6e-107; I=89%
14639_0	Transketolase (CBB)	*Bacteria – Proteobacteria – Zetaproteobacteria*	*Mariprofundus*	*Mariprofundus ferrooxydans* PV-1	E=6e-107; I=86%
13562_0	Succinate dehydrogenase or fumarate reductase (TCA)	*Bacteria – Proteobacteria – Zetaproteobacteria*	*Mariprofundus*	*Mariprofundus ferrooxydans* PV-1	E=2e-18; I=83%
00721_1	Acetyl-CoA carboxylase (TCA)	*Bacteria – Proteobacteria – Zetaproteobacteria*	*Mariprofundus*	*Mariprofundus ferrooxydans* PV-1	E=1e-90; I=76%
00721_0	Oxaloacetate decarboxylase	*Bacteria – Proteobacteria – Zetaproteobacteria*	*Mariprofundus*	*Mariprofundus ferrooxydans* PV-1	E=3e-82; I=90%
01448_0	Oxaloacetate decarboxylase	*Bacteria – Proteobacteria – Zetaproteobacteria*	*Mariprofundus*	*Mariprofundus ferrooxydans* PV-1	E=5e-35; I=84%
12359_0	Carbonic anhydrase	*Bacteria – Proteobacteria – Zetaproteobacteria*	*Mariprofundus*	*Mariprofundus ferrooxydans* PV-1	E=3e-40; I=69%
17317_0	Carbonic anhydrase	*Bacteria – Proteobacteria – Zetaproteobacteria*	*Mariprofundus*	*Mariprofundus ferrooxydans* PV-1	E=9e-19; I=76%

2. Sulfur metabolism

Gene name (contig)	Gene product name	Organism (class)	Genus (best hit)	Strain (best hit)	Expected Value (E) and Identities (I)
11627_0	Sulfur oxidation protein SoxA	*Bacteria – Proteobacteria – Epsilonproteobacteria*	*Sulfurovum*	*Sulfurovum sp.* NBC37-1	E=1e-58; I=80%
16784_0	Sulfur oxidation protein SoxA	*Bacteria – Proteobacteria – Epsilonproteobacteria*	*Sulfurovum*	*Sulfurovum sp.* NBC37-1	E=2e-34; I=70%
24932_1	Sulfur oxidation protein SoxA	*Bacteria – Proteobacteria – Epsilonproteobacteria*	*Sulfurovum*	*Sulfurovum sp.* NBC37-1	E=8e-43; I=83%
26799_0	Sulfur oxidation protein SoxA	*Bacteria – Proteobacteria – Epsilonproteobacteria*	*Sulfurovum*	*Sulfurovum sp.* NBC37-1	E=2e-43; I=82%
11627_1	Sulfur oxidation protein SoxB	*Bacteria – Proteobacteria – Epsilonproteobacteria*	*Sulfurovum*	*Sulfurovum sp.* NBC37-1	E=1e-58; I=80%
00205_5	Sulfur oxidation protein SoxD	*Bacteria – Proteobacteria – Epsilonproteobacteria*	*Sulfurovum*	*Sulfurovum sp.* NBC37-1	E=1e-89; I=51%
15003_1	Sulfur oxidation protein SoxD	*Bacteria – Proteobacteria – Epsilonproteobacteria*	*Sulfurovum*	*Sulfurovum sp.* NBC37-1	E=2e-51; I=59%
05861_12	Sulfur oxidation protein SoxX	*Bacteria – Proteobacteria – Epsilonproteobacteria*	*Sulfurovum*	*Sulfurovum sp.* NBC37-1	E=1e-65; I=69%
00205_3	Sulfur oxidation protein SoxY	*Bacteria – Proteobacteria – Epsilonproteobacteria*	*Sulfurimonas*	*Sulfurimonas denitrificans* DSM 1251	E=3e-38; I=51%
24932_2	Sulfur oxidation protein SoxY	*Bacteria – Proteobacteria – Epsilonproteobacteria*	*Sulfurovum*	*Sulfurovum sp.* NBC37-1	E=1e-29; I=76%
26799_2	Sulfur oxidation protein SoxY	*Bacteria – Proteobacteria – Epsilonproteobacteria*	*Sulfurovum*	*Sulfurovum sp.* NBC37-1	E=1e-29; I=76%
24932_0	Sulfur oxidation protein SoxZ	*Bacteria – Proteobacteria – Epsilonproteobacteria*	*Sulfurovum*	*Sulfurovum sp.* NBC37-1	E=2e-28; I=90%
26799_1	Sulfur oxidation protein SoxZ	*Bacteria – Proteobacteria – Epsilonproteobacteria*	*Sulfurovum*	*Sulfurovum sp.* NBC37-1	E=1e-28; I=90%
00220_4	Beta-lactamase SoxH	*Bacteria – Proteobacteria – Epsilonproteobacteria*	*Sulfurimonas*	*Sulfurimonas denitrificans* DSM 1251	E=5e-102; I=60%
03940_0	Sulfate adenylyltransferase (ATP-sulfurylase)	*Bacteria – Proteobacteria – Epsilonproteobacteria*	*Sulfurovum*	*Sulfurovum sp.* NBC37-1	E=4e-40; I=66%
11490_1	Sulfate adenylyltransferase (ATP-sulfurylase)	*Bacteria – Proteobacteria – Epsilonproteobacteria*	*Sulfurovum*	*Sulfurovum sp.* NBC37-1	E=6e-176; I=77%
11490_2	Sulfate adenylyltransferase (ATP-sulfurylase)	*Bacteria – Proteobacteria – Epsilonproteobacteria*	*Sulfurovum*	*Sulfurovum sp.* NBC37-1	E=7e-122; I=90%
17898_0	Sulfate adenylyltransferase (ATP-sulfurylase)	*Bacteria – Proteobacteria – Epsilonproteobacteria*	*Sulfurovum*	*Sulfurovum sp.* NBC37-1	E=1e-80; I=71%
23145_6	Sulfate adenylyltransferase (ATP-sulfurylase)	*Bacteria – Proteobacteria – Epsilonproteobacteria*	*Sulfurovum*	*Sulfurovum sp.* NBC37-1	E=8e-141; I=62%
11095_1	Sulfide-quinone reductase/oxidoreductase (SQR)	*Bacteria – Proteobacteria – Epsilonproteobacteria*	*Sulfurovum*	*Sulfurovum sp.* NBC37-1	E=1e-101; I=85%
14086_2	Sulfide-quinone reductase/oxidoreductase (SQR)	*Bacteria – Proteobacteria – Epsilonproteobacteria*	*Sulfurovum*	*Sulfurovum sp.* NBC37-1	E=4e-175; I=74%
15666_2	Sulfide-quinone reductase/oxidoreductase (SQR)	*Bacteria – Proteobacteria – Epsilonproteobacteria*	*Sulfurovum*	*Sulfurovum sp.* NBC37-1	E=6e-32; I=85%
15292_1	NrfD-like polysulphide reductases	*Bacteria – Proteobacteria – Epsilonproteobacteria*	*Sulfurovum*	*Sulfurovum sp.* NBC37-1	E=1e-121; I=73%
20433_5/6	Sulfur oxidation protein SoxA	*Bacteria – Proteobacteria – Gammaproteobacteria*	*Neptuniibacter*	*Neptuniibacter caesariensis*	E=6e-46; I=56%
01126_1	Sulfur oxidation protein SoxB	*Bacteria – Proteobacteria – Gammaproteobacteria*	*Thiomicrospira*	*Thiomicrospira crunogena* XCL-2	E=0.0; I=57%
20433_2	Sulfur oxidation protein SoxX	*Bacteria – Proteobacteria – Gammaproteobacteria*	Unclassified	uncultured SUP05 cluster bacterium	E=4e-21; I=65%
20433_3	Sulfur oxidation protein SoxY	*Bacteria – Proteobacteria – Gammaproteobacteria*	*Allochromatium*	*Allochromatium vinosum* DSM 180	E=3e-18; I=51%
20433_4	Sulfur oxidation protein SoxZ	*Bacteria – Proteobacteria – Gammaproteobacteria*	*Thioalkalivibrio*	*Thioalkalivibrio sulfidophilus* HL-EbGr7	E=1e-24; I=52%
12022_1	Sulfur oxidation protein SoxW	*Bacteria – Proteobacteria – Gammaproteobacteria*	*Leptothrix*	*Leptothrix cholodnii* SP-6	E=7e-65; I=34%
05247_0	Sulfate adenylyltransferase (ATP-sulfurylase)	*Bacteria – Proteobacteria – Gammaproteobacteria*	Unclassified	*Candidatus Ruthia magnifica str. Cm*	E=2e-62; I=79%
05247_1	Sulfate adenylyltransferase (ATP-sulfurylase)	*Bacteria – Proteobacteria – Gammaproteobacteria*	*Thiobacillus*	*Thiobacillus denitrificans* ATCC 25259	E=1e-89; I=75%
00614_0	Sulfite reductase subunit A (dsrA)	*Bacteria – Proteobacteria – Gammaproteobacteria*	*Allochromatium*	*Allochromatium vinosum* DSM 180	E=3e-154; I=82%
03171_0	Sulfite reductase subunit A (dsrA)	*Bacteria – Proteobacteria – Gammaproteobacteria*	*Allochromatium*	*Allochromatium vinosum* DSM 180	E=5e-120; I=75%
00614_1	Sulfite reductase subunit B (dsrB)	*Bacteria – Proteobacteria – Gammaproteobacteria*	*Allochromatium*	*Allochromatium vinosum* DSM 180	E=4e-54; I=79%
03104_14	Sulfite reductase subunit L (dsrL)	*Bacteria – Proteobacteria – Gammaproteobacteria*	*Allochromatium*	*Allochromatium vinosum* DSM 180	E=2e-93; I=68%
03104_12	Sulfite reductase subunit J (dsrJ)	*Bacteria – Proteobacteria – Gammaproteobacteria*	*Beggiatoa*	*Beggiatoa sp.* PS	E=7e-16; I=47%
03104_11	Sulfite reductase subunit O (dsrO)	*Bacteria – Proteobacteria – Gammaproteobacteria*	*Thioalkalivibrio*	*Thioalkalivibrio sulfidophilus* HL-EbGr7	E=8e-76; I=59%

03104_9	Sulfite reductase subunit P (dsrP)	*Bacteria – Proteobacteria – Gammaproteobacteria*	*Thioalkalivibrio*	*Thioalkalivibrio sulfidophilus* HL-EbGr7	E=4e-90; I=55%
26680_0	Sulfite reductase subunit F (dsrF)	*Bacteria – Proteobacteria – Gammaproteobacteria*	Unclassified	uncultured SUP05 cluster bacterium	E=2e-46; I=73%
26680_1	Sulfite reductase subunit H (dsrH)	*Bacteria – Proteobacteria – Gammaproteobacteria*	*Beggiatoa*	*Beggiatoa sp.* PS	E=3e-24; I=54%
26680_2	Sulfite reductase subunit C (dsrC)	*Bacteria – Proteobacteria – Gammaproteobacteria*	Unclassified	uncultured bacterium BAC13K9BAC	E=5e-44; I=74%
26680_3/4	Sulfite reductase subunit M (dsrM)	*Bacteria – Proteobacteria – Gammaproteobacteria*	*Thioalkalivibrio*	*Thioalkalivibrio sulfidophilus* HL-EbGr7	E=4e-28; I=65%
26680_5	Sulfite reductase subunit K (dsrK)	*Bacteria – Proteobacteria – Gammaproteobacteria*	*Beggiatoa*	*Beggiatoa sp.* PS	E=0.0; I=73%
27759_0	Sulfite reductase subunit E (dsrE)	*Bacteria – Proteobacteria – Gammaproteobacteria*	*Thioalkalivibrio*	*Thioalkalivibrio sulfidophilus* HL-EbGr7	E=2e-49; I=67%
11918_0	Sulfite reductase subunit E (dsrE)	*Bacteria – Proteobacteria – Gammaproteobacteria*	*Thioalkalivibrio*	*Thioalkalivibrio sulfidophilus* HL-EbGr7	E=2e-48; I=66%
11082_5	Sulfite reductase subunit S (dsrS)	*Bacteria – Proteobacteria – Gammaproteobacteria*	*Beggiatoa*	*Beggiatoa sp.* PS	E=3e-52; I=49%
04014_0	Flavocytochrome c sulfide dehydrogenase *fcc*B	*Bacteria – Proteobacteria – Gammaproteobacteria*	*Beggiatoa*	*Beggiatoa sp.* PS	E=1e-60; I=56%
09909_1	Flavocytochrome c sulfide dehydrogenase *fcc*A	*Bacteria – Proteobacteria – Gammaproteobacteria*	*Allochromatium*	*Allochromatium vinosum* DSM 180	E=2e-49; I=60%
09909_0	Flavocytochrome c sulfide dehydrogenase *fcc*B	*Bacteria – Proteobacteria – Gammaproteobacteria*	Unclassified	Endoriftia persephone 'Hot96_1+Hot96_2'	E=2e-35; I=54%
12605_3	NrfD-like polysulphide reductases	*Bacteria – Proteobacteria – Gammaproteobacteria*	*Shewanella*	*Shewanella sp.* ANA-3	E=8e-18; I=35%
03226_0	Sulfate adenylyltransferase (ATP-sulfurylase)	*Bacteria – Proteobacteria – Zetaproteobacteria*	*Mariprofundus*	*Mariprofundus ferrooxydans* PV-1	E=5e-82; I=85%
15812_0	Sulfate adenylyltransferase (ATP-sulfurylase)	*Bacteria – Proteobacteria – Zetaproteobacteria*	*Mariprofundus*	*Mariprofundus ferrooxydans* PV-1	E=4e-52; I=78%
02918_1	NrfD-like polysulphide reductases	*Bacteria – Proteobacteria – Zetaproteobacteria*	*Mariprofundus*	*Mariprofundus ferrooxydans* PV-1	E=1e-98; I=78%
11594_1	NrfD-like polysulphide reductases	*Bacteria – Proteobacteria – Zetaproteobacteria*	*Mariprofundus*	*Mariprofundus ferrooxydans* PV-1	E=2e-92; I=88%
16533_0	Sulfite reductase	*Bacteria – Proteobacteria – Zetaproteobacteria*	*Mariprofundus*	*Mariprofundus ferrooxydans* PV-1	E=4e-51; I=79%
00425_2	Sulfide-quinone reductase/oxidoreductase (SQR)	*Bacteria – Proteobacteria – Zetaproteobacteria*	*Mariprofundus*	*Mariprofundus ferrooxydans* PV-1	E=7e-99; I=77%
16798_1	Sulfide-quinone reductase/oxidoreductase (SQR)	*Bacteria – Proteobacteria – Betaproteobacteria*	*Sideroxydans*	*Sideroxydans lithotrophicus* ES-1	E=7e-64; I=82%

3. Nitrogen metabolism

Gene name (contig)	Gene product name	Organism (class)	Genus (best hit)	Strain (best hit)	Expected Value (E) and Identities (I)
12234_4	Periplasmic nitrate reductase Nap subunit A	*Bacteria – Proteobacteria – Epsilonproteobacteria*	*Sulfurovum*	*Sulfurovum sp*. NBC37-1	E=0.0; I=85%
12234_3	Periplasmic nitrate reductase Nap subunit A	*Bacteria – Proteobacteria – Epsilonproteobacteria*	*Sulfurovum*	*Sulfurovum sp*. NBC37-1	E=0.0; I=93%
12234_2	Periplasmic nitrate reductase Nap subunit G	*Bacteria – Proteobacteria – Epsilonproteobacteria*	*Sulfurovum*	*Sulfurovum sp*. NBC37-1	E=5e-131; I=81%
12234_1	Periplasmic nitrate reductase Nap subunit H	*Bacteria – Proteobacteria – Epsilonproteobacteria*	*Sulfurovum*	*Sulfurovum sp*. NBC37-1	E=2e-132; I=87%
12234_0	Periplasmic nitrate reductase Nap subunit B	*Bacteria – Proteobacteria – Epsilonproteobacteria*	*Sulfurovum*	*Sulfurovum sp*. NBC37-1	E=8e-72; I=62%
17790_11	Periplasmic nitrate reductase Nap subunit F	*Bacteria – Proteobacteria – Epsilonproteobacteria*	*Sulfurovum*	*Sulfurovum sp*. NBC37-1	E=2e-57; I=71%
17790_9	Periplasmic nitrate reductase Nap subunit L	*Bacteria – Proteobacteria – Epsilonproteobacteria*	*Sulfurovum*	*Sulfurovum sp*. NBC37-1	E=9e-131; I=72%
17790_8	Periplasmic nitrate reductase Nap subunit D	*Bacteria – Proteobacteria – Epsilonproteobacteria*	*Sulfurovum*	*Sulfurovum sp*. NBC37-1	E=3e-26; I=72%
01109_1	Cytoplasmic assimilatory nitrate reductase Nas	*Bacteria – Proteobacteria – Epsilonproteobacteria*	*Sulfurovum*	*Sulfurovum sp*. NBC37-1	E=0.0; I=77%
04308_0	Cytoplasmic assimilatory nitrate reductase Nas	*Bacteria – Proteobacteria – Epsilonproteobacteria*	*Sulfurovum*	*Sulfurovum sp*. NBC37-1	E=3e-33; I=85%
04498_1	Cytoplasmic assimilatory nitrate reductase Nas	*Bacteria – Proteobacteria – Epsilonproteobacteria*	*Sulfurovum*	*Sulfurovum sp*. NBC37-1	E=8e-18; I=66%
09741_6	Periplasmic cytochrome c nitrite reductase Nrf subunit H	*Bacteria – Proteobacteria – Epsilonproteobacteria*	*Sulfurospirillum*	*Sulfurospirillum deleyianum* DSM 6946	E=1e-63; I=69.5%
09741_7	Periplasmic cytochrome c nitrite reductase Nrf subunit A	*Bacteria – Proteobacteria – Epsilonproteobacteria*	*Sulfurospirillum*	*Sulfurospirillum deleyianum* DSM 6946	E=5e-53; I=67%
00053_0/4	Periplasmic cytochrome c nitrite reductase Nrf subunit A	*Bacteria – Proteobacteria – Epsilonproteobacteria*	*Sulfurospirillum*	*Sulfurospirillum deleyianum* DSM 6946	E=3e-83; I=71%
00053_1	Periplasmic cytochrome c nitrite reductase Nrf subunit I	*Bacteria – Proteobacteria – Epsilonproteobacteria*	*Sulfurospirillum*	*Sulfurospirillum deleyianum* DSM 6946	E=0.0; I=51%

