ARTICLES

Symbiosis insights through metagenomic analysis of a microbial consortium

Tanja Woyke[1,2], Hanno Teeling[3], Natalia N. Ivanova[1], Marcel Huntemann[3], Michael Richter[3], Frank Oliver Gloeckner[3,4], Dario Boffelli[1,2], Iain J. Anderson[1], Kerrie W. Barry[1], Harris J. Shapiro[1], Ernest Szeto[1], Nikos C. Kyrpides[1], Marc Mussmann[3], Rudolf Amann[3], Claudia Bergin[3], Caroline Ruehland[3], Edward M. Rubin[1,2] & Nicole Dubilier[3]

Symbioses between bacteria and eukaryotes are ubiquitous, yet our understanding of the interactions driving these associations is hampered by our inability to cultivate most host-associated microbes. Here we use a metagenomic approach to describe four co-occurring symbionts from the marine oligochaete *Olavius algarvensis*, a worm lacking a mouth, gut and nephridia. Shotgun sequencing and metabolic pathway reconstruction revealed that the symbionts are sulphur-oxidizing and sulphate-reducing bacteria, all of which are capable of carbon fixation, thus providing the host with multiple sources of nutrition. Molecular evidence for the uptake and recycling of worm waste products by the symbionts suggests how the worm could eliminate its excretory system, an adaptation unique among annelid worms. We propose a model that describes how the versatile metabolism within this symbiotic consortium provides the host with an optimal energy supply as it shuttles between the upper oxic and lower anoxic coastal sediments that it inhabits.

Symbiosis has a major role in shaping the evolution and diversity of eukaryotic organisms[1]. Remarkably, it is only in recent times that there has been an emerging recognition that most eukaryotic organisms are intimately associated with a complex community of beneficial microbes that are essential for their development, health and interactions with the environment[2]. This renaissance in symbiosis research stems from advances in molecular approaches that have enabled the study of natural microbial consortia using cultivation-independent methods[3–5]. Metagenomic analyses have provided a new dimension in the study of community organization and metabolism in natural microbial communities[6–10]. So far, however, genomic analyses of symbiotic microbes from eukaryotes have been confined to individual species (for the only exception, see ref. 11), limiting our ability to understand the intricate interactions involving communication, competition and resource partitioning that shape symbiotic microbial communities.

Here we use random shotgun sequencing and nucleotide-signature-based binning to study the symbiotic community in *Olavius algarvensis*. This marine worm belongs to a group of oligochaetes (phylum Annelida) that lack a mouth, gut and anus, and are unique among annelid worms in having reduced their nephridial excretory system[12]. They live in obligate and species-specific associations with multiple extracellular bacterial endosymbionts located just below the worm cuticle[12]. As the symbionts have yet to be grown in culture, their phylogeny has only been accessible through 16S ribosomal RNA analysis and fluorescence *in situ* hybridization (FISH)[13,14]. *O. algarvensis* lives in coastal Mediterranean sediments and harbours a chemoautotrophic sulphur-oxidizing gammaproteobacterium (γ1 symbiont) and a deltaproteobacterial sulphate reducer (δ1 symbiont), recently shown to be engaged in an endosymbiotic sulphur cycle[14]. An additional gamma- and delta-proteobacterial symbiont (γ3 and δ4 symbionts) of unknown function occur consistently in these hosts, and, in some individuals, a spirochaete has been observed as a minor part of the symbiotic consortium (see Supplementary Fig. 1a)[12].

Given that most chemosynthetic hosts harbour only one or two bacteria, the association of gutless oligochaetes with multiple symbionts is remarkable, raising a series of questions about their interactions with each other, their host and the environment. How does the symbiosis compensate for the loss of digestive and excretory systems in the host, what do the various partners gain from this relationship, and is it mutually obligate? What is the selective advantage for *O. algarvensis* in harbouring multiple symbiotic partners? Our metagenomic analyses explain how the bacterial consortium meets the energy and waste management needs of its oligochaete host. We describe how resource partitioning between the phylogenetically diverse symbionts benefits both the symbionts and the worm in the heterogeneous environment. Finally, we propose a model showing that the selective advantage of harbouring multiple symbionts lies in their ability to supply their host with energy from an abundant and diverse supply of reducing equivalents and electron acceptors as it shuttles between the oxidized and reduced sediment layers.

Metagenomic data analysis and binning

Pooled samples of 200 *O. algarvensis* specimens per library were shotgun-sequenced (see Supplementary Information, Supplementary Fig. 1b) and the sequences assembled using the whole-genome shotgun assembler JAZZ[15] (Supplementary Methods, Supplementary Fig. 2). To assign the metagenomic scaffolds to their phylotype origin, we used a combinatory binning approach based on intrinsic DNA signatures (see Supplementary Methods). Binning of the *Olavius* symbionts' metagenome resulted in the formation of four distinct clusters (Fig. 1). The presence of the corresponding rRNA operons, which represented the only rRNA operons found in these bins and the assembly, enabled us to identify them as the *O. algarvensis* symbionts δ1, δ4, γ1 and γ3 (Fig. 1, Supplementary Table 1;

[1]DOE Joint Genome Institute, Walnut Creek, California 94598, USA. [2]Lawrence Berkeley National Laboratory, Genomics Division, Berkeley, California 94720, USA. [3]Max Planck Institute for Marine Microbiology, 28359 Bremen, Germany. [4]International University Bremen, 28759 Bremen, Germany.

