

Two intracellular and cell type-specific bacterial symbionts in the placozoan *Trichoplax* H2

Harald R. Gruber-Vodicka^{1,7*}, Nikolaus Leisch^{1,7}, Manuel Kleiner², Tjorven Hinzke^{3,4,5},
Manuel Liebeke¹, Margaret McFall-Ngai⁶, Michael G. Hadfield^{1,6*} and Nicole Dubilier^{1*}

Placozoa is an enigmatic phylum of simple, microscopic, marine metazoans^{1,2}. Although intracellular bacteria have been found in all members of this phylum, almost nothing is known about their identity, location and interactions with their host^{3–6}. We used metagenomic and metatranscriptomic sequencing of single host individuals, plus metaproteomic and imaging analyses, to show that the placozoan *Trichoplax* sp. H2 lives in symbiosis with two intracellular bacteria. One symbiont forms an undescribed genus in the Midichloriaceae (Rickettsiales)^{7,8} and has a genomic repertoire similar to that of rickettsial parasites^{9,10}, but does not seem to express key genes for energy parasitism. Correlative image analyses and three-dimensional electron tomography revealed that this symbiont resides in the rough endoplasmic reticulum of its host's internal fibre cells. The second symbiont belongs to the Margulisbacteria, a phylum without cultured representatives and not known to form intracellular associations^{11–13}. This symbiont lives in the ventral epithelial cells of *Trichoplax*, probably metabolizes algal lipids digested by its host and has the capacity to supplement the placozoan's nutrition. Our study shows that one of the simplest animals has evolved highly specific and intimate associations with symbiotic, intracellular bacteria and highlights that symbioses can provide access to otherwise elusive microbial dark matter.

Placozoa are marine invertebrates at the base of the animal tree and are considered among the simplest animals known. These millimetre-sized benthic animals can be easily cultured and are considered key models for understanding metazoan evolution, developmental biology and tissue formation^{1,14–16}. Electron microscopy studies as early as in the 1970s revealed the presence of intracellular bacteria in these animals^{3–6}. Remarkably, nearly five decades later, only very little is known about the biology of these symbionts and their interactions with their hosts.

The phylum Placozoa encompasses 19 cryptic species, on the basis of mitochondrial haplotypes^{2,6}. These benthic animals do not have a mouth or gut and feed on algae and bacterial biofilms by external digestion and subsequent uptake via their ventral epithelium^{17,18}. All placozoans have three cell layers and six morphologically differentiated cell types^{3,6,19,20}. The thick ventral epidermis consists of ciliated epithelial cells in which glandular and lipophilic cells are irregularly interspersed^{17,18}. Ciliated epithelial cells make up the thin dorsal epidermis in which crystal cells occasionally occur. An internal meshwork of fibre cells, sandwiched between the

two epidermal layers, connects the ventral and dorsal body walls²⁰. Intracellular symbionts were first described in these fibre cells^{3,5,20}. The bacteria were present in all seven haplotypes examined, independent of sampling site or time, and were hypothesized to reside in the lumen of the rough endoplasmic reticulum (rER)^{3,5,6,20}. Persistent and stable residence of a bacterium in the rER of a host would be remarkable as the vast majority of intracellular symbionts live in the cytoplasm or vacuoles, and the few known exceptions inhabit the nucleus or mitochondria^{21–23}.

In this study, we focused on the *Trichoplax* sp. haplotype H2 (*Trichoplax* H2), previously reported to host two bacterial morphotypes⁵. Sequencing of placozoan genomes consistently yielded rickettsial and other bacterial sequences^{6,24,25}. However, as thousands of host individuals were pooled for these analyses, it was neither clear whether these bacterial sequences originated from contaminants or symbionts nor whether they were present in all host individuals. Our recent advances in high-throughput sequencing of single placozoan individuals, together with correlative imaging analyses and three-dimensional (3D) reconstruction, allowed us to explore the patterns, structure and function of the placozoan symbiosis at the individual and cellular level.

The *Trichoplax* H2 microbiome is dominated by two bacterial symbionts. We isolated a placozoan H2 haplotype lineage from a seawater tank at the Kewalo Marine Laboratory, University of Hawai'i (Supplementary Fig. 1). To characterize the microbiome of this *Trichoplax* H2, we combined highly sensitive DNA and RNA extraction and library preparation protocols to sequence the metagenomes and metatranscriptomes of microscopic single individuals that have an estimated biovolume of 0.02 μl and from which we could isolate 0.5 to 4 ng of nucleic acids ($n=5$). All five individuals had similar microbial communities based on 16S ribosomal RNA (rRNA) gene reads, but only two taxa were consistently dominant in all five host individuals (Supplementary Fig. 2 and Supplementary Table 1).

The first and most abundant 16S rRNA phylotype was an alphaproteobacterium from the family Midichloriaceae (Rickettsiales)⁷ (Fig. 1a). Midichloriaceae are obligate intracellular, often pathogenic, bacteria found in protists and animals, including humans⁸. In 16S rRNA analyses, the *Trichoplax* H2 midichloriacean phylotype formed an unnamed lineage that consisted of sequences recovered from diverse invertebrate hosts and sequences from subsurface sediment samples (98.4–99.4% pairwise identity; Fig. 1a). We recovered a high-quality²⁶ 1.26 Mb metagenome-assembled genome that

¹Max Planck Institute for Marine Microbiology, Bremen, Germany. ²Department of Plant and Microbial Biology, North Carolina State University, Raleigh, NC, USA. ³Department of Pharmaceutical Biotechnology, Institute of Pharmacy, University of Greifswald, Greifswald, Germany. ⁴Institute of Marine Biotechnology, Greifswald, Germany. ⁵Department of Geoscience, University of Calgary, Calgary, Alberta, Canada. ⁶Kewalo Marine Laboratory, Pacific Biosciences Research Center, University of Hawai'i at Mānoa, Honolulu, HI, USA. ⁷These authors contributed equally: Harald R. Gruber-Vodicka, Nikolaus Leisch. *e-mail: hgruber@mpi-bremen.de; hadfield@hawaii.edu; ndubillie@mpi-bremen.de

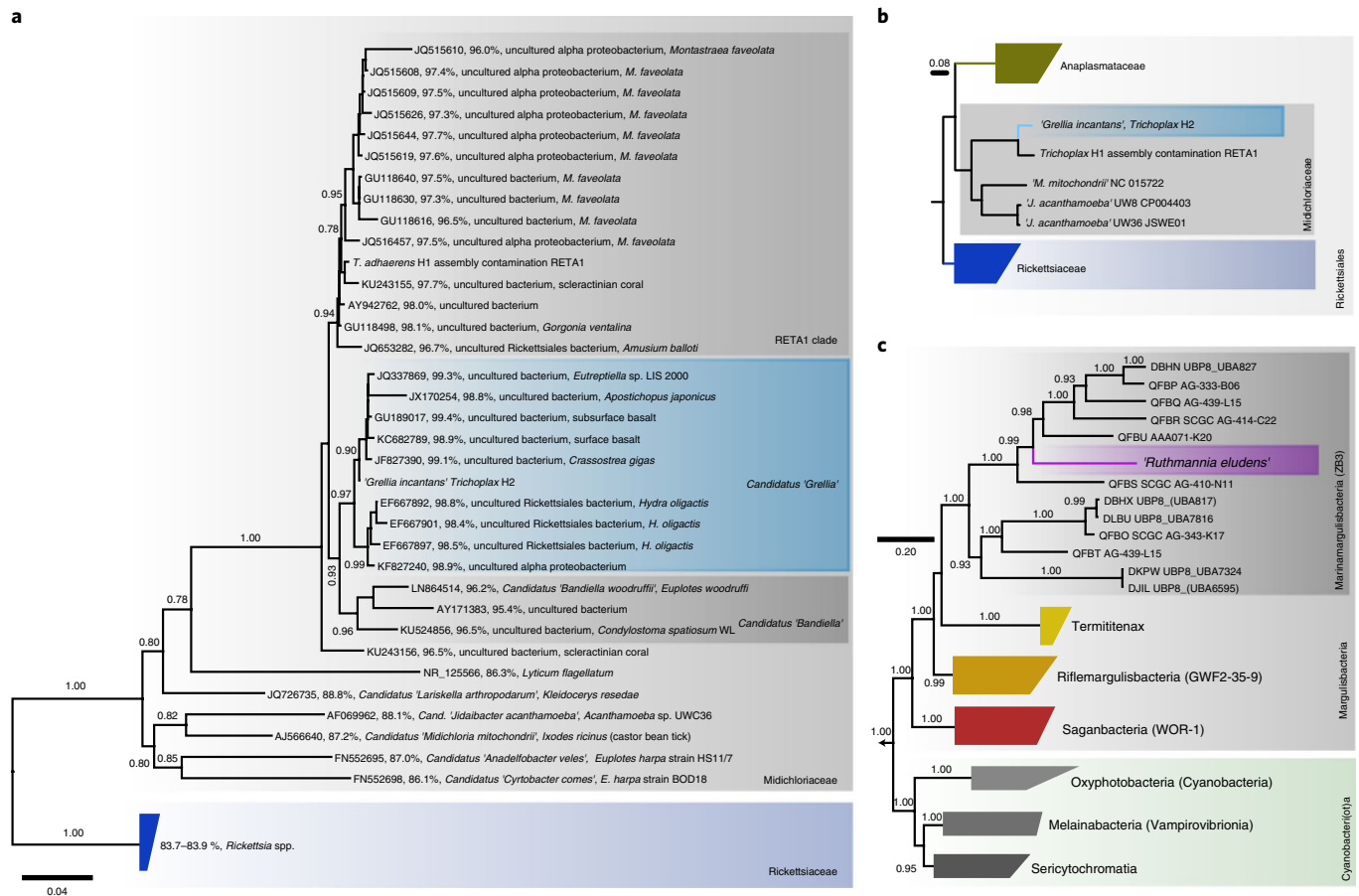


Fig. 1 | Phylogenetic analyses of the *Trichoplax* H2 symbionts. 'G. incantans' represents an undescribed genus in the Midichloriaceae (Rickettsiales) and 'R. eludens' is a Marinamargulisbacteria (Margulisbacteria). Bootstrap support values below 0.5 are not shown. Scale bars indicate substitutions per site. Colours and shades of grey indicate taxonomic groups. **a, A 16S rRNA tree of 'G. incantans' and related Midichloriaceae. For each sequence, the accession number, the percentage identity to 'G. incantans' and the published taxonomic names and hosts (where available) are indicated. **b,c**, Phylogenomic analyses using 43 conserved marker genes based on metagenome-assembled genomes and reference genomes: 'G. incantans' and Midichloriaceae are placed in the Rickettsiales (**b**) and 'R. eludens' in the Margulisbacteria (**c**). Phylum-level classification follows the Genome Taxonomy Database. Taxon names from the Genome Taxonomy Database are indicated in parentheses where available.**

included the midichloriacean 16S rRNA phylotype. Sequences from the *Trichoplax adhaerens* haplotype H1 genome project¹⁵ included a midichloriacean 16S rRNA gene fragment and a partial genome of a rickettsial phylotype (RETA1) was also recovered²⁵. Phylogenetic analyses based on the 16S rRNA gene and phylogenomic analyses based on 43 conserved marker genes placed the *Trichoplax* H2 phylotype and *Trichoplax* H1 RETA1 in the Midichloriaceae. The *Trichoplax* H2 and H1 phylotypes were phylogenetically distinct and, according to amino acid sequence identity, these two symbionts belong to two separate but undescribed genera, with *Candidatus 'Bandiella'*^{27,28} as the closest characterized genus (Fig. 1a,b, Supplementary Note 1). We propose the *Candidatus* taxon 'Grellia incantans' for the midichloriacean phylotype from our haplotype H2 isolate (see Supplementary Note 1 for description and etymology).