00053_2	Periplasmic cytochrome c nitrite reductase Nrf subunit J	*Bacteria – Proteobacteria – Epsilonproteobacteria*	*Sulfurospirillum*	*Sulfurospirillum deleyianum* DSM 6946	E=2e-13; I=47%
00053_3	Heme d1 biosynthesis protein NirJ	*Bacteria – Proteobacteria – Epsilonproteobacteria*	*Sulfurovum*	*Sulfurovum sp*. NBC37-1	E=6e-59; I=79%
19498_1	Heme d1 biosynthesis protein NirJ	*Bacteria – Proteobacteria – Epsilonproteobacteria*	*Sulfurovum*	*Sulfurovum sp*. NBC37-1	E=6e-71; I=89%
14523_0	Cytochrome cd1 nitrite reductase NirS	*Bacteria – Proteobacteria – Epsilonproteobacteria*	*Sulfurovum*	*Sulfurovum sp*. NBC37-1	E=2e-23; I=96%
18507_0	Heme d1 biosynthesis protein NirF	*Bacteria – Proteobacteria – Epsilonproteobacteria*	*Sulfurovum*	*Sulfurovum sp*. NBC37-1	E=2e-97; I=85%
13983_1	Cytochrom c nitric oxide reductase cNor subunit C	*Bacteria – Proteobacteria – Epsilonproteobacteria*	*Sulfurovum*	*Sulfurovum sp*. NBC37-1	E=2e-87; I=87%
13983_0	Cytochrom c nitric oxide reductase cNor subunit B	*Bacteria – Proteobacteria – Epsilonproteobacteria*	*Sulfurovum*	*Sulfurovum sp*. NBC37-1	E=1e-28; I=85%
05963_0/1/2	Cytochrom c nitrous oxide reductase cNos subunit NosZ	*Bacteria – Proteobacteria – Epsilonproteobacteria*	*Sulfurovum*	*Sulfurovum sp*. NBC37-1	E=7e-38; I=65%
00943_0	Cytochrom c nitrous oxide reductase cNos subunit NosZ	*Bacteria – Proteobacteria – Epsilonproteobacteria*	*Sulfurovum*	*Sulfurovum sp*. NBC37-1	E=7e-129; I=66%
09321_2/3	Cytochrom c nitrous oxide reductase cNos subunit NosZ	*Bacteria – Proteobacteria – Epsilonproteobacteria*	*Sulfurovum*	*Sulfurovum sp*. NBC37-1	E=3e-14; I=54%
00275_4	Nitrogen fixing protein NifU	*Bacteria – Proteobacteria – Epsilonproteobacteria*	*Sulfurovum*	*Sulfurovum sp*. NBC37-1	E=5e-161; I=85%
02580_1	Nitrogen fixing protein NifU	*Bacteria – Proteobacteria – Epsilonproteobacteria*	*Sulfurovum*	*Sulfurovum sp*. NBC37-1	E=1e-58; I=92%
10531_1	Nitrogen fixing protein NifU	*Bacteria – Proteobacteria – Epsilonproteobacteria*	*Sulfurovum*	*Sulfurovum sp*. NBC37-1	E=2e-29; I=65%
10998_1	Nitrogen fixing protein NifU	*Bacteria – Proteobacteria – Epsilonproteobacteria*	*Sulfurovum*	*Sulfurovum sp*. NBC37-1	E=4e-18; I=90%
15116_5	Nitrogen fixing protein NifU	*Bacteria – Proteobacteria – Epsilonproteobacteria*	*Sulfurovum*	*Sulfurovum sp*. NBC37-1	E=9e-32; I=68%
00526_11	Glutamine synthetase type I	*Bacteria – Proteobacteria – Epsilonproteobacteria*	*Sulfurovum*	*Sulfurovum sp*. NBC37-1	E=2e-20; I=59%
09709_10	Glutamine synthetase type I	*Bacteria – Proteobacteria – Epsilonproteobacteria*	*Sulfurovum*	*Sulfurovum sp*. NBC37-1	E=2e-178; I=77%
14203_0	Glutamine synthetase type I	*Bacteria – Proteobacteria – Epsilonproteobacteria*	*Sulfurovum*	*Sulfurovum sp*. NBC37-1	E=1e-76; I=75%
16005_0	Glutamate dehydrogenase	*Bacteria – Proteobacteria – Epsilonproteobacteria*	*Sulfurovum*	*Sulfurovum sp*. NBC37-1	E=3e-103; I=78%
00879_5	Signal transduction histidine kinase nitrate/nitrite-specific NarX	*Bacteria – Proteobacteria – Gammaproteobacteria*	*Thioalkalivibrio*	*Thioalkalivibrio sulfidophilus* HL-EbGr7	E=4e-89; I=44%
10758_4/5	Signal transduction histidine kinase nitrate/nitrite-specific NarX	*Bacteria – Proteobacteria – Gammaproteobacteria*	Unclassified	uncultured SUP05 cluster bacterium	E=0.0; I=71%
10758_0	Periplasmic nitrate reductase Nap subunit B	*Bacteria – Proteobacteria – Gammaproteobacteria*	*Aeromonas*	*Aeromonas hydrophila ssp. hydrophila*	E=3e-21; I=44%
10758_3	Periplasmic nitrate reductase Nap subunit G	*Bacteria – Proteobacteria – Gammaproteobacteria*	Unclassified	uncultured SUP05 cluster bacterium	E=2e82; I=55%
10758_1/2	Periplasmic nitrate reductase Nap subunit H	*Bacteria – Proteobacteria – Gammaproteobacteria*	*Aeromonas*	*Aeromonas hydrophila ssp. Hydrophila*	E=6e-41; I=55%
30108_0	Periplasmic nitrate reductase Nap subunit F	*Bacteria – Proteobacteria – Gammaproteobacteria*	*Haemophilus*	*Haemophilus parainfluenzae* T3T1	E=2e-35; I=42%
10758_6	Periplasmic nitrate reductase Nap subunit D	*Bacteria – Proteobacteria – Gammaproteobacteria*	*Vibrio*	*Vibrio harveyi* HY01	E=5e-6; I=40%
13891_0	Cytoplasmic assimilatory nitrate reductase Nas	*Bacteria – Proteobacteria – Gammaproteobacteria*	*Saccharophagus*	*Saccharophagus degradans* 2-40	E=0.0; I=68%
06214_0	Nitrite reductase copper-containing NirK	*Bacteria – Proteobacteria – Gammaproteobacteria*	Unclassified	uncultured bacterium 1062	E=9e-22; I=71%
01148_1	Nitrite reductase copper-containing NirK	*Bacteria – Proteobacteria – Gammaproteobacteria*	*Nitrosococcus*	*Nitrosococcus oceani* ATCC 19707	E=1e-45; I=69%
03044_0/1	Cytochrom c nitric oxide reductase cNor subunit B	*Bacteria – Proteobacteria – Gammaproteobacteria*	*Photobacterium*	*Photobacterium profundum* SS9	E=3e-147; I=76%
03044_2	Cytochrom c nitric oxide reductase cNor subunit C	*Bacteria – Proteobacteria – Gammaproteobacteria*	Unclassified	Endoriftia persephone 'Hot96_1+Hot96_2'	E=3e-40; I=85%
11620_6	Cytochrom c nitric oxide reductase cNor subunit C	*Bacteria – Proteobacteria – Gammaproteobacteria*	Unclassified	Endoriftia persephone 'Hot96_1+Hot96_2'	E=4e-47; I=57%
11221_2	Nitrogen regulatory protein PII	*Bacteria – Proteobacteria – Zetaproteobacteria*	*Mariprofundus*	*Mariprofundus ferrooxydans* PV-1	E=7e-42; I=72%
11221_0/1	Ammonium transporter	*Bacteria – Proteobacteria – Zetaproteobacteria*	*Mariprofundus*	*Mariprofundus ferrooxydans* PV-1	E=3e-68; I=78%
09041_1	Glutamine synthetase type III	*Bacteria – Proteobacteria – Zetaproteobacteria*	*Mariprofundus*	*Mariprofundus ferrooxydans* PV-1	E=0.0; I=73%
22266_1	Glutamine synthetase type I	*Bacteria – Proteobacteria – Zetaproteobacteria*	*Mariprofundus*	*Mariprofundus ferrooxydans* PV-1	E=3e-64; I=76%
02616_0	NAD-glutamate dehydrogenase	*Bacteria – Proteobacteria – Zetaproteobacteria*	*Mariprofundus*	*Mariprofundus ferrooxydans* PV-1	E=6e-111; I=60%
14032_1	NAD-glutamate dehydrogenase	*Bacteria – Proteobacteria – Zetaproteobacteria*	*Mariprofundus*	*Mariprofundus ferrooxydans* PV-1	E=9e-54; I=31%

4. Oxygen metabolism & oxidative stress response

Gene name (contig)	Gene product name	Organism (class)	Genus (best hit)	Strain (best hit)	Expected Value (E) and Identities (I)
00802_1	cbb3-type cytochrome c oxidase subunit I CooN	*Bacteria – Proteobacteria – Epsilonproteobacteria*	*Arcobacter*	*Arcobacter nitrofigilis* DSM 7299	E=1e-148; I=84%
01805_4	cbb3-type cytochrome c oxidase subunit I CooN	*Bacteria – Proteobacteria – Epsilonproteobacteria*	*Arcobacter*	*Arcobacter nitrofigilis* DSM 7299	E=3e-51; I=83%
12977_0	cbb3-type cytochrome c oxidase subunit I CooN	*Bacteria – Proteobacteria – Epsilonproteobacteria*	*Sulfurovum*	*Sulfurovum sp*. NBC37-1	E=1e-118; I=79%
14645_0/2	cbb3-type cytochrome c oxidase subunit I CooN	*Bacteria – Proteobacteria – Epsilonproteobacteria*	*Sulfurovum*	*Sulfurovum sp*. NBC37-1	E=4e-40; I=80%
19623_0	cbb3-type cytochrome c oxidase subunit I CooN	*Bacteria – Proteobacteria – Epsilonproteobacteria*	*Sulfurovum*	*Sulfurovum sp*. NBC37-1	E=2e-33; I=83%
00802_0	cbb3-type cytochrome c oxidase subunit II CooO	*Bacteria – Proteobacteria – Epsilonproteobacteria*	*Arcobacter*	*Arcobacter nitrofigilis* DSM 7299	E=2e-100; I=75%
00978_0	cbb3-type cytochrome c oxidase subunit II CooO	*Bacteria – Proteobacteria – Epsilonproteobacteria*	*Sulfurovum*	*Sulfurovum sp*. NBC37-1	E=7e-78; I=78%
13945_0	cbb3-type cytochrome c oxidase subunit II CooO	*Bacteria – Proteobacteria – Epsilonproteobacteria*	*Sulfurimonas*	*Sulfurimonas denitrificans* DSM 1251	E=2e-70; I=81%
14645_1	cbb3-type cytochrome c oxidase subunit II CooO	*Bacteria – Proteobacteria – Epsilonproteobacteria*	*Sulfurovum*	*Sulfurovum sp*. NBC37-1	E=6e-83; I=80%
17792_0	cbb3-type cytochrome c oxidase subunit II CooO	*Bacteria – Proteobacteria – Epsilonproteobacteria*	*Sulfurovum*	*Sulfurovum sp*. NBC37-1	E=6e-38; I=66%
00974_4	cbb3-type cytochrome c oxidase subunit III CooP	*Bacteria – Proteobacteria – Epsilonproteobacteria*	*Arcobacter*	*Arcobacter nitrofigilis* DSM 7299	E=1e-107; I=68%
18034_1	cbb3-type cytochrome c oxidase subunit III CooP	*Bacteria – Proteobacteria – Epsilonproteobacteria*	*Sulfurovum*	*Sulfurovum sp*. NBC37-1	E=1e-35; I=78%
00832_1	Methionine sulfoxide reductase	*Bacteria – Proteobacteria – Epsilonproteobacteria*	*Arcobacter*	*Arcobacter nitrofigilis* DSM 7299	E=5e-91; I=58%
05361_0	Methionine sulfoxide reductase	*Bacteria – Proteobacteria – Epsilonproteobacteria*	*Sulfurimonas*	*Sulfurimonas denitrificans* DSM 1251	E=9e-26; I=62%
13165_0	Methionine sulfoxide reductase	*Bacteria – Proteobacteria – Epsilonproteobacteria*	*Arcobacter*	*Arcobacter butzleri* RM4018	E=1e-42; I=69%
01163_0	Alkylhydroperoxide reductase	*Bacteria – Proteobacteria – Epsilonproteobacteria*	*Sulfurovum*	*Sulfurovum sp*. NBC37-1	E=3e-69; I=79%
01728_4	Alkylhydroperoxide reductase	*Bacteria – Proteobacteria – Epsilonproteobacteria*	*Sulfurovum*	*Sulfurovum sp*. NBC37-1	E=2e-94; I=89%
03003_1	Alkylhydroperoxide reductase	*Bacteria – Proteobacteria – Epsilonproteobacteria*	*Sulfurovum*	*Sulfurovum sp*. NBC37-1	E=5e-94; I=83%
15402_0	Alkylhydroperoxide reductase	*Bacteria – Proteobacteria – Epsilonproteobacteria*	*Sulfurovum*	*Sulfurovum sp*. NBC37-1	E=1e-53; I=86%
16054_4	Alkylhydroperoxide reductase	*Bacteria – Proteobacteria – Epsilonproteobacteria*	*Arcobacter*	*Arcobacter nitrofigilis* DSM 7299	E=3e-93; I=83%
01728_5	Peroxide stress regulator	*Bacteria – Proteobacteria – Epsilonproteobacteria*	*Sulfurovum*	*Sulfurovum sp*. NBC37-1	E=8e-32; I=50%
03003_0	Peroxide stress regulator	*Bacteria – Proteobacteria – Epsilonproteobacteria*	*Sulfurovum*	*Sulfurovum sp*. NBC37-1	E=3e-8; I=59%
04022_14	Cytochrome c peroxidase	*Bacteria – Proteobacteria – Epsilonproteobacteria*	*Sulfurovum*	*Sulfurovum sp*. NBC37-1	E=2e-110; I=60%
12151_0	Cytochrome c peroxidase	*Bacteria – Proteobacteria – Epsilonproteobacteria*	*Arcobacter*	*Arcobacter nitrofigilis* DSM 7299	E=2e-19; I=65%
12152_2	Cytochrome c peroxidase	*Bacteria – Proteobacteria – Epsilonproteobacteria*	*Arcobacter*	*Arcobacter nitrofigilis* DSM 7299	E=8e-25; I=72%
16475_0	Cytochrome c peroxidase	*Bacteria – Proteobacteria – Epsilonproteobacteria*	*Nitratiruptor*	*Nitratiruptor sp*. SB155-2	E=7e-93; I=67%
11262_0	Thiol peroxidase	*Bacteria – Proteobacteria – Epsilonproteobacteria*	Unclassified	*Campylobacterales bacterium* GD 1	E=4e-40; I=84%
15641_3	Thiol peroxidase	*Bacteria – Proteobacteria – Epsilonproteobacteria*	Unclassified	*Campylobacterales bacterium* GD 1	E=2e-83; I=90%
08666_1	Superoxide dismutase (Nickel-type)	*Bacteria – Proteobacteria – Epsilonproteobacteria*	*Alteromonas*	*Alteromonas macleodii* ATCC 27126	E=3e-93; I=83%
00101_3	Thioredoxin	*Bacteria – Proteobacteria – Epsilonproteobacteria*	*Sulfurovum*	*Sulfurovum sp*. NBC37-1	E=3e-47; I=85%
00101_4	Thioredoxin reductase	*Bacteria – Proteobacteria – Epsilonproteobacteria*	*Sulfurovum*	*Sulfurovum sp*. NBC37-1	E=2e-136; I=80%
00314_11	Thioredoxin	*Bacteria – Proteobacteria – Epsilonproteobacteria*	*Sulfurovum*	*Sulfurovum sp*. NBC37-1	E=4e-37; I=52%
00480_0	Thioredoxin reductase	*Bacteria – Proteobacteria – Epsilonproteobacteria*	*Sulfurovum*	*Sulfurovum sp*. NBC37-1	E=2e-56; I=71%
00483_0	Thioredoxin reductase	*Bacteria – Proteobacteria – Epsilonproteobacteria*	*Sulfurovum*	*Sulfurovum sp*. NBC37-1	E=4e-36; I=71%
00845_13	Thioredoxin	*Bacteria – Proteobacteria – Epsilonproteobacteria*	*Sulfurovum*	*Sulfurovum sp*. NBC37-1	E=3e41; I=74%
01805_5	Thioredoxin redcutase	*Bacteria – Proteobacteria – Epsilonproteobacteria*	*Sulfurovum*	*Sulfurovum sp*. NBC37-1	E=2e-52; I=72%
03429_0	Thioredoxin reductase	*Bacteria – Proteobacteria – Epsilonproteobacteria*	*Nautilia*	*Nautilia profundicola AmH*	E=3e-91; I=63%

04105_4	Thioredoxin	*Bacteria – Proteobacteria – Epsilonproteobacteria*	*Sulfurovum*	*Sulfurovum sp*. NBC37-1	E=2e-49; I=64%
05970_0	Thioredoxin	*Bacteria – Proteobacteria – Epsilonproteobacteria*	*Sulfurovum*	*Sulfurovum sp*. NBC37-1	E=3e-24; I=44%
10827_2	Thioredoxin	*Bacteria – Proteobacteria – Epsilonproteobacteria*	*Sulfurovum*	*Sulfurovum sp*. NBC37-1	E=1e-25; I=40%
11323_0	Thioredoxin	*Bacteria – Proteobacteria – Epsilonproteobacteria*	*Sulfurovum*	*Sulfurovum sp*. NBC37-1	E=4e-45; I=61%
11769_1	Thioredoxin	*Bacteria – Proteobacteria – Epsilonproteobacteria*	*Sulfurovum*	*Sulfurovum sp*. NBC37-1	E=9e-28; I=67%
13561_0	Thioredoxin	*Bacteria – Proteobacteria – Epsilonproteobacteria*	*Sulfurovum*	*Sulfurovum sp*. NBC37-1	E=9e-37; I=89%
14590_0	Thioredoxin reductase	*Bacteria – Proteobacteria – Epsilonproteobacteria*	*Sulfurovum*	*Sulfurovum sp*. NBC37-1	E=2e-49; I=94%
04342_0	Endonuclease type III restriction enzyme	*Bacteria – Proteobacteria – Epsilonproteobacteria*	*Helicobacter*	*Helicobacter pylori 35A*	E=6e-58; I=51%
05140_0	Endonuclease type III	*Bacteria – Proteobacteria – Epsilonproteobacteria*	*Sulfurovum*	*Sulfurovum sp*. NBC37-1	E=3e-85; I=83%
05143_7	Endonuclease type IV	*Bacteria – Proteobacteria – Epsilonproteobacteria*	*Sulfurovum*	*Sulfurovum sp*. NBC37-1	E=6e-134; I=82%
05861_9	Endonuclease type III	*Bacteria – Proteobacteria – Epsilonproteobacteria*	*Sulfurovum*	*Sulfurovum sp*. NBC37-1	E=3e-80; I=77%
09926_5	Endonuclease type IV	*Bacteria – Proteobacteria – Epsilonproteobacteria*	*Sulfurovum*	*Sulfurovum sp*. NBC37-1	E=1e-140; I=81%
10699_1	Endonuclease type VII recombination	*Bacteria – Proteobacteria – Epsilonproteobacteria*	*Sulfurovum*	*Sulfurovum sp*. NBC37-1	E=3e-74; I=71%
15886_0	Exodeoxyrobonuclease type III	*Bacteria – Proteobacteria – Epsilonproteobacteria*	*Arcobacter*	*Arcobacter nitrofigilis* DSM 7299	E=9e-102; I=80%
00289_0	DNA mismatch repair protein MutS	*Bacteria – Proteobacteria – Epsilonproteobacteria*	*Sulfurospirillum*	*Sulfurospirillum deleyianum* DSM 6946	E=4e-161; I=60%
00873_2	DNA mismatch repair protein MutS	*Bacteria – Proteobacteria – Epsilonproteobacteria*	*Sulfurospirillum*	*Sulfurospirillum deleyianum* DSM 6946	E=5e-161; I=61%
03559_0	MutS2 family protein	*Bacteria – Proteobacteria – Epsilonproteobacteria*	*Sulfurovum*	*Sulfurovum sp*. NBC37-1	E=1e-29; I=60%
036478_0	DNA mismatch repair protein MutS	*Bacteria – Proteobacteria – Epsilonproteobacteria*	Unclassified	*Campylobacterales bacterium* GD 1	E=4e-37; I=75%
03706_0	DNA mismatch repair protein MutS	*Bacteria – Proteobacteria – Epsilonproteobacteria*	*Sulfurospirillum*	*Sulfurospirillum deleyianum* DSM 6946	E=2e-104; I=57%
11098_0	DNA mismatch repair protein MutS	*Bacteria – Proteobacteria – Epsilonproteobacteria*	*Sulfurospirillum*	*Sulfurospirillum deleyianum* DSM 6946	E=2e-149; I=60%
13963_2	cbb3-type cytochrome c oxidase subunit I CooN	*Bacteria – Proteobacteria – Gammaproteobacteria*	*Thioalkalivibrio*	*Thioalkalivibrio sp.* K90mix	E=0.0; I=79%
15049_1/2	cbb3-type cytochrome c oxidase subunit I CooN	*Bacteria – Proteobacteria – Gammaproteobacteria*	*Thiomicrospira*	*Thiomicrospira crunogena* XCL-2	E=3e-60; I=63%
15049_3	cbb3-type cytochrome c oxidase subunit I CooN	*Bacteria – Proteobacteria – Gammaproteobacteria*	*Methylophaga*	*Methylophaga thiooxydans* DMS010	E=5e-58; I=51%
13963_3	cbb3-type cytochrome c oxidase subunit II CooO	*Bacteria – Proteobacteria – Gammaproteobacteria*	*Thiobacillus*	*Thiobacillus denitrificans* ATCC 25259	E=8e-59; I=77%
24977_0	cbb3-type cytochrome c oxidase subunit II CooO	*Bacteria – Proteobacteria – Gammaproteobacteria*	*Thioalkalivibrio*	*Thioalkalivibrio sp.* HL-EbGr7	E=2e-24; I=65%
10652_0	cbb3-type cytochrome c oxidase subunit III CooP	*Bacteria – Proteobacteria – Gammaproteobacteria*	*Thioalkalivibrio*	*Thioalkalivibrio sulfidophilus* HL-EbGr7	E=6e-42; I=50%
11620_1	cbb3-type cytochrome c oxidase subunit III CooP	*Bacteria – Proteobacteria – Gammaproteobacteria*	*Colwellia*	*Colwellia psychrerythraea* 34H	E=5e-59; I=56%
25697_0	cbb3-type cytochrome c oxidase subunit III CooP	*Bacteria – Proteobacteria – Gammaproteobacteria*	*Thioalkalivibrio*	*Thioalkalivibrio sp.* K90mix	E=3e-23; I=42%
19205_0	cbb3-type cytochrome c oxidase accessory protein CooG	*Bacteria – Proteobacteria – Gammaproteobacteria*	*Thioalkalivibrio*	*Thioalkalivibrio sp.* K90mix	E=1e-133; I=53%
09805_8	Methionine sulfoxide reductase	*Bacteria – Proteobacteria – Gammaproteobacteria*	*Thioalkalivibrio*	*Thioalkalivibrio sp.* HL-EbGr7	E=7e-37; I=62%
01608_1	Cytochrome c peroxidase	*Bacteria – Proteobacteria – Gammaproteobacteria*	*Vibrio*	*Vibrio sp.* MED22	E=2e-130; I=59%
01608_2	Cytochrome c peroxidase	*Bacteria – Proteobacteria – Gammaproteobacteria*	*Vibrio*	*Vibrio splendidus* LGP32	E=3e-129; I=66%
02464_2	Cytochrome c peroxidase	*Bacteria – Proteobacteria – Gammaproteobacteria*	*Shewanella*	*Shewanella woodyi* ATCC 51908	E=9e-99; I=45%
04649_0	Glutathione peroxidase	*Bacteria – Proteobacteria – Gammaproteobacteria*	*Shewanella*	*Shewanella violacea* DSS12	E=5e-56; I=64%
14481_4	Peroxiredoxin	*Bacteria – Proteobacteria – Gammaproteobacteria*	*Vibrio*	*Vibrio cholerae* 623-39	E=6e-85; I=71%
03570_1	Cytochrome c peroxidase	*Bacteria – Proteobacteria – Zetaproteobacteria*	*Mariprofundus*	*Mariprofundus ferrooxydans* PV-1	E=4e-32; I=64%
00225_9	Endonuclease DNA/RNA non specific type	*Bacteria – Proteobacteria – Zetaproteobacteria*	*Mariprofundus*	*Mariprofundus ferrooxydans* PV-1	E=2e-47; I=47%