©2006 Nature Publishing Group

Supplementary Table 2 provides a comparison to other symbiont genomes). This study illustrates the usefulness of our nucleotide frequency method for the assignment of metagenomic scaffolds from a natural microbial community to a phylotype when using shotgun sequencing. It would not have been possible to reconstruct genome assemblies of the four symbionts purely on the basis of GC contents and scaffold read depth, and without a closely related, fully sequenced reference genome.

Symbiont bin assignments based on 16S rRNA genes were confirmed by phylogenetic analysis of predicted proteins within each cluster of scaffolds (Supplementary Fig. 3), and the clusters were furthermore supported by the distribution of 49 single-copy genes. Although the populations are not clonal, the frequencies of polymorphic sites found in the four symbiont bins, ranging from 0.01–0.1% (Supplementary Table 4), were rather low compared with environmental microbes such as those from acid mine drainage[8]. The following discussion describing the metabolism of the symbionts is based on the genes found in each symbiont bin. With a focus on capturing the core metabolic pathways present in the community, it is of minor importance if the genes within each bin originated from a single strain or represent a pan-genome of several very closely related strains[16].

Carbon and energy metabolism

Gammaproteobacterial symbionts. Chemoautotrophic symbionts feed their hosts by providing them with organic carbon from autotrophic CO_2 fixation driven by oxidation of reduced inorganic compounds such as sulphide. In agreement with previous studies indicating the chemoautotrophic, sulphur-oxidizing nature of the *O. algarvensis* γ1 symbiont[12], our analysis of the γ1 bin revealed the presence of genes required for autotrophic CO_2 fixation via the Calvin–Benson–Bassham cycle using type I ribulose 1,5-bisphosphate carboxylase–oxygenase (Rubisco; *cbb*), the oxidation of reduced sulphur compounds (genes such as *dsr*, *fcc* and *sox*), and the storage of sulphur in globules (*sgpB*, encoding one of three known sulphur-globule proteins).

Unexpectedly, our metagenomic analyses revealed that the γ3 symbiont of *O. algarvensis* is also a sulphur-oxidizing chemoautotroph. Several gutless oligochaete species are known to harbour γ3 symbionts but the metabolism of these bacteria was previously unknown and the benefit of harbouring additional Gammaproteobacteria is unclear. The nearly complete genomic sequence for the *O. algarvensis* γ3 symbiont obtained in this study is notable, as this is the only genome from a chemoautotrophic symbiont sequenced thus far. The γ3 bin carries all the genes required for a thiotrophic (sulphur-oxidizing) metabolism including those needed for the oxidation of reduced sulphur compounds (including *dsr*, *apr*, *sat*, *fcc* and *sox*) as well as autotrophic CO_2 fixation by means of genes closely related to but phylogenetically distinct from the γ1 symbiont (Fig. 2). The absence of genes encoding sulphur-globule proteins in the near-complete γ3 bin indicates that these symbionts do not store sulphur, supporting transmission electron microscopy analyses showing that only γ1 symbionts contain sulphur globules (O. Giere, personal communication). In addition to using oxygen as an electron acceptor, the presence of *nap* and *nir* gene clusters indicates that the γ3 symbionts couple oxidation of reduced sulphur compounds to dissimilatory nitrate reduction under oxygen-limiting conditions (Fig. 2). In deeper sediment layers, with neither oxygen nor nitrate, both Gammaproteobacteria have the ability to use fumarate as an electron acceptor for the oxidation of reduced sulphur compounds (Fig. 2).

Deltaproteobacterial symbionts. The presence of genes characteristic of dissimilatory sulphate reduction (such as *dsr*, *qmo* and *apr*) in both the δ1 and δ4 bins indicates that these symbionts are sulphate-reducing bacteria that use oxidized sulphur compounds such as sulphate as an electron acceptor, thereby producing sulphide (Fig. 2). In addition to the syntrophic cycling of sulphate and sulphide between the gamma- and delta-proteobacterial *O. algarvensis* symbionts[14], intermediate sulphur compounds such as tetrathionate and thiosulphate may also be cycled between the symbionts. The δ4 symbiont seems to be able to reduce sulphur compounds of intermediate oxidation states, as suggested by the presence of a multi-haem cytochrome most closely related to tetrathionate reductase of *Shewanella oneidensis*[17] and located in a chromosomal cluster with molybdopterin-dependent dehydrogenase related to thiosulphate reductase of *Wolinella succinogenes*. Cycling of intermediate sulphur compounds is energetically more favourable than the exchange of sulphide and sulphate, as shown previously in experiments with free-living sulphate reducers and sulphur oxidizers[18].

Heterotrophy is important in sulphate-reducing bacteria and correspondingly we found in the δ1 bin genes coding for the transport and utilization of a large variety of carbohydrate substrates, including uronic acids (glucuronate, galacturonate and fructuronate), xylose, fructose, dihydroxyacetone and polyols (mannitol, sorbitol and glycerol). Although all sulphate-reducing bacteria are heterotrophic only some fix CO_2, and it is intriguing that both deltaproteobacterial symbionts are capable of autotrophic carbon fixation via the reductive acetyl-coenzyme-A (CoA) pathway, as well as via the reductive tricarboxylic acid (TCA) cycle (Fig. 2). Thus, *O. algarvensis* has established an association with four symbionts that are all capable of providing it with organic carbon through three different autotrophic pathways.