The second most abundant and consistently present bacterial taxon in the *Trichoplax* H2 metagenomes belonged to the Margulisbacteria, a phylum without isolated representatives that forms the sister clade to Cyanobacteriota^{11–13,29}. No 16S rRNA gene sequences with >90% identity to this bacterium were found in public sequence databases, warranting the establishment of a taxon at the genus or even family level. We therefore propose the *Candidatus* taxon 'Ruthmannia eludens' for this bacterium (see Supplementary Note 2 for a detailed description and etymology).

Using metagenomics binning, we recovered a high-quality 1.51 Mb metagenome-assembled genome for 'R. eludens'. Our phylogenomic analyses confirmed our 16S rRNA gene analysis and placed 'R. eludens' in the Marinamargulisbacteria (Margulisbacteria) (Fig. 1c). Marinamargulisbacteria are aquatic bacteria that occur worldwide^{11,13}. 'R. eludens' is distantly related to single-cell amplified genomes and metagenome-assembled genomes from marine pelagic samples¹³ (Fig. 1c). Marinamargulisbacteria are known only from sequence-based studies, with recovered draft genomes of 0.5–2.0 Mb, and all genomes are classified as medium to low quality²⁶. Despite the small genome size, our metagenome-assembled genome was classified as a high-quality draft genome (Supplementary Note 2).

Both symbionts are intracellular, spatially segregated and specific to the host cell type. We used fluorescence in situ hybridization (FISH; Fig. 2a; Supplementary Table 2) to link the bacterial sequences to their morphotypes and visualize the distribution of the two symbionts in *Trichoplax*. No bacteria except the two symbionts 'G. incantans' and 'R. eludens' were detected in all the placozan individuals examined (Fig. 2a; Supplementary Fig. 4). 'G. incantans' was thin and rod-shaped ($\leq 1.2 \times 0.30 \mu\text{m}^2$; Fig. 2a and Supplementary Fig. 4). In contrast, 'R. eludens' had a wider and stouter rod-shaped morphology ($\leq 1.2 \times 0.47 \mu\text{m}^2$; Fig. 2a and Supplementary Fig. 4; for details see Supplementary Notes 1 and 2). Our correlative FISH and transmission electron microscopy (TEM) analyses of five *Trichoplax*

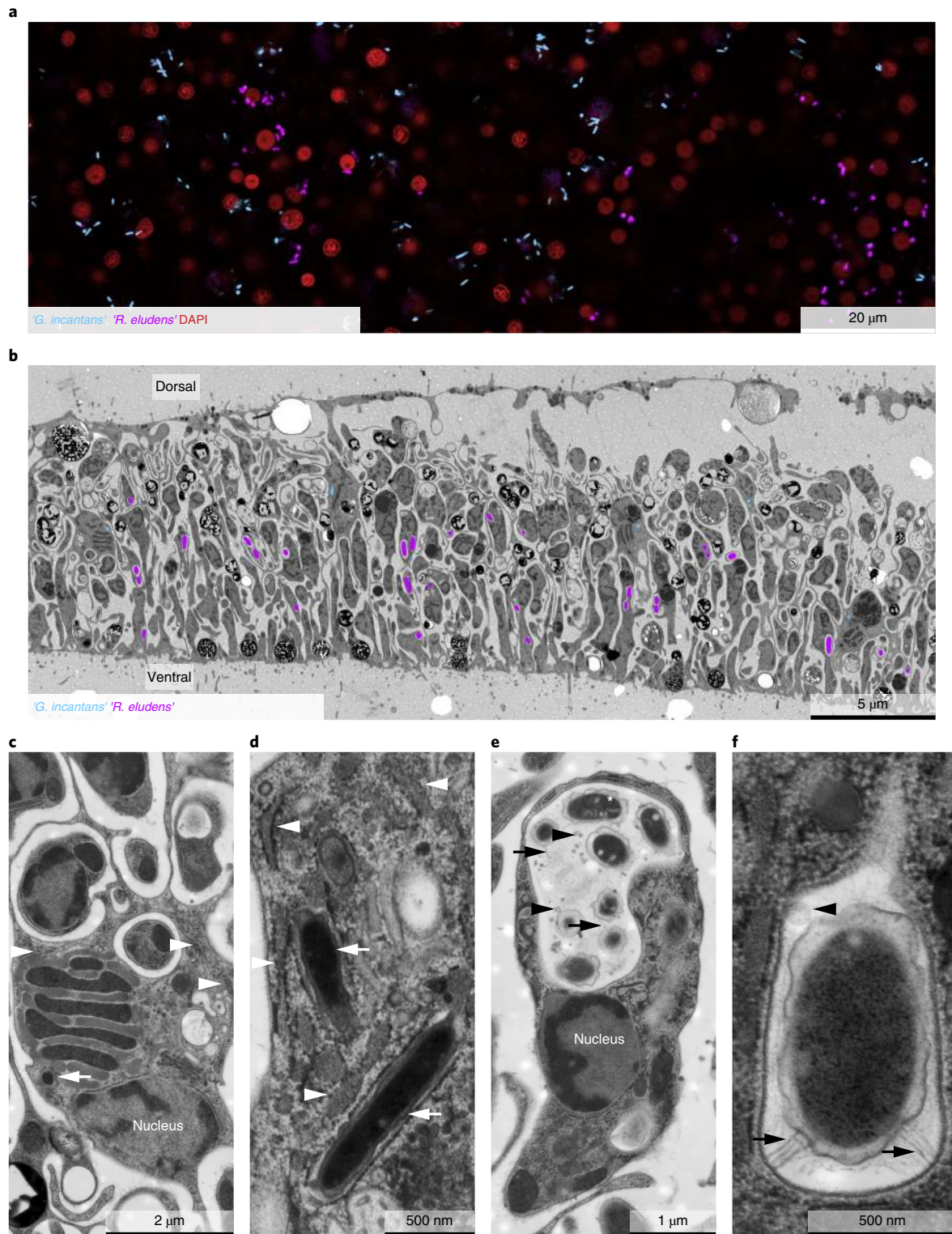


Fig. 2 | '*R. eludens*' and '*G. incantans*' are specific to two spatially segregated host cell types. **a**, A false-coloured FISH image using probes specific for '*G. incantans*' (GRIN-61-2, Atto-647) and '*R. eludens*' (RUEL-846-22, Atto 594); host nuclei are stained with 4,6-diamidino-2-phenylindole (DAPI). The results are representative of five independent experiments. **b**, A TEM image of a cross-section of *Trichoplax* H2 with '*G. incantans*' and '*R. eludens*' indicated in false colour (for raw image data see Supplementary Fig. 5). **c, d**, TEM images of fibre cells. '*G. incantans*' is indicated with white arrows and the rER with white arrowheads. **e, f**, TEM images of ventral epithelial cells containing '*R. eludens*'. Outer membrane vesicles are indicated with black arrowheads, fimbriae-like structures with black arrows and internal structures by a white asterisk. Results in **b–f** are representative of three independent experiments.

H2 individuals revealed that the two bacterial symbionts were always intracellular, spatially segregated and specific to one of the six host cell types (Fig. 2b and Supplementary Figs. 5–7). '*G. incantans*' was

observed only in fibre cells and was the only bacterium located in these cells (Fig. 2b and Supplementary Figs. 5 and 6). All '*G. incantans*' cells were surrounded by a host membrane that was densely

covered with ribosomes (Fig. 2c,d and Supplementary Fig. 6; $n = 49$ symbiont cells in 9 specimens). Similar host structures surrounding the bacteria in other *Trichoplax* lineages were interpreted as indicating that the bacteria reside inside the host's rER³. An alternative interpretation for such host membrane structures was shown for the human intracellular pathogens *Brucella* and *Legionella*, as well as the amoebal midichloriacean parasite *Candidatus 'Jidaibacter'*. These bacteria remodel the phagosomal surfaces of their hosts so that they become covered by host ribosomes as an effective strategy for avoiding digestion by their hosts^{21,30}.

To resolve the subcellular architecture of '*G. incantans*' symbiosis, we used high-resolution 3D TEM tomography to determine whether the structures surrounding the symbiont cells were remodelled phagosomes or rER. Our 3D electron tomographic reconstructions revealed that the ribosome-covered membranes, in which '*G. incantans*' occurred, formed networks that were connected to the nuclear envelope³¹. This indicates that the structure in which '*G. incantans*' is embedded is in fact rER. '*G. incantans*' symbionts were only observed in the rER, some even within the same rER lumen, and never in other host structures (Fig. 3; Supplementary Fig. 8; Supplementary Video 1). These analyses suggest that '*G. incantans*' persistently resides in the rER of its host. The second symbiont, '*R. eludens*', colonized only the ventral epithelial cells. These symbionts were always found within cytoplasmic vacuoles of the host (Fig. 2e,f). The vacuoles contained numerous membrane-bound vesicles, presumably outer membrane vesicles produced by '*R. eludens*' (Supplementary Fig. 7). Thin, tubular structures that resemble fimbriae appeared to connect the bacterial cells to the host vacuole membrane (Fig. 2f; Supplementary Fig. 7).

Bacteria that live inside animal cells are known from only 6 of the 114 recognized bacterial phyla³². The number of bacterial phyla with representatives that can live as intracellular symbionts has not increased since the characterization of *Mycoplasmatales* in the early 1960s, despite huge advances in the sequencing of animals from a wide range of phyla and environments that have led to the discovery of numerous lineages of microbiota^{11,32}. *Marinamargulisbacteria* is one of the most phylogenetically remote clades of bacteria, discovered through high-throughput sequencing of environmental samples³³. The remote position of the placozoans in the animal tree of life has probably contributed to this late discovery of Margulisbacteria as the seventh bacterial phylum with intracellular symbionts of animals. Our study of the *Trichoplax* microbiome highlights how bacteria captured by eukaryotes provide a route for studying bacterial groups that are otherwise known only from sequences found in water or sediment samples.

'*R. eludens*' gains nutrition from lipids degraded by its host. We sequenced the metatranscriptomes of the same single placozoan individuals that were used for metagenomic analyses ($n = 3$) and generated metaproteomes from pooled samples of 10 to 30 individuals ($n = 3$) to investigate the physiology of '*R. eludens*'. Physiological modelling of these expression data revealed that '*R. eludens*' is an aerobic chemoorganoheterotroph, with a complete tricarboxylic acid (TCA) cycle that generates energy and biomass from glycerol and the β -oxidation of fatty acids (Fig. 4a; Supplementary Table 3). The source of the glycerol and fatty acids is probably lipids from the algal diet of the host. Our analyses of the host's transcriptome revealed that *Trichoplax* H2 expressed several lipases, most probably for the digestion of the algae it feeds on (Supplementary Table 4). These host lipases hydrolyze lipids to glycerol and fatty acids. The genome of '*R. eludens*' also encodes lipases that would allow '*R. eludens*' to digest lipids independently of its host. Interestingly, we found neither transcripts nor peptides for these symbiont lipases, suggesting that '*R. eludens*' relies on the lipases expressed by its host (Supplementary Table 3).

The transfer of glycerol and even-chain fatty acids from the host to '*R. eludens*' probably occurs passively, as they can easily diffuse

through cell membranes. We predict that the fatty acids are taken up and activated by '*R. eludens*' on the basis of its high expression of a long-chain fatty acid coenzyme A (CoA) ligase (among the top 25% of expressed genes; Fig. 4a; Supplementary Table 3). The fatty acids are then probably catabolized to acetyl-CoA and respired, as indicated by the expression of all the genes needed for β -oxidation and the oxidative TCA cycle. However, the anabolic incorporation of fatty acids is unlikely, as we could not detect the genes for the glyoxylate shunt.