5. Other genes of metabolic and environmental relevance

Gene name (contig)	Gene product name	Organism (class)	Genus (best hit)	Strain (best hit)	Expected Value (E) and Identities (I)
01373_7	Ferrochelatase	*Bacteria – Proteobacteria – Epsilonproteobacteria*	*Sulfurovum*	*Sulfurovum sp.* NBC37-1	E=8e-124; I=71%
00480_1	Methyl-accepting chemotaxis protein	*Bacteria – Proteobacteria – Epsilonproteobacteria*	Unclassified	Campylobacterales *bacterium* GD 1	E=5e-45; I=46%
13170_8	Chemotaxis protein MotB	*Bacteria – Proteobacteria – Epsilonproteobacteria*	*Sulfurospirillum*	*Sulfurospirillum deleyianum* DSM 6946	E=3e-63; I=41%
00068_0	Citrate transporter	*Bacteria – Proteobacteria – Epsilonproteobacteria*	*Sulfurovum*	*Sulfurovum sp.* NBC37-1	E=1e-166; I=82%
00344_1	Amino acid transporter	*Bacteria – Proteobacteria – Epsilonproteobacteria*	*Sulfurovum*	*Sulfurovum sp.* NBC37-1	E=1e-121; I=74%
00810_0	TRAP transporter	*Bacteria – Proteobacteria – Epsilonproteobacteria*	*Sulfurospirillum*	*Sulfurospirillum deleyianum* DSM 6946	E=2e-55; I=59%
00810_0	TRAP transporter	*Bacteria – Proteobacteria – Epsilonproteobacteria*	*Nautilia*	*Nautilia profundicola* AmH	E=4e-55; I=61%
00889_1	Periplasmic phosphate binding protein	*Bacteria – Proteobacteria – Epsilonproteobacteria*	*Sulfurospirillum*	*Sulfurospirillum deleyianum* DSM 6946	E=5e-17; I=63%
04714_0	Molybdenum ABC transporter, periplasmic molybdate-binding protein	*Bacteria – Proteobacteria – Epsilonproteobacteria*	*Arcobacter*	*Arcobacter nitrofigilis* DSM 7299	E=5e-30; I=61%
05765_0	Oligopeptide/dipeptide ABC transporter, ATPase subunit	*Bacteria – Proteobacteria – Epsilonproteobacteria*	*Sulfurovum*	*Sulfurovum sp.* NBC37-1	E=9e-46; I=70%
05765_1/2	Oligopeptide ABC transporter ATP-binding protein	*Bacteria – Proteobacteria – Epsilonproteobacteria*	*Sulfurovum*	*Sulfurovum sp.* NBC37-1	E=8e-18; I=59%
00192_3	ABC transporter, ATP-binding protein	*Bacteria – Proteobacteria – Epsilonproteobacteria*	*Sulfurovum*	*Sulfurovum sp.* NBC37-1	E=6e-76; I=61%
00216_0	ABC transporter, ATP-binding protein	*Bacteria – Proteobacteria – Epsilonproteobacteria*	*Sulfurovum*	*Sulfurovum sp.* NBC37-1	E=1e-75; I=83%
00279_1	ABC transporter, ATP-binding protein	*Bacteria – Proteobacteria – Epsilonproteobacteria*	*Sulfurovum*	*Sulfurovum sp.* NBC37-1	E=2e-96; I=76%
00314_14	Magnesium transporter	*Bacteria – Proteobacteria – Epsilonproteobacteria*	Unclassified	Campylobacterales *bacterium* GD 1	E=6e-149; I=65%
00346_2	ABC transporter, permease protein	*Bacteria – Proteobacteria – Epsilonproteobacteria*	*Nitratiruptor*	*Nitratiruptor sp.* SB155-2	E=2e-48; I=63%
00387_0	ABC transporter, ATP-binding protein	*Bacteria – Proteobacteria – Epsilonproteobacteria*	*Sulfurovum*	*Sulfurovum sp.* NBC37-1	E=5e-71; I=76%
00387_1	ABC transporter, ATPase component	*Bacteria – Proteobacteria – Epsilonproteobacteria*	*Sulfurovum*	*Sulfurovum sp.* NBC37-1	E=7e-62; I=76%
00391_2/3	ABC transporter, ATP-binding protein	*Bacteria – Proteobacteria – Epsilonproteobacteria*	*Arcobacter*	*Arcobacter butzleri* RM4018	E=4e-40; I=58%
00430_0	ABC transporter, ATP-binding protein	*Bacteria – Proteobacteria – Epsilonproteobacteria*	*Sulfurovum*	*Sulfurovum sp.* NBC37-1	E=1e-147; I=80%
00599_1	Phosphate ABC transporter	*Bacteria – Proteobacteria – Epsilonproteobacteria*	*Sulfurovum*	*Sulfurovum sp.* NBC37-1	E=2e-149; I=85%
00610_2	O-antigen transporter	*Bacteria – Proteobacteria – Epsilonproteobacteria*	*Caminibacter*	*Caminibacter mediatlanticus* TB-2	E=4e-69; I=45%
00641_0	TRAP dicarboxylate transporter, DctP subunit	*Bacteria – Proteobacteria – Epsilonproteobacteria*	*Caminibacter*	*Caminibacter mediatlanticus* TB-2	E=9e-72; I=77%
00644_0	Phosphate ABC transporter, permease protein PstC	*Bacteria – Proteobacteria – Epsilonproteobacteria*	*Sulfurovum*	*Sulfurovum sp.* NBC37-1	E=7e-54; I=82%
00728_4	ABC transporter, ATP-binding protein	*Bacteria – Proteobacteria – Epsilonproteobacteria*	*Sulfurovum*	*Sulfurovum sp.* NBC37-1	E=3e-100; I=70%
00845_14	Multidrug ABC transporter, ATP-binding protein	*Bacteria – Proteobacteria – Epsilonproteobacteria*	*Sulfurovum*	*Sulfurovum sp.* NBC37-1	E=0.0; I=58%
00845_16	Cation ABC transporter, ATP-binding protein	*Bacteria – Proteobacteria – Epsilonproteobacteria*	*Sulfurovum*	*Sulfurovum sp.* NBC37-1	E=3e-92; I=69%
00845_17	Zinc ABC transporter, permease protein	*Bacteria – Proteobacteria – Epsilonproteobacteria*	*Sulfurovum*	*Sulfurovum sp.* NBC37-1	E=3e-107; I=75%
00940_24	Ammonium transporter	*Bacteria – Proteobacteria – Epsilonproteobacteria*	*Sulfurovum*	*Sulfurovum sp.* NBC37-1	E=2e-141; I=86%
00951_0	Magnesium transporter	*Bacteria – Proteobacteria – Epsilonproteobacteria*	Unclassified	Campylobacterales *bacterium* GD 1	E=8e-105; I=65%
00968_0	ABC transporter, ATP-binding protein	*Bacteria – Proteobacteria – Epsilonproteobacteria*	*Sulfurovum*	*Sulfurovum sp.* NBC37-1	E=6e-82; I=76%
01073_6	ABC transporter, ATP-binding protein	*Bacteria – Proteobacteria – Epsilonproteobacteria*	*Sulfurovum*	*Sulfurovum sp.* NBC37-1	E=0.0; I=73%
01114_0	Biopolymer transport protein exbB	*Bacteria – Proteobacteria – Epsilonproteobacteria*	*Sulfurovum*	*Sulfurovum sp.* NBC37-1	E=5e-27; I=68%
01352_2	Transporter protein	*Bacteria – Proteobacteria – Epsilonproteobacteria*	*Sulfurovum*	*Sulfurovum sp.* NBC37-1	E=1e-14; I=61%
01396_0	Sodium/sulfate symporter	*Bacteria – Proteobacteria – Epsilonproteobacteria*	*Sulfurovum*	*Sulfurovum sp.* NBC37-1	E=1e-17; I=56%
01414_7	Biopolymer transport protein, ExbD/TolR family	*Bacteria – Proteobacteria – Epsilonproteobacteria*	*Helicobacter*	*Helicobacter winghamensis* ATCC BAA-430	E=2e-30; I=49%
01414_8	Exbb-olq family transport protein	*Bacteria – Proteobacteria – Epsilonproteobacteria*	*Sulfurovum*	*Sulfurovum sp.* NBC37-1	E=8e-49; I=62%

01417_1	Cation efflux system protein	*Bacteria – Proteobacteria – Epsilonproteobacteria*	*Campylobacter*	*Campylobacterales* bacterium GD 1	E=1e-49; I=66%
01495_0	Na$^+$-transporting methylmalonyl-CoA/oxaloacetate decarboxylase	*Bacteria – Proteobacteria – Epsilonproteobacteria*	*Arcobacter*	*Arcobacter nitrofigilis* DSM 7299	E=4e-24; I=58%
01593_1	Amino acid ABC transporter, permease protein	*Bacteria – Proteobacteria – Epsilonproteobacteria*	*Arcobacter*	*Arcobacter nitrofigilis* DSM 7299	E=2e-51; I=83%
01716_8	Transport system inner membrane component	*Bacteria – Proteobacteria – Epsilonproteobacteria*	*Sulfurospirillum*	*Sulfurospirillum deleyianum* DSM 6946	E=3e-85; I=64%
01727_1	ABC transporter, ATP-binding protein	*Bacteria – Proteobacteria – Epsilonproteobacteria*	*Sulfurovum*	*Sulfurovum sp.* NBC37-1	E=3e-24; I=93%
01869_0	Multidrug-efflux transporter	*Bacteria – Proteobacteria – Epsilonproteobacteria*	*Nitratiruptor*	*Nitratiruptor sp.* SB155-2	E=1e-58; I=67%
02031_0	Transporter AcrB/D/F family protein	*Bacteria – Proteobacteria – Epsilonproteobacteria*	*Sulfurovum*	*Sulfurovum sp.* NBC37-1	E=0.0; I=72%
02034_1	Cation ABC transporter, permease protein	*Bacteria – Proteobacteria – Epsilonproteobacteria*	*Sulfurovum*	*Sulfurovum sp.* NBC37-1	E=2e-95; I=79%
02526_0	ABC transporter, ATP-binding protein	*Bacteria – Proteobacteria – Epsilonproteobacteria*	*Sulfurovum*	*Sulfurovum sp.* NBC37-2	E=2e-77; I=73%
02759_0	ABC transporter metal-binding lipoprotein	*Bacteria – Proteobacteria – Epsilonproteobacteria*	*Sulfurovum*	*Sulfurovum sp.* NBC37-1	E=7e-51; I=60%
02849_0	Molybdenum ABC transporter, permease protein	*Bacteria – Proteobacteria – Epsilonproteobacteria*	*Caminibacter*	*Caminibacter mediatlanticus* TB-2	E=4e-29; I=62%
02903_0	Citrate transporter	*Bacteria – Proteobacteria – Epsilonproteobacteria*	*Sulfurovum*	*Sulfurovum sp.* NBC37-1	E=0.0; I=77%
02920_0	Glutamine ABC transporter, ATP-binding protein	*Bacteria – Proteobacteria – Epsilonproteobacteria*	*Wolinella*	*Wolinella succinogenes* DSM 1740	E=3e-67; I=72%
03785_0	Formate/nitrite transporter	*Bacteria – Proteobacteria – Epsilonproteobacteria*	*Caminibacter*	*Caminibacter mediatlanticus* TB-2	E=3e-34; I=64%
04022_12	Anion transporter	*Bacteria – Proteobacteria – Epsilonproteobacteria*	*Arcobacter*	*Arcobacter nitrofigilis* DSM 7299	E=4e-106; I=52%
04195_0	Ferrous iron transport protein B	*Bacteria – Proteobacteria – Epsilonproteobacteria*	*Sulfurovum*	*Sulfurovum sp.* NBC37-1	E=4e-97; I=85%
04354_1	Multidrug ABC transporter, ATP-binding protein	*Bacteria – Proteobacteria – Epsilonproteobacteria*	*Sulfurovum*	*Sulfurovum sp.* NBC37-1	E=3e-35; I=51%
04791_1	Transporter, NadC family	*Bacteria – Proteobacteria – Epsilonproteobacteria*	*Nitratiruptor*	*Nitratiruptor sp.* SB155-2	E=4e-11; I=67%
04920_0	Ammonium transporter	*Bacteria – Proteobacteria – Epsilonproteobacteria*	*Sulfurovum*	*Sulfurovum sp.* NBC37-1	E=6e-46; I=81%
05389_0	ABC transporter	*Bacteria – Proteobacteria – Epsilonproteobacteria*	*Arcobacter*	*Arcobacter nitrofigilis* DSM 7299	E=4e-13; I=43%
05805_2	Capsule polysaccharide export system	*Bacteria – Proteobacteria – Epsilonproteobacteria*	*Campylobacter*	*Campylobacter upsaliensis* RM3195	E=7e-124; I=46%
05861_2	Ferrous iron transport protein B	*Bacteria – Proteobacteria – Epsilonproteobacteria*	*Sulfurovum*	*Sulfurovum sp.* NBC37-1	E=0.0; I=64%
05861_3	Ferrous iron transporter, FeoA subunit	*Bacteria – Proteobacteria – Epsilonproteobacteria*	*Sulfurovum*	*Sulfurovum sp.* NBC37-1	E=1e-14; I=59%
05904_0/1	Dipeptide ABC transporter, permease component	*Bacteria – Proteobacteria – Epsilonproteobacteria*	*Helicobacter*	*Helicobacter cetorum*	E=4e-90; I=70%
06043_0	ABC transporter, ATP-binding protein	*Bacteria – Proteobacteria – Epsilonproteobacteria*	*Nitratiruptor*	*Nitratiruptor sp.* SB155-2	E=1e-81; I=63%
06234_3	ABC transporter, permease protein	*Bacteria – Proteobacteria – Epsilonproteobacteria*	*Nitratiruptor*	*Nitratiruptor sp.* SB155-2	E=2e-87; I=58%
06396_7	MFS metabolite transporter	*Bacteria – Proteobacteria – Epsilonproteobacteria*	*Sulfurovum*	*Sulfurovum sp.* NBC37-1	E=1e-93; I=57%
08666_2	Nickel/cobalt transporter, high-affinity	*Bacteria – Proteobacteria – Epsilonproteobacteria*	*Nitratiruptor*	*Nitratiruptor sp.* SB155-2	E=1e-22; I=41%
09587_3	Transporter	*Bacteria – Proteobacteria – Epsilonproteobacteria*	*Sulfurovum*	*Sulfurovum sp.* NBC37-1	E=3e-107; I=53%
09765_2	Phosphate ABC transporter, permease protein PstC	*Bacteria – Proteobacteria – Epsilonproteobacteria*	*Sulfurovum*	*Sulfurovum sp.* NBC37-1	E=2e-101; I=82%
09765_3	Phosphate transport system permease protein	*Bacteria – Proteobacteria – Epsilonproteobacteria*	*Sulfurovum*	*Sulfurovum sp.* NBC37-1	E=1e-87; I=81%
09765_4	Phosphate ABC transporter, ATPase subunit	*Bacteria – Proteobacteria – Epsilonproteobacteria*	*Sulfurovum*	*Sulfurovum sp.* NBC37-1	E=4e-117; I=79%
09927_3	Major facilitator superfamily multidrug efflux transporter	*Bacteria – Proteobacteria – Epsilonproteobacteria*	*Sulfurovum*	*Sulfurovum sp.* NBC37-1	E=1e-144; I=64%
09954_0/1	Molybdenum ABC transporter, permease protein	*Bacteria – Proteobacteria – Epsilonproteobacteria*	*Sulfurovum*	*Sulfurovum sp.* NBC37-1	E=5e-58; I=74%
10270_0	Polysialic acid transport protein kpsD precursor	*Bacteria – Proteobacteria – Epsilonproteobacteria*	*Helicobacter*	*Helicobacter canadensis* MIT 98-5491	E=2e-32; I=42%
10309_1	Nickel/cobalt transporter, high-affinity	*Bacteria – Proteobacteria – Epsilonproteobacteria*	*Sulfurovum*	*Sulfurovum sp.* NBC37-1	E=3e-83; I=76%
10827_3	ABC transporter, ATP-binding/permease protein	*Bacteria – Proteobacteria – Epsilonproteobacteria*	*Nitratiruptor*	*Nitratiruptor sp.* SB155-2	E=3e-102; I=43%
10887_1	Mercuric transport protein MerT	*Bacteria – Proteobacteria – Epsilonproteobacteria*	*Caminibacter*	*Caminibacter mediatlanticus* TB-2	E=2e-23; I=62%
10887_21	Heavy metal transport/detoxification protein	*Bacteria – Proteobacteria – Epsilonproteobacteria*	*Caminibacter*	*Caminibacter mediatlanticus* TB-2	E=4e-12; I=51%
11337_4	TonB system transport protein ExbD type-2	*Bacteria – Proteobacteria – Epsilonproteobacteria*	*Sulfurimonas*	*Sulfurimonas denitrificans* DSM 1251	E=2e-19; I=53%
11378_3	ABC transporter ATP-binding protein	*Bacteria – Proteobacteria – Epsilonproteobacteria*	*Sulfurovum*	*Sulfurovum sp.* NBC37-1	E=2e-18; I=67%

