Discovery of chemoautotrophic symbiosis in the giant shipworm *Kuphus polythalamia* (Bivalvia: Teredinidae) extends wooden-steps theory

Daniel L. Distel[a,1], Marvin A. Altamia[b], Zhenjian Lin[c], J. Reuben Shipway[a], Andrew Han[d], Imelda Forteza[b], Rowena Antemano[b], Ma. Gwen J. Peñaflor Limbaco[b], Alison G. Tebo[e], Rande Dechavez[f], Julie Albano[f], Gary Rosenberg[g], Gisela P. Concepcion[b,h], Eric W. Schmidt[c], and Margo G. Haygood[c,1]

[a]Ocean Genome Legacy Center, Department of Marine and Environmental Science, Northeastern University, Nahant, MA 01908; [b]Marine Science Institute, University of the Philippines, Diliman, Quezon City 1101, Philippines; [c]Department of Medicinal Chemistry, University of Utah, Salt Lake City, UT 84112; [d]Second Genome, South San Francisco, CA 94080; [e]Pasteur, Département de Chimie, École Normale Supérieure, PSL Research University, Sorbonne Universités, Pierre and Marie Curie University Paris 06, CNRS, 75005 Paris, France; [f]Sultan Kudarat State University, Tacurong City 9800, Sultan Kudarat, Philippines; [g]Academy of Natural Sciences of Drexel University, Philadelphia, PA 19103; and [h]Philippine Genome Center, University of the Philippines System, Diliman, Quezon City 1101, Philippines

Edited by Margaret J. McFall-Ngai, University of Hawaii at Manoa, Honolulu, HI, and approved March 21, 2017 (received for review December 15, 2016)

The "wooden-steps" hypothesis [Distel DL, et al. (2000) *Nature* 403:725–726] proposed that large chemosynthetic mussels found at deep-sea hydrothermal vents descend from much smaller species associated with sunken wood and other organic deposits, and that the endosymbionts of these progenitors made use of hydrogen sulfide from biogenic sources (e.g., decaying wood) rather than from vent fluids. Here, we show that wood has served not only as a stepping stone between habitats but also as a bridge between heterotrophic and chemoautotrophic symbiosis for the giant mud-boring bivalve *Kuphus polythalamia*. This rare and enigmatic species, which achieves the greatest length of any extant bivalve, is the only described member of the wood-boring bivalve family Teredinidae (shipworms) that burrows in marine sediments rather than wood. We show that *K. polythalamia* harbors sulfur-oxidizing chemoautotrophic (thioautotrophic) bacteria instead of the cellulolytic symbionts that allow other shipworm species to consume wood as food. The characteristics of its symbionts, its phylogenetic position within Teredinidae, the reduction of its digestive system by comparison with other family members, and the loss of morphological features associated with wood digestion indicate that *K. polythalamia* is a chemoautotrophic bivalve descended from wood-feeding (xylotrophic) ancestors. This is an example in which a chemoautotrophic endosymbiosis arose by displacement of an ancestral heterotrophic symbiosis and a report of pure culture of a thioautotrophic endosymbiont.

symbiosis | shipworm | thioautotrophy | Teredinidae | chemoautotrophy

Although *Kuphus polythalamia* (Fig. 1 *A*–*C*) is a member of the common wood-boring and wood-feeding bivalve family Teredinidae (commonly known as shipworms), it is among the most infrequently observed and least understood of extant bivalves. Its enormous size—specimens may reach 155 cm in length and 6 cm in diameter—its unusual anatomy, and the fact that adult specimens are found burrowing in marine sediments rather than in wood, set it apart from other members of the family (1, 2). Because living specimens of *K. polythalamia* have not been described in scientific literature, and no peer-reviewed analysis of its anatomy has been published, the life history, habits, habitat, anatomy, and mode of nutrition of this remarkable species remain to be fully described and explained.

Typical shipworms are small worm-like bivalves with elongate bodies and rasp-like shells used to burrow in wood (Fig. 1*D*) (2). As they burrow, they ingest and digest the excavated wood particles with the aid of enzymes produced by heterotrophic bacterial endosymbionts maintained in specialized cells in the shipworm's gills (Fig. 1*D*) (3–6). These bacteria include *Teredinibacter turnerae* (7, 8) and other closely related but as-yet-unnamed endosymbionts (3–5, 9). At least one shipworm species has been shown to grow and reproduce normally with wood as its sole food source (10).

Although few other marine invertebrates are known to consume wood as food, an increasing number are believed to use waste products associated with microbial degradation of wood on the seafloor. Among these waste products is hydrogen sulfide, produced by bacteria that respire sulfate (the fourth most abundant solute in seawater) as an alternative to oxygen under anaerobic conditions (11, 12). A number of marine invertebrates can use this hydrogen sulfide with the aid of sulfide-oxidizing chemoautotrophic (thioautotrophic) symbionts. The symbionts reoxidize this waste sulfide back to sulfate, using the resulting energy to fix inorganic carbon into organic compounds that can be used by the host. In this way, wood deposits on the seafloor may support significant sulfur-based chemoautotrophic communities like those found at hydrothermal vents and cold water seeps (13–19). Indeed, it has been suggested that wood deposits may have acted as ecological stepping-stones, introducing some thioautotrophic animals to deep-sea hydrothermal vents and facilitating their migration between vent sites (13, 15,

Significance

Certain marine invertebrates harbor chemosynthetic bacterial symbionts, giving them the remarkable ability to consume inorganic chemicals such as hydrogen sulfide (H_2S) rather than organic matter as food. These chemosynthetic animals are found near geochemical (e.g., hydrothermal vents) or biological (e.g., decaying wood or large animal carcasses) sources of H_2S on the seafloor. Although many such symbioses have been discovered, little is known about how or where they originated. Here, we demonstrate a new chemosynthetic symbiosis in the giant teredinid bivalve (shipworm) *Kuphus polythalamia* and show that this symbiosis arose in a wood-eating ancestor via the displacement of ancestral cellulolytic symbionts by sulfur-oxidizing invaders. Here, wood served as an evolutionary stepping stone for a dramatic transition from heterotrophy to chemoautotrophy.

Author contributions: D.L.D., G.R., E.W.S., and M.G.H. designed research; D.L.D., M.A.A., Z.L., J.R.S., A.H., I.F., R.A., M.G.J.P.L., R.D., J.A., G.P.C., E.W.S., and M.G.H. performed research; Z.L., A.G.T., E.W.S., and M.G.H. contributed new reagents/analytic tools; D.L.D., Z.L., A.G.T., E.W.S., and M.G.H. analyzed data; and D.L.D. wrote the paper.

The authors declare no conflict of interest.

This article is a PNAS Direct Submission.

Freely available online through the PNAS open access option.

Data deposition: The DNA sequences and projects reported in this paper have been deposited in the GenBank database (accession nos. KY643660–KY643671, CP019936, MWSE00000000, MWSD00000000, MWSC00000000, MWSB00000000, and MWSA00000000) (*SI Appendix*, Table S6).

[1]To whom correspondence should be addressed. Email: disteld@gmail.com or margo.haygood@gmail.com.