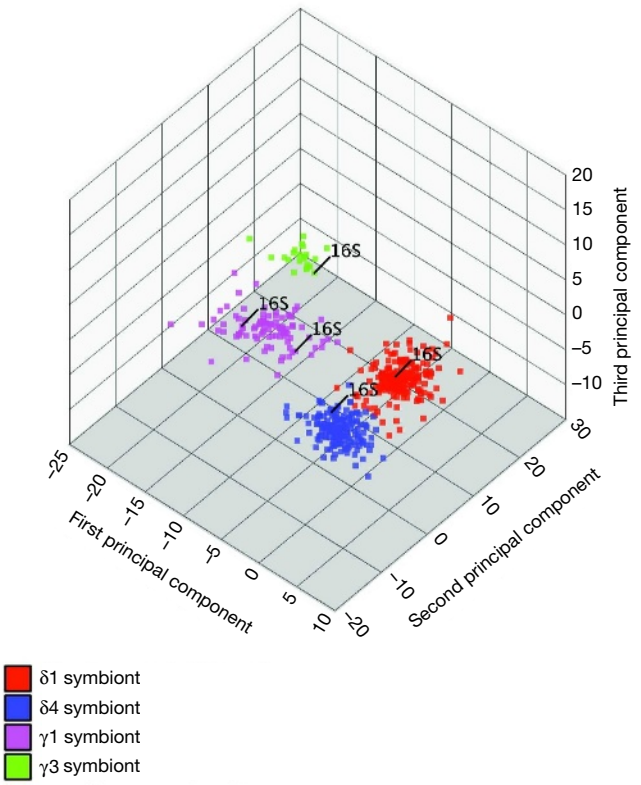


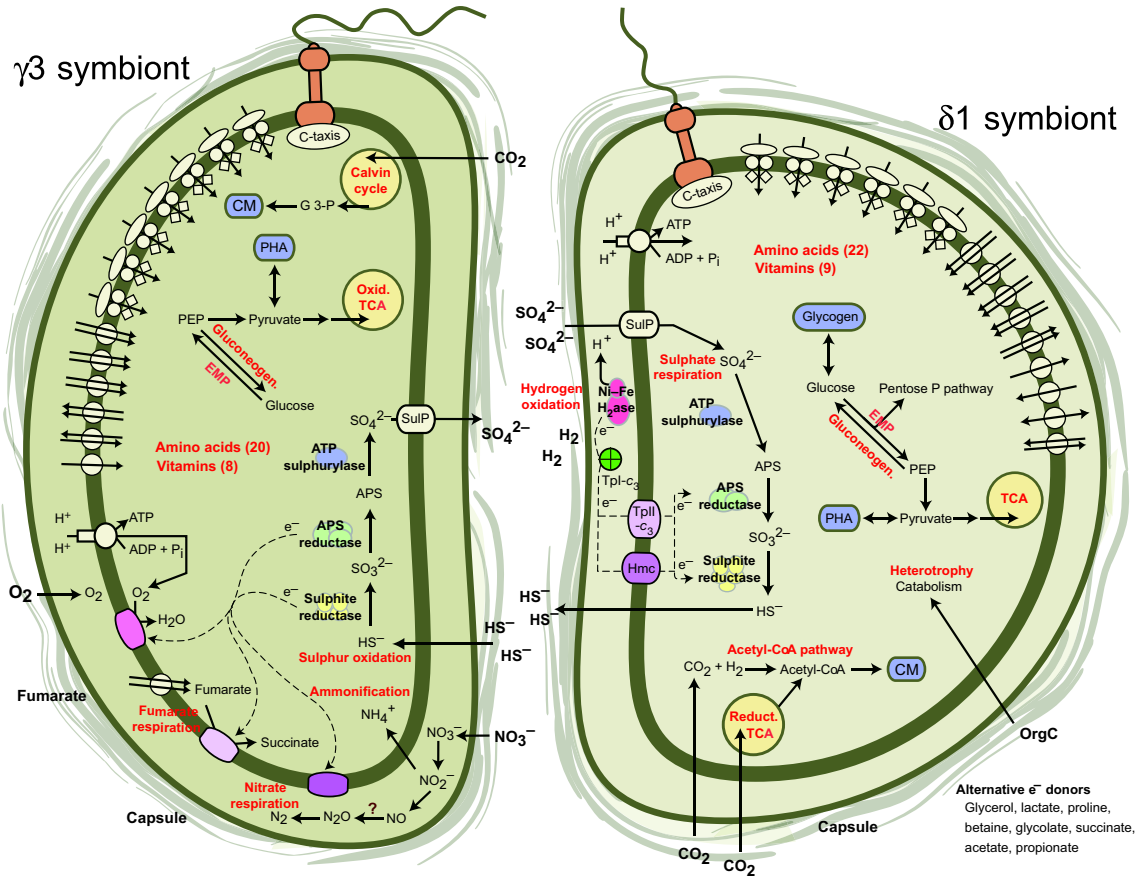
Figure 1 | Clustering of the *O. algarvensis* symbiont scaffolds. Visualization of the first three components of a principal component analysis, in which GC-content, net-read-depth *z*-scores for all possible 64 trinucleotides and 256 tetranucleotides were incorporated with equal weight (*z*-scores calculated with TETRA[39], and normalized by length). The colours represent the four clusters of scaffolds (calculated with MetaClust[35]) that were binned based on GC-content, dinucleotide relative abundance, Markov-model-based statistical evaluations of tri-, tetra- and penta-mer over- and under-representation, and normalized chaos game representations for tri- to hexamers; sequences <5 kilobases (kb) in length are not represented. Scaffolds containing 16S rRNA genes are indicated.