11652_3	TRAP transporter solute receptor, TAXI family	*Bacteria – Proteobacteria – Epsilonproteobacteria*	*Arcobacter*	*Arcobacter nitrofigilis* DSM 7299	E=3e-40; I=32%
12008_3	Transporter AcrB/D/F family	*Bacteria – Proteobacteria – Epsilonproteobacteria*	*Sulfurovum*	*Sulfurovum sp*. NBC37-1	E=0.0; I=65%
12014_2	Ferrous iron transport protein B	*Bacteria – Proteobacteria – Epsilonproteobacteria*	*Sulfurospirillum*	*Sulfurospirillum deleyianum* DSM 6946	E=4e-27; I=69%
12574_0	peptide ABC transporter, permease protein	*Bacteria – Proteobacteria – Epsilonproteobacteria*	*Sulfurovum*	*Sulfurovum sp*. NBC37-1	E=8e-11; I=62%
13876_0	Magnesium transporter	*Bacteria – Proteobacteria – Epsilonproteobacteria*	*Sulfurovum*	*Sulfurovum sp*. NBC37-1	E=1e-24; I=81%
13876_1	Magnesium transporter	*Bacteria – Proteobacteria – Epsilonproteobacteria*	*Sulfurovum*	*Sulfurovum sp*. NBC37-1	E=5e-31; I=67%
13994_0	ABC transporter, ATP-binding protein	*Bacteria – Proteobacteria – Epsilonproteobacteria*	*Sulfurovum*	*Sulfurovum sp*. NBC37-1	E=1e-52; I=91%
13994_1	N-terminal fragment of an ABC transporter	*Bacteria – Proteobacteria – Epsilonproteobacteria*	*Sulfurovum*	*Sulfurovum sp*. NBC37-1	E=9e-19; I=92%
14247_0/1	Secretion protein HlyD	*Bacteria – Proteobacteria – Epsilonproteobacteria*	Unclassified	*Campylobacterales* bacterium GD 1	E=2e-33; I=55%
14453_9	Ribonucleotide-transport ATP-binding protein ABC transporter	*Bacteria – Proteobacteria – Epsilonproteobacteria*	*Nitratiruptor*	*Nitratiruptor sp.* SB155-2	E=4e-87; I=64%
14574_0	Heavy-metal transporting P-type ATPase	*Bacteria – Proteobacteria – Epsilonproteobacteria*	*Arcobacter*	*Arcobacter butzleri* RM4018	E=2e-19; I=61%
14577_0	ABC transporter, ATP-binding protein	*Bacteria – Proteobacteria – Epsilonproteobacteria*	*Sulfurimonas*	*Sulfurimonas denitrificans* DSM 1251	E=1e-19; I=63%
14577_2	ABC transporter, ATP binding protein	*Bacteria – Proteobacteria – Epsilonproteobacteria*	*Sulfurimonas*	*Sulfurimonas denitrificans* DSM 1251	E=7e-19; I=79%
14661_0	ABC transporter, ATP-binding protein	*Bacteria – Proteobacteria – Epsilonproteobacteria*	*Sulfurovum*	*Sulfurovum sp*. NBC37-1	E=7e-43; I=76%
14661_1	N-terminal fragment of an ABC transporter	*Bacteria – Proteobacteria – Epsilonproteobacteria*	*Sulfurovum*	*Sulfurovum sp*. NBC37-1	E=3e-21; I=82%
14866_0	Ammonium transporter	*Bacteria – Proteobacteria – Epsilonproteobacteria*	Unclassified	Campylobacterales *bacterium* GD 1	E=6e-38; I=86%
14901_0/1	Sulfate transporter	*Bacteria – Proteobacteria – Epsilonproteobacteria*	*Sulfurovum*	*Sulfurovum sp*. NBC37-1	E=5e-20; I=93%
15390_0	ABC transporter, ATP-binding/permease protein	*Bacteria – Proteobacteria – Epsilonproteobacteria*	*Nitratiruptor*	*Nitratiruptor sp.* SB155-2	E=1e-20; I=41%
15641_0/1	Molybdenum ABC transporter, periplasmic protein	*Bacteria – Proteobacteria – Epsilonproteobacteria*	*Sulfurovum*	*Sulfurovum sp*. NBC37-1	E=5e-86; I=71%
15641_2	ABC spermidine/putrescine transporter, ATPase subunit	*Bacteria – Proteobacteria – Epsilonproteobacteria*	*Arcobacter*	*Arcobacter nitrofigilis* DSM 7299	E=4e-76; I=69%
15712_0	TRAP dicarboxylate transporter, DctM subunit	*Bacteria – Proteobacteria – Epsilonproteobacteria*	*Sulfurovum*	*Sulfurovum sp*. NBC37-1	E=9e-83; I=82%
15741_0	Cation effux transport protein	*Bacteria – Proteobacteria – Epsilonproteobacteria*	*Sulfurovum*	*Sulfurovum sp*. NBC37-1	E=6e-36; I=78%
15902_0	Cation ABC transporter, permease protein	*Bacteria – Proteobacteria – Epsilonproteobacteria*	*Sulfurovum*	*Sulfurovum sp*. NBC37-1	E=2e-82; I=83%
15938_0	Phosphate ABC transporter, periplasmic protein	*Bacteria – Proteobacteria – Epsilonproteobacteria*	*Sulfurovum*	*Sulfurovum sp*. NBC37-1	E=6e-81; I=87%
16003_0	Cation efflux system protein	*Bacteria – Proteobacteria – Epsilonproteobacteria*	Unclassified	Campylobacterales *bacterium* GD 1	E=3e-57; I=57%
16072_0	ABC transporter, ATP-binding protein	*Bacteria – Proteobacteria – Epsilonproteobacteria*	*Sulfurovum*	*Sulfurovum sp*. NBC37-1	E=2e-22; I=68%
16074_0	ABC transporter, ATP-binding protein	*Bacteria – Proteobacteria – Epsilonproteobacteria*	*Sulfurovum*	*Sulfurovum sp*. NBC37-1	E=7e-16; I=52%
16380_0	TonB system transport protein ExbD type-2	*Bacteria – Proteobacteria – Epsilonproteobacteria*	*Sulfurimonas*	*Sulfurimonas denitrificans* DSM 1251	E=4e-23; I=61%
16386_0	ABC transporter, ATP-binding protein	*Bacteria – Proteobacteria – Epsilonproteobacteria*	*Sulfurovum*	*Sulfurovum sp*. NBC37-1	E=1e-45; I=64%
16459_2	Transport ATP-binding protein msbA	*Bacteria – Proteobacteria – Epsilonproteobacteria*	*Sulfurovum*	*Sulfurovum sp*. NBC37-1	E=3e-108; I=58%
16683_1	Ammonium transporter	*Bacteria – Proteobacteria – Epsilonproteobacteria*	*Sulfurovum*	*Sulfurovum sp*. NBC37-1	E=7e-11; I=74%
17446_0	ABC transporter, ATP-binding protein	*Bacteria – Proteobacteria – Epsilonproteobacteria*	*Sulfurovum*	*Sulfurovum sp*. NBC37-1	E=7e-44; I=59%
17484_0	Molybdenum ABC transporter, ATP-binding protein	*Bacteria – Proteobacteria – Epsilonproteobacteria*	*Sulfurovum*	*Sulfurovum sp*. NBC37-1	E=1e-46; I=60%
18433_0	Iron compound ABC transporter, permease protein	*Bacteria – Proteobacteria – Epsilonproteobacteria*	*Sulfurimonas*	*Sulfurimonas denitrificans* DSM 1251	E=3e-8; I=47%
19535_1	Cobalt transport protein CbiM	*Bacteria – Proteobacteria – Epsilonproteobacteria*	*Sulfurimonas*	*Sulfurimonas denitrificans* DSM 1251	E=3e-61; I=62%
20121_4	ABC-transporter, ATP-binding protein	*Bacteria – Proteobacteria – Epsilonproteobacteria*	*Sulfurovum*	*Sulfurovum sp*. NBC37-1	E=2e-100; I=73%
20151_0	ABC-transporter, ATP-binding protein	*Bacteria – Proteobacteria – Epsilonproteobacteria*	*Arcobacter*	*Arcobacter nitrofigilis* DSM 7299	E=2e-27; I=53%
20287_0	ABC transporter, ATP-binding protein	*Bacteria – Proteobacteria – Epsilonproteobacteria*	*Sulfurovum*	*Sulfurovum sp*. NBC37-1	E=9e-34; I=78%
20445_0	Polysialic acid transport protein KpsD precursor	*Bacteria – Proteobacteria – Epsilonproteobacteria*	*Campylobacter*	*Campylobacter lari* RM2100	E=4e-16; I=62%
25002_0	Cation efflux system protein	*Bacteria – Proteobacteria – Epsilonproteobacteria*	*Campylobacter*	*Campylobacterales* bacterium GD 1	E=2e-53; I=54%
25044_0	Iron compound ABC transporter	*Bacteria – Proteobacteria – Epsilonproteobacteria*	*Arcobacter*	*Arcobacter nitrofigilis* DSM 7299	E=1e-26; I=49%
26966_1	Sulfate transporter family protein	*Bacteria – Proteobacteria – Epsilonproteobacteria*	*Sulfurovum*	*Sulfurovum sp*. NBC37-1	E=0.0; I=73%

28290_0	Cation efflux system protein	*Bacteria – Proteobacteria – Epsilonproteobacteria*	Unclassified	Campylobacterales bacterium GD 1	E=2e-56; I=56%
28976_1	Metal ion ABC transporter, permease	*Bacteria – Proteobacteria – Epsilonproteobacteria*	*Sulfurovum*	*Sulfurovum sp*. NBC37-1	E=2e-20; I=66%
29795_0	Metal ion ABC transporter, permease	*Bacteria – Proteobacteria – Epsilonproteobacteria*	*Sulfurovum*	*Sulfurovum sp*. NBC37-1	E=1e-35; I=84%
00914_1	Tellurite resistance protein TehB	*Bacteria – Proteobacteria – Epsilonproteobacteria*	*Sulfurovum*	*Sulfurovum sp*. NBC37-1	E=1e-54; I=57%
01073_0	Camphor resistance CrcB protein	*Bacteria – Proteobacteria – Epsilonproteobacteria*	*Sulfurovum*	*Sulfurovum sp*. NBC37-1	E=8e-16; I=51%
02353_0	Acriflavin resistance protein AcrB	*Bacteria – Proteobacteria – Epsilonproteobacteria*	*Sulfurovum*	*Sulfurovum sp*. NBC37-1	E=4e-29; I=66%
04130_0	Acriflavin resistance protein AcrB	*Bacteria – Proteobacteria – Epsilonproteobacteria*	*Sulfurovum*	*Sulfurovum sp*. NBC37-1	E=1e-38; I=54%
05554_0	Acriflavin resistance protein AcrB	*Bacteria – Proteobacteria – Epsilonproteobacteria*	*Sulfurovum*	*Sulfurovum sp*. NBC37-1	E=2e-55; I=86%
05823_1	Tellurite resistance protein TehB	*Bacteria – Proteobacteria – Epsilonproteobacteria*	*Sulfurovum*	*Sulfurovum sp*. NBC37-1	E=2e-34; I=51%
09152_2	Mitomycin resistance protein	*Bacteria – Proteobacteria – Epsilonproteobacteria*	*Sulfurovum*	*Sulfurovum sp*. NBC37-1	E=1e-28; I=63%
13096_1	Acriflavin resistance AcrA/AcrE protein family	*Bacteria – Proteobacteria – Epsilonproteobacteria*	Unclassified	*Alvinella pompejana* epibiont 7G3	E=8e-82; I=77%
14247_2	Acriflavin resistance protein AcrB	*Bacteria – Proteobacteria – Epsilonproteobacteria*	*Sulfurovum*	*Sulfurovum sp*. NBC37-1	E=5e-45; I=79%
27245_0	Acriflavin resistance protein AcrB	*Bacteria – Proteobacteria – Epsilonproteobacteria*	Unclassified	*Alvinella pompejana* epibiont 7G3	E=2e-25; I=85%
00454_5	Type IV pilin peptidase	*Bacteria – Proteobacteria – Epsilonproteobacteria*	*Sulfurovum*	*Sulfurovum sp*. NBC37-1	E=2e-35; I=57%
05258_0/1	Prepilin-type cleavage/methylation, N-terminal domain	*Bacteria – Proteobacteria – Epsilonproteobacteria*	*Sulfurovum*	*Sulfurovum sp*. NBC37-1	E=8e-22; I=39%
05626_2	Type IV pilin peptidase	*Bacteria – Proteobacteria – Epsilonproteobacteria*	*Sulfurovum*	*Sulfurovum sp*. NBC37-1	E=8e-42; I=65%
13420_1	Prepilin-type cleavage/methylation, N-terminal domain	*Bacteria – Proteobacteria – Epsilonproteobacteria*	*Sulfurovum*	*Sulfurovum sp*. NBC37-1	E=3e-33; I=33%
26905_0	Prepilin-type cleavage/methylation, N-terminal domain	*Bacteria – Proteobacteria – Epsilonproteobacteria*	*Campylobacter*	*Campylobacter hominis* ATCC BAA-381	E=2e-7; I=56%
00192_9	Invasion antigen B CiaB	*Bacteria – Proteobacteria – Epsilonproteobacteria*	*Sulfurovum*	*Sulfurovum sp*. NBC37-1	E=0.0; I=62%
00596_1	Surface antigen (D15)	*Bacteria – Proteobacteria – Epsilonproteobacteria*	*Sulfurovum*	*Sulfurovum sp*. NBC37-1	E=4e-106; I=38%
00610_2	O-antigen transporter	*Bacteria – Proteobacteria – Epsilonproteobacteria*	*Caminibacter*	*Caminibacter mediatlanticus* TB-2	E=4e-69; I=45%
01687_0	Bacterial surface antigen (D15)	*Bacteria – Proteobacteria – Epsilonproteobacteria*	*Sulfurovum*	*Sulfurovum sp*. NBC37-1	E=4e-78; I=42%
02088_0	Type II secretion system protein E	*Bacteria – Proteobacteria – Epsilonproteobacteria*	*Sulfurovum*	*Sulfurovum sp*. NBC37-1	E=2e-42; I=53%
09820_2	Surface antigen (D15)	*Bacteria – Proteobacteria – Epsilonproteobacteria*	*Sulfurovum*	*Sulfurovum sp*. NBC37-1	E=6e-42; I=28%
10220_1	Surface antigen (D15)	*Bacteria – Proteobacteria – Epsilonproteobacteria*	*Sulfurovum*	*Sulfurovum sp*. NBC37-1	E=1e-33; I=35%
28233_3	Acid membrane antigen A	*Bacteria – Proteobacteria – Epsilonproteobacteria*	*Sulfurovum*	*Sulfurovum sp*. NBC37-1	E=6e-95; I=29%
02741_0	Type IV secretion system protein	*Bacteria – Proteobacteria – Epsilonproteobacteria*	*Campylobacter*	*Campylobacter coli* RM2228	E=4e-6; I=56%
03571_4/10	Type II secretion system gspD	*Bacteria – Proteobacteria – Epsilonproteobacteria*	*Sulfurovum*	*Sulfurovum sp*. NBC37-1	E=0.0; I=58%
04763_1	Type II and III secretion system protein	*Bacteria – Proteobacteria – Epsilonproteobacteria*	*Sulfurovum*	*Sulfurovum sp*. NBC37-1	E=1e-24; I=48%
04886_4	Type II secretion system protein F	*Bacteria – Proteobacteria – Epsilonproteobacteria*	*Sulfurovum*	*Sulfurovum sp*. NBC37-1	E=6e-9; I=73%
05258_2	Type II secretion system protein G	*Bacteria – Proteobacteria – Epsilonproteobacteria*	*Sulfurovum*	*Sulfurovum sp*. NBC37-1	E=4e-27; I=53%
06396_4	Type II secretion system protein F	*Bacteria – Proteobacteria – Epsilonproteobacteria*	*Sulfurovum*	*Sulfurovum sp*. NBC37-1	E=5e-144; I=65%
09618_4	Type II secretion system protein F	*Bacteria – Proteobacteria – Epsilonproteobacteria*	*Sulfurovum*	*Sulfurovum sp*. NBC37-1	E=3e-156; I=67%
09618_8	Type II secretion system protein D	*Bacteria – Proteobacteria – Epsilonproteobacteria*	*Sulfurovum*	*Sulfurovum sp*. NBC37-1	E=7e-44; I=63%
11318_1	Type II secretion system protein F	*Bacteria – Proteobacteria – Epsilonproteobacteria*	*Sulfurovum*	*Sulfurovum sp*. NBC37-1	E=1e-26; I=56%
11334_2	Type II secretion system protein E	*Bacteria – Proteobacteria – Epsilonproteobacteria*	*Sulfurovum*	*Sulfurovum sp*. NBC37-1	E=4e-17; I=74%
00905_4	Exopolysaccharide biosynthesis	*Bacteria – Proteobacteria – Epsilonproteobacteria*	*Caminibacter*	*Caminibacter mediatlanticus* TB-2	E=3e-63; I=65%
09338_0	Inorganic polyphosphate kinase ppnK	*Bacteria – Proteobacteria – Gammaproteobacteria*	*Nitrococcus*	*Nitrococcus mobilis* Nb-23	E=4e-23; I=54%
23609_3	Polyphosphate kinase 2	*Bacteria – Proteobacteria – Gammaproteobacteria*	Unclassified	*Candidatus Vesicomyosocius okutanii* HA	E=1e-106; I=64%
03104_3	Dinucleoside polyphosphate hydrolase	*Bacteria – Proteobacteria – Gammaproteobacteria*	*Shewanella*	*Shewanella violacea* DSS12	E=4e-54; I=61%
02555_0	Exopolyphosphatase	*Bacteria – Proteobacteria – Gammaproteobacteria*	*Nitrosococcus*	*Nitrosococcus oceani* ATCC 19707	E=5e-57; I=43%
13663_2	Ferrochelatase	*Bacteria – Proteobacteria – Gammaproteobacteria*	*Allochromatium*	*Allochromatium vinosum* DSM 180	E=8e-108; I=58%