This article contains supporting information online at www.pnas.org/lookup/suppl/doi:10.1073/pnas.1620470114/-/DCSupplemental.

www.pnas.org/cgi/doi/10.1073/pnas.1620470114

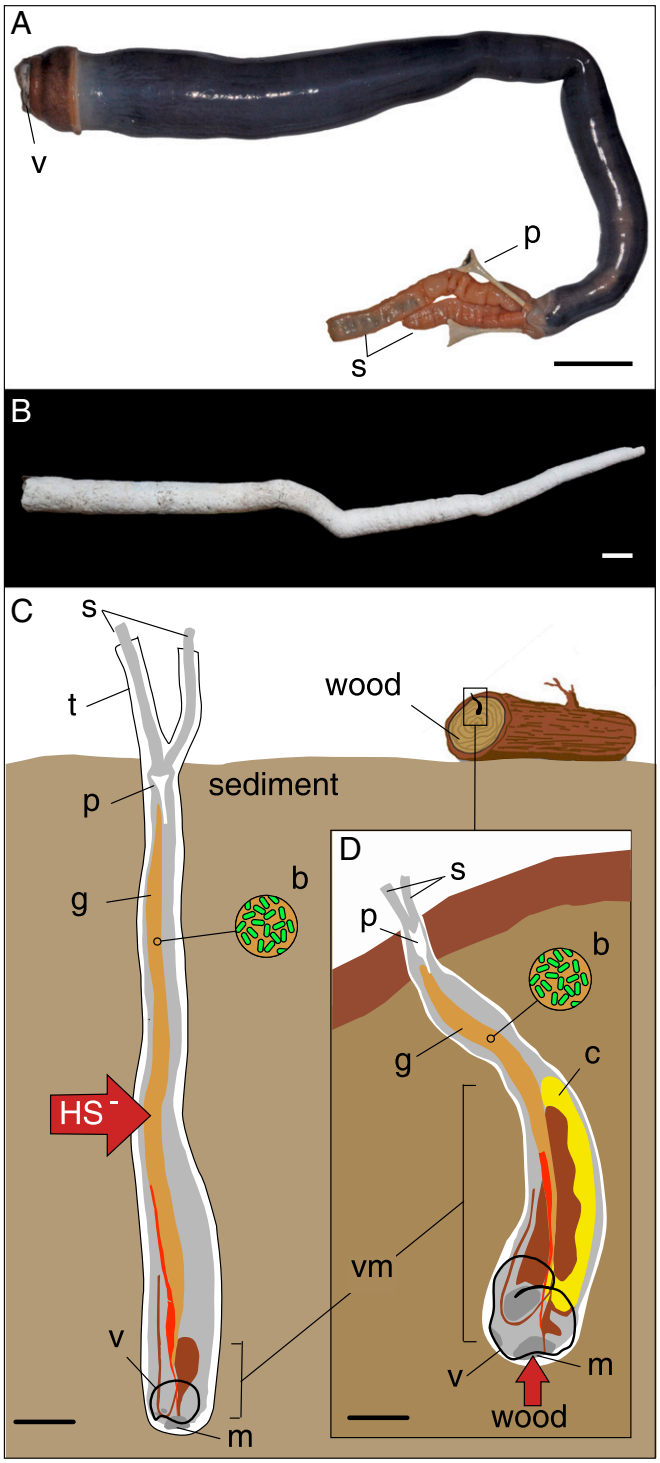

Fig. 1. Comparative anatomy and life position of *Kuphus polythalamia* and *Lyrodus pedicellatus*. (*A*) Fresh specimen of *K. polythalamia* (PMS-1672Y) removed from its calcareous tube, (*B*) calcareous tube of *K. polythalamia* (PMS-1674K) removed from sediment, (*C*) diagram depicting the anatomy and life position of *K. polythalamia* in sediment, and (*D*) *Inset* from box in *C* depicting the anatomy and life position of the wood-feeding shipworm *Lyrodus pedicellatus* in wood. (Scale bars: *A*–*C*, 5.0 cm; *D*, 0.5 cm.) b, bacteria; c, cecum; g, gill; HS^-, hydrogen sulfide; m, mouth; p, pallet; s, siphon; t, calcareous tube; v, valve (shell); vm, visceral mass. Movie S1 shows a specimen of *K. polythalamia* being removed from its tube and dissected.

19–21). Here, we ask whether *Kuphus polythalamia* might use a similar thioautotrophic strategy to harvest energy from wood, and whether wood has acted as an evolutionary stepping-stone to thioautotrophic endosymbiosis in this species.

Results and Discussion

Unlike other Teredinidae, which are thought to be obligate wood borers and wood feeders (22), *K. polythalamia* has been variously reported to burrow into mud and/or rotting wood. The specimens that we describe here were found burrowing in black, organic-rich sediment in a marine bay with high input of wood and woody plant debris. Only trace quantities of fecal matter were seen in the digestive systems of these specimens, and, as others have observed (23–26), the anterior ends of the tubes were sealed by a calcareous cap that covers the excavation face of the burrow. Although this cap must be resorbed periodically to allow the animal to grow and extend its tube, when present, it covers the mouth and prevents excavation and ingestion of sediments. These specimens also lack the large sac-like cecum, found in all other Teredinidae, in which wood particles accumulate during digestion (Fig. 1*D*). Moreover, the visceral mass and the posterior adductor muscles, which respectively house the digestive organs and power the burrowing action of the shells, are small and poorly developed compared with other shipworm species (2). Taken together, these morphological observations suggest that neither ingestion of wood, sediments, nor filtered particulate food likely explains the large body mass achieved by adults of this species.

If not wood, what then is the primary nutrient source for *K. polythalamia*? We reasoned that *K. polythalamia* might harbor thioautotrophic endosymbionts rather than, or in addition to, the heterotrophic gill endosymbionts that aid wood digestion in other Teredinidae, allowing them to exploit biogenic hydrogen sulfide produced by decaying wood and other organic materials.

To address this hypothesis, we first used scanning and transmission electron microscopy to demonstrate that the gills of *K. polythalamia*, like those of previously examined shipworm species (4, 5, 27, 28), contain abundant Gram-negative bacteria (Fig. 2 *A*–*E*). These bacteria are similar in size to those of other Teredinidae but contain features, including apparent sulfur globules and carboxysomes, that are commonly observed in thioautotrophic bacteria (Fig. 2*E*). Prominent intracytoplasmic membranes reminiscent of those found in methylotrophic and ammonia-oxidizing bacteria were also observed, although their function remains unknown.

These features suggested that the endosymbionts of *K. polythalamia* might be phylogenetically distinct from those of other Teredinidae. To explore this possibility, we used PCR with bacterial domain-selective primers (*SI Appendix*, Table S1) to amplify bacterial 16S rRNA genes from anterior, middle, and posterior regions of the gill and sequenced the resulting products by the Sanger method. A single, largely unambiguous sequence was obtained from all gill regions, suggesting a homogeneous bacterial community in which one or a few closely related bacterial types predominate. Phylogenetic analyses place this dominant symbiont type within a well-supported clade that includes the as-yet-uncultivated thioautotrophic symbionts of the marine ciliate protozoan *Zoothamnium niveum* (29), and those of the hydrothermal vent gastropods *Crysomallon squamiferum* and *Alvinoconcha hessleri* (30) (Fig. 3*A*). This clade is most closely related to the sulfur-oxidizing family Chromatiaceae and is nested within a larger clade containing most known thioautotrophic symbionts and many cultivated and uncultivated thioautotrophic gammaproteobacteria (*SI Appendix*, Fig. S2) (31). In contrast, *Teredinibacter turnerae* and other shipworm endosymbionts fall within the family Alteromonadaceae, a group largely composed of obligate aerobic heterotrophs (32).