'*R. eludens*' encoded genes for synthesizing all nucleotides and amino acids, including the nine amino acids considered essential for animals. However, we found no genomic or transcriptomic indications that '*R. eludens*' exports nutrients to its host, for example via amino acid exporters (see Fig. 4a and Supplementary Note 3). Moreover, in our TEM analyses, we found no evidence for the intracellular, lysosomal digestion of '*R. eludens*', such as lamellar bodies or tertiary lysosomes commonly observed in other nutritional symbioses^{34,35}. Our ultrastructural analyses did, however, reveal large numbers of putative outer membrane vesicles in the host vacuole surrounding '*R. eludens*' (Fig. 2e,f and Supplementary Fig. 7). It is tempting to speculate that the host takes up outer membrane vesicles produced by '*R. eludens*' via phagocytosis and thus supplements its diet, as the host lacks synthesis pathways for essential amino acids. However, the beneficial effects of such putative amino acid provisioning by '*R. eludens*' are not clear, given that the animal's algal diet may contain sufficient amounts of essential amino acids.

'*G. incantans*' has the genes for energy parasitism but does not express them: it lives in the rER of fibre cells and seems to be a typical Rickettsiales based on genomic features alone, namely a heterotroph that relies on its host for biomass and energy generation (Fig. 4b). The '*G. incantans*' genome encodes the hallmark feature for intracellular energy parasites that is present in all Rickettsiales genomes: a fully functional ADP/ATP-translocase for importing ATP from its host⁹. In contrast to all other known energy parasites, we found no transcripts or respective peptides of the ADP/ATP-translocase in '*G. incantans*' (Supplementary Table 5). Instead, '*G. incantans*' generated ATP with an ATP synthase, and the subunits a and b were highly expressed in the bacterium's proteome (Supplementary Table 6). Compared to the typical energy-parasitic lifestyle of cytosolic Rickettsiales that rely on ATP imported from their hosts¹⁰, the ability of '*G. incantans*' to synthesize ATP by itself likely lowers its detrimental impact on its host considerably.

High transcription of key genes of the oxidative TCA cycle and the presence of a complete electron transport chain in the genome, with some of the subunits of the electron transport chain among the most highly transcribed genes, suggests that the proton gradient for ATP synthesis is fuelled by oxidative phosphorylation (Fig. 4b and Supplementary Table 5). An incomplete glycolysis pathway and several importers for α -ketoacids and C4-dicarboxylates suggest that the metabolites respired in the TCA cycle are imported from the host (Fig. 4b).

The genome and transcriptome of '*G. incantans*' revealed a strong host dependence on both amino acid and nucleotide supply (Fig. 4b; see Supplementary Note 4 for details). In contrast, the transcription profile of '*G. incantans*' suggested that it could supply its host with riboflavin (vitamin B₂), an essential vitamin that cannot be synthesized by most metazoans. Our analyses of the transcriptomic data of *Trichoplax* H2, as well as the genome and proteome of the closely related haplotype H1^{24,36}, revealed that both seem to lack the known genes for synthesizing riboflavin (Supplementary Fig. 9) and rely on an external source of riboflavin (Supplementary Table 4). This suggests that when riboflavin availability is limiting for the host, '*G. incantans*' could supplement the nutrition of its host.

'*G. incantans*' does not seem to be detrimental to *Trichoplax* H2, despite the fact that it has to import most of the compounds it needs for generating energy and biomass from its host. Our metagenomic,

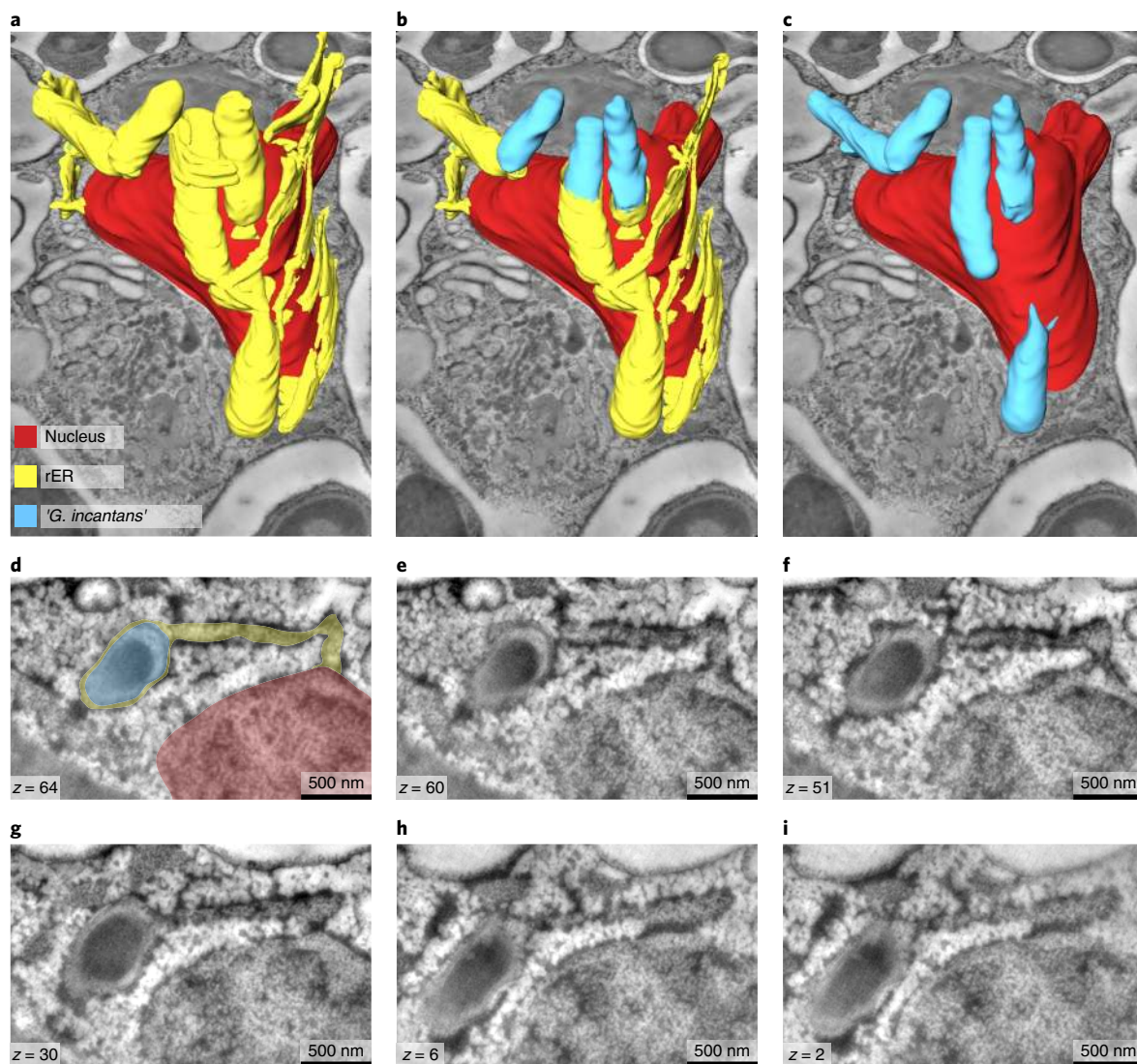


Fig. 3 | 'G. incantans' lives in the rER of *Trichoplax* H2. All panels are the results from a single experiment. **a–c**, 3D volume rendering of reconstructed 'G. incantans', the rER and the nucleus of a fibre cell, superimposed on a virtual slice of the 3D TEM tomography stack; the rER has been virtually removed to partially (**b**) and fully (**c**) show the symbionts within the rER lumen. No scale bar is shown as the scale varies with perspective. **d–i**, Selected tomography slices (on which the 3D reconstruction is based) at various depths (see *z* values indicated) show the connection between the nucleus, the rER and the bacteria. For ease of interpretation, **d** has been false-coloured using the colour key from **a**. For raw data, see Supplementary Fig. 8.

FISH and TEM data revealed 2–20 symbiont cells per fibre cell, so that the total number of 'G. incantans' cells per host individual is roughly the same as the number of host cells (Supplementary Note 5). This indicates closely regulated control of symbiont growth by the symbiont, the host or both partners. Pathogen abundances are typically orders of magnitude higher per host cell and often result in rapid exploitation and destruction of host cells and the impairment of host reproduction³⁷. The relatively low abundance of 'G. incantans' in *Trichoplax* H2 together with the rapid doubling rates of these hosts (2–3 d in our aquaria) are in contrast to virulent pathogenic infections. Unlike all other known energy parasites, 'G. incantans' seems to generate its own ATP and might even modulate its host immune response to prevent apoptosis (Supplementary Note 4).

Bacterial phylotypes highly similar or identical to 'G. incantans' occur worldwide in aquatic environments. To assess how widespread the two *Trichoplax* symbionts are in other environments and hosts, we surveyed the ~300,000 publicly available amplicon-based 16S rRNA sequence libraries using the IMNGS pipeline.

We did not find any sequences related to '*R. eludens*', using a cut-off of 99% identity. In contrast, sequences highly similar or identical to 'G. incantans' were present in aquatic environments, both marine and limnic, from across the globe (Supplementary Table 7). Of the 8,026 libraries from aquatic environments, we found sequences that were at least 99% identical to 'G. incantans' in almost 10% of these libraries ($n=845$). One third of the sequences were identical to 'G. incantans' and almost all were attributed to the genus *Grellia* on the basis of evolutionary placement analysis (Supplementary Fig. 10). This is remarkable for Midichloriaceae, because all other genera are much rarer and were found in only 0–55 libraries, depending on the genus (Supplementary Table 7). The presence of *Grellia* phylotypes in such a wide range of environments, including limnic ones, indicates that these bacteria have host ranges beyond placozoans. Indeed, our phylogenetic 16S rRNA analyses showed that sequences that group with the genus *Grellia* have been found in marine protists (*Eutreptiella*), sea cucumbers (*Apostichopus*) and oysters (*Crassostrea*), as well as in the limnic cnidarian *Hydra oligactis* (see Fig. 1a). The *Hydra* sequences came from specimens