©2006 Nature Publishing Group

One of the most common electron donors for autotrophic sulphate-reducing bacteria is hydrogen. We found gene clusters encoding periplasmic Ni–Fe hydrogenases, a transmembrane high-molecular-weight cytochrome *c* (hmc) complex, and a type II tetrahaem cytochrome c_3 (TpII-c_3) in the bins of both sulphate-reducing symbionts, as well as a type I tetrahaem cytochrome c_3 (TpI-c_3) in the δ1 bin (Fig. 2). This is a compelling indication for the uptake and oxidation of molecular hydrogen using sulphate as an electron acceptor[19]. It is not clear if hydrogen is provided by the γ-symbionts. Within the γ3 bin, we found genes encoding a pyruvate ferredoxin oxidoreductase (POR), typically used in an alternative route for pyruvate oxidation[20] and indicative of hydrogen release from low-potential ferredoxins. Released hydrogen could subsequently be taken up by the sulphate-reducing symbionts leading to hydrogen syntrophy within the microbial consortium. Alternative electron donors to hydrogen include glycerol, lactate, proline and betaine, and potentially glycolate and other 2-hydroxy acids, as well as succinate, acetate and propionate.

Symbiont–host interactions

'Feeding' of the host. In other chemosynthetic associations, the symbionts provide their hosts with nutrition using either reduced



ABC transporters
Amino acids, amides
Peptides/nickel/opines/sugars
Polyamines/opines
Nitrate/cyanate/bicarbonate
Osmolytes
Taurine/sulphonates
Phosphonate
Phosphate
Molybdate
Iron
Vitamin B12
Manganese/zinc
Urea

Uniporters
Ammonium/ammonia (bidirectional)
Arsenite/antimonite (export)

Antiporters
Drugs (export); H^+ (import)
Amino acids (export); H^+ (import)

Symporters
TRAP
Betaine/carnitine/choline; H^+
Di/tricarboxylates; Na^+/H^+
Phosphate/sulphate; Na^+/H^+
Aromatic acids; H^+
Folate; H^+
Benzoate; H^+

ABC transporters
Peptides/nickel/opines/sugars
Amino acids, amides
Polyamines/opines
Nitrate/cyanate/bicarbonate
Mono/disaccharides
Phosphonate
Iron
Manganese/zinc
Vitamin B12
Cobalt
Phosphate
Osmolytes
Taurine/sulphonates

Antiporters
Drugs (export); H^+ (import)
Formate (export); oxalate (import)
Aspartate; alanine (reversible)
ATP; ADP (reversible)
Amino acids (export); H^+ (import)

Uniporters
Ammonium/ammonia (bidirectional)
Phenylpropionate (import)
Arsenite/antimonite (export)

Symporters
TRAP
Amino acids; H^+
Mono/di/tricarboxylates; Na^+/H^+
Phosphate/sulphate; Na^+/H^+
Organic acids; H^+
Aromatic acids
Folate; H^+

Figure 2 | Reconstruction of the symbionts' physiology. APS, adenosine 5′-phosphosulphate; CM, cell material; CoA, coenzyme A; C-taxis, chemotaxis; EMP, Embden–Meyerhof pathway; G 3-P, glyceraldehyde 3-phosphate; H_2ase, hydrogenase; Hmc, high-molecular-weight cytochrome *c*; OrgC, organic compounds; PEP, phosphoenolpyruvate; PHA, polyhydroxyalkanoates; TCA, tricyclic acid; TpI/II-c_3, type I/II tetrahaem cytochrome c_3; TRAP, tripartite ATP-independent periplasmic. SulP, sulphate permease. A question mark (?) indicates the lack of nitric oxide reductase in the γ3 genome bin; numbers in parentheses indicate the numbers of amino acids/vitamin biosynthesis pathways found (Supplementary Table 5).

©2006 Nature Publishing Group

sulphur compounds or methane as their energy source[21]. In the *O. algarvensis* symbiosis, both reduced sulphur compounds and hydrogen can be used as energy sources, and all four symbionts have the potential to fix CO_2 into organic carbon via autotrophy. In addition, the sulphate-reducing symbionts can feed the oligochaete host through heterotrophy by taking up dissolved organic carbon compounds from the environment. Almost all amino acids and a variety of vitamins can be synthesized by the symbionts to provide their host with these required nutrients (Supplementary Table 5). Nutrient transfer to the host is likely to occur via intracellular uptake and digestion of the bacteria, as the number of genes encoding amino-acid exporters was not elevated in the symbionts when compared to those of free-living bacteria, and the only known family of sugar exporters[22] was not encoded in any of the symbiont bins. This conclusion is supported by morphological analyses showing symbiont lysis in the basal region of the worm's epidermis[23].

Host waste recycling. The reduction of nephridia in the oligochaete host, used for the excretion of nitrogenous waste compounds and osmoregulation, suggests that its symbionts have adapted to carry out these functions. Most aquatic organisms excrete ammonium, but urea—which also functions as a common organic osmolyte—has also been found in marine worms with symbionts, such as the hydrothermal vent worm *Riftia pachyptila*[24]. The genome bins for the δ1, δ4 and γ3 symbionts encode a bidirectional uniporter (Amt family transporter[25]) for the counter-ion-independent and energy-independent uptake of ammonium. A probable urea ABC transporter for urea uptake is present in the γ3 bin, adjacent to a urease operon that encodes genes involved in urea hydrolysis. Ammonium and urea uptake by the symbionts would not only aid the host in the removal of these toxic waste products, but also lead to the conservation of valuable nitrogen by the symbionts (Fig. 3).

