

Roger Williams University

DOCS@RWU

Arts & Sciences Faculty Publications

Arts and Sciences

2019

**Phylogenetic, genomic, and biogeographic characterization of
a novel and ubiquitous marine invertebrate-associated
Rickettsiales parasite, *Candidatus Aquarickettsia rohweri*, gen.
nov., sp. nov**

J. Grace Klinges
Oregon State University

Stephanie M. Rosales
Oregon State University

Ryan McMinds
Oregon State University

Koty H. Sharp
Roger Williams University, ksharp@rwu.edu

Follow this and additional works at: https://docs.rwu.edu/fcas_fp



Part of the [Biology Commons](#), and the [Marine Biology Commons](#)

Recommended Citation

Klinges, JG, Rosales, SM, McMinds, R, Shaver, EC, Shantz, AA, Peters, EC, Burkepile, DE, Eitel, M, Wörheide G, Silliman, BR, Sharp, KH, and RL Vega Thurber. (2019). Phylogenetic, genomic, and biogeographic characterization of a novel and ubiquitous marine invertebrate-associated Rickettsiales parasite, *Candidatus Aquarickettsia rohweri*, gen. nov., sp. nov. *The ISME Journal*.

This Article is brought to you for free and open access by the Arts and Sciences at DOCS@RWU. It has been accepted for inclusion in Arts & Sciences Faculty Publications by an authorized administrator of DOCS@RWU. For more information, please contact mwu@rwu.edu.



Phylogenetic, genomic, and biogeographic characterization of a novel and ubiquitous marine invertebrate-associated Rickettsiales parasite, *Candidatus Aquarickettsia rohweri*, gen. nov., sp. nov

J. Grace Klings¹ · Stephanie M. Rosales^{1,2,3} · Ryan McMinds^{1,4} · Elizabeth C. Shaver⁵ · Andrew A. Shantz^{6,7} · Esther C. Peters⁸ · Michael Eitel⁹ · Gert Wörheide^{9,10,11} · Koty H. Sharp¹² · Deron E. Burkepile⁶ · Brian R. Silliman⁵ · Rebecca L. Vega Thurber¹

Received: 22 October 2018 / Revised: 10 July 2019 / Accepted: 13 July 2019
© The Author(s) 2019. This article is published with open access

Abstract

Bacterial symbionts are integral to the health and homeostasis of invertebrate hosts. Notably, members of the Rickettsiales genus *Wolbachia* influence several aspects of the fitness and evolution of their terrestrial hosts, but few analogous partnerships have been found in marine systems. We report here the genome, phylogenetics, and biogeography of a ubiquitous and novel Rickettsiales species that primarily associates with marine organisms. We previously showed that this bacterium was found in scleractinian corals, responds to nutrient exposure, and is associated with reduced host growth and increased mortality. This bacterium, like other Rickettsiales, has a reduced genome indicative of a parasitic lifestyle. Phylogenetic analysis places this Rickettsiales within a new genus we define as “*Candidatus Aquarickettsia*.” Using data from the Earth Microbiome Project and SRA databases, we also demonstrate that members of “*Ca. Aquarickettsia*” are found globally in dozens of invertebrate lineages. The coral-associated “*Candidatus A. rohweri*” is the first finished genome in this new clade. “*Ca. A. rohweri*” lacks genes to synthesize most sugars and amino acids but possesses several genes linked to pathogenicity including Tlc, an antiporter that exchanges host ATP for ADP, and a complete Type IV secretion system. Despite its inability to metabolize nitrogen, “*Ca. A. rohweri*” possesses the NtrY-NtrX two-component system involved in sensing and responding to extracellular nitrogen. Given these data, along with visualization of the parasite in host tissues, we hypothesize that “*Ca. A. rohweri*” reduces coral health by consuming host nutrients and energy, thus weakening and eventually killing host cells. Last, we hypothesize that nutrient enrichment, which is increasingly common on coral reefs, encourages unrestricted growth of “*Ca. A. rohweri*” in its host by providing abundant N-rich metabolites to be scavenged.