We used selective culture media to isolate and grow thioautotrophic and xylotrophic bacteria from freshly homogenized gill tissue (*SI Appendix*, Table S4). Strains isolated in thioautotrophic medium include several that are >99% identical in 16S rRNA

sequence to that obtained by direct PCR amplification from the gill tissue of *K. polythalamia*. One of these, designated strain 2141T, was selected to represent these strains for further study. Similarly, multiple xylotrophic strains isolated from *K. polythalamia* are >99% identical in 16S rRNA sequence to known strains of *T. turnerae*.

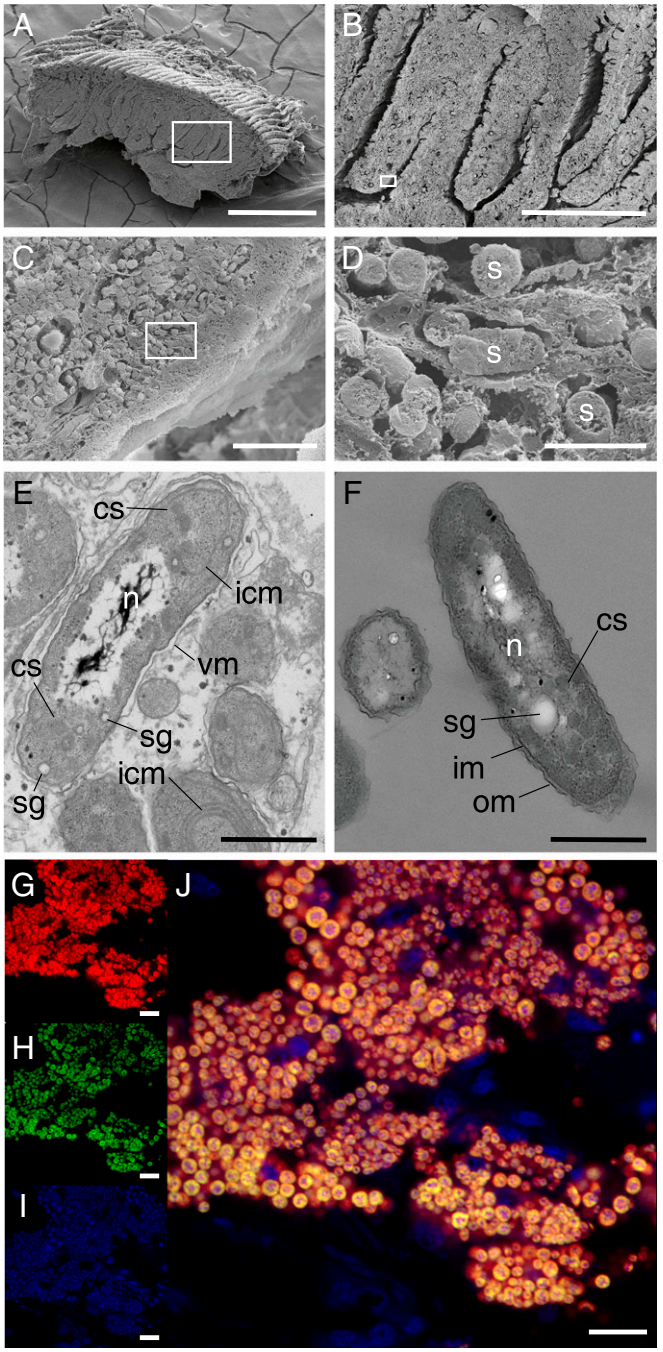

Fig. 2. Endosymbiont ultrastructure and localization. (*A*) Scanning electron micrograph of freeze-fractured gill; (*B*) *Inset* from *A*; (*C*) *Inset* from *B*; (*D*) *Inset* from *C*; (*E*) transmission electron micrograph (TEM) of endosymbionts in gill; (*F*) TEM of isolate 2141T. cs, carboxysomes; icm, intracytoplasmic membranes; im, inner membrane; om, outer membrane; s, symbionts; sg, sulfur globules; vm, vacuole membrane. (Scale bars: *A*, 500 μm; *B*, 100 μm; *C*, 10 μm; *D*, 2 μm; *E* and *F*, 0.5 μm.) (*G–J*) FISH images showing (*G*) bacterial probe Eub338 (Cy5, red); (*H*) symbiont probe Kp1 (Cy3, green); (*I*) nucleic acid stain (SYTOX, blue); and (*J*) merged images *G–I*. (Scale bars: *G–J*, 10 μm.) Negative controls are in *SI Appendix*, Fig. S1.

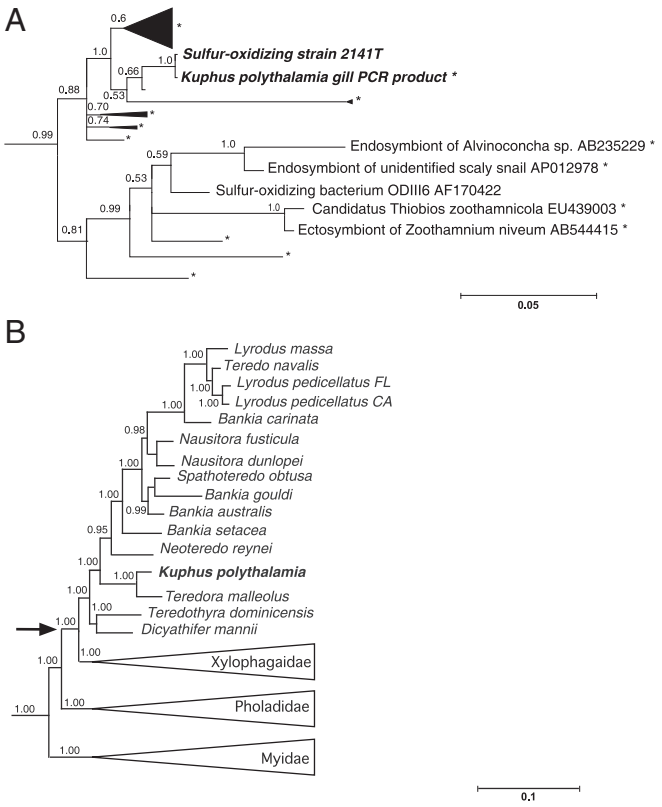

Fig. 3. Endosymbiont and host phylogeny. (*A*) Bayesian inference subtree for symbionts of *K. polythamia* and closely related bacteria (excerpted from *SI Appendix*, Fig. S2) based on partial 16S rRNA sequences (*uncultivated environmental sample). (*B*) Bayesian inference tree for the bivalve family Teredinidae and related families within the order Myida based on partial 18S and 28S rRNA sequences, excerpted from previously published work (22). Reproduced from ref. 22, with permission from Elsevier. Numbers at nodes indicate posterior probabilities. Scale bars denote nucleotide substitutions per site. Arrow in *B* indicates proposed acquisition of xylotrophic symbionts.