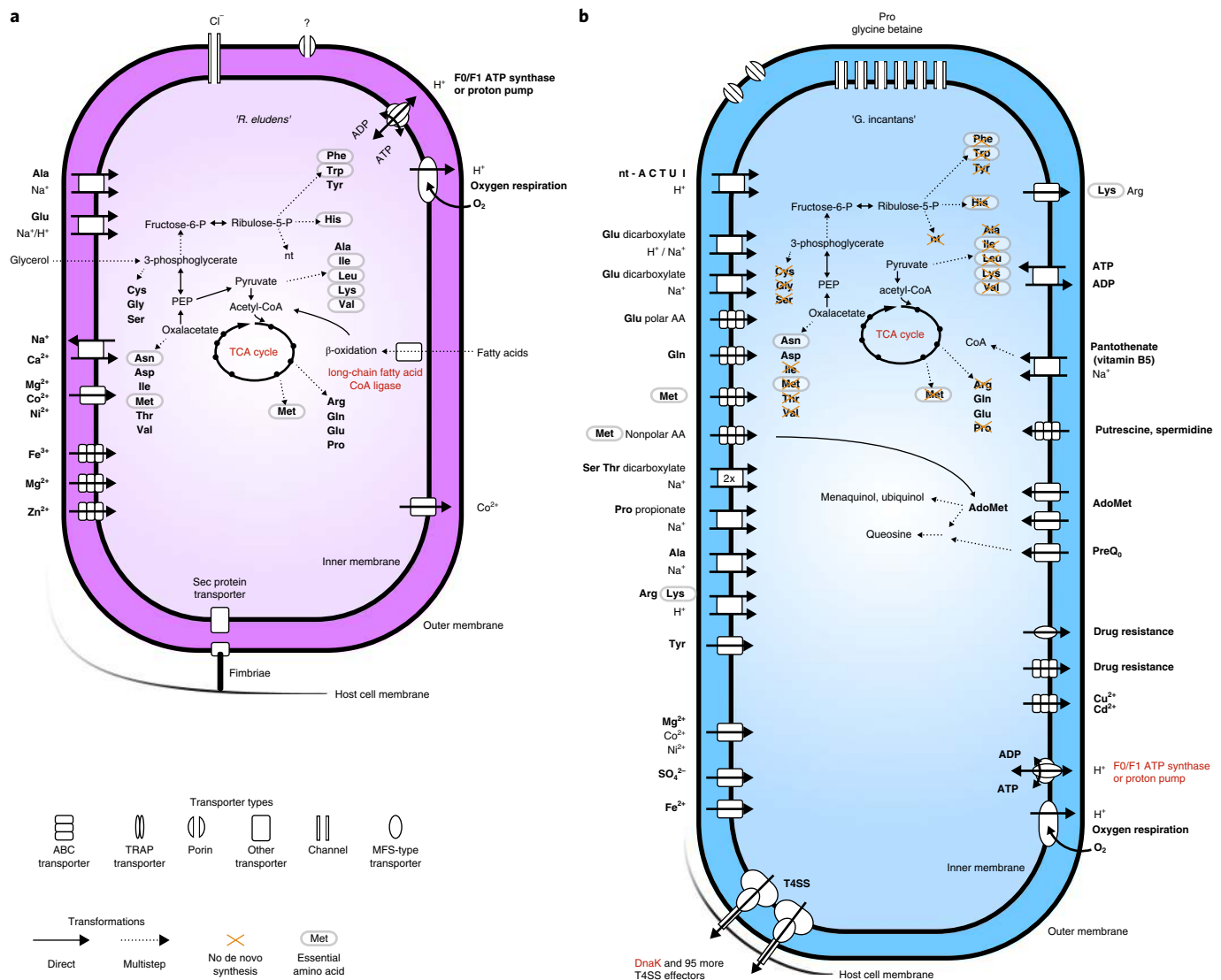


Fig. 4 | *R. eludens* has versatile biosynthesis pathways, whereas *G. incantans* depends on the import of most nutrients from its host. Physiological reconstructions based on RAST annotations and Pathway Tools metabolic modelling. Functions that are discussed in the text and highly expressed are indicated in red. **a**, *R. eludens*. **b**, *G. incantans*. Bold font indicates primary function. ABC, ATP-binding cassette; AdoMet, S-adenosyl-L-methionine; MFS, major facilitator superfamily; nt, nucleotide; nt-ACTUI, the nucleotides a cell can import (all but guanine); P, phosphate; PEP, phosphoenolpyruvate; T4SS, type IV secretion system; TRAP, tripartite ATP-independent periplasmic.

collected freshly from their natural environments and animals reared in the laboratory for more than 30 yrs, indicating the stability of this association in these hosts³⁸.

The recent realization that human pathogens such as Chlamydiae, Legionellales and Rickettsiales have close relatives that live in hosts ranging from protists to fish and from aquatic and soil habitats has led to a paradigm shift in our view of the ecology and evolution of intracellular bacteria^{27,39,40}. *G. incantans* extends our conceptual understanding of the pervasiveness of such bacteria and shows that a single environmental rickettsial genus occurs worldwide in marine and limnic habitats. This remarkable distribution raises the question of whether all *Grellia* are host-associated. If *G. incantans* had a free-living stage, this would be in contrast to all other known Rickettsiales that infect animals²⁷.

Unlike other animals at the base of the animal tree, such as sponges, cnidarians or ctenophores, Placozoa is the only phylum in which intracellular bacteria have been observed in all individuals and haplotypes investigated. Intracellular symbiosis thus seems to be an invariant trait across this phylum. Our study identifies these

bacteria in *Trichoplax* H2, shows that they are found in every specimen examined and defines the specificity and fidelity to the host cell type in which the symbionts reside.

How might the *Trichoplax* symbionts be transmitted within a growing individual and to its offspring? Within a host individual, the symbiont-containing cells could pass on their bacteria during division, or the symbionts could continuously reinfect host cells derived from aposymbiotic cells. However, little is known about cell turnover and proliferation in placozoa and it remains to be determined whether they even have stem cells. Similarly, we can only speculate on transmission during asexual reproduction (the main mode of reproduction in placozoa). In *Trichoplax* H2, which has been reproducing asexually in our aquaria for several years, the symbionts are transmitted with high fidelity, as all host individuals had both symbionts. Information on sexual reproduction, which is much rarer and has not been observed in nature, is too limited to allow us to know whether the symbionts are incorporated into resulting embryos. If not, the symbionts must be obtained from the environment. Symbiont uptake from the environment could explain

why the midichloriacean symbionts of *Trichoplax* H1 and H2 do not belong to the same clade, although their hosts are very closely related and separated only a few decades ago²⁴. This split could have been caused by their midichloriacean symbionts, as Rickettsiales are well known to induce reproductive incompatibility in insects⁴¹. Future studies of the microbiomes of the large number of extant haplotypes are needed to fully understand the ecology and evolution of symbioses between placozoans and their bacterial symbionts.

Methods

Isolation and cultivation. The placozoans were isolated from a coral tank at the Kewalo Marine Laboratory, University of Hawai'i at Mānoa in October 2015 by placing glass slides, mounted in plastic slide boxes that had the top and bottom cut out, into the tank for 10 d (ref. 17). Placozoans were identified under a dissection microscope, transferred to 400 ml glass beakers with 34.5% artificial seawater and fed weekly with 2×10^6 cells ml⁻¹ of *Isochrysis galbana* from a log-phase culture. Doubling times were 2–3 d at 25 °C in 34.5% artificial seawater and with a 16:8 h light:dark regime.

Nucleic acid extractions. DNA was extracted from two single individuals from the *Trichoplax* H2 cultures using the DNeasy Blood & Tissue Kit (Qiagen) and DNA and RNA from three additional single individuals were extracted using the AllPrep DNA/RNA Micro Kit (Qiagen), according to the manufacturer's protocols for both kits except for the following modifications. Proteinase K digests were performed overnight. Elution volumes were halved and all samples were eluted twice, reusing the first eluate. Elutions were carried out with a 10-min-long waiting step before centrifugation.

DNA and RNA sequencing. Illumina-library preparation and sequencing were performed by the Max Planck Genome Centre. In brief, DNA/RNA quality was assessed with the Agilent 2100 Bioanalyzer (Agilent) and genomic DNA was fragmented to an average fragment size of 500 base pairs (bp). For the DNA samples, the concentration was increased (MinElute PCR Purification Kit; Qiagen) and an Illumina-compatible library was prepared using the Ovation Ultralow Library Systems Kit (NuGen) according to the manufacturer's protocol. For the RNA samples, the Ovation RNA-seq System V2 (NuGen) was used to synthesize complementary DNA and sequencing libraries were then generated with the DNA Library Prep Kit for Illumina (BioLABS). All libraries were size selected by agarose gel electrophoresis and the recovered fragments quality-assessed and quantified by fluorometry. For each DNA library, 14–22 million 150 bp paired-end reads were sequenced on a HiSeq 4000 (Illumina) and, for the RNA libraries, 150 bp single-end reads were sequenced to a depth of 42–44 million.

Host mitochondrial 16S rRNA gene phylogenetic analyses. The metagenomic assembly was screened for the contig containing the host mitochondrial 16S rRNA gene (*m16S*) using BLAST v2.7.1 as implemented in Geneious R11⁴². The gene was extracted from the contig and aligned together with a database of publicly available *m16S* sequences using MAFFT v7.394 in G-Ins mode. The phylogenetic tree was reconstructed using FastTree v2.1.5 (ref. 43) with a GTR model, 20 rate categories and Gamma20 likelihood optimization, generating approximate likelihood-ratio-test values for node support. The tree was drawn with Geneious⁴². The tree was rooted with clade A placozoans².

Bacterial diversity 16S rRNA gene phylogenetic analyses. For the 16S rRNA gene database of all phylotypes recovered, the phyloFlash v3.0 beta1 pipeline (<https://github.com/HRGV/phyloFlash>) assembled full-length SSU genes for all samples. The dataset was aligned and phylogenetic trees were calculated and visualized as for the host *m16S* dataset above. The tree was rooted with the Eukarya and only the bacterial part of the tree is shown in this Letter.

Genome analyses. Full-length 16S rRNA gene sequences were reconstructed for each metagenomic and metatranscriptomic library using phyloFlash v3.0 beta1 (<https://github.com/HRGV/phyloFlash>) from raw reads.

For assembly, adapters and low-quality reads were removed with bbduk v37.9 (<https://sourceforge.net/projects/bbmap/>) with a minimum quality value of 2 and a minimum length of 36; single reads were excluded from the analysis. Each library was error corrected using BayesHammer v3.62⁴⁴. A combined assembly of all the libraries was performed using SPAdes 3.62 (ref. 45) with standard parameters and *k*-mers 21, 33, 55, 77 and 99.

The reads of each library were mapped back to the assembled scaffolds using bbmap v37.9 (<https://sourceforge.net/projects/bbmap/>) with the option *fast = t*. Scaffolds were binned on the basis of the mapped read data using MetaBAT v1.0⁴⁶. The binning was refined using Bandage v0.8.1⁴⁷ by collecting all contigs linked to the contig that contained the full-length 16S rRNA gene of the target organism. The bin quality metrics were computed with QUAST v5.0.2⁴⁸ and the completeness for all bins was estimated using checkM v1.07 (ref. 49).

Annotation of the symbiont draft genomes was performed using RAST⁵⁰ and verified with PSI-BLAST v2.7.1⁵¹ for selected genes discussed. Average nucleotide

and amino acid identities between genomes^{52,53} were calculated with the ANI/AAI matrix calculator (<http://enve-omics.ce.gatech.edu/g-matrix>). Comparative analyses were conducted using the PATRIC database and services⁵⁴. Pathway Tools v22.0⁵⁵, in combination with the BioCyc database⁵⁶, was used to analyse the metabolic capacities of '*G. incantans*' and '*R. eludens*'. The genomes were screened for secretion systems and effectors using EffectiveDB⁵⁷.

Transcriptomic analyses. Adapters and rRNA gene reads were removed from the RNA-seq reads using bbduk v37.9. The gene expression for each symbiont genome bin and of the host (based on the published predicted proteome of *T. adhaerens* H1) was calculated from RNA-seq libraries using kallisto v0.45.0 with default settings⁵⁸. Transcription levels were mapped onto metabolic pathways using Pathway Tools v22.0⁵⁵.

Proteomic analyses. Peptide samples for proteomics were prepared and quantified from two samples of 10 *Trichoplax* each and one sample of 30 *Trichoplax* specimens, as described by Kleiner et al.⁵⁹ and according to the filter-aided sample preparation protocol described by Wisniewski et al.⁶⁰. In addition to minor modifications described in Hamann et al.⁶¹, we did not clear the lysate by centrifugation after boiling the sample in lysis buffer. Instead, as the sample size was extremely limited (10 *Trichoplax* specimens = 0.2 µl), we loaded the whole lysate onto the filter units used for the filter-aided sample preparation procedure. Centrifugation times before column washes with 100 µl UA (8 M urea in 0.1 M Tris/HCl pH 8.5) were halved as compared to Hamann et al.⁶¹. Peptides were not desalted. Peptide concentrations were determined with the Pierce Micro BCA assay (Thermo Fisher Scientific) following the manufacturer's instructions.