Introduction

Symbioses are common in marine systems, and the stability of these relationships among hosts and symbionts are often essential to each member’s health [1]. Yet symbioses can easily be disrupted by pathogens or abiotic stressors [2–4]. For example, coral reefs are habitats highly dependent on symbioses, and stressors like increased temperatures shift

animal microbiomes and drive disease events [5–7], which are major causes of mortality in tropical reefs worldwide [8]. The complex nature of microbiomes and the contextual nature of disease, however, makes it difficult to incontrovertibly identify the causative agents of many epizootics. It is likely that, in addition to identifiably virulent pathogens, other bacterial species make their hosts more vulnerable to harm from subtle environmental changes. Recent research efforts have focused on identifying conditions that lead to dysbiosis between coral host and symbiont, yet few studies have been able to test the contextual roles of the host, microbe, and environment in disease susceptibility [9]. Understanding how these microorganisms are involved in host responses to environmental conditions may provide insight into the elusive causes of marine invertebrate disease.

Our recent study of the endangered Caribbean staghorn coral, *Acropora cervicornis*, identified a Rickettsiales-

Supplementary information The online version of this article (<https://doi.org/10.1038/s41396-019-0482-0>) contains supplementary material, which is available to authorized users.

✉ Rebecca L. Vega Thurber
rvegathurber@gmail.com
Rebecca.Vega-Thurber@oregonstate.edu

Extended author information available on the last page of the article.

annotated OTU that increased in relative abundance from 11.4 to 87.9% of the total coral bacterial community after exposure to elevated inorganic nutrients [3]. Furthermore, we found a strong negative correlation between the relative abundance of this Rickettsiales OTU and coral growth [3]. We also showed that members of this taxonomic group are associated with increased tissue loss and mortality in three other species of corals after chronic exposure to elevated nitrate and phosphate [2].

The order Rickettsiales is one of several microbial groups being investigated as a disease agent of corals, but the high prevalence of these bacteria in both visually healthy and diseased corals renders their etiological role unclear [10, 11]. Rickettsiales spp. have been identified in corals suffering from white-band disease type I (WBD I), an epizootic which resulted in the near extirpation of acroporid corals from the Caribbean [7, 8, 10], with up to 95% of the population destroyed since the disease was first documented in the 1980s [8, 10]. Histological samples of both healthy and WBD-afflicted acroporid corals are often dominated by Rickettsiales (or *Rickettsia*)-like organisms (RLOs) [10, 12], but Di Lauro [13] found that the number of RLO aggregates in WBD-infected tissue was positively correlated with worsening coral tissue conditions.

Members of the order Rickettsiales are typically obligate intracellular parasites, often associated with invertebrate vectors and sometimes pathogenic to humans. They have highly reduced genomes, averaging 1.1–1.5 Mb [14–17], with varying amounts of pseudogenization and repetitive regions [16, 18, 19]. Three families are currently defined within Rickettsiales: *Rickettsiaceae*, *Anaplasmataceae*, and “*Candidatus* Midichloriaceae” [14, 20–22]. *Rickettsiaceae* consists of bacteria that induce spotted fever rickettsiosis in humans [19, 23], while members of “*Ca.* Midichloriaceae” have been shown to infect protists and the mitochondria of metazoans [24–26]. The family *Anaplasmataceae* consists of four genera: *Anaplasma*, *Ehrlichia*, *Neorickettsia*, and *Wolbachia* [20].

We report here the genome assembly, phylogenetics, and biogeography of the Rickettsiales OTU identified in our previous studies. For this organism, we propose the novel genus and species “*Candidatus* Aquarickettsia rohweri,” which is named for the association of members of this genus with aquatic organisms and Dr. Forest Rohwer, whose group first described this taxon in corals. This study places “*Ca.* *A. rohweri*” in the family “*Ca.* Midichloriaceae”, and assesses genome features related to its host dependency and potential virulence mechanisms. The presence of genes for secretion and metabolite acquisition as well as genes involved in sensing extracellular nutrient concentrations suggests that “*Ca.* *Aquarickettsia*” exploits its host for necessary nutrients, energy, and amino acids. Last, we report that members of “*Ca.* *Aquarickettsia*” are found not only in corals but also in other aquatic organisms, including

sea anemones, sponges, hydrozoans, placozoans, ctenophores, and protists worldwide.