To localize these strains in host tissue, we performed fluorescence in situ hybridization (FISH) on paraffin-embedded gill tissue sections using 16S rRNA directed probes (*SI Appendix*, Table S2) selective for strain 2141T and *T. turnerae* in combination with a bacterial-domain–selective probe (EUB338). The probe selective for strain 2141T colocalized with all bacterial cells identified by the EUB338 probe (Fig. 2 *G–J*).

To better characterize the gill endosymbiont community, we sequenced the genome of strain 2141T and bacterial metagenomes from the gills of *K. polythalamia*. The closed circular genome of strain 2141T comprised 4,790,451 bp, had a G+C content of 60.1%, and contained 4,508 predicted protein-coding genes. Gill metagenome sequences were determined directly from bulk DNA extracted from five gill tissue samples from two specimens of *K. polythalamia* without enrichment for bacterial cells (*SI Appendix*, Table S3). These were separately sequenced, yielding a total of 647 million reads after quality filtering. Of these, 9.4% were bacterial; the remainder derived from the host (*SI Appendix*, Table S3). Assembly-assisted binning (*Materials and Methods*) was used to show that the vast majority of bacterial reads (99.3%) originated from a small number of genomes closely related to strain 2141T, with two dominant genomes accounting for >70% of this total (Fig. 4*A*). These bacterial reads show >96% average nucleotide identity with each other and with strain 2141T (*SI Appendix*, Fig. S4). The assembled contigs are largely syntenic with, and one is potentially identical to, the genome of 2141T (Fig. 4*B* and *SI Appendix*, Fig. S5). The remaining 0.7% represent other bacteria inconsistently associated with the samples.

To further explore the diversity and relative abundance of 2141T-like genomes present in each sample, we examined single-nucleotide polymorphisms (SNPs) in a conserved gene, DNA gyrase subunit B (*gyrB*), which is present in a single copy in the genomes of most bacteria (33). In agreement with binning results, SNP analysis suggests that the two examined specimens of *K. polythalamia* each harbor two numerically dominant 2141T-like genomes (*SI Appendix*, Fig. S6). At least four additional 2141T-like genomes represent <10% of the total bacterial metagenome.

Interestingly, sequences originating from *T. turnerae* or other cellulolytic bacteria were not detected in the metagenomes, nor were cells of this species detected by FISH in the examined gill sections (*SI Appendix*, Fig. S3). Nonetheless, we were able cultivate this species and to amplify two genes specific to *T. turnerae* from gill tissue of *K. polythalamia* (*SI Appendix*, Table S1). These results suggest that *T. turnerae* is likely present in *K. polythalamia*, but at very low abundance.

All metabolic genes found in strain 2141T were also present in the symbiont metagenome of *K. polythalamia*, indicating that the cultivated isolate is representative of this endosymbiont community and likely has very similar, if not identical, metabolic capacity. Based on these genome and metagenome data, the dominant endosymbiont type in *K. polythalamia* has all genes necessary to predict its ability to oxidize hydrogen sulfide, thiosulfate, and tetrathionate as energy sources for oxidative phosphorylation and to fix carbon dioxide autotrophically via a carboxysome-associated Calvin–Benson–Bassham (CBB) cycle. In agreement, strain 2141T grows autotrophically with bicarbonate and thiosulfate as sole carbon and energy sources, accumulates elemental sulfur in intracellular globules (*SI Appendix*, Fig. S7), bears carboxysome-like structures (*SI Appendix* and Fig. 2*F*), and expresses activity of ribulose bisphosphate carboxylase oxygenase (RuBisCO), a diagnostic enzyme of the CBB cycle (*SI Appendix*, Fig. S8*A* and Table S5). We also detected RuBisCO activity in gill tissue of *K. polythalamia* at levels comparable to those observed in the symbiont-containing tissues of known thioautotrophic animals (*SI Appendix*, Fig. S8*B* and Table S5). Taken together, these micrographic, genomic, and biochemical data demonstrate unequivocally that *K. polythalamia* harbors thioautotrophic endosymbionts, raising to six the number of bivalve families (Lucinidae, Mytilidae, Solemyidae, Teredinidae, Thyasiridae, and Vesicomyidae) known to do so.

These data also shed light on the evolution of thioautotrophic endosymbiosis in Teredinidae and on the extent to which changes in symbiont community composition may influence host biology. Previously published molecular phylogenetic analyses (Fig. 3*B*) have established that wood boring, wood feeding, the wood-storing cecum, and xylotrophic gill endosymbionts were already established in the common ancestor of Teredinidae and its wood-feeding sister family Xylophagaidae (22). Thus, *K. polythalamia* is thought to have lost adaptations to wood feeding that were present in its evolutionary ancestors (22). By the same reasoning, the gill endosymbiont community of *K. polythalamia* likely reflects an evolutionary transition from an ancestral xylotrophic community to one dominated by thioautotrophic bacteria today.

This transition not only led to spectacular changes in the size and anatomy of this species, but also to a fundamental change in the host–symbiont relationship. In other Teredinidae, the heterotrophic endosymbionts acquire organic carbon from the host and in return provide digestive enzymes that enable the host to consume an otherwise-indigestible diet of wood. In *K. polythalamia*, the endosymbionts require no organic carbon from the host, but instead provide the host with chemosynthetically derived carbon.

Finally, we note that many chemosynthetic eukaryotes, including the bivalve genera *Bathymodiolus*, *Idas*, *Thyasira*, *Solemya*, and *Acharax*, the siboglinid tubeworm *Sclerolinum*, and several vestimentiferan tubeworms colonize hydrothermal vents, cold-seeps, and sunken wood (13, 31). Could wood, which can simultaneously serve both as a substrate for heterotrophic and thioautotrophic metabolism, also have helped to bridge the gap between heterotrophic and thioautotrophic metabolism for other chemosynthetic animals?

Materials and Methods

Specimen Collection. Over the course of two field seasons, five specimens of *Kuphus polythalamia* were collected in Mindanao, Philippines, at a depth of ~3 m in a marine bay formerly used as a log storage pond. Specimens were transported alive to the Marine Science Institute, University of the Philippines, Diliman, Quezon City, Philippines, where they were dissected, and tissues were