14815_0	Ferrochelatase	*Bacteria – Proteobacteria – Gammaproteobacteria*	*Thioalkalivibrio*	*Thioalkalivibrio sp.* K90mix	E=7e-55; I=57%
18905_0	Ferrochelatase	*Bacteria – Proteobacteria – Gammaproteobacteria*	*Congregibacter*	*Congregibacter litoralis* KT71	E=6e-25; I=63%
00240_0	ABC transporter, ATP-binding protein	*Bacteria – Proteobacteria – Gammaproteobacteria*	*Methylovulum*	*Methylovulum miyakonense* HT12	E=0.0; I=61%
00264_1	Co/Zn/Cd cation transporter	*Bacteria – Proteobacteria – Gammaproteobacteria*	*Psychromonas*	*Psychromonas ingrahamii* 37	E=5e-112; I=63%
00465_1	DMT superfamily drug/metabolite transporter	*Bacteria – Proteobacteria – Gammaproteobacteria*	*Neptuniibacter*	*Neptuniibacter caesariensis*	E=6e-35; I=43%
00469_1	Molybdenum transport system permease protein modB	*Bacteria – Proteobacteria – Gammaproteobacteria*	*Beggiatoa*	*Beggiatoa sp.* PS	E=1e-77; I=76%
00759_9	Mg^{2+} and Co^{2+} transporter CorB	*Bacteria – Proteobacteria – Gammaproteobacteria*	*Thioalkalivibrio*	*Thioalkalivibrio sulfidophilus* HL-EbGr7	E=4e-109; I=56%
01009_0	Amino acid transporter LysE	*Bacteria – Proteobacteria – Gammaproteobacteria*	*Colwellia*	*Colwellia psychrerythraea* 34H	E=4e-58; I=67%
01152_0	TRAP transporter	*Bacteria – Proteobacteria – Gammaproteobacteria*	*Marinomonas*	*Marinomonas sp.* MWYL1	E=2e-117; I=63%
01152_3	TRAP transporter	*Bacteria – Proteobacteria – Gammaproteobacteria*	*Marinomonas*	*Marinomonas sp.* MWYL1	E=2e-155; I=71%
01152_6	Thiamine transporter membrane protein	*Bacteria – Proteobacteria – Gammaproteobacteria*	*Vibrio*	*Vibrio parahaemolyticus* AN-5034	E=5e-31; I=41%
01155_2	ABC transporter, ATP-binding protein	*Bacteria – Proteobacteria – Gammaproteobacteria*	*Photobacterium*	*Photobacterium profundum* SS9	E=1e-59; I=47%
01263_3	TonB system transport protein ExbD type-2	*Bacteria – Proteobacteria – Gammaproteobacteria*	*Halothiobacillus*	*Halothiobacillus neapolitanus* c2	E=1e-19; I=52%
01274_0	Sugar ABC transporter, permease protein	*Bacteria – Proteobacteria – Gammaproteobacteria*	*Psychromonas*	*Psychromonas ingrahamii* 37	E=1e-28; I=65%
01280_5	Membrane-fusion protein	*Bacteria – Proteobacteria – Gammaproteobacteria*	*Saccharophagus*	*Saccharophagus degradans* 2-40	E=1e-43; I=32%
01280_6	Multidrug resistance protein (AcrB/AcrD/AcrF family)	*Bacteria – Proteobacteria – Gammaproteobacteria*	*Colwellia*	*Colwellia psychrerythraea* 34H	E=2e-144; I=56%
01381_0	TRAP-type C4-dicarboxylate transporter	*Bacteria – Proteobacteria – Gammaproteobacteria*	*Vibrio*	*Vibrio vulnificus* YJ016	E=3e-69; I=74%
01381_2	TRAP-type C4-dicarboxylate transporter	*Bacteria – Proteobacteria – Gammaproteobacteria*	*Neptuniibacter*	*Neptuniibacter caesariensis*	E=1e-107; I=60%
01403_6/7	TRAP transporter, 4TM/12TM fusion protein	*Bacteria – Proteobacteria – Gammaproteobacteria*	*Moritella*	*Moritella sp.* PE36	E=8e-136; I=57%
01503_0	Molybdenum transport system permease protein modB	*Bacteria – Proteobacteria – Gammaproteobacteria*	*Allochromatium*	*Allochromatium vinosum* DSM 180	E=3e-51; I=52%
01730_3/4	Phosphonate ABC transporter	*Bacteria – Proteobacteria – Gammaproteobacteria*	*Marinobacter*	*Marinobacter sp.* ELB17	E=2e-95; I=71%
01764_1	ABC transporter, ATP-binding protein	*Bacteria – Proteobacteria – Gammaproteobacteria*	*Grimontia*	*Grimontia hollisae* CIP 101886	E=7e-124; I=62%
01764_2	ABC transporter permease	*Bacteria – Proteobacteria – Gammaproteobacteria*	*Grimontia*	*Grimontia hollisae* CIP 101886	E=3e-128; I=67%
01897_0	Acriflavin resistance protein	*Bacteria – Proteobacteria – Gammaproteobacteria*	*Thioalkalivibrio*	*Thioalkalivibrio sulfidophilus* HL-EbGr7	E=0.0; I=66%
01897_1	Efflux transporter, RND family	*Bacteria – Proteobacteria – Gammaproteobacteria*	*Thioalkalivibrio*	*Thioalkalivibrio sp.* K90mix	E=6e-70; I=41%
01972_0	Amino acid ABC transporter, permease protein	*Bacteria – Proteobacteria – Gammaproteobacteria*	*Vibrio*	*Vibrio cholerae* MZO-2	E=7e-30; I=64%
02222_3	Transport protein (ABC superfamily, membrane)	*Bacteria – Proteobacteria – Gammaproteobacteria*	*Vibrio*	*Vibrio mimicus* MB451	E=7e-79; I=63%
02328_2	Copper transporter, P-type ATPase	*Bacteria – Proteobacteria – Gammaproteobacteria*	*Vibrio*	*Vibrio parahaemolyticus* AQ3810	E=1e-16; I=39%
02462_0	Amino acid transporter, LysE family	*Bacteria – Proteobacteria – Gammaproteobacteria*	*Pseudomonas*	*Pseudomonas stutzeri* A1501	E=1e-47; I=59%
02505_0	Cation transporting ATPase	*Bacteria – Proteobacteria – Gammaproteobacteria*	*Beggiatoa*	*Beggiatoa sp.* PS	E=3e-31; I=52%
02547_0	Electron transport protein SCO1/ SenC	*Bacteria – Proteobacteria – Gammaproteobacteria*	Unclassified	Uncultured SUP05 cluster bacterium	E=3e-13; I=58%
02965_1	Glycerol-3-phosphate transporter ATP-binding subunit	*Bacteria – Proteobacteria – Gammaproteobacteria*	*Marinomonas*	*Marinomonas sp.* MWYL1	E=7e-105; I=71%
03104_11	Electron transport protein DsrJ	*Bacteria – Proteobacteria – Gammaproteobacteria*	*Beggiatoa*	*Beggiatoa sp.* PS	E=7e-16; I=47%
03254_0	Copper transporter	*Bacteria – Proteobacteria – Gammaproteobacteria*	*Shewanella*	*Shewanella putrefaciens* CN-32	E=2e-101; I=54%
03480_3/4	Potassium transporter peripheral membrane component	*Bacteria – Proteobacteria – Gammaproteobacteria*	*Thioalkalivibrio*	*Thioalkalivibrio sulfidophilus* HL-EbGr7	E=3e-115; I=62%
03480_6	TrkH family potassium uptake protein	*Bacteria – Proteobacteria – Gammaproteobacteria*	*Thioalkalivibrio*	*Thioalkalivibrio sulfidophilus* HL-EbGr7	E=6e-25; I=72%
03480_8	Zinc ABC transporter, permease protein	*Bacteria – Proteobacteria – Gammaproteobacteria*	*Thioalkalivibrio*	*Thioalkalivibrio sp.* K90mix	E=5e-66; I=56%
03771_1	ABC transporter, ATP-binding protein	*Bacteria – Proteobacteria – Gammaproteobacteria*	*Thioalkalivibrio*	*Thioalkalivibrio sulfidophilus* HL-EbGr7	E=2e-93; I=68%
03887_2	Peptide ABC transporter, permease protein	*Bacteria – Proteobacteria – Gammaproteobacteria*	*Thiomicrospira*	*Thiomicrospira crunogena* XCL-2	E=4e-115; I=85%
04020_0	Biopolymer transport protein ExbD/TolR	*Bacteria – Proteobacteria – Gammaproteobacteria*	*Nitrosococcus*	*Nitrosococcus halophilus* Nc4	E=5e-17; I=59%
04020_1	MotA/TolQ/ExbB proton channel	*Bacteria – Proteobacteria – Gammaproteobacteria*	*Nitrosococcus*	*Nitrosococcus halophilus* Nc4	E=1e-24; I=47%
04384_0	Outer membrane receptor for transport of vitamin B	*Bacteria – Proteobacteria – Gammaproteobacteria*	*Xanthomonas*	*Xanthomonas oryzae pv. oryzae* KACC 10331	E=4e-71; I=38%

04540_0	Phosphate ABC transporter, inner membrane subuni t PstC	*Bacteria – Proteobacteria – Gammaproteobacteria*	*Pseudoalteromonas*	*Pseudoalteromonas atlantica* T6c	E=3e-139; I=66%
04540_1	Phosphate ABC transporter, permease protein PstA	*Bacteria – Proteobacteria – Gammaproteobacteria*	Unclassified	*gamma proteobacterium* HTCC5015	E=1e-150; I=65%
04540_2	Phosphate ABC transporter, ATPase subunit	*Bacteria – Proteobacteria – Gammaproteobacteria*	*Teredinibacter*	*Teredinibacter turnerae* T7901	E=1e-117; I=77%
04540_5	Phosphate ABC transporter, periplasmic phosphate-binding protein	*Bacteria – Proteobacteria – Gammaproteobacteria*	*Saccharophagus*	*Saccharophagus degradans* 2-40	E=5e-25; I=74%
04696_0	Heavy metal transport/detoxification protein	*Bacteria – Proteobacteria – Gammaproteobacteria*	Unclassified	Unknown	Unknown
04774_0	Phosphate ABC transporter periplasmic binding protein	*Bacteria – Proteobacteria – Gammaproteobacteria*	Unclassified	*gamma proteobacterium* HTCC2207	E=3e-84; I=69%
05124_0	Lysophospholipid transporter	*Bacteria – Proteobacteria – Gammaproteobacteria*	*Methylococcus*	*Methylococcus capsulatus* str. Bath	E=6e-15; I=51%
05577_12	ABC transporter, inner membrane subunit	*Bacteria – Proteobacteria – Gammaproteobacteria*	*Nitrosococcus*	*Nitrosococcus halophilus* Nc4	E=4e-102; I=58%
05577_13/14	ABC transporter, inner membrane subunit	*Bacteria – Proteobacteria – Gammaproteobacteria*	*Thioalkalivibrio*	*Thioalkalivibrio sulfidophilus* HL-EbGr7	E=2e-98; I=78%
05577_15	ABC transporter, ATP-binding protein	*Bacteria – Proteobacteria – Gammaproteobacteria*	*Nitrosococcus*	*Nitrosococcus watsonii* C-113	E=0.0; I=63%
05724_1	ABC trasnporter, ATP-binding protein	*Bacteria – Proteobacteria – Gammaproteobacteria*	*Methylococcus*	*Methylococcus capsulatus* str. Bath	E=5e-34; I=54%
06023_0	MotA/TolQ/ExbB proton channel	*Bacteria – Proteobacteria – Gammaproteobacteria*	*Thioalkalivibrio*	*Thioalkalivibrio sulfidophilus* HL-EbGr7	E=4e-39; I=53%
06023_1	Biopolymer transport protein ExbD/TolR	*Bacteria – Proteobacteria – Gammaproteobacteria*	*Thioalkalivibrio*	*Thioalkalivibrio sulfidophilus* HL-EbGr7	E=2e-27; I=54%
08756_0	Peptide ABC transporter, permease protein	*Bacteria – Proteobacteria – Gammaproteobacteria*	*Hahella*	*Hahella chejuensis* KCTC 2396	E=4e-91; I=56%
08781_1	TRAP dicarboxylate trasnporter, DctP subunit	*Bacteria – Proteobacteria – Gammaproteobacteria*	*Marinomonas*	*Marinomonas sp.* MWYL1	E=6e-79; I=48%
08851_0	TRAP dicarboxylate transporter, DctP subunit	*Bacteria – Proteobacteria – Gammaproteobacteria*	*Hahella*	*Hahella chejuensis* KCTC 2396	E=5e-115; I=65%
08870_0	Glutamate/aspartate ABC transporter, ATP-binding protein	*Bacteria – Proteobacteria – Gammaproteobacteria*	*Photorhabdus*	*Photorhabdus luminescens ssp. laumondii*	E=2e-48; I=84%
08872_1	Electron transport complex, RnfB	*Bacteria – Proteobacteria – Gammaproteobacteria*	*Teredinibacter*	*Teredinibacter turnerae* T7901	E=7e-51; I=70%
08872_2/3	Electron transport complex, RnfABCDGE type, C subunit	*Bacteria – Proteobacteria – Gammaproteobacteria*	*Thioalkalivibrio*	*Thioalkalivibrio sulfidophilus* HL-EbGr7	E=2e-95; I=53%
08872_4	Electron transport complex, RnfABCDGE type, D subunit	*Bacteria – Proteobacteria – Gammaproteobacteria*	*Thioalkalivibrio*	*Thioalkalivibrio sulfidophilus* HL-EbGr7	E=2e-40; I=51%
09004_5	TRAP-type C4-dicarboxylate transporter	*Bacteria – Proteobacteria – Gammaproteobacteria*	*Hahella*	*Hahella chejuensis* KCTC 2396	E=7e-65; I=70%
09059_0	C4-dicarboxylate transport system	*Bacteria – Proteobacteria – Gammaproteobacteria*	*Neptuniibacter*	*Neptuniibacter caesariensis*	E=0.0; I=78%
09059_2	Tripartite ATP-independent periplasmic transporter	*Bacteria – Proteobacteria – Gammaproteobacteria*	*Neptuniibacter*	*Neptuniibacter caesariensis*	E=4e-16; I=65%
09059_3	TRAP transporter solute receptor, DctP family protein	*Bacteria – Proteobacteria – Gammaproteobacteria*	*Neptuniibacter*	*Neptuniibacter caesariensis*	E=2e-135; I=75%
09103_2	Binding-protein-dependent transport systems	*Bacteria – Proteobacteria – Gammaproteobacteria*	*Marinobacter*	*Marinobacter aquaeolei* VT8	E=3e-62; I=59%
09103_4	Phosphonate ABC transporter, ATPase subunit	*Bacteria – Proteobacteria – Gammaproteobacteria*	*Marinobacter*	*Marinobacter aquaeolei* VT8	E=5e-52; I=48%
09103_5	ABC phosphonate transporter periplasmic	*Bacteria – Proteobacteria – Gammaproteobacteria*	*Photobacterium*	*Photobacterium profundum* SS9	E=4e-94; I=63%
09109_0	Thiamin transporter, substrate binding subunit	*Bacteria – Proteobacteria – Gammaproteobacteria*	*Yersinia*	*Yersinia pseudotuberculosis* YPIII	E=4e-106; I=61%
09187_1	ABC-type multidrug transport system, permease component	*Bacteria – Proteobacteria – Gammaproteobacteria*	*Photobacterium*	*Photobacterium sp.* SKA34	E=6e-96; I=66%
09187_2	ATP-binding component of ABC transporter	*Bacteria – Proteobacteria – Gammaproteobacteria*	*Shewanella*	*Shewanella amazonensis* SB2B	E=3e-102; I=68%
09211_1	Sulfate transport protein CysZ	*Bacteria – Proteobacteria – Gammaproteobacteria*	*Nitrosococcus*	*Nitrosococcus halophilus* Nc4	E=6e-22; I=50%
09211_2	Sulfate transport protein CysZ	*Bacteria – Proteobacteria – Gammaproteobacteria*	*Allochromatium*	*Allochromatium vinosum* DSM 180	E=1e-6; I=49%
09214_0	ABC transporter, permease protein	*Bacteria – Proteobacteria – Gammaproteobacteria*	*Methylococcus*	*Methylococcus capsulatus* str. Bath	E=2e-31; I=48%
09237_0	Multidrug resistance ABC transporter ATP binding protein	*Bacteria – Proteobacteria – Gammaproteobacteria*	Unclassified	candidate division TM7 *genomosp.* GTL1	E=2e-106; I=45%
09276_5	Transporter	*Bacteria – Proteobacteria – Gammaproteobacteria*	*Colwellia*	*Colwellia psychrerythraea* 34H	E=3e-10; I=35%
09298_0	Ferrous iron transport protein B FeoB	*Bacteria – Proteobacteria – Gammaproteobacteria*	Unclassified	Uncultured SUP05 cluster bacterium	E=2e-53; I=52%
09304_2	Ferrous iron transport protein B FeoB	*Bacteria – Proteobacteria – Gammaproteobacteria*	*Allochromatium*	*Allochromatium vinosum* DSM 180	E=6e-44; I=69%
09325_0	TRAP dicarboxylate transporter DctM subunit	*Bacteria – Proteobacteria – Gammaproteobacteria*	*Reinekea*	*Reinekea blandensis* MED297	E=0.0; I=72%
09334_0	Biopolymer transport protein ExbD/TolR	*Bacteria – Proteobacteria – Gammaproteobacteria*	*Nitrosococcus*	*Nitrosococcus halophilus* Nc4	E=3e-24; I=45%
09408_1	Magnesium transporter	*Bacteria – Proteobacteria – Gammaproteobacteria*	*Marinomonas*	*Marinomonas sp.* MED121	E=9e-78; I=54%
09413_0	Amino acid ABC transporter, ATP-binding protein	*Bacteria – Proteobacteria – Gammaproteobacteria*	*Vibrio*	*Vibrio vulnificus* CMCP6	E=9e-61; I=79%
09414_0	Amino acid ABC transporter, ATP-binding protein	*Bacteria – Proteobacteria – Gammaproteobacteria*	*Neptuniibacter*	*Neptuniibacter caesariensis*	E=2e-32; I=85%

09584_2	ABC transporter, urea, ATP-binding protein, UrtE	*Bacteria – Proteobacteria – Gammaproteobacteria*	*Vibrio*	*Vibrio shilonii* AK1	E=1e-70; I=72%
09615_0	ABC-type tungstate transport system, periplasmic component	*Bacteria – Proteobacteria – Gammaproteobacteria*	*Neptuniibacter*	*Neptuniibacter caesariensis*	E=2e-65; I=61%
09615_1	ABC transporter, ATP-binding protein	*Bacteria – Proteobacteria – Gammaproteobacteria*	*Shewanella*	*Shewanella piezotolerans* WP3	E=2e-26; I=44%
09641_4	Transport ATP-binding protein msbA	*Bacteria – Proteobacteria – Gammaproteobacteria*	*Thioalkalivibrio*	*Thioalkalivibrio sulfidophilus* HL-EbGr7	E=2e-179; I=52%
09801_0	Sugar ABC transporter, permease protein	*Bacteria – Proteobacteria – Gammaproteobacteria*	*Ahrensia*	*Ahrensia sp.* R2A130	E=1e-22; I=74%
09830_2	ABC transporter, ATP-binding protein	*Bacteria – Proteobacteria – Gammaproteobacteria*	*Methylophaga*	*Methylophaga thiooxydans* DMS010	E=0.0; I=58%
09830_5	Cell division ABC transporter, permease protein FtsX	*Bacteria – Proteobacteria – Gammaproteobacteria*	Unclassified	*gamma proteobacterium* HTCC5015	E=3e-53; I=46%
09888_0	TRAP dicarboxylate transporter, DctM subunit	*Bacteria – Proteobacteria – Gammaproteobacteria*	*Grimontia*	*Grimontia hollisae* CIP 101886	E=1e-36; I=70%
09889_0	TRAP dicarboxylate transporter, DctM subunit	*Bacteria – Proteobacteria – Gammaproteobacteria*	*Marinomonas*	*Marinomonas sp.* MWYL1	E=1e-84; I=51%
09917_2	ABC-type tungstate transport system, permease component	*Bacteria – Proteobacteria – Gammaproteobacteria*	*Vibrio*	*Vibrio furnissii* CIP 102972	E=6e-73; I=57%
09928_6	Sulfate permease family protein	*Bacteria – Proteobacteria – Gammaproteobacteria*	*Methylophaga*	*Methylophaga thiooxydans* DMS010	E=1e-176; I=65%
10384_3	Capsular polysaccharide transport protein	*Bacteria – Proteobacteria – Gammaproteobacteria*	*Idiomarina*	*Idiomarina baltica* OS145	E=9e-20; I=58%
10596_0	TRAP dicarboxylate transporter DctP subunit	*Bacteria – Proteobacteria – Gammaproteobacteria*	*Reinekea*	*Reinekea blandensis* MED297	E=6e-24; I=68%
10602_0	Transport-associated OB	*Bacteria – Proteobacteria – Gammaproteobacteria*	*Agrobacterium*	*Agrobacterium tumefaciens* str. C58	E=3e-18; I=38%
10826_0	Sulfate anion transporter	*Bacteria – Proteobacteria – Gammaproteobacteria*	*Beggiatoa*	*Beggiatoa sp.* PS	E=2e-137; I=58%
10978_1	Ferric iron-catecholate outer membrane transporter	*Bacteria – Proteobacteria – Gammaproteobacteria*	Unclassified	*gamma proteobacterium* HdN1	E=8e-108; I=39%
11111_0	urea ABC transporter, ATP-binding protein UrtD	*Bacteria – Proteobacteria – Gammaproteobacteria*	*Moritella*	*Moritella sp.* PE36	E=6e-104; I=70%
11111_1	Urea ABC transporter, permease protein UrtC	*Bacteria – Proteobacteria – Gammaproteobacteria*	*Cellvibrio*	*Cellvibrio japonicus* Ueda107	E=2e-138; I=75%
11111_2	Urea ABC transporter, permease protein UrtB	*Bacteria – Proteobacteria – Gammaproteobacteria*	*Vibrio*	*Vibrio coralliilyticus* ATCC BAA-450	E=3e-162; I=61%
11146_0	Transporter	*Bacteria – Proteobacteria – Gammaproteobacteria*	*Magnetococcus*	*Magnetococcus sp.* MC-1	E=2e-65; I=80%
11223_0/1	Ammonium transporter	*Bacteria – Proteobacteria – Gammaproteobacteria*	*Desulfonatronospira*	*Desulfonatronospira thiodismutans* ASO3-1	E=2e-61; I=74%
11386_3	Arginine transporter protein ArgO	*Bacteria – Proteobacteria – Gammaproteobacteria*	*Acinetobacter*	*Acinetobacter radioresistens* SK82	E=1e-22; I=51%
11504_2	Cation/multidrug efflux pump	*Bacteria – Proteobacteria – Gammaproteobacteria*	*Hahella*	*Hahella chejuensis* KCTC 2396	E=0.0; I=49%
11660_0	ABC transporter membrane, ATP binding component	*Bacteria – Proteobacteria – Gammaproteobacteria*	*Thiomonas*	*Thiomonas intermedia* K12	E=1e-53; I=56%
11835_2	Thiamine ABC transporter, permease protein	*Bacteria – Proteobacteria – Gammaproteobacteria*	*Vibrio*	*Vibrio mimicus* MB451	E=5e-18; I=48%
11883_0	Amino acid transporter, inner membrane protein	*Bacteria – Proteobacteria – Gammaproteobacteria*	*Marinomonas*	*Marinomonas sp.* MWYL1	E=4e-12; I=43%
12408_3	Electron transport protein SCO1/SenC	*Bacteria – Proteobacteria – Gammaproteobacteria*	*Neptuniibacter*	*Neptuniibacter caesariensis*	E=3e-46; I=53%
12780_0	Electron transport complex, Rnf/Nqr	*Bacteria – Proteobacteria – Gammaproteobacteria*	*Neptuniibacter*	*Neptuniibacter caesariensis*	E=1e-13; I=59%
12780_1	Electron transport complex, RnfG	*Bacteria – Proteobacteria – Gammaproteobacteria*	Unclassified	marine *gamma proteobacterium* HTCC2148	E=1e-42; I=47%
12787_0	Electron transport complex, Rnf/Nqr	*Bacteria – Proteobacteria – Gammaproteobacteria*	*Marinomonas*	*Marinomonas sp.* MWYL1	E=2e-25; I=69%
13212_8/9	Nitrate/nitrite transporter	*Bacteria – Proteobacteria – Gammaproteobacteria*	Unclassified	uncultured SUP05 cluster *bacterium*	E=1e-167; I=74%
13213_0	Sugar ABC transporter, ATP-binding protein	*Bacteria – Proteobacteria – Gammaproteobacteria*	*Marinomonas*	*Marinomonas sp.* MED121	E=2e-110; I=60%
13213_1	Sugar ABC transporter, permease protein	*Bacteria – Proteobacteria – Gammaproteobacteria*	*Marinomonas*	*Marinomonas sp.* MWYL1	E=1e-104; I=64%
13214_0	Sugar ABC transporter, ATP-binding protein	*Bacteria – Proteobacteria – Gammaproteobacteria*	*Agrobacterium*	*Agrobacterium radiobacter* K84	E=1e-66; I=72%
13215_0	Sugar ABC transporter ATP binding protein	*Bacteria – Proteobacteria – Gammaproteobacteria*	*Agrobacterium*	*Agrobacterium radiobacter* K84	E=3e-48; I=71%
13282_0	Sugar ABC transporter, periplasmic sugar-binding protein	*Bacteria – Proteobacteria – Gammaproteobacteria*	*Leptothrix*	*Leptothrix cholodnii* SP-6	E=2e-117; I=66%
13373_0	protein containing ABC transporter-like domain	*Bacteria – Proteobacteria – Gammaproteobacteria*	*Vibrio*	*Vibrio vulnificus* CMCP6	E=3e-58; I=54%
13410_0	Histidine/lysine/arginine/ornithine transporter	*Bacteria – Proteobacteria – Gammaproteobacteria*	*Moritella*	*Moritella sp.* PE36	E=7e-31; I=57%
13415_0	Sugar ABC transporter, ATP-binding protein	*Bacteria – Proteobacteria – Gammaproteobacteria*	Unclassified	*Rhodobacteraceae bacterium* HTCC2083	E=5e-85; I=53%
13415_4	Sugar ABC transporter, ATP-binding protein	*Bacteria – Proteobacteria – Gammaproteobacteria*	*Azospirillum*	*Azospirillum sp.* B510	E=9e-53; I=65%
13602_0	Cation transporting ATPase	*Bacteria – Proteobacteria – Gammaproteobacteria*	*Beggiatoa*	*Beggiatoa sp.* PS	E=8e-61; I=61%
13796_0	Nucleoside/H^+ symporter, major facilitator superfamily	*Bacteria – Proteobacteria – Gammaproteobacteria*	*Nitrococcus*	*Nitrococcus mobilis* Nb-23	E=1e-20; I=53%