Materials and methods

Coral RLO enrichment experiment and coral histology

Following the methods of Shaver et al. [3], we performed nutrient enrichment on fragments of *A. cervicornis* in the field. In brief, we obtained 40 fragments of *A. cervicornis* from the Coral Restoration Foundation’s coral nursery in Key Largo, Florida. Fragments were transported by boat to Pickles Reef, Florida, and secured to mounting structures using All Fix Marine Epoxy®. Corals were allowed to acclimate for 1 week before 20 fragments were exposed to 70 g of Osmocote 19-6-12 Smart-Release® Plant fertilizer in mesh bags that were attached ~5 cm below corals. The remaining 20 fragments were maintained under ambient conditions at the site to serve as unenriched controls. Enrichment was maintained for 4 weeks with fertilizer bags replaced after 2 weeks. Shaver et al. [3] demonstrated effective nutrient enrichment using this treatment for up to 8 weeks.

Due to mortality or breakage, 34 coral fragments of the initial 40 were sampled for histological analysis. Fragments were placed in 50-mL centrifuge tubes filled with a formaldehyde-based fixative composed of 1:4 parts Z-Fix Concentrate (Z-Fix Concentrate, Anatech, Ltd.) and Instant Ocean® seawater. Centrifuge tubes were capped, tightly sealed using Parafilm, and shipped to the Histology Laboratory at George Mason University for analysis using methods from Miller et al. [12], and Price and Peters [27]. In brief, after decalcification with Formical-2000, trimmed subsamples were processed into paraffin blocks. Sections (at 5- μ m thickness) were mounted on glass slides, stained with Harris’s hematoxylin and 1% alcoholic eosin and a Giemsa procedure that distinguishes rickettsia from other Gram-negative bacteria [27], and were then visually assessed using an Olympus BX43 compound microscope [12]. Tissue parameters, including cellular integrity, *Symbiodiniaceae* abundance, and tissue components (e.g. epidermal and mesenterial mucocytes) were assessed using a previously established numerical system [12]. Image analysis with Olympus cellSens software was conducted to count numbers of dividing algal symbionts and to obtain gastrodermis thickness in pixel lengths.

Fluorescence in situ hybridization

The protocol for fluorescence in situ hybridization (FISH) was adapted from a previously described method [28]. Three paraffin-embedded *A. cervicornis* specimens from the

high-nutrient treatments were sectioned onto glass slides at a thickness of 10–15 µm. Sections were deparaffinized 2× for 10 min in xylene, 2× for 10 min in 95% ethanol, rinsed once in distilled water, and air dried. FISH with end-labeled oligonucleotide probes was performed in a humidity chamber in hybridization buffer (0.9 M NaCl, 20 mM Tris-HCl [pH 7.4], 0.01% sodium dodecyl sulfate) with 40% formamide for 2–4 h. After hybridization, the slides were incubated at 48 °C in wash buffer (0.7 M NaCl, 20 mM Tris-HCl [pH 7.4], 50 mM EDTA, 0.01% sodium dodecyl sulfate) for 20–40 min, depending on the hybridization incubation period. The wash buffer was rinsed away with distilled water, and slides were air dried and mounted in VectaShield (Vector Labs, Burlingame, CA). Slides were visualized on a Nikon Eclipse E800 epifluorescence microscope. All probes, including the general bacterial probe suite (EUB338-I, EUB338-II, and EUB338-III) that together target most known bacteria, the bacterial probe negative control (EUB338-NEG), and the sequence-specific probes designed in this study (RICK, RICKNEG), were used at a final concentration of 10 ng/µl in hybridization buffer.

Design and application of species-specific probes

A suite of probes (EUB338-I, EUB338-II, EUB338-III) targeting most bacteria 16S rRNA sequences (EUB338-I: 5'-CTGCCTCCCGTAGGAGT-3'; EUB338-II: 5'-CAGCCACCCGTAGGTGT-3'; EUB338-III: 5'-CTGCCACCCGTAGGTGT-3') was used for FISH [29, 30]. Probes were designed targeting the “*Candidatus Aquarickettsia rohweri*” 16S rRNA gene sequence obtained from the draft genome for use, as confirmation of their presence in *A. cervicornis* specimens. In order to ensure specificity, the probe was designed to target a hypervariable region of the 16S rRNA, and for efficiency, it targets a region of extremely high accessibility on the 16S rRNA molecule [31]. The specific oligonucleotide probe, RICK (5'-CCTCCAATTCTCCATTGG-3'), was 5'-labeled with either Alexa546 or Alexa488. Inquiry on ProbeMatch (RDP; <http://rdp.cme.msu.edu/probematch/search.jsp>) using 100% sequence match suggests that the probe only matches 16S sequences previously assigned to *Rickettsia* species or unclassified alphaproteobacteria in the database. A single-mismatch probe, RICKNEG (5'-CCTCAA-TACTCCATTGG-3'), was designed as a negative control to confirm probe specificity and matches 0 bacterial sequences in the RDP database. Optimal formamide concentration (40%) for the specific probe was empirically determined by using the highest formamide concentration that simultaneously exhibited presence of mixed EUB338 binding and RICK binding, and absence of RICKNEG binding.