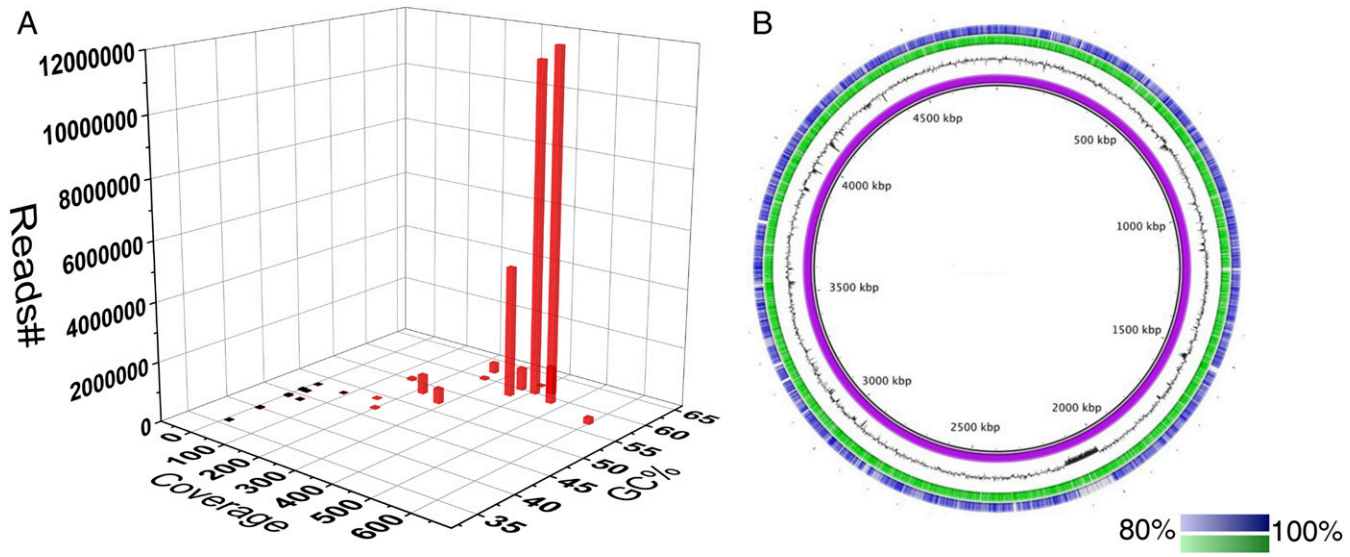

Fig. 4. Endosymbiont metagenome and isolate 2141T genome. (*A*) Metagenome read clusters plotted as a function of GC content, coverage, and read counts (2141T-like clusters are shown in red). Note that the majority of reads fall into two clusters with similar read counts, GC content, and coverage. (*B*) BLASTn alignment of the 2141T-like bacterial sequences from metagenomes of two individual specimens of *K. polythalamia*, specimen 2132W, sample 2249p (green), and specimen 2133X, sample 2110w (blue) to the genome of strain 2141T (purple). Color shades correspond to the identity range of 80–100% compared with the reference genome (2141T). The GC content of the genome of strain 2141T is plotted in black. See *SI Appendix*, Table S3 for additional specimen, sample, and metagenome data.

preserved for microscopy or were frozen at −80 °C until use (*SI Appendix*, Table S3).

Bacterial Cultivation. Approximately 10- to 30-mg slices of gill tissue were rinsed with sterile seawater and homogenized manually in 1 mL of shipworm homogenization buffer (filter-sterilized natural seawater diluted to 75% final concentration with sterile distilled deionized water (ddH_2O), 0.025% cysteine hydrochloride, 20 mM Hepes buffer, pH 8.0). The homogenate was then streaked on STBD agar plates [66% (vol/vol) natural filter-sterilized seawater, 0.025% (wt/vol) NH_4Cl, 20 mM Hepes buffer, pH 8.0, 1.5% (vol/vol) metals and mineral mix (8), 20 mM $Na_2S_2O_3$, 10 mM $NaHCO_3$, 10% (vol/vol) DMEM/Nutrient F-12 Ham (Sigma; D6421), and 1.0% (wt/vol) Bacto agar, final concentrations]. Antimycotics [10 μg/mL cycloheximide and 125 units/mL of nystatin (Sigma; N1638)] were added to the agar medium to suppress fungal growth. Plates were incubated in microaerobic conditions using GasPaK EZ Campy container at 25 °C until small opaque white colonies appeared on the surface of the agar. Colonies were subsequently transferred to aerobic STB50 medium, which is similar to STBD medium but without the DMEM/Nutrient F-12 Ham component; and with concentrations of sodium thiosulfate and sodium bicarbonate increased to 30 and 15 mM, respectively. For long-term storage, freezer stocks of the isolates were prepared by adding 40% glycerol to broth culture at 1:1 ratio and storing at −80 °C. To produce bacterial cells for use in subsequent experiments, a single colony of *K. polythalamia* symbiont strain 2141T, grown from frozen stock on an STB50 agar plate, was used to inoculate a 6-mL starter culture in STB50 broth. After 3 d, 1 L of STB50 broth was inoculated with 0.1% (vol/vol) of the starter culture. The broth cultures were then incubated at 30 °C in a shaker incubator (Series 26; New Brunswick Scientific Co., Inc.) at 100–150 rpm for 3–5 d. To harvest cells, the culture flask was placed in an ice bath for 1 h, and then the bacterial pellets were harvested by centrifugation (3,500–10,000 × *g*, 4 °C, 30 min).

Transmission Electron Microscopy. Bacterial pellets or tissue samples were fixed for 2 h using 2.5% glutaraldehyde in a 0.1 M sodium cacodylate buffer, followed by two 15-min washes in a 0.1 M sodium cacodylate buffer at 4 °C. Samples were postfixed for 2 h in 1% osmium tetroxide in a 0.1 M sodium cacodylate buffer, again followed by two 15-min washes in a 0.1 M sodium cacodylate buffer at 4 °C. Samples were then dehydrated through an ethanol series (30, 50, 70, 85, and 95% ethanol for 15 min at each stage), with a final dehydration stage in absolute ethanol for 1 h. Dehydrated samples were embedded in Spurr resin overnight and polymerized the following day at 60 °C for 24 h. Embedded samples were sectioned (50- to 200-nm thickness), mounted on TEM square mesh copper grids, and stained with uranyl acetate and lead citrate for ∼10–15 min each. Resulting grids were visualized on a JEOL JEM 1010 transmission electron microscope.

SEM and Energy-Dispersive X-Ray Analysis. For SEM imaging of tissues, samples were frozen in liquid nitrogen, fractured with a razor blade, critical point dried using the SAMDRI-PVT-3D Critical Point Dryer (Tousumis), and mounted on a standard aluminum SEM stub and coated with platinum to a thickness of 5 nm using the Cressington 208 HR High Resolution Sputter Coater (Cressington Scientific Instruments). For SEM imaging of bacteria, cells of *K. polythalamia* symbiont strain 2141T or *K. polythalamia* gill homogenates were placed in 70% ethanol, pipetted directly onto a TEM square mesh copper grid, and were allowed to air dry before mounting onto a standard aluminum SEM stub. Samples were imaged on the Hitachi S-4800 field emission scanning electron microscope. Elemental analysis was conducted using the EDAX Element EDS Analysis System and accompanying Element EDS Analysis Software Suite.

DNA Preparation. Cross-sectional tissue samples were taken from multiple locations (anterior, middle, and posterior) along the gills of three adult specimens of *Kuphus polythalamia* (*SI Appendix*, Table S3). Approximately 25 mg of tissue was excised from each section, rinsed thoroughly with filter-sterilized seawater, and homogenized in 4.5 M guanidinium thiocyanate, 2% *N*-lauroylsarcosine (wt/vol), 50 mM EDTA, pH 8.0, 25 mM Tris·HCl. Genomic DNA was extracted from each tissue homogenate using a QiaCube DNA extractor (Qiagen) and Qiagen DNeasy Blood and Tissue Kit followed by the Zymo DNA Clean and Concentrate Kit, according to the respective manufacturer's recommended protocols.