13797_1	Nucleoside/H^+ symporter, major facilitator superfamily	*Bacteria – Proteobacteria – Gammaproteobacteria*	Unclassified	*Endoriftia* persephone 'Hot96_1+Hot96_2'	E=1e-13; I=51%
13891_5	Transporter, major facilitator family	*Bacteria – Proteobacteria – Gammaproteobacteria*	*Thioalkalivibrio*	*Thioalkalivibrio sulfidophilus* HL-EbGr7	E=3e-52; I=54%
13891_6	Transport protein	*Bacteria – Proteobacteria – Gammaproteobacteria*	*Thioalkalivibrio*	*Thioalkalivibrio sulfidophilus* HL-EbGr8	E=7e-15; I=50%
13893_1	ABC-type phosphate transport system, periplasmic component	*Bacteria – Proteobacteria – Gammaproteobacteria*	*Nitrosococcus*	*Nitrosococcus oceani* ATCC 19707	E=2e-79; I=55%
13911_1	Sulphate transporter	*Bacteria – Proteobacteria – Gammaproteobacteria*	Unclassified	Uncultured marine *bacterium* 578	E=4e-113; I=46%
14013_0	Glyceine betaine transporter membrane protein	*Bacteria – Proteobacteria – Gammaproteobacteria*	*Arcobacter*	*Arcobacter nitrofigilis* DSM 7299	E=7e-16; I=57%
14193_0	Efflux transporter, RND family	*Bacteria – Proteobacteria – Gammaproteobacteria*	*Shewanella*	*Shewanella sp.* MR-7	E=2e-11; I=27%
14210_0	Transport ATP-binding protein CydC	*Bacteria – Proteobacteria – Gammaproteobacteria*	*Allochromatium*	*Allochromatium vinosum* DSM 180	E=2e-59; I=46%
14692_0/1/2	Formate/nitrite transporter	*Bacteria – Proteobacteria – Gammaproteobacteria*	*Psychromonas*	*Psychromonas ingrahamii* 37	E=1e-25; I=90%
15118_5	MIT, microtubule interacting and transport	*Bacteria – Proteobacteria – Gammaproteobacteria*	*Marinobacter*	*Marinobacter aquaeolei* VT8	E=2e-58; I=67%
15202_0	Ammonium transporter	*Bacteria – Proteobacteria – Gammaproteobacteria*	Unclassified	Uncultured SUP05 cluster *bacterium*	E=1e-162; I=80%
15401_2	Sulfate transporter family protein	*Bacteria – Proteobacteria – Gammaproteobacteria*	*Grimontia*	*Grimontia hollisae* CIP 101886	E=1e-168; I=69%
15565_0	ABC transporter, ATP-binding/permease protein	*Bacteria – Proteobacteria – Gammaproteobacteria*	*Nitrosococcus*	*Nitrosococcus halophilus* Nc4	E=1e-132; I=55%
15643_1/2	Urea/short-chain amide ABC transporter, periplasmic protein	*Bacteria – Proteobacteria – Gammaproteobacteria*	*Arcobacter*	*Arcobacter nitrofigilis* DSM 7299	E=2e-130; I=81%
16227_3	ABC transporter, ATP-binding protein	*Bacteria – Proteobacteria – Gammaproteobacteria*	*Thiomicrospira*	*Thiomicrospira crunogena* XCL-2	E=0.0; I=74%
16340_1	Magnesium transporter	*Bacteria – Proteobacteria – Gammaproteobacteria*	*Xanthomonas*	*Xanthomonas campestris pv. Vasculorum*	E=8e-18; I=41%
17365_0	TRAP C4-dicarboxylate transport system permease DctM	*Bacteria – Proteobacteria – Gammaproteobacteria*	*Neptuniibacter*	*Neptuniibacter caesariensis*	E=6e-29; I=74%
17581_3	RND family efflux transporter	*Bacteria – Proteobacteria – Gammaproteobacteria*	Unclassified	*gamma proteobacterium* HTCC5015	E=4e-46; I=66%
17581_4	RND family efflux transporter	*Bacteria – Proteobacteria – Gammaproteobacteria*	*Moritella*	*Moritella sp.* PE36	E=7e-76; I=46%
18834_0	MFS family major facilitator transporter	*Bacteria – Proteobacteria – Gammaproteobacteria*	*Mannheimia*	*Mannheimia haemolytica* PHL213	E=1e-14; I=79%
19205_2	P-type cation-transporting ATPase	*Bacteria – Proteobacteria – Gammaproteobacteria*	*Thioalkalivibrio*	*Thioalkalivibrio sulfidophilus* HL-EbGr7	E=0.0; I=55%
20433_7	ABC-type Fe^{3+} transport system, periplasmic component	*Bacteria – Proteobacteria – Gammaproteobacteria*	*Neptuniibacter*	*Neptuniibacter caesariensis*	E=2e-74; I=44%
24388_4	Transporter, major facilitator family	*Bacteria – Proteobacteria – Gammaproteobacteria*	*Beggiatoa*	*Beggiatoa sp.* PS	E=3e-6; I=54%
24732_0	Ion transport protein	*Bacteria – Proteobacteria – Gammaproteobacteria*	*Thiomicrospira*	*Thiomicrospira crunogena* XCL-2	E=1e-27; I=46%
24983_0	Polysialic acid transport protein KpsD precursor	*Bacteria – Proteobacteria – Gammaproteobacteria*	*Pseudoalteromonas*	*Pseudoalteromonas atlantica* T6c	E=3e-34; I=51%
26680_5	Electron transport protein DsrK	*Bacteria – Proteobacteria – Gammaproteobacteria*	*Allochromatium*	*Allochromatium vinosum* DSM 180	E=0.0; I=68%
28213_3	Phosphate ABC transporter, ATPase subunit	*Bacteria – Proteobacteria – Gammaproteobacteria*	*Serratia*	*Serratia odorifera* 4Rx13	E=2e-41; I=38%
01203_0/1	Copper resistance protein	*Bacteria – Proteobacteria – Gammaproteobacteria*	*Shewanella*	Shewanella denitrificans OS217	E=1e-65; I=65%
03000_0	Acriflavin resistance protein AcrB	*Bacteria – Proteobacteria – Gammaproteobacteria*	*Neptuniibacter*	*Neptuniibacter caesariensis*	E=6e-47; I=68%
03044_4/5	Multiple antibiotic resistance (MarC)-related protein	*Bacteria – Proteobacteria – Gammaproteobacteria*	*Acidithiobacillus*	*Acidithiobacillus caldus* ATCC 51756	E=8e-20; I=53%
04074_0	Acriflavin resistance protein AcrB	*Bacteria – Proteobacteria – Gammaproteobacteria*	*Neptuniibacter*	*Neptuniibacter caesariensis*	E=9e-62; I=52%
05828_1	Acriflavin resistance protein AcrB	*Bacteria – Proteobacteria – Gammaproteobacteria*	Unclassified	marine *gamma proteobacterium* HTCC2148	E=2e-35; I=76%
08875_0	Copper resistance protein	*Bacteria – Proteobacteria – Gammaproteobacteria*	*Shewanella*	*Shewanella benthica* KT99	E=6e-60; I=63%
09253_6/7	Mercury reistance protein MerC	*Bacteria – Proteobacteria – Gammaproteobacteria*	*Alteromonas*	*Alteromonas macleodii* 'Deep ecotype'	E=9e-17; I=66%
15802_3	Acriflavin resistance protein AcrB	*Bacteria – Proteobacteria – Gammaproteobacteria*	*Thioalkalivibrio*	*Thioalkalivibrio sp.* HL-EbGR7	E=0.0; I=63%
05105_0	Type IV pilin biosynthesis protein	*Bacteria – Proteobacteria – Gammaproteobacteria*	*Marinobacter*	*Marinobacter aquaeolei* VT8	E=2e-21; I=46%
06024_1	Type IV pilin biosynthesis protein	*Bacteria – Proteobacteria – Gammaproteobacteria*	*Alteromonas*	*Alteromonas macleodii* 'Deep ecotype'	E=6e-50; I=31%
09482_5	Type IV fimbrial pilin related signal peptide protein	*Bacteria – Proteobacteria – Gammaproteobacteria*	*Nitrosomonas*	*Nitrosomonas* sp. AL212	E=4e-21; I=35%
15118_3	Type IV fimbrial pilin related signal peptide protein	*Bacteria – Proteobacteria – Gammaproteobacteria*	*Nitrosomonas*	*Nitrosomonas* sp. AL212	E=9e-18; I=50%
17310_0	Type IV pilin biosynthesis protein	*Bacteria – Proteobacteria – Gammaproteobacteria*	*Alteromonas*	*Alteromonas macleodii* 'Deep ecotype'	E=5e-74; I=32%
00608_0	MSHA biosynthesis protein MshQ (pilus type IV)	*Bacteria – Proteobacteria – Gammaproteobacteria*	Unclassified	marine *gamma proteobacterium* HTCC2143	E=2e-20; I=42%

01324_0	MSHA biosynthesis protein MshQ (pilus type IV)	*Bacteria – Proteobacteria – Gammaproteobacteria*	*Pseudoalteromonas*	*Pseudoalteromonas haloplanktis* TAC 125	E=1e-16; I=48%
02564_2	O-antigen transporter	*Bacteria – Proteobacteria – Gammaproteobacteria*	*Sulfurihydrogenibium*	*Sulfurihydrogenibium* sp. YO3AOP1	E=3e-28; I=37%
10735_3/5	Surface antigen (D15)	*Bacteria – Proteobacteria – Gammaproteobacteria*	*Thioalkalivibrio*	*Thioalkalivibrio sp.* HL-EbGr7	E=6e-28; I=44%
16322_0	Surface antigen (D15)	*Bacteria – Proteobacteria – Gammaproteobacteria*	*Thioalkalivibrio*	*Thioalkalivibrio sp.* HL-EbGr7	E=2e-85; I=43%
00460_1	Type II secretion system protein	*Bacteria – Proteobacteria – Gammaproteobacteria*	*Xanthomonas*	*Xanthomonas campestris* str.B100	E=2e-33; I=53%
00670_2	General secretion pathway protein G	*Bacteria – Proteobacteria – Gammaproteobacteria*	*Halothiobacillus*	*Halothiobacillus neapolitanus* c2	E=6e-28; I=56%
00670_3	General secretion pathway protein H	*Bacteria – Proteobacteria – Gammaproteobacteria*	*Beggiatoa*	*Beggiatoa sp.* PS	E=3e-9; I=34%
00670_4	General secretion pathway protein I	*Bacteria – Proteobacteria – Gammaproteobacteria*	Unclassified	marine *gamma proteobacterium* HTCC2143	E=1e-13; I=38%
00670_5	General secretion pathway protein J	*Bacteria – Proteobacteria – Gammaproteobacteria*	*Pseudomonas*	*Pseudomonas stutzeri* A1501	E=2e-26; I=37%
02283_0/1	General secretion pathway protein E	*Bacteria – Proteobacteria – Gammaproteobacteria*	*Moritella*	*Moritella sp.* PE36	E=4e-27; I=60%
03098_0	Type II secretion system protein F	*Bacteria – Proteobacteria – Gammaproteobacteria*	*Shewanella*	*Shewanella amazonensis* SB2B	E=4e-17; I=38%
03102_1	General secretion pathway protein B	*Bacteria – Proteobacteria – Gammaproteobacteria*	*Dickeya*	*Dickeya dadantii* Ech703	E=2e-7; I=36%
03190_0	Type II secretion system protein F	*Bacteria – Proteobacteria – Gammaproteobacteria*	*Psychromonas*	*Psychromonas ingrahamii* 37	E=7e-10; I=39%
04095_0	Type II secretion system protein D	*Bacteria – Proteobacteria – Gammaproteobacteria*	Unclassified	Candidatus *Poribacteria* sp. WGA-A3	E=7e-21; I=40%
04283_0	Type II and III secretion system protein	*Bacteria – Proteobacteria – Gammaproteobacteria*	*Nitrosospira*	*Nitrosospira multiformis* ATCC 25196	E=4e-25; I=62%
04606_2	General secretion pathway protein A	*Bacteria – Proteobacteria – Gammaproteobacteria*	*Hahella*	*Hahella chejuensis* KCTC 2396	E=3e-22; I=61%
04975_0	General secretion pathway protein D	*Bacteria – Proteobacteria – Gammaproteobacteria*	Unclassified	*gamma proteobacterium* HTCC5015	E=5e-25; I=45%
09255_2	General secretion pathway protein C	*Bacteria – Proteobacteria – Gammaproteobacteria*	*Grimontia*	*Grimontia hollisae* CIP 101886	E=3e-8; I=40%
09481_0/1	General secretion pathway protein F	*Bacteria – Proteobacteria – Gammaproteobacteria*	*Thioalkalivibrio*	*Thioalkalivibrio sp.* HL-EbGr7	E=8e-74; I=58%
11080_2	General secretion pathway protein D	*Bacteria – Proteobacteria – Gammaproteobacteria*	*Beggiatoa*	*Beggiatoa sp.* PS	E=6e-34; I=58%
11081_2	General secretion pathway protein D	*Bacteria – Proteobacteria – Gammaproteobacteria*	*Shewanella*	*Shewanella loihica* PV-4B	E=2e-23; I=50%
11489_1	Type II secretion system protein D	*Bacteria – Proteobacteria – Gammaproteobacteria*	*Vibrio*	*Vibrio harveyi* 1DA3	E=3e-25; I=26%
13086_1	Type IV pilus secretin PilQ	*Bacteria – Proteobacteria – Gammaproteobacteria*	*Allochromatium*	*Allochromatium vinosum* DSM 180	E=6e-51; I=44%
14298_0	Type II secretion system protein E	*Bacteria – Proteobacteria – Gammaproteobacteria*	*Xanthomonas*	*Xanthomonas campestris* str.B100	E=6e-19; I=69%
16484_0	Type I secretion outer membrane protein TolC	*Bacteria – Proteobacteria – Gammaproteobacteria*	*Thioalkalivibrio*	*Thioalkalivibrio sp.* HL-EbGr7	E=2e-75; I=43%
19578_3	General secretion pathway protein K	*Bacteria – Proteobacteria – Gammaproteobacteria*	*Thiomicrospira*	*Thiomicrospira crunogena* XCL-2	E=1e-12; I=31%
19578_4	General secretion pathway protein L	*Bacteria – Proteobacteria – Gammaproteobacteria*	*Beggiatoa*	*Beggiatoa sp.* PS	E=4e-16; I=37%
29325_0	General secretion pathway protein E	*Bacteria – Proteobacteria – Gammaproteobacteria*	*Acinetobacter*	*Acinetobacter* sp. ADP1	E=7e-18; I=92%
12758_1	Exopolysaccharide biosynthesis protein	*Bacteria – Proteobacteria – Gammaproteobacteria*	*Nitrosococcus*	*Nitrosococcus oceani* AFC7	E=3e-62; I=39%
09016_0	Polyphosphate kinase II	*Bacteria – Proteobacteria – Zetaproteobacteria*	*Mariprofundus*	*Mariprofundus ferrooxydans* PV-1	E=4e-110; I=73%
12872_0	Polyphosphate kinase	*Bacteria – Proteobacteria – Zetaproteobacteria*	*Mariprofundus*	*Mariprofundus ferrooxydans* PV-1	E=6e-18; I=80%
12875_0	Polyphosphate kinase	*Bacteria – Proteobacteria – Zetaproteobacteria*	*Mariprofundus*	*Mariprofundus ferrooxydans* PV-1	E=2e-17; I=70%
17302_0	Polyphosphate kinase	*Bacteria – Proteobacteria – Zetaproteobacteria*	*Mariprofundus*	*Mariprofundus ferrooxydans* PV-1	E=2e-46; I=66%
16261_1	inorganic polyphosphate/ATP-NAD kinase	*Bacteria – Proteobacteria – Zetaproteobacteria*	*Mariprofundus*	*Mariprofundus ferrooxydans* PV-1	E=4e-45; I=67%
11594_0	4Fe-4S ferredoxin, iron-sulfur binding protein	*Bacteria – Proteobacteria – Zetaproteobacteria*	*Mariprofundus*	*Mariprofundus ferrooxydans* PV-1	E=1e-10; I=67%
11594_1	Polysulphide reductase, NrfD	*Bacteria – Proteobacteria – Zetaproteobacteria*	*Mariprofundus*	*Mariprofundus ferrooxydans* PV-1	E=2e-92; I=88%
11595_0	4Fe-4S ferredoxin, iron-sulfur binding protein	*Bacteria – Proteobacteria – Zetaproteobacteria*	*Mariprofundus*	*Mariprofundus ferrooxydans* PV-1	E=5e-78; I=73%
12064_1	Ferredoxin	*Bacteria – Proteobacteria – Zetaproteobacteria*	*Mariprofundus*	*Mariprofundus ferrooxydans* PV-1	E=1e-44; I=84%
12727_0	Ferredoxin	*Bacteria – Proteobacteria – Zetaproteobacteria*	*Mariprofundus*	*Mariprofundus ferrooxydans* PV-1	E=9e-27; I=50%
12827_0/1	Glutamate synthase (ferredoxin)	*Bacteria – Proteobacteria – Zetaproteobacteria*	*Mariprofundus*	*Mariprofundus ferrooxydans* PV-1	E=1e-73; I=73%
12510_0	FeS assembly ATPase SufC	*Bacteria – Proteobacteria – Zetaproteobacteria*	*Mariprofundus*	*Mariprofundus ferrooxydans* PV-1	E=6e-63; I=79%
11100_1	iron-sulfur cluster insertion protein ErpA	*Bacteria – Proteobacteria – Zetaproteobacteria*	*Mariprofundus*	*Mariprofundus ferrooxydans* PV-1	E=2e-36; I=74%