All samples were analysed by one-dimensional LC-MS/MS as described in Kleiner et al.⁵⁹ with the modification that a 75 cm analytical column was used. Briefly, the sample containing 30 specimens was measured in technical replicate, for the others the whole sample was used in one analysis. The peptide (0.8–3 µg) was loaded with an UltiMate 3000 RSLCnano Liquid Chromatograph (Thermo Fisher Scientific) in loading solvent A (2% acetonitrile, 0.05% trifluoroacetic acid) onto a 5 mm × 300 µm ID C18 Acclaim PepMap100 pre-column (Thermo Fisher Scientific). Elution and separation of peptides on the analytical column (75 cm × 75 µm analytical EASY-Spray column packed with PepMap RSLC C18, 2 µm material, Thermo Fisher Scientific; heated to 60 °C) was performed at a flow rate of 225 nl min⁻¹ using a 460 min gradient going from 98% buffer A (0.1% formic acid) to 31% buffer B (0.1% formic acid, 80% acetonitrile) in 363 min, then to 50% B in 70 min, to 99% B in 1 min and ending with 99% B. The analytical column was connected to a Q Exactive Plus Hybrid Quadrupole-Orbitrap mass spectrometer (Thermo Fisher Scientific) via an Easy-Spray source. Eluting peptides were ionized via electrospray ionization. Carry-over was reduced by two wash runs (injection of 20 µl acetonitrile, 99% eluent B) between samples. Data acquisition in the Q Exactive Plus was performed as in Petersen et al.⁶².

A database containing protein sequences from the *Trichoplax* host as well as the two symbionts was used. Sequences of common laboratory contaminants were included by appending the cRAP protein sequence database (<http://www.thegpm.org/crap/>). The final database contained 13,801 protein sequences. Searches of the MS/MS spectra against this database were performed with the Sequest HT node in Proteome Discoverer v2.2.0.388 (Thermo Fisher Scientific) as in Petersen et al.⁶². For protein quantification, normalized spectral abundance factors⁶³ were calculated per species and multiplied by 100, to give the relative protein abundance as a percentage.

Phylogenetic and phylogenomic analyses. A 16S rRNA gene database for '*G. incantans*' was constructed using the assembled 16S rRNA gene sequence from each metagenomic library, the 20 best BLAST⁶⁴ hits in the nr database and all other sequences of described *Candidatus* taxa in the Midichloriaceae. We added the five type strains with the best BLAST hit score (five species of *Rickettsia*) as an outgroup. We also screened the trace reads from the *Trichoplax* H1 genome project for reads containing Midichloriaceae 16S rRNA gene fragments using BLAST v2.7.1⁶⁴, assembled them in Geneious R9 (<http://www.geneious.com>)⁴² and added the resulting sequence to the database. A similar search for margulisbacterial 16S rRNA fragments yielded no hits.

The 16S rRNA gene dataset was aligned using MAFFT v7.394⁴⁵ and the phylogenetic tree was calculated using FastTree v2.1.10⁴³ with a GTR model for nucleotide substitution. The tree was drawn with Geneious R9⁴².

For '*G. incantans*', the database of genomes for phylogenetic analysis was compiled from all available genomes from the Midichloriaceae as well as representatives for all genera of the Anaplasmataceae and Rickettsiaceae. We also screened the assembly of the *Trichoplax* H1 genome project for contigs that belong to Midichloriaceae contamination using BLAST v2.7.1⁶⁴ with the '*G. incantans*' genome as implemented in Geneious R9 (<http://www.geneious.com>)⁴². The identified set of contigs corresponded to the set found by Driscoll et al.²⁵ and was added to the database. We similarly searched for sequences related to '*R. eludens*' in the H1 genome project, but no hits were detected.

For genome-based alignments of the amino acids of 43 conserved phylogenetic marker genes, the tree workflow as implemented in CheckM v1.0.11 was used⁴⁹. For *Ruthmannia*, the genome bin data were integrated into a taxonomically

selected part of the alignment from Hug et al.¹¹ that covered all Melainabacteria and Cyanobacteria, WOR-1 and RBX-1 (Margulisbacteria), as well as five short branching Firmicutes as an outgroup. The phylogenetic reconstructions of the concatenated alignments were calculated using FastTree v2.1.10 with the WAG model for amino acid substitutions and visualized and analysed using iTOL⁶⁶.

Tag-sequence data analysis. The 16S rRNA gene sequences from *G. incantans*, as well as representative sequences from all characterized midichloriacean *Candidatus* taxa were used as query sequences to search the global collection of the microbial tag-sequencing library. The search was carried out using the IMGNS service⁶⁷ with a minimal alignment length of 200 bp and a minimal identity of 99%. Identified amplicon libraries were grouped according to their deposited metadata. For the top 10% of libraries with the highest number of sequences from *G. incantans*, the habitat type (limnic or marine) and geolocation were manually collected from the deposited metadata and related publications. The detected 16S rRNA reads were aligned to the Rickettsiales dataset using MAFFT—addfragments and the evolutionary placements in the tree were performed using raxml v8.2.12⁶⁸.

TEM. Live specimens were high-pressure frozen with a HPM 100 (Leica Microsystems) in 3 mm aluminium sample holders, using hexane as filler as needed. The samples were transferred onto frozen acetone containing 1% osmium tetroxide and processed using a very quick freeze-substitution method⁶⁹. After reaching room temperature, the samples were washed three times with acetone and infiltrated using centrifugation, modified after McDonald⁷⁰, in 2 ml tubes sequentially with 25%, 50%, 75% and 2 × 100% Agar Low Viscosity Resin (Agar Scientific). For this process, the samples were placed on top of the resin and centrifuged for 30 s with a benchtop centrifuge (Heathrow Scientific) at 2,000g for each step. After the second pure resin step, they were transferred into fresh resin in embedding moulds and polymerized at 60 °C for 12 h.

Ultrathin (70 nm) sections were cut with an Ultracut UC7 (Leica Microsystems) and mounted on formvar-coated slot grids (Agar Scientific). They were contrasted with 0.5% aqueous uranyl acetate (Science Services) for 20 min and with 2% Reynold's lead citrate for 6 min before imaging them at 20–30 kV with a Quanta FEG 250 transmission electron microscope (FEI Company) equipped with a scanning TEM detector using the xT microscope control software v6.2.6.3123.

For electron tomography, 300 nm serial sections were placed on formvar-coated 2 × 1 mm² slot grids and stained with uranyl acetate and lead citrate. 30 nm gold fiducials were applied on both sides of the slot grid. Dual-axis tilt series (±60°, step size 1°) were acquired with a FEI Tecnai F30 300 kV electron microscope equipped with an Axial Gatan US1000 CCD camera. SerialEM software was used for the automated acquisition of tomographic tilt series⁷¹. Alignment and reconstruction of the tilt series were carried out with IMOD v4.9⁷². The serial tomograms were aligned with TrakEM2 v1.01⁷³ in Fiji⁷⁴ and visualization and segmentation were carried out using the software Amira 3D v6.5.

FISH. We used ARB—SILVA database 128 (ref. ⁷⁵) and the ARB PROBE_DESIGN tool (the ARB software package v6.0.6)⁷⁶ to design two FISH probes for each symbiont that were specific to their 16S rRNA sequences (Supplementary Table 2). We confirmed the specificity of the probes by comparing their sequences to all available sequences in the ARB—SILVA 128 database and Ribosomal Database Project release 11.5 (ref. ⁷⁷). The most specific probe for *R. eludens* had two mismatches to first non-target hit sequences; the most specific probe for *G. incantans* also matched the six most closely related *Grellia* sequences; detailed results are presented in Supplementary Table 2.

Specimens were fixed on coverslips with 2% formaldehyde and 0.1% glutaraldehyde in 1.5 × PIPES, HEPES, EGTA and MgCl₂ (PHEM) buffers modified from Montanaro et al.⁷⁸ at 4 °C for 12 h. After three washing steps in 1.5 × PHEM buffer, the samples were stored in 70% ethanol until use. Samples were rehydrated in PBS and hybridization was performed according to Manz et al.⁷⁹. Mono-labelled-, DOPE-⁸⁰ or MIL-⁸¹ probes (Supplementary Table 2) at a concentration of 8.4 pmol μl⁻¹ were diluted with hybridization buffer containing 35% formamide, 900 mM NaCl, 20 mM Tris/HCl and 0.01% SDS at a ratio of 15:1. Whole animals were incubated in 30 μl of the probe/hybridization buffer mix at 46 °C in 250 μl PCR tubes for 3–4 h, followed by a 30-min-long washing step in washing buffer containing 700 mM NaCl, 20 mM Tris/HCl, 5 mM EDTA and 0.1% SDS. After a 10-min-long washing step in PBS, the animals were stained with DAPI for 30 min, washed twice again in PBS and mounted on glass slides in Vectashield mounting medium.

To test the probes designed for this study, 30 clonal individuals of *Trichoplax* H2 were pooled, fixed as described above, homogenized by sonication and applied to a filter. The parts of the filter were then tested with different formamide concentrations and the optimal formamide concentration was determined.

Fluorescence images were taken with a Zeiss LSM 780 equipped with a GaAsP detector or an Airyscan detector and a Plan-Apochromat 63×/1.4 and a Plan-Apochromat 100×/1.46 oil immersion objective using the ZEN software (black edition, 64bits, v14.0.1.201; Carl Zeiss Microscopy GmbH).

Reporting Summary. Further information on research design is available in the Nature Research Reporting Summary linked to this article.

Data availability

The metagenomic and metatranscriptomic raw reads and assembled symbiont genomes are available in the European Nucleotide Archive under Study Accession Number PRJEB30343. The mass spectrometry metaproteomics data and protein sequence databases were deposited in the ProteomeXchange Consortium⁸² via the PRIDE partner repository with the dataset PXD012106. The TEM 3D reconstruction data were deposited in figshare; the aligned tomography slices used for the reconstruction shown in Fig. 4 are available at <https://figshare.com/s/886b869a9ada0264ffb2> (ref. ³¹).

Code availability

The script used for the assembly graph-based binning is available at https://github.com/HRGV/tools_and_scripts.

Received: 31 December 2018; Accepted: 26 April 2019;