Marine invertebrates are typical osmoconformers, maintaining their cell volume primarily with organic osmoregulatory compounds such as amino acids, taurine, glycine betaine, trimethylamine *N*-oxide (TMAO) and urea, as well as polyols and sugars[26]. The presence of gene clusters encoding proteins used for taurine and glycine betaine import and catabolism in the γ3 bin may indicate the use of these osmolytes as both a carbon and nitrogen source[27] (Fig. 3). Furthermore, we found pathways for TMAO degradation in the *O. algarvensis* symbionts. Microbial TMAO degradation includes its conversion to trimethylamine by TMAO reductase, and further demethylation of trimethylamine either by trimethylamine and dimethylamine dehydrogenases found in Bacteria[28] or by trimethylamine, dimethylamine and monomethylamine methyltransferases found mostly in Archaea[29]. Genes coding for the key enzymes in both pathways were present in the γ3 and δ1 symbionts, suggestive of their TMAO catabolic activity (Fig. 3). The availability of the osmolyte TMAO would, furthermore, be particularly advantageous for the sulphur-oxidizing symbionts that could use this organic carbon compound as an alternate electron acceptor in the absence of oxygen or nitrate. Although the γ3 bin encodes enzymes likely to be involved in the reduction of TMAO, their specificity for this substrate could not be clearly identified.

Polyamines are essential in all organisms for DNA stabilization, DNA replication and cell proliferation[30], and represent additional products of host protein breakdown. We found abundant gene clusters encoding ABC transporters for the uptake of the polyamines putrescine and spermidine in the δ1 bin, and putrescine in the γ3 bin (Fig. 3).

Finally, evidence is provided for the recycling of host fermentation waste products such as the dicarboxylate succinate, as well as the monocarboxylates acetate and propionate. Pathways for their utilization and a variety of potential dicarboxylate transporters, and probable monocarboxylate transporters, were found in all four symbiont bins. The δ1 bin encodes 23 tripartite ATP-independent periplasmic (TRAP)-T family dicarboxylate transporters[31], some of which are likely to be involved in monocarboxylate and dicarboxylate transport (Fig. 3).

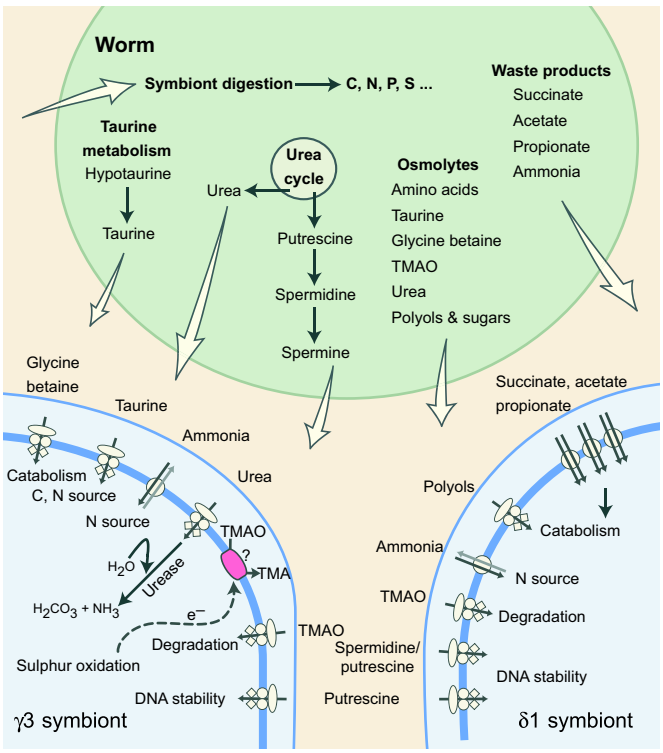


Figure 3 | Reconstruction of symbiont–host interactions. The metagenomic data uncovered pathways for the uptake and recycling of organic osmolytes and excretion products of the worm by its symbionts. TMA, trimethylamine; TMAO, trimethylamine *N*-oxide.

A mutually obligate relationship?

The lack of a digestive and excretory system in *O. algarvensis* means that its symbionts are crucial for its survival. But is this relationship mutually obligate or can the symbionts survive outside of the host in a free-living stage? Several pieces of evidence from the sequence data and the metabolic reconstruction analyses suggest a lack of evidence for obligate host-dependence of any of the symbionts. This includes the observation that the genomes of the extracellular bacteria do not show AT-bias or genome size reduction (Supplementary Table 2) and there was no obvious loss of essential metabolic pathways in the δ1 or γ3 bins suggestive of host-dependence, as is the case for many obligate host-associated bacteria[32,33]. Finally, we found genes required for cell motility via a flagellum in the δ1, δ4 and γ3 bins. Although this evidence suggests that the symbionts associated with the worm may have a free-living stage, the presence of a remarkable number of transposases in the γ3 and γ1 symbiont bins (7.5% and 20.5%, respectively; Supplementary Tables 4 and 6) indicates that these symbionts may be in transition to an obligate symbiotic lifestyle. Bacteria that have recently evolved into obligate symbionts show an increase in frequency of mobile elements, representing a source for chromosomal rearrangements and gene inactivation[34] (for symbiont transmission hypotheses, see Supplementary Information).