Metagenome preparation, sequencing, and sequence quality control pipeline

An *Acropora cervicornis* fragment from ref. [3] with a relative abundance of 87.9% for one Rickettsiales OTU was selected for genome assembly. DNA extraction was performed using EZNA Tissue DNA Kit (Omega Bio-Tek), and DNA yield was quantified using the Qubit dsDNA HS Assay and analyzed on a Qubit 3.0 Fluorometer. A barcoded DNA fragment library was prepared using the Nextera XT (Illumina) sample preparation kit (Supplementary Table S1A). This sample was shotgun sequenced on an Illumina Miseq with a v3 reagent kit at OSU's Center for Genome Research and Biotechnology (CGRB). A total of 13,785,074 metagenomic reads were sequenced with an average sequence length of 301 bp and a mean phred score of 36. Further details about the sample used for genome assembly are accessible as National Center for Biotechnology Information (NCBI) BioSample SAMN10490412.

Genome assembly and quality assessment

The filtering and trimming of low-quality sequences and removal of adaptor sequences were conducted with FQtrim (Supplementary Table S2). The program PEAR [32] was used with default settings to merge high-quality paired-end sequences. The contaminating reads were filtered out by mapping to the human genome [33], and genomes of *Acropora digitifera* [34] (as there was no genome available for *A. cervicornis*), and the algal symbiont *Symbiodiniaceae* spp. [35], using Bowtie2 [36] (Supplementary Table S3). The remaining 13,688,870 reads (99.46%) were assembled with the de novo meta-assembler IDBA-UD, a de novo assembler [37], and SPAdes [38] were subsequently used to improve assembly. Unsupervised binning of contigs was performed with MaxBin v.2.2.1 [39] using default parameters, and VizBin was implemented to visually bin contigs further. Sequence bins were analyzed by BLAST searches [40] to the NCBI nonredundant database (NR) and to all complete bacterial genomes in RefSeq [41]. Bins with sequences identified as *Trichoplax adhaerens* were retained due to previous observations that genomic scaffolds of this organism contained sequences from Rickettsial endosymbionts that may have been horizontally transferred to the placozoan host [42]. We used SSPACE [43] to merge contigs and create scaffolds, and MaxBin was re-run (with minimum contig length 500) to confirm that all sequences belonged to a single bin. We conducted manual quality control on contigs in this final bin using BLAST, and removed contigs with identity to *Symbiodiniaceae* spp. that had erroneously been retained. CheckM [44] was used to assess assembly quality and contamination (Supplementary Table S4, Fig. S1) using marker genes that are ubiquitous

and single copy within a given lineage. We elected to use marker genes from all of Alphaproteobacteria rather than exclusively Rickettsiales due to the large amount of divergence between the three families of Rickettsiales and an overrepresentation of the comparatively well-studied *Rickettsiaceae* in the marker gene database. The program Pseudo-Finder [45] was used to identify potential pseudo-genes, which were left in the genome assembly but removed from calculations of coding density (see below).

Genome annotation and comparative genomics

Prokka [46] was used to annotate the genome, and annotations were uploaded to the KAAS [47] server. KEGG orthology (KO) identifiers were assigned using BLAST+ against the KEGG GENES [48] database with the BBH (bi-directional best hit) method to assign orthologs and enable the reconstruction of KEGG pathways and BRITE hierarchies (Supplementary Table S5). KO numbers linked to the KEGG pathway maps and BRITE functional hierarchies were used to group genes into modules and assess the metabolic pathway completion. OrthoFinder 2.2.3 [49] was used to identify orthologous clusters of genes that were compared between species (listed in Supplementary Table S6) to assess shared and unique gene clusters. CGView [50] was used to generate a circular representation of the genome by providing the program with a Mauve [51] alignment of the draft genome to a reference genome (“*Ca. Midichloria mitochondrii*” strain IricVA) and gene coordinates generated by Prodigal [52] (Supplementary Table S7). EggNOG-mapper [53] was used (DIAMOND mapping, default parameters) to assign eggNOG orthologous groups (OGs), GO terms, and COG functional categories inferred from best matching OGs.