Published online: 10 June 2019

References

- Laumer, C. E. et al. Support for a clade of Placozoa and Cnidaria in genes with minimal compositional bias. *eLife* **7**, e36278 (2018).
- Eitel, M., Osigus, H. J., DeSalle, R. & Schierwater, B. Global diversity of the Placozoa. *PLoS ONE* **8**, e57131 (2013).
- Grell, K. G. & Benwitz, G. Die Ultrastruktur von *Trichoplax adhaerens* F.E. Schulze. *Cytobiologie* **4**, 216–240 (1971).
- Eitel, M., Guidi, L., Hadrys, H., Balsamo, M. & Schierwater, B. New insights into placozoan sexual reproduction and development. *PLoS ONE* **6**, e19639 (2011).
- Guidi, L., Eitel, M., Cesarini, E., Schierwater, B. & Balsamo, M. Ultrastructural analyses support different morphological lineages in the phylum Placozoa Grell, 1971. *J. Morphol.* **272**, 371–378 (2011).
- Eitel, M. et al. Comparative genomics and the nature of placozoan species. *PLoS Biol.* **16**, e2005359 (2018).
- Montagna, M. et al. “*Candidatus* Midichloriaceae” fam. nov. (*Rickettsiales*), an ecologically widespread clade of intracellular alphaproteobacteria. *Appl. Environ. Microbiol.* **79**, 3241–3248 (2013).
- Castelli, M., McCarthy, U., Petroni, G. & Bazzocchi, C. in *Rickettsiales: Biology, Molecular Biology, Epidemiology, and Vaccine Development* (ed. Thomas, S.) 283–292 (Springer, 2016).
- Schmitz-Esser, S. et al. ATP/ADP translocases: a common feature of obligate intracellular amoebal symbionts related to Chlamydiae and Rickettsiae. *J. Bacteriol.* **186**, 683–691 (2004).
- Driscoll, T. P. et al. Wholly *Rickettsia*! Reconstructed metabolic profile of the quintessential bacterial parasite of eukaryotic cells. *mBio* **8**, e00859-17 (2017).
- Hug, L. A. et al. A new view of the tree of life. *Nat. Microbiol.* **1**, 16048 (2016).
- Anantharaman, K. et al. Thousands of microbial genomes shed light on interconnected biogeochemical processes in an aquifer system. *Nat. Commun.* **7**, 13219 (2016).
- Matheus Carnevali, P. B. et al. Hydrogen-based metabolism as an ancestral trait in lineages sibling to the Cyanobacteria. *Nat. Commun.* **10**, 463 (2019).
- Simion, P. et al. A large and consistent phylogenomic dataset supports sponges as the sister group to all other animals. *Curr. Biol.* **27**, 958–967 (2017).
- Srivastava, M. et al. The *Trichoplax* genome and the nature of placozoans. *Nature* **454**, 955–960 (2008).
- Sebé-Pedrós, A. et al. Early metazoan cell type diversity and the evolution of multicellular gene regulation. *Nat. Ecol. Evol.* **2**, 1176–1188 (2018).
- Pearse, V. B. & Voigt, O. Field biology of placozoans (*Trichoplax*): distribution, diversity, biotic interactions. *Integr. Comp. Biol.* **47**, 677–692 (2007).
- Smith, C. L., Pivovarova, N. & Reese, T. S. Coordinated feeding behavior in *Trichoplax*, an animal without synapses. *PLoS ONE* **10**, e0136098 (2015).
- Grell, K. G. & Benwitz, G. Ergänzende Untersuchungen zur Ultrastruktur von *Trichoplax adhaerens* F.E. Schulze (Placozoa). *Zoomorphologie* **98**, 47–67 (1981).
- Smith, C. L. et al. Novel cell types, neurosecretory cells, and body plan of the early-diverging metazoan *Trichoplax adhaerens*. *Curr. Biol.* **24**, 1565–1572 (2014).
- Martinez, E., Siadous, F. A. & Bonazzi, M. Tiny architects: biogenesis of intracellular replicative niches by bacterial pathogens. *FEMS Microbiol. Rev.* **42**, 425–447 (2018).
- Schulz, F. & Horn, M. Intracellular bacteria: inside the cellular control center of eukaryotes. *Trends Cell Biol.* **25**, 339–346 (2015).
- Sassera, D. et al. ‘*Candidatus* Midichloria mitochondrii’, an endosymbiont of the tick *Ixodes ricinus* with a unique intramitochondrial lifestyle. *Int. J. Syst. Evol. Bacteriol.* **56**, 2535–2540 (2006).

24. Kamm, K., Osigus, H. J., Stadler, P. F., DeSalle, R. & Schierwater, B. Trichoplax genomes reveal profound admixture and suggest stable wild populations without bisexual reproduction. *Sci. Rep.* **8**, 11168 (2018).
25. Driscoll, T., Gillespie, J. J., Nordberg, E. K., Azad, A. F. & Sobral, B. W. Bacterial DNA sifted from the *Trichoplax adhaerens* (Animalia: Placozoa) genome project reveals a putative rickettsial endosymbiont. *Genome Biol. Evol.* **5**, 621–645 (2013).
26. Bowers, R. M. et al. Minimum information about a single amplified genome (MISAG) and a metagenome-assembled genome (MIMAG) of bacteria and archaea. *Nat. Biotechnol.* **35**, 725–731 (2017).
27. Castelli, M., Sasser, D. & Petroni, G. In *Rickettsiales: Biology, Molecular Biology, Epidemiology, and Vaccine Development* (ed. Thomas, S.) 59–91 (Springer, 2016).
28. Senra, M. V. et al. A house for two—double bacterial infection in *Euplotes woodruffi* Sq1 (Ciliophora, Euplotia) sampled in southeastern Brazil. *Microb. Ecol.* **71**, 505–517 (2016).
29. Soo, R. M., Hemp, J., Parks, D. H., Fischer, W. W. & Hugenholtz, P. On the origins of oxygenic photosynthesis and aerobic respiration in Cyanobacteria. *Science* **355**, 1436–1440 (2017).
30. Schulz, F. et al. A Rickettsiales symbiont of amoebae with ancient features. *Environ. Microbiol.* **18**, 2326–2342 (2016).
31. Leisch, N. Electron tomography dataset of a serial sectioned *Trichoplax* haplotype H2 fiber cell. *Figshare* <https://figshare.com/s/886b869a9ada0264ffb2> (2018).
32. Parks, D. H. et al. A standardized bacterial taxonomy based on genome phylogeny substantially revises the tree of life. *Nat. Biotechnol.* **36**, 996–1004 (2018).
33. Parks, D. H. et al. Recovery of nearly 8,000 metagenome-assembled genomes substantially expands the tree of life. *Nat. Microbiol.* **2**, 1533–1542 (2017).
34. Bright, M. & Sorgo, A. Ultrastructural reinvestigation of the trophosome in adults of *Riftia pachyptila* (Annelida, Siboglinidae). *Invertebr. Biol.* **122**, 347–368 (2003).
35. Fiala-Médioni, A., Michalski, J.-C., Jollès, J., Alonso, C. & Montreuil, J. Lysosomal and lysozyme activities in the gill of bivalves from deep hydrothermal vents. *C. R. Acad. Sci. III* **317**, 239–244 (1994).
36. Ringrose, J. H. et al. Deep proteome profiling of *Trichoplax adhaerens* reveals remarkable features at the origin of metazoan multicellularity. *Nat. Commun.* **4**, 1408 (2013).
37. Sedzicki, J. et al. 3D correlative electron microscopy reveals continuity of Brucella-containing vacuoles with the endoplasmic reticulum. *J. Cell Sci.* **131**, jcs210799 (2018).
38. Fraune, S. & Bosch, T. C. G. Long-term maintenance of species-specific bacterial microbiota in the basal metazoan *Hydra*. *Proc. Natl Acad. Sci. USA* **104**, 13146–13151 (2007).
39. Horn, M. et al. Illuminating the evolutionary history of Chlamydiae. *Science* **304**, 728–730 (2004).
40. Duron, O., Doublet, P., Vavre, F. & Bouchon, D. The importance of revisiting Legionellales diversity. *Trends Parasitol.* **34**, 1027–1037 (2018).
41. Werren, J. H., Baldo, L. & Clark, M. E. *Wolbachia*: master manipulators of invertebrate biology. *Nat. Rev. Microbiol.* **6**, 741–751 (2008).
42. Kearse, M. et al. Geneious Basic: an integrated and extendable desktop software platform for the organization and analysis of sequence data. *Bioinformatics* **28**, 1647–1649 (2012).
43. Price, M. N., Dehal, P. S. & Arkin, A. P. FastTree 2—approximately maximum-likelihood trees for large alignments. *PLoS ONE* **5**, e9490 (2010).
44. Nikolenko, S. I., Korobeynikov, A. I. & Alekseyev, M. A. BayesHammer: Bayesian clustering for error correction in single-cell sequencing. *BMC Genom.* **14**, S7 (2013).
45. Bankevich, A. et al. SPAdes: a new genome assembler and its applications to single cell sequencing. *J. Comput. Biol.* **19**, 455–477 (2012).
46. Kang, D. D., Froula, J., Egan, R. & Wang, Z. MetaBAT, an efficient tool for accurately reconstructing single genomes from complex microbial communities. *PeerJ* **3**, e1165 (2015).
47. Wick, R. R., Schultz, M. B., Zobel, J. & Holt, K. E. Bandage: interactive visualization of de novo genome assemblies. *Bioinformatics* **31**, 3350–3352 (2015).
48. Gurevich, A., Saveliev, V., Vyahhi, N. & Tesler, G. QUAST: quality assessment tool for genome assemblies. *Bioinformatics* **29**, 1072–1075 (2013).
49. Parks, D. H., Imelfort, M., Skennerton, C. T., Hugenholtz, P. & Tyson, G. W. CheckM: assessing the quality of microbial genomes recovered from isolates, single cells, and metagenomes. *Genome Res.* **25**, 1043–1055 (2015).
50. Aziz, R. K. et al. The RAST server: rapid annotations using subsystems technology. *BMC Genom.* **9**, 75 (2008).
51. Johnson, M. et al. NCBI BLAST: a better web interface. *Nucleic Acids Res.* **36**, W5–W9 (2008).
52. Richter, M. & Rossello-Mora, R. Shifting the genomic gold standard for the prokaryotic species definition. *Proc. Natl Acad. Sci. USA* **106**, 19126–19131 (2009).
53. Goris, J. et al. DNA–DNA hybridization values and their relationship to whole-genome sequence similarities. *Int. J. Syst. Evol. Bacteriol.* **57**, 81–91 (2007).
54. Wattam, A. R. et al. Improvements to PATRIC, the all-bacterial bioinformatics database and analysis resource center. *Nucleic Acids Res.* **45**, D535–D542 (2017).
55. Karp, P. D. et al. Pathway Tools version 13.0: integrated software for pathway/genome informatics and systems biology. *Brief. Bioinform.* **11**, 40–79 (2010).
56. Caspi, R. et al. The MetaCyc database of metabolic pathways and enzymes and the BioCyc collection of pathway/genome databases. *Nucleic Acids Res.* **44**, D471–D480 (2016).
57. Eichinger, V. et al. EffectiveDB—updates and novel features for a better annotation of bacterial secreted proteins and Type III, IV, VI secretion systems. *Nucleic Acids Res.* **44**, D669–D674 (2016).
58. Bray, N. L., Pimentel, H., Melsted, P. & Pachter, L. Near-optimal probabilistic RNA-seq quantification. *Nat. Biotechnol.* **34**, 525–527 (2016).
59. Kleiner, M. et al. Assessing species biomass contributions in microbial communities via metaproteomics. *Nat. Commun.* **8**, 1558 (2017).
60. Wiśniewski, J. R., Zougman, A., Nagaraj, N. & Mann, M. Universal sample preparation method for proteome analysis. *Nat. Methods* **6**, 359–362 (2009).
61. Hamann, E. et al. Environmental Breviatea harbour mutualistic *Arcobacter* epibionts. *Nature* **534**, 254–258 (2016).
62. Petersen, J. M. et al. Chemosynthetic symbionts of marine invertebrate animals are capable of nitrogen fixation. *Nat. Microbiol.* **2**, 16195 (2016).
63. Zybailov, B. et al. Statistical analysis of membrane proteome expression changes in *Saccharomyces cerevisiae*. *J. Proteome Res.* **5**, 2339–2347 (2006).
64. Altschul, S. F. et al. Gapped BLAST and PSI-BLAST: a new generation of protein database search programs. *Nucleic Acids Res.* **25**, 3389–3402 (1997).
65. Katoh, K. & Standley, D. M. MAFFT multiple sequence alignment software version 7: improvements in performance and usability. *Mol. Biol. Evol.* **30**, 772–780 (2013).
66. Letunic, I. & Bork, P. Interactive tree of life (iTOL) v3: an online tool for the display and annotation of phylogenetic and other trees. *Nucleic Acids Res.* **44**, W242–W245 (2016).
67. Lagkouvardos, I. et al. IMNGS: a comprehensive open resource of processed 16S rRNA microbial profiles for ecology and diversity studies. *Sci. Rep.* **6**, 33721 (2016).
68. Berger, S. A., Krompass, D. & Stamatakis, A. Performance, accuracy, and web server for evolutionary placement of short sequence reads under maximum likelihood. *Syst. Biol.* **60**, 291–302 (2011).
69. McDonald, K. L. & Webb, R. I. Freeze substitution in 3 hours or less. *J. Microsc.* **243**, 227–233 (2011).
70. McDonald, K. L. Rapid embedding methods into epoxy and LR white resins for morphological and immunological analysis of cryofixed biological specimens. *Microsc. Microanal.* **20**, 152–163 (2014).
71. Mastronarde, D. N. Automated electron microscope tomography using robust prediction of specimen movements. *J. Struct. Biol.* **152**, 36–51 (2005).
72. Kremer, J. R., Mastronarde, D. N. & McIntosh, J. R. Computer visualization of three-dimensional image data using IMOD. *J. Struct. Biol.* **116**, 71–76 (1996).
73. Cardona, A. et al. TrakEM2 software for neural circuit reconstruction. *PLoS ONE* **7**, e38011 (2012).
74. Schindelin, J. et al. Fiji: an open-source platform for biological-image analysis. *Nat. Methods* **9**, 676–682 (2012).
75. Quast, C. et al. The SILVA ribosomal RNA gene database project: improved data processing and web-based tools. *Nucleic Acids Res.* **41**, D590–D596 (2013).
76. Ludwig, W. et al. ARB: a software environment for sequence data. *Nucleic Acids Res.* **32**, 1363–1371 (2004).
77. Cole, J. R. et al. Ribosomal Database Project: data and tools for high throughput rRNA analysis. *Nucleic Acids Res.* **42**, D633–D642 (2014).
78. Montanaro, J., Gruber, D. & Leisch, N. Improved ultrastructure of marine invertebrates using non-toxic buffers. *PeerJ* **4**, e1860 (2016).
79. Manz, W., Amann, R., Ludwig, W., Wagner, M. & Schleifer, K.-H. Phylogenetic oligodeoxynucleotide probes for the major subclasses of Proteobacteria: problems and solutions. *Syst. Appl. Microbiol.* **15**, 593–600 (1992).
80. Stoecker, K., Dorninger, C., Daims, H. & Wagner, M. Double labeling of oligonucleotide probes for fluorescence in situ hybridization (DOPE–FISH) improves signal intensity and increases rRNA accessibility. *Appl. Environ. Microbiol.* **76**, 922–926 (2010).
81. Schimak, M. P. et al. MiL–FISH: multilabeled oligonucleotides for fluorescence in situ hybridization improve visualization of bacterial cells. *Appl. Environ. Microbiol.* **82**, 62–70 (2015).
82. Vizcaino, J. A. et al. 2016 update of the PRIDE database and its related tools. *Nucleic Acids Res.* **44**, D447–D456 (2016).