Phylogenetic Characterization of Bacteria. To establish the phylogenetic identity of bacterial isolates in culture, and uncultivated bacteria in gill tissue, genes encoding 16S rRNA (SSU), a multifunctional cellulase (*celAB*), and a rhodanese-like thiotransferase (*sseA*) were PCR amplified from genomic DNA extracts and sequenced using primers and PCR conditions described in *SI Appendix*, Table S1. PCR products were purified using GenCatch PCR Purification Kit and sequenced using an ABI 3130xl Genetic Analyzer (Macrogen). Similarity to previously published sequences in GenBank (National Center for Biotechnology Information) was determined using BLASTn or BLASTp as appropriate. Phylogenetic analyses of 16S rRNA sequences was performed using programs implemented in Geneious, version 8.1.2. Sequences were aligned using MAFFT (version 7.017) by using the auto algorithm. The aligned sequences were trimmed manually, resulting in a final aligned dataset of 1,172 nucleotide positions. Phylogenetic analysis was performed using MrBayes (version 3.2.6) using the GTR + I + Γ substitution model. Chain length was set to 2 million, subsampling every 2,000 generations, and discarding the first 20% of analytical results (burn-in).

FISH. FISH was performed as in ref. 34. Briefly, all samples were fixed in 4% paraformaldehyde in filter-sterilized seawater at 4 °C overnight. Samples were transferred to 70% ethanol and stored at −20 °C. Fixed samples were embedded in paraffin and sectioned at 10-μm thickness. Sections were deparaffinized with a 5-min treatment in xylenes, followed by a 5-min treatment in ethanol, and then rinsed in ddH_2O. Slides were air-dried. Sections were hybridized with 5 ng/μL probe (final concentration) in hybridization buffer [35% formamide, 0.9 M NaCl, 20 mM Tris·HCl (pH 7.4), 0.01% SDS] for 2 h at 46 °C. After hybridization, the slides were incubated at 48 °C in wash buffer [35% formamide, 0.7 M NaCl, 20 mM Tris·HCl (pH 7.4), 50 mM EDTA, 0.01% SDS] for 20 min, rinsed with ddH_2O, and air-dried. After drying, 30 μL of 500 mM Sytox Green was added and incubated for 15 min, and then washed with ddH_2O. Slides were then mounted in a 4:1 Citifluor (Citifluor, Ltd.)/Vectashield (Vector Labs) mounting medium. Slides were visualized on a Zeiss AxioImager laser-scanning confocal microscope with LSM 5 Pascal, version 4.0, imaging software, with imaging parameters held constant between specific and control probes.

Genome Sequencing of *K. polythalamia* Symbiont Strain 2141T. Genomic DNA was extracted from the bacterial pellets using CTAB/phenol/chloroform as detailed at www.pacb.com/wp-content/uploads/2015/09/DNA-extraction-chlamy-CTAB-JGI.pdf, except that Phase-Lock gel tubes were not used. The purity of the genomic DNA was assessed spectrophotometrically using NanoDrop ND-1000; size and quantity were estimated by agarose gel electrophoresis. A high-quality draft genome sequence was determined for *K. polythalamia* symbiont strain 2141T at the US Department of Energy (DOE) Joint Genome Institute (JGI) using the Pacific Biosciences (PacBio) RS platform, which generated 201,156 filtered subreads totaling 827.2 Mbp. All general aspects of library construction and sequencing performed at the JGI can be found at jgi.doe.gov/. The raw reads were assembled using HGAP (version: 2.0.0) (35). The final draft assembly contained 1 contig in 1 scaffold, totaling 4.8 Mbp. The input read coverage was 204×.

Metagenome Sequencing and Assembly. Illumina libraries were prepared as ∼350-bp inserts and sequenced using an Illumina HiSeq 2000 sequencer with 125-bp paired-end runs at the Huntsman Cancer Institute's High-Throughput Genomics Center at the University of Utah. Illumina fastq reads were trimmed using Sickle (36) with the parameters (pe sanger -q 30 –l 125). The trimmed FASTQ files were converted to FASTA files and merged using the Perl script 'fq2fq' in IBDA_ud package (37). Merged FASTA files were assembled using IDBA_ud with standard parameters in Futuregrid (https://portal.futuresystems.org/) or the Center for High Performance Computing at the University of Utah.

Identification of Bacterial Sequences in Metagenome Data. Assembly-assisted binning was used to sort and analyze trimmed reads and assembled contigs into clusters putatively representing single genomes using MetaAnnotator (38). Each binned cluster was retrieved using Samtool (39, 40). To identify bacterial clusters, a BLASTp search was performed using as a query the longest protein sequence coded in the longest contig of each cluster. The longest contig in each cluster was extracted using the command 'cat cluster.fasta | awk -v RS='>' 'NR>1*i++~i>=0&&i<=1*print ">" $0^' | sed '/^>$/d''. The protein coded in each longest contig was recalled using webAUGUSTUS (bioinf.uni-greifswald.de/webaugustus/). The length of each protein sequence was calculated using Samtool (39, 40). Command 'awk 'BEGIN *max = 0~ *if ($2+0>max+0) *max=$2; content=$0~~ END *print content~'' was used to extract the sequence head corresponding to the longest protein in each contig. Each of the longest protein sequences was extracted using Samtool (39, 40). A BLASTp search was performed using the command 'blastp -max_target_ seqs 1 -word_size 6 -evalue 10 -matrix BLOSUM62 -query longest.faa -outfmt '7 qseqid sseqid pident qlen slen length mismatch gapopen evalue bitscore staxids sscinames scomnames sskingdoms sblastnames stitle' -db nr -remote -out kuphus1_longest_blastp_remote'. The bacterial hits were extracted with a threshold of e-value < 2.00E-19. The identified bacterial clusters were checked manually using BLASTx search to remove false positives originating in the animal genome. BLASTn search was used to identify clusters representing 2141T-like genomes. Contigs from each cluster were used as query to do BLASTn search against the genome of

sulfur-oxidizing strain 2141T. Clusters were considered to be related to sulfur-oxidizing strain 2141T if >70% of the contained contigs aligned to the 2141T genome with a threshold of identity of >85%.

Metagenome and Genome Sequence Comparison. Contigs from all 2141T-like bacterial clusters from each sample were combined in a single FASTA file and converted to a pseudomolecule. The BLASTn search comparison was performed by BLAST ring image generator (BRIG) (41) using the genome of sulfur-oxidizing strain 2141T against the pseudomolecule with the standard BLASTn search options. Synteny comparison (*SI Appendix*, Fig. S5) was done by MUMmer 3.0 (42), using the NUCmer alignment tool with parameters (-maxmatch -c 400). The alignment result was plotted by 'mummerplot' with a layout option (–l). A single FASTA file containing all bacterial contigs from 2141T-like clusters from each metagenome sample was used in the synteny comparison. The identity distributions of the average nucleotide identity (ANI) between the genome of strain 2141T and the pooled metagenome sequences (*SI Appendix*, Fig. S4) was calculated by online tool (enve-omics.ce.gatech.edu/ani/) with the following parameters: minimum length, 700 bp; minimum identity, 70%; minimum alignments, 50; window size, 1,000 bp; step size, 200 bp.