Module completion analysis with MAPLE

The KEGG pathway completion was examined with module completion ratios of each KEGG functional module using MAPLE-2.3.0 [54]. The completion of each module is directly related to the ability of the organism to perform the function related to the module, and MAPLE corrects for existence of KOs that are shared between independent modules to provide accurate statistics for completion. We provided MAPLE with unannotated coding sequences from Prodigal, allowing us to assess different gene abundances as a measure of module robustness.

Placozoan-associated bacterial 16S rRNA gene sequencing

A set of primers was selected to amplify the Rickettsial endosymbionts from a diversity of placozoans. The full-

length Rickettsial 16S rRNA gene sequence was identified in the draft metagenomes of three placozoan species to select conserved regions to amplify an almost complete bacterial 16S fragment. Draft genomes were from *Trichoplax adhaerens* (“Grell” strain) [55], *Trichoplax* sp. H2 [56], and *Hoilungia hongkongensis* [57]. To amplify and Sanger-sequence the cleaned PCR product, two nested sets of primers were used (Supplementary Table S1b). PCR was conducted using GoTaq (#M3175; Promega) with 25 mM MgCl₂, 10 mM each dNTP, 5 μM each primer, 1 μ Taq polymerase. The PCR program included 3-min pre-heating (95 °C) followed by 35 cycles of 30 s heating (95 °C), 30 s annealing (63 °C), and 1 min 30 s extension (72 °C). After the cycles, a final extension step of 3 min (72 °C) completed the PCR. The Rickettsial 16S was amplified from DNA isolations of either newly isolated single individuals (see ref. [58]) or from established clonal cultures (see ref. [57]). For details on isolates see Supplementary Table S8.

Phylogenetic analysis of coral-associated Rickettsiales

The full-length 16S rRNA region of the genome was identified using BLASTn searches against 16S rRNA regions from other Rickettsiales species from Greengenes (Version 13.8) [59] as a partial 16S rRNA sequence from this OTU already existed in the Greengenes database [3]. The trimmed sequence was aligned to 16S rRNA sequences from 79 members of Rickettsiales and 11 Alphaproteobacteria outgroups (Supplementary Table S9) using cmlign (part of the INFERNAL 1.1.1 package) [60] with a covariance model built using the Rfam seed alignment for bacterial small subunit ribosomal RNA. PhyML [61] was used for phylogenetic analysis, with 1000 bootstrap replicates and parameters selected using jModelTest [62] (Supplementary Table S10). These replicates were summarized using the Geneious 11.1.4 Consensus Tree Builder command [63]. Sequences were manually BLASTed to the NCBI RefSeq database [41], and metazoan hosts were identified from 100% identical sequences in the database with reported hosts from NCBI sequence description information. Pairwise BLAST was performed between the 16S rRNA sequence of “*Ca. A. rohweri*” and closely related sequences from “*Ca. Midichloriaceae*.”

A concatenated marker protein-based phylogeny was generated using a previously curated set of 92 orthologous, single-copy marker genes present in all bacteria [64] (Supplementary Table S11). Hmmssearch 3.1b2 [65] was used to identify these marker genes in 20 genomes within Rickettsiales and four outgroups. Sequences were aligned using MAFFT 7.310 [66] and were concatenated, excluding alignment positions that contained a gap in more than 50% of sequences. Phylogenetic reconstruction was performed