Acknowledgements

This study was funded by the Max Planck Society, a Gordon and Betty Moore Foundation Marine Microbial Initiative Investigator Award (grant no. GBMF3811 to N.D.), grants to M.G.H. from the Gordon and Betty Moore Foundation (no. 5009) and the US Office of Naval Research (no. N00014-15-1-2658), the German Academic Exchange Service DAAD (T.H.) and the NC State Chancellor's Faculty Excellence Program Cluster on Microbiomes and Complex Microbial Communities (M.K.). We thank M. Strous for access to proteomic equipment and A. Kouris for LC-MS/MS operation. The purchase of the proteomics equipment was supported by a grant from the Canadian Foundation for Innovation to M. Strous. We thank the Electron Microscopy Facility of the MPI-CBG, the Max Planck Genome Centre and the Electron Microscopy Facility at the MPI-PZ Cologne and the Cell Imaging and Ultrastructure Research Core Facility of the University of Vienna for technical support. We thank C. Peters for nucleic acid extractions, C. Peters, M. Meyer and W. Ruschmeier for support with *Trichoplax* cultivation, B. Nedved for support in the field and G. Bennett, T. Erb, L. Schada von Borzyskowski and P. A. Chakkiath for discussions on symbiont physiology.

Author contributions

M.G.H., N.D., M.M.-N. and H.R.G.-V. conceived the study. H.R.G.-V. sampled and cultivated the organisms and performed the assemblies, genome and transcriptome analyses, tag-sequencing analyses and phylogenetic analyses. H.R.G.-V. reconstructed the symbiont physiology with the help of M.L. and M.K. M.K. and T.H. generated the proteomic data and H.R.G.-V. and M.K. analysed the proteomic data. N.L. performed the fluorescence microscopy, electron microscopy and electron tomography, subsequent data analysis and 3D reconstruction. H.R.G.-V. and N.L. wrote the manuscript with support from N.D., M.M.-N. and M.G.H. All authors revised the manuscript and approved the final version.

Competing interests

The authors declare no competing interests.

Additional information

Supplementary information is available for this paper at <https://doi.org/10.1038/s41564-019-0475-9>.

Reprints and permissions information is available at www.nature.com/reprints.

Correspondence and requests for materials should be addressed to H.R.G., M.G.H. or N.D.

Publisher's note: Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

© The Author(s), under exclusive licence to Springer Nature Limited 2019



Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons license, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons license and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this license, visit <http://creativecommons.org/licenses/by/4.0/>.

Reporting Summary

Nature Research wishes to improve the reproducibility of the work that we publish. This form provides structure for consistency and transparency in reporting. For further information on Nature Research policies, see [Authors & Referees](#) and the [Editorial Policy Checklist](#).

Statistics

For all statistical analyses, confirm that the following items are present in the figure legend, table legend, main text, or Methods section.

n/a Confirmed

- The exact sample size (n) for each experimental group/condition, given as a discrete number and unit of measurement
- A statement on whether measurements were taken from distinct samples or whether the same sample was measured repeatedly
- The statistical test(s) used AND whether they are one- or two-sided
Only common tests should be described solely by name; describe more complex techniques in the Methods section.
- A description of all covariates tested
- A description of any assumptions or corrections, such as tests of normality and adjustment for multiple comparisons
- A full description of the statistical parameters including central tendency (e.g. means) or other basic estimates (e.g. regression coefficient) AND variation (e.g. standard deviation) or associated estimates of uncertainty (e.g. confidence intervals)
- For null hypothesis testing, the test statistic (e.g. F , t , r) with confidence intervals, effect sizes, degrees of freedom and P value noted
Give P values as exact values whenever suitable.
- For Bayesian analysis, information on the choice of priors and Markov chain Monte Carlo settings
- For hierarchical and complex designs, identification of the appropriate level for tests and full reporting of outcomes
- Estimates of effect sizes (e.g. Cohen's d , Pearson's r), indicating how they were calculated

Our web collection on [statistics for biologists](#) contains articles on many of the points above.

Software and code

Policy information about [availability of computer code](#)

Data collection

Sequencing data were produced on an Illumina HiSeq 2500 instrument with manufacturer's software. For the 16S tag sequencing analysis, the IMNGS web-platform was used to gather data; see <https://www.imngs.org/> tomography tilt-series were acquired using SerialEM. Transmission electron micrographs were acquired with the xT microscope control software ver. 6.2.6.3123

Data analysis

PhyloFlash 3.0 beta1- <https://github.com/HRGV/phyloFlash>
 BBmap suite v37.9 from Bushnell B. - sourceforge.net/projects/bbmap
 BayesHammer 3.6.2
 SPAdes 3.62
 MetaBAT 1.0
 Bandage 0.8.1
 QUAST 5.0.2
 checkM 1.0.11
 Rast <http://rast.nmpdr.org/>
 ANI/AAI matrix calculator - <http://enve-omics.ce.gatech.edu/g-matrix/>
 PATRIC <https://www.patricbrc.org>
 Pathway Tools 22.0
 EffectiveDB <https://effectors.csb.univie.ac.at>
 kallisto v0.45.0
 Proteome Discoverer version 2.2.0.388
 Geneious R11 <http://www.geneious.com>
 iTOL <https://itol.embl.de>
 arb software package 6.0.6.
 Ribosomal Database Project <https://rdp.cme.msu.edu/>
 raxml 8.2.12

mafft 7.394
 fasttree 2.1.10
 IMOD 4.9
 TrakEM2 1.0i
 FIJI
 Amira 3D 6.5
 ZEN black 14.0.1.201

For manuscripts utilizing custom algorithms or software that are central to the research but not yet described in published literature, software must be made available to editors/reviewers. We strongly encourage code deposition in a community repository (e.g. GitHub). See the Nature Research [guidelines for submitting code & software](#) for further information.

Data

Policy information about [availability of data](#)

All manuscripts must include a [data availability statement](#). This statement should provide the following information, where applicable:

- Accession codes, unique identifiers, or web links for publicly available datasets
- A list of figures that have associated raw data
- A description of any restrictions on data availability

The metagenomic and metatranscriptomic raw reads and assembled symbiont genomes are available in the European Nucleotide Archive under Study Accession Number PRJEB30343

The mass spectrometry metaproteomics data and protein sequence database were deposited in the ProteomeXchange Consortium via the PRIDE partner repository with the dataset PXD012106 (access for reviewers: <https://www.ebi.ac.uk/pride/archive/login>, email/Username: reviewer71196@ebi.ac.uk, password: RSqmy9cK)

The aligned tomography slices used for the 3D-reconstruction shown in Figure 4 and the Supplementary video were deposited in figshare and are available at <https://figshare.com/s/886b869a9ada0264ffb2> or under the doi 10.6084/m9.figshare.7429793

Field-specific reporting

Please select the one below that is the best fit for your research. If you are not sure, read the appropriate sections before making your selection.

Life sciences Behavioural & social sciences Ecological, evolutionary & environmental sciences

For a reference copy of the document with all sections, see [nature.com/documents/nr-reporting-summary-flat.pdf](https://www.nature.com/documents/nr-reporting-summary-flat.pdf)

Life sciences study design

All studies must disclose on these points even when the disclosure is negative.

Sample size	<i>Describe how sample size was determined, detailing any statistical methods used to predetermine sample size OR if no sample-size calculation was performed, describe how sample sizes were chosen and provide a rationale for why these sample sizes are sufficient.</i>
Data exclusions	<i>Describe any data exclusions. If no data were excluded from the analyses, state so OR if data were excluded, describe the exclusions and the rationale behind them, indicating whether exclusion criteria were pre-established.</i>
Replication	<i>Describe the measures taken to verify the reproducibility of the experimental findings. If all attempts at replication were successful, confirm this OR if there are any findings that were not replicated or cannot be reproduced, note this and describe why.</i>
Randomization	<i>Describe how samples/organisms/participants were allocated into experimental groups. If allocation was not random, describe how covariates were controlled OR if this is not relevant to your study, explain why.</i>
Blinding	<i>Describe whether the investigators were blinded to group allocation during data collection and/or analysis. If blinding was not possible, describe why OR explain why blinding was not relevant to your study.</i>

Behavioural & social sciences study design

All studies must disclose on these points even when the disclosure is negative.

Study description	<i>Briefly describe the study type including whether data are quantitative, qualitative, or mixed-methods (e.g. qualitative cross-sectional, quantitative experimental, mixed-methods case study).</i>
Research sample	<i>State the research sample (e.g. Harvard university undergraduates, villagers in rural India) and provide relevant demographic information (e.g. age, sex) and indicate whether the sample is representative. Provide a rationale for the study sample chosen. For studies involving existing datasets, please describe the dataset and source.</i>
Sampling strategy	<i>Describe the sampling procedure (e.g. random, snowball, stratified, convenience). Describe the statistical methods that were used to predetermine sample size OR if no sample-size calculation was performed, describe how sample sizes were chosen and provide a rationale for why these sample sizes are sufficient. For qualitative data, please indicate whether data saturation was considered, and what criteria were used to decide that no further sampling was needed.</i>

Data collection	<i>Provide details about the data collection procedure, including the instruments or devices used to record the data (e.g. pen and paper, computer, eye tracker, video or audio equipment) whether anyone was present besides the participant(s) and the researcher, and whether the researcher was blind to experimental condition and/or the study hypothesis during data collection.</i>
Timing	<i>Indicate the start and stop dates of data collection. If there is a gap between collection periods, state the dates for each sample cohort.</i>
Data exclusions	<i>If no data were excluded from the analyses, state so OR if data were excluded, provide the exact number of exclusions and the rationale behind them, indicating whether exclusion criteria were pre-established.</i>
Non-participation	<i>State how many participants dropped out/declined participation and the reason(s) given OR provide response rate OR state that no participants dropped out/declined participation.</i>
Randomization	<i>If participants were not allocated into experimental groups, state so OR describe how participants were allocated to groups, and if allocation was not random, describe how covariates were controlled.</i>

Ecological, evolutionary & environmental sciences study design

All studies must disclose on these points even when the disclosure is negative.