SNP Analysis. A blast database was produced from the merged FASTA file used in IDBA_ud assembly with the command "makeblastdb -input_type fasta -dbtype nucl -parse_seqids". A BLASTn search (-perc_identity 90 -max_target_seqs 100000000 -outfmt "6 qseqid sseqid pident length evalue qcovs qlen slen") was performed to find Illumina reads that are homologs of the *gyrB* gene sequence in the genome of sulfur-oxidizing strain 2141T. Hits were defined as "length/slen > 0.95"; the sequence heads of the hit reads were extracted using Samtool (39, 40) and were aligned to the reference *gyraseB* gene using the CLC genomics workbench (version 9.0). The resulting alignment was equally split into 29 fragments (starting from the 85th base pair, 85 bp for each) by shell script (alignment_split.sh; for code, see Dataset S1). The alignment in each fragment were also removed by the shell script. SNP categories of the reads in each small alignment were retrieved by unambiguous assembly (Word Length = 4, Word Length = 4, Maximum Gaps Per Reads = 0, Maximum Mismatches = 0, Maximum Ambiguity = 1) using Geneious (R9).

Cell and Tissue Preparation for RuBisCO Assays. Analyses on bacterial samples were performed on previously frozen cell pellets. To confirm that cells were intact, a small amount of the frozen pellet was resuspended in STB_{50} and examined by light microscope. A 10× volume of RuBisCO Assay Buffer (RAB) [50 mM bicine, pH 8, 5 mM DTT (freshly added from a frozen stock), 20 mM $MgCl_2$, 10 mM $NaHCO_3$] was added to the remaining frozen cell pellet, and then sonicated on ice with a BRANSON Digital Sonifier 250 at 50% amplitude for 10 s twice, with 10 s between pulses, and the extract was examined microscopically to confirm lysis. Gill and siphon (symbiont-free control) tissues were dissected, kept at –80 °C, transported on dry ice, and stored at –80 °C until assayed. Tissues were chopped with a razor blade, a 10× volume of RAB added, and sonicated as above.

RuBisCO Assay. Ribulose-1,5-bisphosphate (RuBP)-stimulated $^{14}CO_2$ incorporation was measured as an indicator of RuBisCO activity as in refs. 43–45. Assays were performed in triplicate (three technical replicates, one biological replicate; only one specimen of this rare species was available for assay), with and without addition of RuBP over a 30-min time course. The rate of RuBP-stimulated incorporation was normalized to total protein (*SI Appendix*, Table S5). Positive controls contained purified spinach RuBisCO (Sigma; R8000-1UN). Additional controls were performed with no isotope and no extract (100 μL of RAB). Before assay, 50 μL of $NaH^{14}CO_3$ (0.75 μCi per reaction) was added to 100 μL of sample and incubated for 20 min. The assay was then initiated by addition of 50 μL of 5 mM RuBP. Reactions were stopped at 10, 20, and 30 min by addition of 200 μL of glacial acetic acid. Unincorporated $^{14}CO_2$ was evaporated by drying in uncapped tubes for 4 h at 90 °C. Acid-stable products were resuspended in 400 μL of water and 7 mL of scintillation fluid, and label incorporation was measured using a Beckman LS6500 scintillation counter. Counts were converted to total micromoles of C incorporated and normalized using the negative controls at the corresponding time point (*SI Appendix*, Table S5).

Protein Assay. The Thermo Scientific Pierce 660-nm Protein Assay was chosen to avoid interference from bicine and DTT present in the RAB. Sonicated extracts were assayed according to manufacturer's instructions and compared with an albumin standard curve. All samples fell within the range of the standard curve.

Data Analysis. All data analysis was carried out in GraphPad Prism 6. For the activity assays, each triplicate was plotted versus time, and a line was fitted to the data (*SI Appendix*, Table S5). The slope of the line corresponds to the activity of the sample. For the protein concentration assays, triplicate samples were interpolated to the standard curve, and the mean and SE were computed for the resulting protein concentration values. The specific activity of the samples was calculated by dividing the activity of the sample by the protein concentration in the sample. The error of the specific activity was calculated by propagating the SEs for the activity and the protein concentration according to standard techniques (46):

For multiplication and division: $xy = z$,

$$\frac{\Delta z}{z} = \frac{\Delta x}{x} + \frac{\Delta y}{y}.$$

ACKNOWLEDGMENTS. The research reported in this publication was supported by Fogarty International Center of the National Institutes of Health Award U19TW008163 (to M.G.H., E.W.S., G.P.C., and D.L.D.), by National Science Foundation Award 1442759 (to D.L.D.), and by the US Department of Energy (DOE) Joint Genome Institute, a DOE Office of Science User Facility supported by the Office of Science of the DOE under Contract DE-AC02-05CH11231 (to M.G.H.). The work was completed under the supervision of the Department of Agriculture–Bureau of Fisheries and Aquatic Resources, Philippines, in compliance with all required legal instruments and regulatory issuances covering the conduct of the research.

1. Huber M (2015) *Compendium of Bivalves 2. A Full-Color Guide to the Remaining Seven Families. A Systematic Listing of 8,500 Bivalve Species and 10,500 Synonyms* (ConchBooks, Hackenheim, Germany).
2. Turner RD (1966) *A Survey and Illustrated Catalogue of the Teredinidae (Mollusca: Bivalvia)* (The Museum of Comparative Zoology, Harvard University, Cambridge, MA).
3. O'Connor RM, et al. (2014) Gill bacteria enable a novel digestive strategy in a wood-feeding mollusk. *Proc Natl Acad Sci USA* 111:E5096–E5104.
4. Betcher MA, et al. (2012) Microbial distribution and abundance in the digestive system of five shipworm species (Bivalvia: Teredinidae). *PLoS One* 7:e45309.
5. Distel DL, DeLong EF, Waterbury JB (1991) Phylogenetic characterization and in situ localization of the bacterial symbiont of shipworms (Teredinidae: Bivalvia) by using 16S rRNA sequence analysis and oligodeoxynucleotide probe hybridization. *Appl Environ Microbiol* 57:2376–2382.
6. Popham JD, Dickson MR (1973) Bacterial associations in the teredo *Bankia australis* (Lamellibranchia, Mollusca). *Mar Biol* 19:338–340.
7. Distel DL, Morrill W, MacLaren-Toussaint N, Franks D, Waterbury J (2002) *Teredinibacter turnerae* gen. nov., sp. nov., a dinitrogen-fixing, cellulolytic, endosymbiotic gamma-proteobacterium isolated from the gills of wood-boring molluscs (Bivalvia: Teredinidae). *Int J Syst Evol Microbiol* 52:2261–2269.
8. Waterbury JB, Calloway CB, Turner RD (1983) A cellulolytic nitrogen-fixing bacterium cultured from the gland of Deshayes in shipworms (Bivalvia: Teredinidae). *Science* 221:1401–1403.
9. Luyten YA, Thompson JR, Morrill W, Polz MF, Distel DL (2006) Extensive variation in intracellular symbiont community composition among members of a single population of the wood-boring bivalve *Lyrodus pedicellatus* (Bivalvia: Teredinidae). *Appl Environ Microbiol* 72:412–417.
10. Gallager SM, Turner RD, Berg CJ (1981) Physiological aspects of wood consumption, growth, and reproduction in the shipworm *Lyrodus pedicellatus* Quatrefages. *J Exp Mar Biol Ecol* 52:63–77.
11. Yücel M, Galand PE, Fagervold SK, Contreira-Pereira L, Le Bris N (2013) Sulfide production and consumption in degrading wood in the marine environment. *Chemosphere* 90:403–409.
12. Fagervold SK, et al. (2012) Sunken woods on the ocean floor provide diverse specialized habitats for microorganisms. *FEMS Microbiol Ecol* 82:616–628.
13. Bienhold C, Pop Ristova P, Wenzhöfer F, Dittmar T, Boetius A (2013) How deep-sea wood falls sustain chemosynthetic life. *PLoS One* 8:e53590.
14. Laurent MCZ, Le Bris N, Gaill F, Gros O (2013) Dynamics of wood fall colonization in relation to sulfide concentration in a mangrove swamp. *Mar Environ Res* 87-88:85–95.
15. Gaudron SM, et al. (2010) Colonization of organic substrates deployed in deep-sea reducing habitats by symbiotic species and associated fauna. *Mar Environ Res* 70:1–12.
16. Lorion J, Duperron S, Gros O, Cruaud C, Samadi S (2009) Several deep-sea mussels and their associated symbionts are able to live both on wood and on whale falls. *Proc Biol Sci* 276:177–185.
17. Laurent MCZ, Gros O, Brulport J-P, Gaill F, Bris NL (2009) Sunken wood habitat for thiotrophic symbiosis in mangrove swamps. *Mar Environ Res* 67:83–88.
18. Duperron S, Laurent MC, Gaill F, Gros O (2008) Sulphur-oxidizing extracellular bacteria in the gills of Mytilidae associated with wood falls. *FEMS Microbiol Ecol* 63: 338–349.
19. Distel DL, et al. (2000) Do mussels take wooden steps to deep-sea vents? *Nature* 403: 725–726.
20. Thubaut J, Puillandre N, Faure B, Cruaud C, Samadi S (2013) The contrasted evolutionary fates of deep-sea chemosynthetic mussels (Bivalvia, Bathymodiolinae). *Ecol Evol* 3:4748–4766.
21. Samadi S, et al. (2007) Molecular phylogeny in mytilids supports the wooden steps to deep-sea vents hypothesis. *C R Biol* 330:446–456.
22. Distel DL, et al. (2011) Molecular phylogeny of Pholadoidea Lamarck, 1809 supports a single origin for xylotrophy (wood feeding) and xylotrophic bacterial endosymbiosis in Bivalvia. *Mol Phylogenet Evol* 61:245–254.