From the metagenome to the environment

The oligochaete symbiosis is inseparably linked to the geochemical properties of the environment, needing access to both reduced and oxidized compounds for energy production and net carbon fixation (Fig. 4). In the upper oxic sediment layers of the *O. algarvensis* habitat, where no sulphide is present, both the γ1 and the γ3 symbiont can use reduced sulphur compounds produced internally by the sulphate-reducing symbionts in a syntrophic sulphur cycle. (We have shown previously that the sulphate-reducing symbionts can produce sulphide in *O. algarvensis* under microaerobic conditions comparable

©2006 Nature Publishing Group

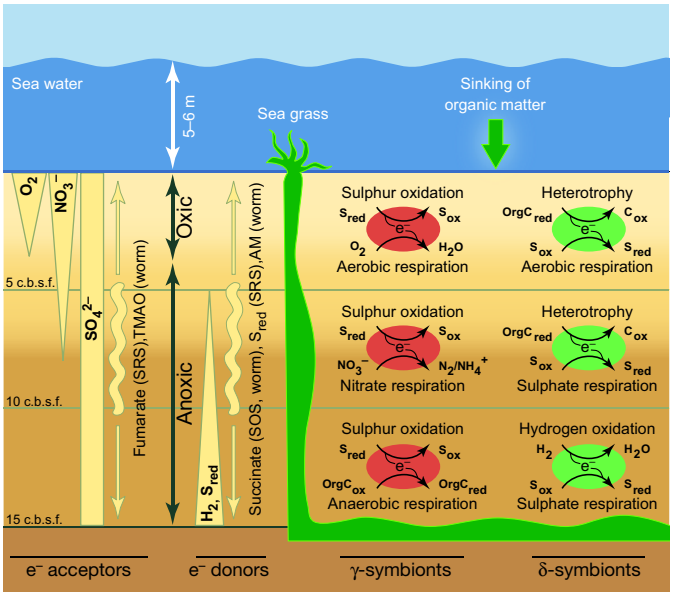


Figure 4 | Model for energy metabolism in the symbiosis. *O. algarvensis* inhabits shallow Mediterranean ocean sediments (5–15-cm depth). Electron acceptors and donors are available to the symbionts in all sediment layers, with some supplied from the environment (shown in the yellow rectangle/triangles) and some internally (shown next to the yellow worms). Carbon is gained autotrophically using reduced sulphur compounds (γ-symbionts) or hydrogen (δ-symbionts), as well as heterotrophically (δ-symbionts). AM, anaerobic metabolites; c.b.s.f., centimetres below sea floor; OrgC, organic compounds; S_{ox}, oxidized sulphur compounds; S_{red}, reduced sulphur compounds; SOS, sulphur-oxidizing symbionts; SRS, sulphate-reducing symbionts; TMAO, trimethylamine *N*-oxide.

to those in the lower oxic zone at 2–5-cm sediment depth[14].) The γ1 symbiont can also gain energy independently of the sulphate-reducing symbionts by oxidizing the large supply of sulphur stored in its cytoplasm. For both gammaproteobacterial symbionts, oxygen would be the energetically most-favourable electron acceptor (Fig. 4).

As the worm migrates downwards, it encounters sediment layers in which oxygen is no longer present. Under these conditions, the γ3 symbiont can use nitrate from the environment for the oxidation of reduced sulphur compounds. Given the extremely low concentrations of sulphide in this layer as well as in the deeper reduced sediment layers of the Elba habitat (in the low nM range[14]), it is likely that the sulphate-reducing symbionts provide a large part of the reduced sulphur compounds for the γ-symbionts under most conditions.

In the deeper sediment layers characterized by reducing conditions and the absence of oxygen or nitrate, hydrogen oxidation by the sulphate-reducing symbionts may occur, in which hydrogen is used as an energy source for the autotrophic fixation of inorganic carbon. As hydrogen concentrations are commonly very low in most marine sediments, heterotrophic pathways should also have an important role for the deltaproteobacterial symbionts under these reducing conditions. Although energetically less favourable than nitrate or oxygen, organic electron acceptors including TMAO (host-derived) and fumarate (produced by the δ-symbionts) are provided internally within the symbiosis and may be used by the γ-symbionts for the oxidation of reduced sulphur compounds. This would enable the γ1 symbiont to replenish its sulphur stores, which could be fully oxidized using the energetically more favourable electron acceptor oxygen when the worm returns to the oxic zone. Fumarate respiration by the γ-symbionts would produce succinate, which could be used by the deltaproteobacterial symbionts as an energy source and reoxidized to fumarate, thus leading to a syntrophic cycling of reductants and oxidants.

Our analysis of the *O. algarvensis* microbial symbiont genomes has provided insights into how resources are used and shared among the different symbionts and with their host, and how different metabolic pathways are used by the symbionts to generate energy as the worm migrates through the chemocline. We have shown how the *O. algarvensis* symbiosis is unique among known chemosynthetic associations, as reductants and oxidants are not only supplied from the environment but also internally to drive energy production. Thus, this comprehensive metagenomic analysis shows that these highly integrated synergistic assemblages of multiple bacterial partners provide their eukaryotic host with an optimal energy supply and waste management through resource partitioning and cooperation during syntrophic cycling of oxidized and reduced compounds.

METHODS

Further details for all methods used in this study are provided in Supplementary Information.