Study description	Placozoans were isolated from a coral tank to study their symbiotic association with intracellular bacteria
Research sample	Placozoans were isolated from a coral tank at the Kewalo Marine Laboratory, University of Hawai'i at Mānoa, Honolulu, Hawai'i by placing glass slides mounted in cut-open plastic slide holders into the tank for 10 days. Placozoans were identified under a dissection microscope, transferred to 400 ml glass beakers with 34.5 ‰ artificial seawater (ASW) and fed weekly with 2×10^6 cells ml ⁻¹ of <i>Isochrysis galbana</i> from a log-phase culture. At 25°C in 34.5 ‰ ASW and with a 16:8 hour light/dark regime, doubling times were 2-3 days.
Sampling strategy	All individuals used in the experiments originated from an asexually dividing clonal lineage of <i>Trichoplax H2</i> , which was established from a single individual. Individual specimens were sampled from the cultures at chance for all experiments conducted. Only intact, healthy-looking animals were used in the experiments. Single individuals were used for metagenomic, metatranscriptomic and imaging analyses. Samples were pooled (10-30) for metaproteomic analyses. Sample size was chosen according to sizes used in comparable studies.
Data collection	The gene expression data were generated by DNA and RNA sequencing at the Max Planck Genome Centre in Cologne under the supervision of Dr. B. Huettel. Metaproteomics data were generated by mass spectrometry at the University of Calgary by Dr. M. Kleiner and Tjorven Hinzke. Imaging data were collected at the MPLs for Marine Microbiology (Bremen) and of Molecular Cell Biology and Genetics (Dresden) by Dr. Nikolaus Leisch.
Timing and spatial scale	Placozoan specimens were collected in Hawai'i in October 2015 and cultivated in the lab until used for this study. Animal specimens for metagenomics and metatranscriptomics were taken from the lab cultures in February 2016, for proteomics 2017 and for microscopy between January 2016 and September 2018
Data exclusions	none
Reproducibility	DNA metagenomic libraries were constructed from five specimens, transcriptomic libraries from 3 of the five metagenomic specimens. Proteomic data were generated from 3 separate pools of 10 to 30 specimens. Microscopy data were generated from three to nine separate specimens as indicated in the text and materials and methods. Using the deposited raw sequencing, proteomic and imaging data, the data analyses that were performed in this study can be easily and repeatedly reproduced. All attempts to repeat the experiments were successful.
Randomization	All experiments were performed on a clonal lineage and specimens were randomly chosen for experiments.
Blinding	Blinding was not performed because it was not relevant to this study. This study was an exploratory survey of microbial diversity without a priori expectations that would influence the analyses.
Did the study involve field work?	<input type="checkbox"/> Yes <input checked="" type="checkbox"/> No

Field work, collection and transport

Field conditions	<i>Describe the study conditions for field work, providing relevant parameters (e.g. temperature, rainfall).</i>
Location	<i>State the location of the sampling or experiment, providing relevant parameters (e.g. latitude and longitude, elevation, water depth).</i>
Access and import/export	<i>Describe the efforts you have made to access habitats and to collect and import/export your samples in a responsible manner and in compliance with local, national and international laws, noting any permits that were obtained (give the name of the issuing authority, the date of issue, and any identifying information).</i>
Disturbance	<i>Describe any disturbance caused by the study and how it was minimized.</i>

Reporting for specific materials, systems and methods

We require information from authors about some types of materials, experimental systems and methods used in many studies. Here, indicate whether each material, system or method listed is relevant to your study. If you are not sure if a list item applies to your research, read the appropriate section before selecting a response.

Materials & experimental systems

- n/a Involved in the study
- Antibodies
- Eukaryotic cell lines
- Palaeontology
- Animals and other organisms
- Human research participants
- Clinical data

Methods

- n/a Involved in the study
- ChIP-seq
- Flow cytometry
- MRI-based neuroimaging

Antibodies

Antibodies used

Describe all antibodies used in the study; as applicable, provide supplier name, catalog number, clone name, and lot number.

Validation

Describe the validation of each primary antibody for the species and application, noting any validation statements on the manufacturer's website, relevant citations, antibody profiles in online databases, or data provided in the manuscript.

Eukaryotic cell lines

Policy information about [cell lines](#)

Cell line source(s)

State the source of each cell line used.

Authentication

Describe the authentication procedures for each cell line used OR declare that none of the cell lines used were authenticated.

Mycoplasma contamination

Confirm that all cell lines tested negative for mycoplasma contamination OR describe the results of the testing for mycoplasma contamination OR declare that the cell lines were not tested for mycoplasma contamination.

Commonly misidentified lines
(See [ICLAC](#) register)

Name any commonly misidentified cell lines used in the study and provide a rationale for their use.

Palaeontology

Specimen provenance

Provide provenance information for specimens and describe permits that were obtained for the work (including the name of the issuing authority, the date of issue, and any identifying information).

Specimen deposition

Indicate where the specimens have been deposited to permit free access by other researchers.

Dating methods

If new dates are provided, describe how they were obtained (e.g. collection, storage, sample pretreatment and measurement), where they were obtained (i.e. lab name), the calibration program and the protocol for quality assurance OR state that no new dates are provided.

Tick this box to confirm that the raw and calibrated dates are available in the paper or in Supplementary Information.

Animals and other organisms

Policy information about [studies involving animals](#); [ARRIVE guidelines](#) recommended for reporting animal research

Laboratory animals

A clonal lineage of haplotype H2 placozoans was isolated from a coral tank at the Kewalo Marine Laboratory, University of Hawai'i at Mānoa, Honolulu, Hawai'i. Initial specimens were sampled by placing glass slides mounted in cut-open plastic slide holders into the tank for 10 days. Placozoans were identified under a dissection microscope, and single individuals were transferred to 400 ml glass beakers with 34.5 % artificial seawater (ASW). The cultures were fed weekly with 2×10^6 cells ml⁻¹ of *Isochrysis galbana* from a log-phase culture. At 25°C in 34.5 % ASW and with a 16:8 hour light/dark regime, doubling times were 2-3 days. The culture is available upon request.

Wild animals

An initial sample of placozoans was used to isolate a clonal lineage that was then used for all experiments - see section on laboratory animals above. The culture is available upon request.

Field-collected samples

No field collected samples were used.

Ethics oversight

Work on placozoans is not subjected to a approval by an ethics committee

Note that full information on the approval of the study protocol must also be provided in the manuscript.

Human research participants

Policy information about [studies involving human research participants](#)

Population characteristics

Describe the covariate-relevant population characteristics of the human research participants (e.g. age, gender, genotypic information, past and current diagnosis and treatment categories). If you filled out the behavioural & social sciences study design questions and have nothing to add here, write "See above."

Recruitment

Describe how participants were recruited. Outline any potential self-selection bias or other biases that may be present and how these are likely to impact results.

Ethics oversight

Identify the organization(s) that approved the study protocol.

Note that full information on the approval of the study protocol must also be provided in the manuscript.

Clinical data

Policy information about [clinical studies](#)

All manuscripts should comply with the ICMJE [guidelines for publication of clinical research](#) and a completed [CONSORT checklist](#) must be included with all submissions.

Clinical trial registration

Provide the trial registration number from [ClinicalTrials.gov](#) or an equivalent agency.

Study protocol

Note where the full trial protocol can be accessed OR if not available, explain why.

Data collection

Describe the settings and locales of data collection, noting the time periods of recruitment and data collection.

Outcomes

Describe how you pre-defined primary and secondary outcome measures and how you assessed these measures.

ChIP-seq

Data deposition

Confirm that both raw and final processed data have been deposited in a public database such as [GEO](#).

Confirm that you have deposited or provided access to graph files (e.g. BED files) for the called peaks.

Data access links

May remain private before publication.

For "Initial submission" or "Revised version" documents, provide reviewer access links. For your "Final submission" document, provide a link to the deposited data.

Files in database submission

Provide a list of all files available in the database submission.

Genome browser session

(e.g. [UCSC](#))

Provide a link to an anonymized genome browser session for "Initial submission" and "Revised version" documents only, to enable peer review. Write "no longer applicable" for "Final submission" documents.

Methodology

Replicates

Describe the experimental replicates, specifying number, type and replicate agreement.

Sequencing depth

Describe the sequencing depth for each experiment, providing the total number of reads, uniquely mapped reads, length of reads and whether they were paired- or single-end.

Antibodies

Describe the antibodies used for the ChIP-seq experiments; as applicable, provide supplier name, catalog number, clone name, and lot number.

Peak calling parameters

Specify the command line program and parameters used for read mapping and peak calling, including the ChIP, control and index files used.

Data quality

Describe the methods used to ensure data quality in full detail, including how many peaks are at FDR 5% and above 5-fold enrichment.

Software

Describe the software used to collect and analyze the ChIP-seq data. For custom code that has been deposited into a community repository, provide accession details.

Flow Cytometry

Plots

Confirm that:

- The axis labels state the marker and fluorochrome used (e.g. CD4-FITC).
- The axis scales are clearly visible. Include numbers along axes only for bottom left plot of group (a 'group' is an analysis of identical markers).
- All plots are contour plots with outliers or pseudocolor plots.
- A numerical value for number of cells or percentage (with statistics) is provided.

Methodology

- Sample preparation
- Instrument
- Software
- Cell population abundance
- Gating strategy
- Tick this box to confirm that a figure exemplifying the gating strategy is provided in the Supplementary Information.

Magnetic resonance imaging

Experimental design

- Design type
- Design specifications
- Behavioral performance measures

Acquisition

- Imaging type(s)
- Field strength
- Sequence & imaging parameters
- Area of acquisition
- Diffusion MRI Used Not used

Preprocessing

- Preprocessing software
- Normalization
- Normalization template
- Noise and artifact removal
- Volume censoring

Statistical modeling & inference

Model type and settings

Specify type (mass univariate, multivariate, RSA, predictive, etc.) and describe essential details of the model at the first and second levels (e.g. fixed, random or mixed effects; drift or auto-correlation).

Effect(s) tested

Define precise effect in terms of the task or stimulus conditions instead of psychological concepts and indicate whether ANOVA or factorial designs were used.

Specify type of analysis: Whole brain ROI-based Both

Statistic type for inference
(See [Eklund et al. 2016](#))

Specify voxel-wise or cluster-wise and report all relevant parameters for cluster-wise methods.

Correction

Describe the type of correction and how it is obtained for multiple comparisons (e.g. FWE, FDR, permutation or Monte Carlo).

Models & analysis

n/a | Involved in the study

- Functional and/or effective connectivity
 Graph analysis
 Multivariate modeling or predictive analysis

Functional and/or effective connectivity

Report the measures of dependence used and the model details (e.g. Pearson correlation, partial correlation, mutual information).

Graph analysis

Report the dependent variable and connectivity measure, specifying weighted graph or binarized graph, subject- or group-level, and the global and/or node summaries used (e.g. clustering coefficient, efficiency, etc.).

Multivariate modeling and predictive analysis

Specify independent variables, features extraction and dimension reduction, model, training and evaluation metrics.