00953_0	glycogen synthase	*Bacteria – Proteobacteria – Zetaproteobacteria*	*Mariprofundus*	*Mariprofundus ferrooxydans* PV-1	E=3e-40; I=73%
13015_2	glycogen synthase	*Bacteria – Proteobacteria – Zetaproteobacteria*	*Mariprofundus*	*Mariprofundus ferrooxydans* PV-1	E=3e-73; I=67%
02125_1	Glycogen/starch/alpha-glucan phosphorylase	*Bacteria – Proteobacteria – Zetaproteobacteria*	*Mariprofundus*	*Mariprofundus ferrooxydans* PV-1	E=1e-45; I=66%
15625_0	Glycogen branching enzyme	*Bacteria – Proteobacteria – Zetaproteobacteria*	*Mariprofundus*	*Mariprofundus ferrooxydans* PV-1	E=4e-110; I=80%
00417_0	Long chain fatty acid transport protein	*Bacteria – Proteobacteria – Zetaproteobacteria*	*Mariprofundus*	*Mariprofundus ferrooxydans* PV-1	E=4e-30; I=44%
00504_0/1	Heavy metal efflux pump CzcA	*Bacteria – Proteobacteria – Zetaproteobacteria*	*Mariprofundus*	*Mariprofundus ferrooxydans* PV-1	E=3e-44; I=77%
01809_1	Transporter	*Bacteria – Proteobacteria – Zetaproteobacteria*	*Mariprofundus*	*Mariprofundus ferrooxydans* PV-1	E=6e-10; I=50%
02422_0	Bacterial transport system permease protein	*Bacteria – Proteobacteria – Zetaproteobacteria*	*Mariprofundus*	*Mariprofundus ferrooxydans* PV-1	E=6e-28; I=72%
03228_2	ABC transporter, ATP-binding protein	*Bacteria – Proteobacteria – Zetaproteobacteria*	*Mariprofundus*	*Mariprofundus ferrooxydans* PV-1	E=2e-35; I=88%
03228_3	ABC transporter, ATP-binding protein	*Bacteria – Proteobacteria – Zetaproteobacteria*	*Mariprofundus*	*Mariprofundus ferrooxydans* PV-1	E=2e-44; I=80%
03228_4	ABC transporter, ATP-binding protein	*Bacteria – Proteobacteria – Zetaproteobacteria*	*Mariprofundus*	*Mariprofundus ferrooxydans* PV-1	E=8e-102; I=80%
03238_2	ABC transporter, ATP-binding protein	*Bacteria – Proteobacteria – Zetaproteobacteria*	*Mariprofundus*	*Mariprofundus ferrooxydans* PV-1	E=2e-35; I=88%
03238_3	ABC transporter, ATP-binding protein	*Bacteria – Proteobacteria – Zetaproteobacteria*	*Mariprofundus*	*Mariprofundus ferrooxydans* PV-1	E=2e-44; I=80%
03238_4	ABC transporter, ATP-binding protein	*Bacteria – Proteobacteria – Zetaproteobacteria*	*Mariprofundus*	*Mariprofundus ferrooxydans* PV-1	E=8e-102; I=80%
03276_4	Drug resistance transporter Bcr/CflA	*Bacteria – Proteobacteria – Zetaproteobacteria*	*Mariprofundus*	*Mariprofundus ferrooxydans* PV-1	E=2e-7; I=41%
03276_4	Drug resistance transporter Bcr/CflA	*Bacteria – Proteobacteria – Zetaproteobacteria*	*Mariprofundus*	*Mariprofundus ferrooxydans* PV-1	E=2e-7; I=41%
03370_0	Amino acid ABC transporter	*Bacteria – Proteobacteria – Zetaproteobacteria*	*Mariprofundus*	*Mariprofundus ferrooxydans* PV-1	E=8e-75; I=73%
03463_0/1	Electron transport complex, RnfABCDGE type	*Bacteria – Proteobacteria – Zetaproteobacteria*	*Mariprofundus*	*Mariprofundus ferrooxydans* PV-1	E=7e-63; I=74%
03999_0	Heavy metal efflux pump CzcA	*Bacteria – Proteobacteria – Zetaproteobacteria*	*Mariprofundus*	*Mariprofundus ferrooxydans* PV-1	E=1e-54; I=68%
04218_0	Potassium transporter peripheral membrane component	*Bacteria – Proteobacteria – Zetaproteobacteria*	*Mariprofundus*	*Mariprofundus ferrooxydans* PV-1	E=2e-45; I=67%
04619_1	ABC-type transport system involved in resistance to organic solvents	*Bacteria – Proteobacteria – Zetaproteobacteria*	*Mariprofundus*	*Mariprofundus ferrooxydans* PV-1	E=3e-49; I=61%
04850_0	Fatty acid transport system	*Bacteria – Proteobacteria – Zetaproteobacteria*	*Mariprofundus*	*Mariprofundus ferrooxydans* PV-1	E=5e-12; I=44%
05109_1	Peptide ABC transporter, permease protein	*Bacteria – Proteobacteria – Zetaproteobacteria*	*Mariprofundus*	*Mariprofundus ferrooxydans* PV-1	E=1e-13; I=73%
05109_1	ABC transporter, permease protein	*Bacteria – Proteobacteria – Zetaproteobacteria*	*Mariprofundus*	*Mariprofundus ferrooxydans* PV-1	E=1e-13; I=73%
05605_0	ABC transporter ATPase component	*Bacteria – Proteobacteria – Zetaproteobacteria*	*Mariprofundus*	*Mariprofundus ferrooxydans* PV-1	E=1e-28; I=59%
05797_0	Mn^{2+}/Zn^{2+} ABC transporter, permease protein	*Bacteria – Proteobacteria – Zetaproteobacteria*	*Mariprofundus*	*Mariprofundus ferrooxydans* PV-1	E=9e-18; I=54%
05797_0/1	Mn2+/Zn2+ ABC transporter, permease protein	*Bacteria – Proteobacteria – Zetaproteobacteria*	*Mariprofundus*	*Mariprofundus ferrooxydans* PV-1	E=1e-9; I=80%
05797_1	Mn^{2+}/Zn^{2+} ABC transporter, permease protein	*Bacteria – Proteobacteria – Zetaproteobacteria*	*Mariprofundus*	*Mariprofundus ferrooxydans* PV-1	E=1e-9; I=80%
05872_1	Magnesium transporter	*Bacteria – Proteobacteria – Zetaproteobacteria*	*Mariprofundus*	*Mariprofundus ferrooxydans* PV-1	E=3e-163; I=71%
05879_1	Magnesium transporter	*Bacteria – Proteobacteria – Zetaproteobacteria*	*Mariprofundus*	*Mariprofundus ferrooxydans* PV-1	E=3e-163; I=71%
10393_0	Binding-protein-dependent transport system	*Bacteria – Proteobacteria – Zetaproteobacteria*	*Mariprofundus*	*Mariprofundus ferrooxydans* PV-1	E=5e-62; I=69%
11100_0	ABC transporter, ATP-binding protein	*Bacteria – Proteobacteria – Zetaproteobacteria*	*Mariprofundus*	*Mariprofundus ferrooxydans* PV-1	E=6e-31; I=72%
11100_0	ABC transporter, ATP-binding protein	*Bacteria – Proteobacteria – Zetaproteobacteria*	*Mariprofundus*	*Mariprofundus ferrooxydans* PV-1	E=6e-31; I=72%
11101_1	ABC-2 type transporter	*Bacteria – Proteobacteria – Zetaproteobacteria*	*Mariprofundus*	*Mariprofundus ferrooxydans* PV-1	E=2e-23; I=68%
11101_1	ABC-2 type transporter	*Bacteria – Proteobacteria – Zetaproteobacteria*	*Mariprofundus*	*Mariprofundus ferrooxydans* PV-1	E=2e-23; I=68%
11221_0	Rh family protein/ammonium transporter	*Bacteria – Proteobacteria – Zetaproteobacteria*	*Mariprofundus*	*Mariprofundus ferrooxydans* PV-1	E=3e-68; I=78%
11221_1	Ammonium transporter	*Bacteria – Proteobacteria – Zetaproteobacteria*	*Mariprofundus*	*Mariprofundus ferrooxydans* PV-1	E=1e-16; I=81%
11491_0	Long chain fatty acid transport protein	*Bacteria – Proteobacteria – Zetaproteobacteria*	*Mariprofundus*	*Mariprofundus ferrooxydans* PV-1	E=5e-31; I=40%
11802_0	metal ion ABC transporter, ATP-binding protein	*Bacteria – Proteobacteria – Zetaproteobacteria*	*Mariprofundus*	*Mariprofundus ferrooxydans* PV-1	E=9e-32; I=54%
11802_0	ABC transporter	*Bacteria – Proteobacteria – Zetaproteobacteria*	*Mariprofundus*	*Mariprofundus ferrooxydans* PV-1	E=9e-32; I=54%
12078_0	Biopolymer transport protein	*Bacteria – Proteobacteria – Zetaproteobacteria*	*Mariprofundus*	*Mariprofundus ferrooxydans* PV-1	E=1e-18; I=58%

13102_1	Co/Zn/Cd cation transporter	*Bacteria – Proteobacteria – Zetaproteobacteria*	*Mariprofundus*	*Mariprofundus ferrooxydans* PV-1	E=1e-9; I=50%
13117_0/1/2	ABC transporter, ATPase component	*Bacteria – Proteobacteria – Zetaproteobacteria*	*Mariprofundus*	*Mariprofundus ferrooxydans* PV-1	E=6e-98; I=70%
13117_0	ABC transporter related protein	*Bacteria – Proteobacteria – Zetaproteobacteria*	*Mariprofundus*	*Mariprofundus ferrooxydans* PV-1	E=6e-98; I=70%
13117_1	ABC transporter related protein	*Bacteria – Proteobacteria – Zetaproteobacteria*	*Mariprofundus*	*Mariprofundus ferrooxydans* PV-1	E=1e-25; I=79%
13117_2	ABC transporter related protein	*Bacteria – Proteobacteria – Zetaproteobacteria*	*Mariprofundus*	*Mariprofundus ferrooxydans* PV-1	E=2e-36; I=63%
13302_0	Electron transport complex, RnfABCDGE type, D subunit	*Bacteria – Proteobacteria – Zetaproteobacteria*	*Mariprofundus*	*Mariprofundus ferrooxydans* PV-1	E=1e-57; I=62%
13302_1	Electron transport complex, RnfABCDGE type, C subunit	*Bacteria – Proteobacteria – Zetaproteobacteria*	*Mariprofundus*	*Mariprofundus ferrooxydans* PV-1	E=3e-53; I=67%
13467_0	metal ion ABC transporter, ATP-binding protein	*Bacteria – Proteobacteria – Zetaproteobacteria*	*Mariprofundus*	*Mariprofundus ferrooxydans* PV-1	E=8e-27; I=65%
13467_0	ABC transporter	*Bacteria – Proteobacteria – Zetaproteobacteria*	*Mariprofundus*	*Mariprofundus ferrooxydans* PV-1	E=8e-27; I=65%
13467_1	metal ion ABC transporter, ATP-binding protein	*Bacteria – Proteobacteria – Zetaproteobacteria*	*Mariprofundus*	*Mariprofundus ferrooxydans* PV-1	E=3e-15; I=75%
13467_1	ABC transporter	*Bacteria – Proteobacteria – Zetaproteobacteria*	*Mariprofundus*	*Mariprofundus ferrooxydans* PV-1	E=3e-15; I=75%
13756_0	Oligopeptide/dipeptide ABC transporter, ATP-binding protein	*Bacteria – Proteobacteria – Zetaproteobacteria*	*Mariprofundus*	*Mariprofundus ferrooxydans* PV-1	E=3e-51; I=74%
14144_0	ABC transporter, ATP-binding protein	*Bacteria – Proteobacteria – Zetaproteobacteria*	*Mariprofundus*	*Mariprofundus ferrooxydans* PV-1	E=1e-85; I=85%
14144_0	ABC transporter, ATP-binding protein	*Bacteria – Proteobacteria – Zetaproteobacteria*	*Mariprofundus*	*Mariprofundus ferrooxydans* PV-1	E=1e-85; I=86%
14145_0	ABC transporter, ATP-binding protein	*Bacteria – Proteobacteria – Zetaproteobacteria*	*Mariprofundus*	*Mariprofundus ferrooxydans* PV-1	E=2e-27; I=80%
14200_0/1	ABC transporter related protein	*Bacteria – Proteobacteria – Zetaproteobacteria*	*Mariprofundus*	*Mariprofundus ferrooxydans* PV-1	E=2e-40; I=80%
14200_0/1	ABC transporter, ATP-binding protein	*Bacteria – Proteobacteria – Zetaproteobacteria*	*Mariprofundus*	*Mariprofundus ferrooxydans* PV-1	E=2e-40; I=80%
14940_1	Electron transport complex, RnfE	*Bacteria – Proteobacteria – Zetaproteobacteria*	*Mariprofundus*	*Mariprofundus ferrooxydans* PV-1	E=3e-54; I=81%
15133_1	Cation transport ATPase	*Bacteria – Proteobacteria – Zetaproteobacteria*	*Mariprofundus*	*Mariprofundus ferrooxydans* PV-1	E=5e-12; I=39%
15999_0	ABC transporter related protein	*Bacteria – Proteobacteria – Zetaproteobacteria*	*Mariprofundus*	*Mariprofundus ferrooxydans* PV-1	E=1e-31; I=91%
15999_0	ABC transporter, ATP-binding protein	*Bacteria – Proteobacteria – Zetaproteobacteria*	*Mariprofundus*	*Mariprofundus ferrooxydans* PV-1	E=1e-31; I=91%
15999_0	ABC transporter	*Bacteria – Proteobacteria – Zetaproteobacteria*	*Mariprofundus*	*Mariprofundus ferrooxydans* PV-1	E=1e-31; I=91%
16086_0	Amino acid transport system, periplasmic component	*Bacteria – Proteobacteria – Zetaproteobacteria*	*Mariprofundus*	*Mariprofundus ferrooxydans* PV-1	E=1e-66; I=40%
16421_0/1/2	Peptide ABC transporter, permease protein	*Bacteria – Proteobacteria – Zetaproteobacteria*	*Mariprofundus*	*Mariprofundus ferrooxydans* PV-1	E=3e-16; I=86%
16421_1	Peptide ABC transporter, permease protein	*Bacteria – Proteobacteria – Zetaproteobacteria*	*Mariprofundus*	*Mariprofundus ferrooxydans* PV-1	E=3e-16; I=86%
16421_2	Binding-protein-dependent transport system	*Bacteria – Proteobacteria – Zetaproteobacteria*	*Mariprofundus*	*Mariprofundus ferrooxydans* PV-1	E=5e-62; I=78%
17284_0	peptide ABC transporter, permease protein	*Bacteria – Proteobacteria – Zetaproteobacteria*	*Mariprofundus*	*Mariprofundus ferrooxydans* PV-1	E=5e-58; I=68%
24419_0	ABC transporter related protein	*Bacteria – Proteobacteria – Zetaproteobacteria*	*Mariprofundus*	*Mariprofundus ferrooxydans* PV-1	E=6e-23; I=83%
24419_0	ABC transporter, ATP-binding protein	*Bacteria – Proteobacteria – Zetaproteobacteria*	*Mariprofundus*	*Mariprofundus ferrooxydans* PV-1	E=6e-23; I=83%
24432_1	Heavy metal efflux pump CzcA	*Bacteria – Proteobacteria – Zetaproteobacteria*	*Mariprofundus*	*Mariprofundus ferrooxydans* PV-1	E=4e-23; I=73%
05259_0/1/2	Acriflavin resistance protein AcrB	*Bacteria – Proteobacteria – Zetaproteobacteria*	*Mariprofundus*	*Mariprofundus ferrooxydans* PV-1	E=0.0; I=70%
11083_0	Heavy metal resistance protein CzcC	*Bacteria – Proteobacteria – Zetaproteobacteria*	*Mariprofundus*	*Mariprofundus ferrooxydans* PV-1	E=3e-16; I=35%
03033_0	Type II secretion system protein	*Bacteria – Proteobacteria – Zetaproteobacteria*	*Mariprofundus*	*Mariprofundus ferrooxydans* PV-1	E=6e-69; I=81%
13867_1	Secretion protein HlyD	*Bacteria – Proteobacteria – Zetaproteobacteria*	*Mariprofundus*	*Mariprofundus ferrooxydans* PV-1	E=5e-32; I=46%
00774_1	ATPase involved in pili biogenesis	*Bacteria – Proteobacteria – Zetaproteobacteria*	*Mariprofundus*	*Mariprofundus ferrooxydans* PV-1	E=2e-45; I=72%
14064_1	Type IV prepilin-like protein leader peptide-processing enzyme	*Bacteria – Proteobacteria – Zetaproteobacteria*	*Mariprofundus*	*Mariprofundus ferrooxydans* PV-1	E=5e-47; I=59%
01455_0	Flagellar assembly protein	*Bacteria – Proteobacteria – Zetaproteobacteria*	*Mariprofundus*	*Mariprofundus ferrooxydans* PV-1	E=1e-8; I=48%
01455_1	Flagellar motor switch protein FliG	*Bacteria – Proteobacteria – Zetaproteobacteria*	*Mariprofundus*	*Mariprofundus ferrooxydans* PV-1	E=5e-18; I=65%
02167_0	Flagellar biosynthesis protein FliR	*Bacteria – Proteobacteria – Zetaproteobacteria*	*Mariprofundus*	*Mariprofundus ferrooxydans* PV-1	E=3e-27; I=53%
02441_0	Flagellar P-ring protein	*Bacteria – Proteobacteria – Zetaproteobacteria*	*Mariprofundus*	*Mariprofundus ferrooxydans* PV-1	E=8e-15; I=73%
02441_1	Flagellar L-ring protein	*Bacteria – Proteobacteria – Zetaproteobacteria*	*Mariprofundus*	*Mariprofundus ferrooxydans* PV-1	E=5e-23; I=64%
02570_0	Flagellar M-ring protein FliF	*Bacteria – Proteobacteria – Zetaproteobacteria*	*Mariprofundus*	*Mariprofundus ferrooxydans* PV-1	E=3e-32; I=54%

03127_1	Flagellar basal body-associated protein FliL	*Bacteria – Proteobacteria – Zetaproteobacteria*	*Mariprofundus*	*Mariprofundus ferrooxydans* PV-1	E=4e-39; I=62%
04108_0	Flagellar hook-associated protein FlgK	*Bacteria – Proteobacteria – Zetaproteobacteria*	*Mariprofundus*	*Mariprofundus ferrooxydans* PV-1	E=3e-33; I=53%
05266_0	Flagellar biosynthesis protein	*Bacteria – Proteobacteria – Zetaproteobacteria*	*Mariprofundus*	*Mariprofundus ferrooxydans* PV-1	E=2e-23; I=55%
11604_0/1	Flagellar biosynthesis protein FlhA	*Bacteria – Proteobacteria – Zetaproteobacteria*	*Mariprofundus*	*Mariprofundus ferrooxydans* PV-1	E=2e-29; I=75%
17328_0	Flagellar biosynthetic protein FliP	*Bacteria – Proteobacteria – Zetaproteobacteria*	*Mariprofundus*	*Mariprofundus ferrooxydans* PV-1	E=7e-21; I=64%
19097_0	Flagellar motor switch protein FliM	*Bacteria – Proteobacteria – Zetaproteobacteria*	*Mariprofundus*	*Mariprofundus ferrooxydans* PV-1	E=3e-27; I=38%