O. algarvensis specimens were collected off Capo di Sant' Andrea, Elba, Italy. Metagenomic libraries were constructed from 200 pooled *O. algarvensis* specimens per library. A small-insert pMCL200 library was made from a nycodenz-separated, symbiont-enriched sample and two pCC1Fos fosmid libraries were constructed using the CopyControl Fosmid Library Construction kit (Epicentre). 16S rRNA polymerase chain reaction (PCR) libraries were created from the DNA sources used for each of the libraries and approximately 384 clones sequenced and analysed. From the shotgun libraries, we created 204 megabases (Mb) of vector- and quality-trimmed sequence. This data was assembled using JAZZ, resulting in a set of 2,286 scaffolds, which were binned using a combinatory approach based on dimer-to-hexamer frequencies using the newly developed MetaClust program[35]. Final clusters with 511 scaffolds were verified by phylogenetic affiliation of each scaffold based on the most common phylogeny of its predicted proteins using a bayesian classifier, and by checking for paralogues of 49 genes that typically occur with only one copy per genome (V. Kunin *et al.*, unpublished data). To assess nucleotide sequence variation within the bins, we analysed the multiple alignment of the JAZZ assembly. Potential open reading frames (ORFs) were identified using 'mORFind' (J. Waldmann, H.T. and F.O.G., unpublished data) and annotation performed with the GenDB v2.2 system[36] and MicHanThi[37]. The annotated symbiont metagenome was loaded into the metagenomics version of Integrated Microbial Genomes/M (IMG/M; http://img.jgi.doe.gov/m)[38].

Received 5 May; accepted 29 August 2006.
Published online 17 September; corrected 26 October 2006.

1. Margulis, L. *Symbiosis in Cell Evolution* (W. H. Freeman, New York, 1993).
2. Ruby, E. G., Henderson, B. & McFall-Ngai, M. We get by with a little help from our (little) friends. *Science* **303**, 1305–1307 (2004).
3. DeLong, E. F. Microbial population genomics and ecology. *Curr. Opin. Microbiol.* **5**, 520–524 (2002).
4. Handelsman, J. Metagenomics: application of genomics to uncultured microorganisms. *Microbiol. Mol. Biol. Rev.* **68**, 669–685 (2004).
5. Riesenfeld, C. S., Schloss, P. D. & Handelsman, J. Metagenomics: genomic analysis of microbial communities. *Annu. Rev. Genet.* **38**, 525–552 (2004).
6. DeLong, E. F. *et al.* Community genomics among stratified microbial assemblages in the ocean's interior. *Science* **311**, 496–503 (2006).
7. Hallam, S. J. *et al.* Reverse methanogenesis: testing the hypothesis with environmental genomics. *Science* **305**, 1457–1462 (2004).
8. Tyson, G. W. *et al.* Community structure and metabolism through reconstruction of microbial genomes from the environment. *Nature* **428**, 37–43 (2004).
9. Venter, J. C. *et al.* Environmental genome shotgun sequencing of the Sargasso Sea. *Science* **304**, 66–74 (2004).
10. Tringe, S. G. *et al.* Comparative metagenomics of microbial communities. *Science* **308**, 554–557 (2005).
11. Wu, D. *et al.* Metabolic complementarity and genomics of the dual bacterial symbiosis of sharpshooters. *PLoS Biol.* **4**, e188 (2006).
12. Dubilier, N., Blazejak, A. & Ruhland, C. Symbioses between bacteria and gutless marine oligochaetes. *Prog. Mol. Subcell. Biol.* **41**, 251–275 (2006).
13. Blazejak, A., Erseus, C., Amann, R. & Dubilier, N. Coexistence of bacterial sulfide oxidizers, sulfate reducers, and spirochetes in a gutless worm (Oligochaeta) from the Peru margin. *Appl. Environ. Microbiol.* **71**, 1553–1561 (2005).
14. Dubilier, N. *et al.* Endosymbiotic sulphate-reducing and sulphide-oxidizing bacteria in an oligochaete worm. *Nature* **411**, 298–302 (2001).
15. Aparicio, S. *et al.* Whole-genome shotgun assembly and analysis of the genome of *Fugu rubripes*. *Science* **297**, 1301–1310 (2002).
16. Medini, D., Donati, C., Tettelin, H., Masignani, V. & Rappuoli, R. The microbial pan-genome. *Curr. Opin. Genet. Dev.* **15**, 589–594 (2005).
17. Mowat, C. G. *et al.* Octaheme tetrathionate reductase is a respiratory enzyme with novel heme ligation. *Nature Struct. Mol. Biol.* **11**, 1023–1024 (2004).