using the concatenated alignments in amino acid format, as well as Dayhoff6 [67] recoded format. All analyses were performed under CAT-GTR using PhyloBayes MPI 1.5a [68]. CAT-GTR was shown to be the best-fit evolutionary model for analyzing multigene phylogenomic alignments [69–71], and Dayhoff recoding was applied to rectify the negative impact of compositional heterogeneity [70]. Two independent Markov chains were run and their convergence assessed using the PhyloBayes tools *bpcomp* and *tracecomp* to monitor maximum discrepancy in clade support (maxdiff), the effective sample size (effsize), and relative difference in posterior mean estimates (rel_diff) as summary statistics of the model. The appropriate number of samples to discard as “burnin” was independently determined first by visual inspection of parameter trace plots, and then by optimizing convergence criteria. For amino acid analyses, the maxdiff statistic was 0, while for Dayhoff6 recoded data the statistic was 0.046. The minimum effective sample size was >50 (>100 for Dayhoff6) and the maximum rel_diff statistic was <0.3 (0.1 for Dayhoff6). Node support was evaluated using posterior probabilities.

Redbiom search for presence of “*Ca. A. rohweri*”

The Earth Microbiome Project (EMP) tool Redbiom [72] was used to search for the presence of “*Ca. A. rohweri*” and other OTUs within the genus “*Ca. Aquarickettsia*” in the Qiita database. The Rickettsiales OTU from the sample selected for “*Ca. A. rohweri*” genome assembly was determined by Shaver et al. [3] to match most closely to Greengenes [59] OTU 150441, so this and the five most closely related Greengenes OTUs (as determined using the Greengenes reference phylogeny and confirmed by BLAST to NR) were considered members of the genus. The command “redbiom search features” was used to search 173,714 16S rRNA libraries from the EMP sequencing context “Pick_closed-reference_OTUs-Greengenes-illumina-16S-v4-100nt”. This context provides the most resolution of the 16S rRNA sequence without losing samples that were not sequenced to 150 bp. Overall, 3168 EMP samples from 83 “sample types” had OTUs assigned by 97% closed-reference picking to “*Ca. A. rohweri*” (Greengenes ID 150441). Many of the categories were, however, either ambiguous (e.g., “fragment” and “sediment”), redundant (e.g., “seawater” and “ocean water”), or miscategorized (e.g., sample type listed as sponge tissue, other metadata fields listed sample as water). For clarity, we manually reannotated these into 23 categories: by compiling the available metadata files, we were able to differentiate samples into more explanatory biomes (e.g., “coral (fragment)” and “reef sediment”). A map of the distribution of “*Ca. A. rohweri*” in corals was generated by plotting the sample collection coordinates and host genus, available as part of

the EMP metadata (EMP Study ID 10895: Global Coral Microbiome Project, and EMP Study ID 10798: Palmyra Atoll Corallimorph and Bleaching Surveys). Analyses through Redbiom were confirmed by analysis using the tool IMNGS [73] to query 303,361 16S rRNA libraries in the Sequence Read Archive (SRA) database. The full-length 16S rRNA sequence was queried against the database at 97% and 99% sequence identity thresholds.

Results and discussion

Rickettsia-like organisms are visually abundant in nutrient-enriched corals

Suspected RLOs were present in the Giemsa-stained coral samples, both control and nutrient-enriched (see the Supplementary Results Section 1, Fig. S1, Fig. S2). Fluorescent probes targeting most bacteria (mixed EUB338-I+EUB338-II+EUB338-III) and the Rickettsiales-like organism (RICK) co-localized on clusters of cells within the epithelia of *A. cervicornis* (Fig. 1b, c). The clusters of RLOs appeared to be intracellular and near *Symbiodiniaceae* cells, but they were not in or on the *Symbiodiniaceae* cells in the gastrodermis. The RLO-targeted probe (RICK) designed in this study is sequence specific: at 40% formamide, RICK hybridizes to cells in the clusters but no signal from RICKNEG is apparent (Fig. 1d). All images shown are from a single specimen, but the images are representative of signal localization in all examined specimens.