6. Quorum-sensing genes

Gene name (contig)	Gene product name	Organism (class)	Genus (best hit)	Strain (best hit)	Expected Value (E) and Identities (I)
00775_11	S-ribosylhomocysteinase, quorum-sensing protein LuxS	*Bacteria – Proteobacteria – Epsilonproteobacteria*	*Sulfurovum*	*Sulfurovum sp*. NBC37-1	E=3e-66; I=82%
15824_0	O-acetylhomoserine sulfhydrylase/aminocarboxypropyltransferase	*Bacteria – Proteobacteria – Epsilonproteobacteria*	*Sulfurimonas*	*Sulfurimonas denitrificans* DSM 1251	E=5e-68; I=52%
02032_0	sodium-driven multidrug efflux pump protein	*Bacteria – Proteobacteria – Epsilonproteobacteria*	*Sulfurovum*	*Sulfurovum sp*. NBC37-1	E=3e-100; I=74%
00192_3	lipoprotein-releasing system ATP-binding protein LolD	*Bacteria – Proteobacteria – Epsilonproteobacteria*	*Campylobacter*	*Campylobacter hominis* ATCC BAA-381	E=3e-67; I=60%
00391_2	ABC transporter, ATP-binding protein	*Bacteria – Proteobacteria – Epsilonproteobacteria*	*Sulfurospirillum*	*Sulfurospirillum deleyianum* DSM 6946	E=4e-11; I=59%
06311_1	quorum-sensing protein luxS	*Bacteria – Proteobacteria – Gammaproteobacteria*	*Xenorhabdus*	*Xenorhabdus nematophila* ATCC 19061	E=1e-16; I=71%
03245_0	Regulatory protein LuxR	*Bacteria – Proteobacteria – Gammaproteobacteria*	*Beggiatoa*	*Beggiatoa sp.* PS	E=1e-6; I=56%
08859_1	homoserine lactone efflux protein	*Bacteria – Proteobacteria – Gammaproteobacteria*	*Moritella*	*Moritella sp.* PE36	E=3e-62; I=64%
14817_	Probable homoserine lactone efflux protein	*Bacteria – Proteobacteria – Gammaproteobacteria*	*Vibrio*	*Vibrio sp.* MED222	E=1e-31; I=52%
13373_0	ABC-type antimicrobial peptide transport system ATPase component	*Bacteria – Proteobacteria – Gammaproteobacteria*	*Vibrio*	*Vibrio mimicus* MB451	E=6e-58; I=55%
08756_0	ABC-type antimicrobial peptide transport system ATPase component	*Bacteria – Proteobacteria – Gammaproteobacteria*	*Pseudomonas*	*Pseudomonas savastanoi pv. savastanoi*	E=1e-89; I=51%
00872_4	Lipoprotein-releasing system ATP-binding protein lolD	*Bacteria – Proteobacteria – Gammaproteobacteria*	Unclassified	Uncultured bacterium 9F08	E=1e-79; I=61%
09461_1	Two-component transcriptional regulator, LuxR family	*Bacteria – Proteobacteria – Betaproteobacteria*	*Cupriavidus*	*Cupriavidus metallidurans* CH34	E=1e-45; I=43%
20687_2	Two-component transcriptional regulator, LuxR family	*Bacteria – Proteobacteria – Alphaproteobacteria*	*Magnetococcus*	*Magnetococcus sp.* MC-1	E=3e-29; I=40%

7. Heat shock proteins

Gene name (contig)	Gene product name	Organism (class)	Genus (best hit)	Strain (best hit)	Expected Value (E) and Identities (I)
00065_7	Chaperonin GroEL	*Bacteria – Proteobacteria – Epsilonproteobacteria*	*Sulfurovum*	*Sulfurovum sp*. NBC37-1	E=0.0; I=84%
00164_0	Chaperonin GroEL	*Bacteria – Proteobacteria – Epsilonproteobacteria*	*Sulfurovum*	*Sulfurovum sp*. NBC37-1	E=5e-161; I=86%
00918_0	Molecular chaperone DnaK	*Bacteria – Proteobacteria – Epsilonproteobacteria*	*Sulfurovum*	*Sulfurovum sp*. NBC37-1	E=0.0; I=89%
03789_0	DnaK suppressor protein DskA	*Bacteria – Proteobacteria – Epsilonproteobacteria*	*Sulfurovum*	*Sulfurovum sp*. NBC37-1	E=2e-21; I=60%
12014_1	DnaK suppressor protein	*Bacteria – Proteobacteria – Epsilonproteobacteria*	*Sulfurovum*	*Sulfurovum sp*. NBC37-1	E=9e-39; I=67%
14996_4	DnaK suppressor protein DskA	*Bacteria – Proteobacteria – Epsilonproteobacteria*	*Sulfurovum*	*Sulfurovum sp*. NBC37-1	E=2e-24; I=59%
15310_1	Molecular chaperone DnaK	*Bacteria – Proteobacteria – Epsilonproteobacteria*	*Helicobacter*	*Helicobacter pylori* 26695	E=3e-86; I=86%
03508_0	Chaperone protein DnaJ	*Bacteria – Proteobacteria – Epsilonproteobacteria*	*Sulfurovum*	*Sulfurovum sp*. NBC37-1	E=1e-60; I=65%
04340_0	Chaperone protein DnaJ	*Bacteria – Proteobacteria – Epsilonproteobacteria*	*Sulfurovum*	*Sulfurovum sp*. NBC37-1	E=5e-23; I=80%

10872_0	Chaperone protein DnaJ	*Bacteria – Proteobacteria – Epsilonproteobacteria*	*Sulfurovum*	*Sulfurovum sp*. NBC37-1	E=3e-72; I=69%
10901_6	Hsp DnaJ	*Bacteria – Proteobacteria – Epsilonproteobacteria*	*Sulfurovum*	*Sulfurovum sp*. NBC37-1	E=4e-88; I=62%
12343_0	Hsp DnaJ	*Bacteria – Proteobacteria – Epsilonproteobacteria*	*Arcobacter*	*Arcobacter butzleri* RM4018	E=5e-16; I=44%
15388_7	Hsp DnaJ-like	*Bacteria – Proteobacteria – Epsilonproteobacteria*	*Sulfurovum*	*Sulfurovum sp*. NBC37-1	E=1e-13; I=77%
18015_0	Chaperone protein DnaJ	*Bacteria – Proteobacteria – Epsilonproteobacteria*	*Sulfurovum*	*Sulfurovum sp*. NBC37-1	E=1e-6; I=63%
00918_1	GrpE nucleotide exchange factor	*Bacteria – Proteobacteria – Epsilonproteobacteria*	*Sulfurovum*	*Sulfurovum sp*. NBC37-1	E=1e-48; I=67%
02498_0	GrpE nucleotide exchange factor	*Bacteria – Proteobacteria – Epsilonproteobacteria*	*Sulfurovum*	*Sulfurovum sp*. NBC37-1	E=2e-37; I=59%
14996_11	Hsp HtpG	*Bacteria – Proteobacteria – Epsilonproteobacteria*	*Sulfurovum*	*Sulfurovum sp*. NBC37-1	E=0.0; I=81%
03020_0	ClpB protein	*Bacteria – Proteobacteria – Epsilonproteobacteria*	*Sulfurovum*	*Sulfurovum sp*. NBC37-1	E=4e-107; I=93%
03229_0	ClpB protein	*Bacteria – Proteobacteria – Epsilonproteobacteria*	*Sulfurovum*	*Sulfurovum sp*. NBC37-1	E=9e-138; I=91%
00566_1	CplX protein	*Bacteria – Proteobacteria – Epsilonproteobacteria*	*Sulfurovum*	*Sulfurovum sp*. NBC37-1	E=3e-124; I=79%
13170_4	CplX protein	*Bacteria – Proteobacteria – Epsilonproteobacteria*	*Sulfurovum*	*Sulfurovum sp*. NBC37-1	E=5e-151; I=79%
15789_0	CplX protein	*Bacteria – Proteobacteria – Epsilonproteobacteria*	*Sulfurovum*	*Sulfurovum sp*. NBC37-1	E=1e-60; I=74%
00233_2	ClpA protein	*Bacteria – Proteobacteria – Epsilonproteobacteria*	*Sulfurovum*	*Sulfurovum sp*. NBC37-1	E=0.0; I=66%
00360_1	ClpA protein	*Bacteria – Proteobacteria – Epsilonproteobacteria*	*Sulfurovum*	*Sulfurovum sp*. NBC37-1	E=6e-57; I=53%
01227_1	ClpA protein	*Bacteria – Proteobacteria – Epsilonproteobacteria*	*Sulfurovum*	*Sulfurovum sp*. NBC37-1	E=9e-47; I=70%
04391_0	ClpA protein	*Bacteria – Proteobacteria – Epsilonproteobacteria*	*Sulfurovum*	*Sulfurovum sp*. NBC37-1	E=4e-90; I=89%
13559_1	ClpA protein	*Bacteria – Proteobacteria – Epsilonproteobacteria*	*Sulfurovum*	*Sulfurovum sp*. NBC37-1	E=8e-35; I=64%
15821_0	ClpA protein	*Bacteria – Proteobacteria – Epsilonproteobacteria*	*Sulfurovum*	*Sulfurovum sp*. NBC37-1	E=2e-36; I=80%
15453_2	Chaperonin GroEL	*Bacteria – Proteobacteria – Gammaproteobacteria*	*Methylococcus*	*Methylococcus capsulatus str. Bath*	E=2e-179; I=80%
00844_14	Chaperone protein DnaK	*Bacteria – Proteobacteria – Gammaproteobacteria*	*Halorhodospira*	*Halorhodospira halophila SL1*	E=0.0; I=56%
01784_0	Molecular chaperone DnaK	*Bacteria – Proteobacteria – Gammaproteobacteria*	*Coxiella*	*Coxiella burnetii* RSA 493	E=0.0; I=82%
01319_0	Chaperone protein DnaJ	*Bacteria – Proteobacteria – Gammaproteobacteria*	*Beggiatoa*	*Beggiatoa sp.* PS	E=4e-114; I=61%
04339_0	DnaJ-like chaperone membrane protein	*Bacteria – Proteobacteria – Gammaproteobacteria*	*Nitrosospira*	*Nitrosospira multiformis* ATCC 25196	E=2e-22; I=47%
05933_0	Hsp DnaJ-like	*Bacteria – Proteobacteria – Gammaproteobacteria*	*Nitrosococcus*	*Nitrosococcus oceani* ATCC 19707	E=8e-67; I=56%
14154_0	DnaJ-class molecular chaperone	*Bacteria – Proteobacteria – Gammaproteobacteria*	*Klebsiella*	*Klebsiella pneumoniae* NTUH-K2044	E=2e-11; I=74%
00844_13	GrpE protein	*Bacteria – Proteobacteria – Gammaproteobacteria*	*Halorhodospira*	*Halorhodospira halophila SL1*	E=5e-23; I=30%
01784_2	GrpE nucleotide exchange factor	*Bacteria – Proteobacteria – Gammaproteobacteria*	*Methylococcus*	*Methylococcus capsulatus str. Bath*	E=4e-37; I=52%
23159_0	ClpX protein	*Bacteria – Proteobacteria – Gammaproteobacteria*	Unclassified	Endoriftia persephone 'Hot96_1+Hot96_2'	E=2e-35; I=54%
09373_0/1	ClpA protein	*Bacteria – Proteobacteria – Gammaproteobacteria*	*Methylococcus*	*Methylococcus capsulatus str. Bath*	E=0.0; I=70%
15698_1	Molecular chaperone DnaK	*Bacteria – Proteobacteria – Zetaproteobacteria*	*Mariprofundus*	*Mariprofundus ferrooxydans* PV-1	E=1e-16; I=84%
17135_1	Chaperone protein DnaJ	*Bacteria – Proteobacteria – Zetaproteobacteria*	*Mariprofundus*	*Mariprofundus ferrooxydans* PV-1	E=2e-148; I=75%
15055_1/2	GrpE nucleotide exchange factor	*Bacteria – Proteobacteria – Zetaproteobacteria*	*Mariprofundus*	*Mariprofundus ferrooxydans* PV-1	E=9e-24; I=73%
12886_0	Putative ClpB protein	*Bacteria – Proteobacteria – Zetaproteobacteria*	*Mariprofundus*	*Mariprofundus ferrooxydans* PV-1	E=1e-56; I=89%
13858_1	ClpX protein	*Bacteria – Proteobacteria – Zetaproteobacteria*	*Mariprofundus*	*Mariprofundus ferrooxydans* PV-1	E=3e-21; I=82%
17468_1	ClpX protein	*Bacteria – Proteobacteria – Zetaproteobacteria*	*Mariprofundus*	*Mariprofundus ferrooxydans* PV-1	E=3e-20; I=82%
11017_0	ClpA protein	*Bacteria – Proteobacteria – Zetaproteobacteria*	*Mariprofundus*	*Mariprofundus ferrooxydans* PV-1	E=4e-111; I=64%
15571_1	ClpA protein	*Bacteria – Proteobacteria – Zetaproteobacteria*	*Mariprofundus*	*Mariprofundus ferrooxydans* PV-1	E=1e-30; I=60%
16861_1	Hsp DnaJ-like	Unclassified	Unclassified	Unclassified	Unknown
23479_0	Hsp DnaJ-like	Unclassified	Unclassified	Unclassified	Unknown

8. CRISPRs sequences

Contig number	CRISPR category	Organism (class)	CRISPR length (bp)	Number of spacer(s)
00225	CRISPR, questionable	*Bacteria – Proteobacteria – Epsilonproteobacteria*	92	1
09690	CRISPR, confirmed	*Bacteria – Proteobacteria – Epsilonproteobacteria*	572	8
09750	CRISPR, questionable	*Bacteria – Proteobacteria – Epsilonproteobacteria*	112	1
09019	CRISPR, questionable	*Bacteria – Proteobacteria – Gammaproteobacteria*	81	1

Gene name (contig)	Gene product name	Organism (class)	Genus (best hit)	Strain (best hit)	Expected Value (E) and Identities (I)
00178_0	CRISPR-associated protein Cas4	*Bacteria – Proteobacteria – Epsilonproteobacteria*	*Arcobacter*	*Arcobacter nitrofigilis* DSM 7299	E=7e-56; I=77%
00178_1	CRISPR-associated protein Cas1	*Bacteria – Proteobacteria – Epsilonproteobacteria*	*Sulfurospirillum*	*Sulfurospirillum deleyianum* DSM 6946	E=8e-101; I=74%
00848_0	CRISPR-associated protein Cas1	*Bacteria – Proteobacteria – Epsilonproteobacteria*	*Campylobacter*	*Campylobacter gracilis* RM3268	E=1e-89; I=32%
00848_1	CRISPR-associated protein Cas2	*Bacteria – Proteobacteria – Epsilonproteobacteria*	*Microcoleus*	*Microcoleus chthonoplastes* PCC 7420	E=9e-7; I=42%
02185_1	CRISPR-associated protein Cas6	*Bacteria – Proteobacteria – Epsilonproteobacteria*	*Campylobacter*	*Campylobacter showae* RM3277	E=2e-31; I=35%
09420_1	CRISPR-associated protein Csh2	*Bacteria – Proteobacteria – Epsilonproteobacteria*	*Campylobacter*	*Campylobacter rectus* RM3267	E=1e-121; I=73%
09420_2	CRISPR-associated protein Cas5	*Bacteria – Proteobacteria – Epsilonproteobacteria*	*Campylobacter*	*Campylobacter rectus* RM3267	E=7e-67; I=52%
09420_3	CRISPR-associated helicase Cas3	*Bacteria – Proteobacteria – Epsilonproteobacteria*	*Campylobacter*	*Campylobacter rectus* RM3267	E=0.0; I=52%
09420_4	CRISPR-associated protein Cas4	*Bacteria – Proteobacteria – Epsilonproteobacteria*	*Campylobacter*	*Campylobacter gracilis* RM3268	E=3e-36; I=56%
09420_5	CRISPR-associated protein Cas2	*Bacteria – Proteobacteria – Epsilonproteobacteria*	*Campylobacter*	*Campylobacter gracilis* RM3268	E=2e-20; I=57%
09420_6	CRISPR-associated protein Cas1	*Bacteria – Proteobacteria – Epsilonproteobacteria*	*Campylobacter*	*Campylobacter showae* RM3277	E=6e-129; I=63%
13174_2	CRISPR-associated protein Csh2	*Bacteria – Proteobacteria – Epsilonproteobacteria*	*Caminibacter*	*Caminibacter mediatlanticus* TB-2	E=8e-111; I=69%
13174_3	CRISPR-associated protein Cas5	*Bacteria – Proteobacteria – Epsilonproteobacteria*	*Campylobacter*	*Campylobacter fetus ssp. fetus* 82-40	E=5e-49; I=53%
13174_4	CRISPR-associated helicase Cas3	*Bacteria – Proteobacteria – Epsilonproteobacteria*	*Caminibacter*	*Caminibacter mediatlanticus* TB-2	E=9e-127; I=49%
02868_1	CRISPR-associated protein, Cmr3 family	*Bacteria – Proteobacteria – Gammaproteobacteria*	*Thioalkalivibrio*	*Thioalkalivibrio sp.* HL-EbGr7	E=1e-21; I=41%
03906_0	CRISPR-associated protein Csy4	*Bacteria – Proteobacteria – Gammaproteobacteria*	*Legionella*	*Legionella pneumophila* str. Lens	E=1e-46; I=48%
04771_0	CRISPR-associated protein Cas1	*Bacteria – Proteobacteria – Gammaproteobacteria*	*Acinetobacter*	*Acinetobacter baumannii* AYE	E=6e-37; I=61%
05235_0	CRISPR-associated protein Csy3	*Bacteria – Proteobacteria – Gammaproteobacteria*	*Acinetobacter*	*Acinetobacter baumannii* AYE	E=6e-26; I=71%
05235_1	CRISPR-associated protein Csy3	*Bacteria – Proteobacteria – Gammaproteobacteria*	*Psychromonas*	*Psychromonas ingrahamii* 37	E=5e-48; I=58%
08937_0	CRISPR-associated protein, Cmr1 family	*Bacteria – Proteobacteria – Gammaproteobacteria*	*Vibrio*	*Vibrio sp.* RC341	E=1e-25; I=47%
09013_1	CRISPR-associated protein Cas1	*Bacteria – Proteobacteria – Gammaproteobacteria*	*Halorhodospira*	*Halorhodospira halophila* SL1	E=1e-14; I=48%
09013_2	CRISPR-associated protein Cas1	*Bacteria – Proteobacteria – Gammaproteobacteria*	*Neissera*	*Neissera cinerea* ATCC 14685	E=1e-6; I=47%
09013_3	CRISPR-associated protein Cas2	*Bacteria – Proteobacteria – Gammaproteobacteria*	*Nitrosococcus*	*Nitrosococcus halophilus* Nc4	E=3e-17; I=45%
09013_4	CRISPR-associated protein	*Bacteria – Proteobacteria – Gammaproteobacteria*	*Nitrosococcus*	*Nitrosococcus halophilus* Nc4	E=3e-102; I=54%
09484_1	CRISPR-associated protein Cas1	*Bacteria – Proteobacteria – Gammaproteobacteria*	*Legionella*	*Legionella pneumophila* str. Lens	E=1e-29; I=80%
09514_2	CRISPR-associated protein Cas1	*Bacteria – Proteobacteria – Gammaproteobacteria*	*Shewanella*	*Shewanella benthica* KT99	E=3e-59; I=65%
09541_0	CRISPR-associated protein Cas6	*Bacteria – Proteobacteria – Gammaproteobacteria*	*Beggiatoa*	*Beggiatoa sp.* PS	E=9e-12; I=51%
09541_1	CRISPR-associated protein Cas6	*Bacteria – Proteobacteria – Gammaproteobacteria*	*Allochromatium*	*Allochromatium vinosum* DSM 180	E=4e-26; I=45%
16675_0	CRISPR-associated protein	*Bacteria – Proteobacteria – Gammaproteobacteria*	*Beggiatoa*	*Beggiatoa sp.* PS	E=5e-21; I=36%
19656_0	CRISPR-associated protein	*Bacteria – Proteobacteria – Gammaproteobacteria*	*Legionella*	*Legionella pneumophila* str. Lens	E=6e-37; I=62%