©2006 Nature Publishing Group

18. van den Ende, F. P., Meier, J. & van Gemerden, H. Syntrophic growth of sulfate-reducing bacteria and colorless sulfur bacteria during oxygen limitation. *FEMS Microbiol. Ecol.* **23**, 65–80 (1997).
19. Matias, P. M., Pereira, I. A., Soares, C. M. & Carrondo, M. A. Sulphate respiration from hydrogen in *Desulfovibrio* bacteria: a structural biology overview. *Prog. Biophys. Mol. Biol.* **89**, 292–329 (2005).
20. Kletzin, A. & Adams, M. W. Molecular and phylogenetic characterization of pyruvate and 2-ketoisovalerate ferredoxin oxidoreductases from *Pyrococcus furiosus* and pyruvate ferredoxin oxidoreductase from *Thermotoga maritima*. *J. Bacteriol.* **178**, 248–257 (1996).
21. Cavanaugh, C. M., McKiness, Z. P., Newton, I. L. G. & Stewart, F. J. Marine chemosynthetic symbioses. In *The Prokaryotes: A Handbook on the Biology of Bacteria* (eds Dworkin, M. *et al.*) (Springer, New York, 2006).
22. Liu, J. Y., Miller, P. F., Gosink, M. & Olson, E. R. The identification of a new family of sugar efflux pumps in *Escherichia coli*. *Mol. Microbiol.* **31**, 1845–1851 (1999).
23. Giere, O. & Erseus, C. Taxonomy and new bacterial symbioses of gutless marine Tubificidae (Annelida, Oligochaete) from the island of Elba (Italy). *Org. Divers. Evol.* **2**, 289–297 (2002).
24. De Cian, M., Regnault, M. & Lallier, F. H. Nitrogen metabolites and related enzymatic activities in the body fluids and tissues of the hydrothermal vent tubeworm *Riftia pachyptila*. *J. Exp. Biol.* **203**, 2907–2920 (2000).
25. Khademi, S. *et al.* Mechanism of ammonia transport by Amt/MEP/Rh: structure of AmtB at 1.35 Å. *Science* **305**, 1587–1594 (2004).
26. Yancey, P. H., Blake, W. R. & Conley, J. Unusual organic osmolytes in deep-sea animals: adaptations to hydrostatic pressure and other perturbants. *Comp. Biochem. Physiol. A* **133**, 667–676 (2002).
27. Denger, K., Ruff, J., Schleheck, D. & Cook, A. M. *Rhodococcus opacus* expresses the *xsc* gene to utilize taurine as a carbon source or as a nitrogen source but not as a sulfur source. *Microbiology* **150**, 1859–1867 (2004).
28. Yang, C. C., Packman, L. C. & Scrutton, N. S. The primary structure of *Hyphomicrobium* X dimethylamine dehydrogenase. Relationship to trimethylamine dehydrogenase and implications for substrate recognition. *Eur. J. Biochem.* **232**, 264–271 (1995).
29. Paul, L., Ferguson, D. J. Jr & Krzycki, J. A. The trimethylamine methyltransferase gene and multiple dimethylamine methyltransferase genes of *Methanosarcina barkeri* contain in-frame and read-through amber codons. *J. Bacteriol.* **182**, 2520–2529 (2000).
30. Cohen, S. S. *A Guide to the Polyamines* (Oxford Univ. Press, New York, 1998).
31. Kelly, D. J. & Thomas, G. H. The tripartite ATP-independent periplasmic (TRAP) transporters of bacteria and archaea. *FEMS Microbiol. Rev.* **25**, 405–424 (2001).
32. Moran, N. A. Microbial minimalism: genome reduction in bacterial pathogens. *Cell* **108**, 583–586 (2002).
33. Moran, N. A. Tracing the evolution of gene loss in obligate bacterial symbionts. *Curr. Opin. Microbiol.* **6**, 512–518 (2003).
34. Moran, N. A. & Plague, G. R. Genomic changes following host restriction in bacteria. *Curr. Opin. Genet. Dev.* **14**, 627–633 (2004).
35. Huntemann, M. *MetaClust—Entwicklung eines modularen Programms zum Clustern von Metagenomfragmenten anhand verschiedener intrinsischer DNA-Signaturen*. Diploma thesis, Univ. Bremen (2006).
36. Meyer, F. *et al.* GenDB—an open source genome annotation system for prokaryote genomes. *Nucleic Acids Res.* **31**, 2187–2195 (2003).
37. Quast, C. *MicHanThi—Design and Implementation of a System for the Prediction of Gene Functions in Genome Annotation Projects*. Diploma thesis, Univ. Bremen (2006).
38. Markowitz, V. M. *et al.* The integrated microbial genomes (IMG) system. *Nucleic Acids Res.* **34**, D344–D348 (2006).
39. Teeling, H., Waldmann, J., Lombardot, T., Bauer, M. & Glockner, F. O. TETRA: a web-service and a stand-alone program for the analysis and comparison of tetranucleotide usage patterns in DNA sequences. *BMC Bioinformatics* **5**, 163 (2004).

Supplementary Information is linked to the online version of the paper at www.nature.com/nature.

Acknowledgements This work was performed under the auspices of the US Department of Energy's Office of Science, Biological and Environmental Research Program, and supported by the University of California Lawrence Livermore National Laboratory, Lawrence Berkeley National Laboratory, and Los Alamos National Laboratory, and the Max Planck Society. We thank members of the Rubin, J. Bristow and P. Hugenholtz laboratories, as well as members of the Joint Genome Institute, for their contributions. We thank V. Markowitz, E. Dalin, N. Putnam, R. Sorek, T. Glavina del Rio, A. Salamov, A. Kobayashi and K. Kellaris for their assistance. At the Max Planck Institute for Marine Microbiology we thank S. Wetzel for technical assistance. We are grateful to C. Lott and the staff of the HYDRA field station at Elba for their generous support and help in sampling the worms.

Author Information The assembled sequences from the *Olavius* symbionts' metagenome have been deposited into the NCBI database under the project accession number AASZ00000000. The annotated *Olavius* symbionts' bins were incorporated into the metagenomics version of the US Department of Energy Joint Genome Institute Integrated Microbial Genomes/M (IMG/M; http://img.jgi.doe.gov/m). Reprints and permissions information is available at www.nature.com/reprints. The authors declare no competing financial interests. Correspondence and requests for materials should be addressed to N.D. (ndubilie@mpi-bremen.de) or E.M.R. (EMRubin@lbl.gov).

©2006 Nature Publishing Group