23. Lee HG (1991) X-ray conchology. *Shell-o-Gram: The Official Publication of the Jacksonville Shell Club* (Jacksonville Shell Club, Inc., Jacksonville, FL). Available at www.jaxshells.org/xray.htm. Accessed June 28, 2016.
24. Chinzei K, Savazzi E, Seilacher A (1982) Adaptational strategies of bivalves living as infaunal secondary soft bottom dwellers. *Neues Jahrb Geol Palaontol Abh* 164:229–244.
25. Hanley S (1885) On the *Teredo utriculus* of Gmelin with remarks upon other ship-worms. *Annals and Magazine of Natural History, including Zoology, Botany and Geology XVI*, eds Gunther ACLG, Dallas WS, Carruthers W, Francis W (Taylor and Francis, London), Vol XVI, pp 25–31.
26. Tryon GW (1862) Monograph of the Family Teredidae. *Proc Acad Nat Sci Philadelphia* 14:453–482.
27. Sipe AR, Wilbur AE, Cary SC (2000) Bacterial symbiont transmission in the wood-boring shipworm *Bankia setacea* (Bivalvia: Teredinidae). *Appl Environ Microbiol* 66: 1685–1691.
28. Popham JD (1974) Further observations of the gland of Deshayes in the teredo *Bankia australis*. *Veliger* 18:55–59.
29. Bright M, Espada-Hinojosa S, Lagkouvardos I, Volland JM (2014) The giant ciliate *Zoothamnium niveum* and its thiotrophic epibiont *Candidatus Thiobios zoothamnicoli*: A model system to study interspecies cooperation. *Front Microbiol* 5:145.
30. Nakagawa S, et al. (2014) Allying with armored snails: The complete genome of gammaproteobacterial endosymbiont. *ISME J* 8:40–51.
31. Dubilier N, Bergin C, Lott C (2008) Symbiotic diversity in marine animals: The art of harnessing chemosynthesis. *Nat Rev Microbiol* 6:725–740.
32. Yang JC, et al. (2009) The complete genome of *Teredinibacter turnerae* T7901: An intracellular endosymbiont of marine wood-boring bivalves (shipworms). *PLoS One* 4:e6085.
33. Santos SR, Ochman H (2004) Identification and phylogenetic sorting of bacterial lineages with universally conserved genes and proteins. *Environ Microbiol* 6:754–759.
34. Kwan JC, et al. (2012) Genome streamlining and chemical defense in a coral reef symbiosis. *Proc Natl Acad Sci USA* 109:20655–20660.
35. Chin CS, et al. (2013) Nonhybrid, finished microbial genome assemblies from long-read SMRT sequencing data. *Nat Methods* 10:563–569.
36. Joshi NA, Fass JN (2011) Sickle: A sliding-window, adaptive, quality-based trimming tool for FASTQ files (version 1.33). Available at https://github.com/najoshi/sickle. Accessed January 2, 2017.
37. Peng Y, Leung HC, Yiu SM, Chin FY (2012) IDBA-UD: A de novo assembler for single-cell and metagenomic sequencing data with highly uneven depth. *Bioinformatics* 28:1420–1428.
38. Wang Y, Leung H, Yiu S, Chin F (2014) MetaCluster-TA: Taxonomic annotation for metagenomic data based on assembly-assisted binning. *BMC Genomics* 15:S12.
39. Li H (2011) A statistical framework for SNP calling, mutation discovery, association mapping and population genetical parameter estimation from sequencing data. *Bioinformatics* 27:2987–2993.
40. Li H, et al.; 1000 Genome Project Data Processing Subgroup (2009) The sequence alignment/map format and SAMtools. *Bioinformatics* 25:2078–2079.
41. Alikhan NF, Petty NK, Ben Zakour NL, Beatson SA (2011) BLAST ring image generator (BRIG): Simple prokaryote genome comparisons. *BMC Genomics* 12:402.
42. Kurtz S, et al. (2004) Versatile and open software for comparing large genomes. *Genome Biol* 5:R12.
43. Scott KM, Boller AJ, Dobrinski KP, Le Bris N (2012) Response of hydrothermal vent vestimentiferan *Riftia pachyptila* to differences in habitat chemistry. *Mar Biol* 159:435–442.
44. Tortell PD, Martin CL, Corkum ME (2006) Inorganic carbon uptake and intracellular assimilation by subarctic Pacific phytoplankton assemblages. *Limnol Oceanogr* 51:2102–2110.
45. Robinson JJ, Polz MF, Fiala-Medioni A, Cavanaugh CM (1998) Physiological and immunological evidence for two distinct C-1-utilizing pathways in *Bathymodiolus puteoserpentis* (Bivalvia: Mytilidae), a dual endosymbiotic mussel from the Mid-Atlantic Ridge. *Mar Biol* 132:625–633.
46. Ku HH (1966) Notes on the use of propagation of error formulas. *J Res Natl Bur Stand C* 70C:263–273.