Genome assembly of “*Ca. A. rohweri*” had low contamination and high completion

To evaluate the function of the nutrient-responsive Rickettsiales OTU in *A. cervicornis*, we conducted metagenomic sequencing, de novo genome assembly, and comparative genomics. The genome was assessed for divergence from related genomes by aligning the draft genome to a reference from “*Ca. Midichloriaceae*” [50] (Fig. 2; Fig. S3). The assembled 1.28 Mb genome was estimated at 97.3% completion, with 1469 predicted genes, a coding density of 87.6% and a low GC content of 28.38% (Fig. 2; Supplementary Table S2). Contamination was low as estimated by CheckM at 0.39% (Fig. S4), and completion was estimated at 97.3%, supported by the presence of tRNAs for all 20 proteinogenic amino acids. These genome features are characteristic of most Rickettsiales [18, 20]. At 155 contigs, an N50 of 10,860 bp, and based on completeness and contamination estimates from CheckM, this genome is classified as a “High-Quality Draft” by MIMAG [74] standards for metagenome-assembled genomes. There were

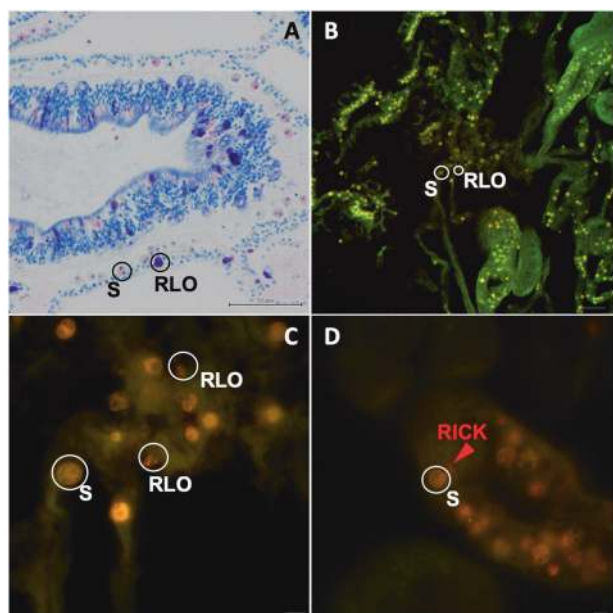


Fig. 1 FISH on *Acropora cervicornis* sections with general and specific probes. **a** Suspected RLOs, evident in Giemsa-stained section of actinopharynx body wall, including mucocyte among *Symbiodiniaceae*-infected cells in the gastrodermis. **b, c** Ten micrometers thick sections, mixed EUB338-I+II+III (Alexa546, red) and RICK (Alexa488, green) probes co-localized to clusters of bacterial cells of similar size, morphology, and location as in the Giemsa-stained sections. **d** RICK probe localizes to cells in an aggregate (red arrow), and there is no signal from single-base mismatch probe RICKNEG (green). RLO Rickettsiales-like organism clusters, S *Symbiodiniaceae*. **a** Bar = 50 μ m; **b** bar = 50 μ m; **c** bar = 10 μ m; **d** bar = 10 μ m

84 potential pseudogenes identified, most of which were annotated as hypothetical genes. Of the identified pseudogenes, 56 were truncations (CDS of <65% of the average length of BLAST hits to this gene, usually caused by premature stop codon), and 28 were predicted fragmentations (genes split into multiple pseudogenes by presence of an early stop and a second start codon). The coding density without pseudogenes is 78.68%. This genome is accessible on NCBI as accession NZ_RXFM000000000.1.

Phylogenetics supports the establishment of a new genus within “*Ca. Midichloriaceae*”

Using the recovered full-length 16S rRNA gene sequence from our coral-associated Rickettsiales, our phylogeny (Fig. 3a; Fig. S5) placed our draft genome, “*Ca. A. rohweri*,” within “*Ca. Midichloriaceae*.” Our 16S rRNA phylogeny is consistent with a recent phylogenomic tree [75], which placed “*Ca. Midichloriaceae*” as a sister clade to *Anaplasmataceae* rather than *Rickettsiaceae*. Our phylogeny disagrees with older phylogenomic trees [20, 42] on this placement but this node is supported by relatively low bootstrap values (65.3). Phylogenomic analysis (Fig. 3b;

Fig. S6) further supports the placement of “*Ca. Aquarickettsia*” in “*Ca. Midichloriaceae*” (uncollapsed branches), but places “*Ca. Midichloriaceae*” as the sister group to both *Anaplasmataceae* and *Rickettsiaceae*. Strikingly, “*Ca. A. rohweri*” was part of a distinct clade along with other uncultured and uncharacterized bacteria associated with marine invertebrates. This distinct clade, present in other phylogenies [76, 77] but never formally defined, contains symbionts of the soft coral *Gorgonia ventalina*, the stony corals *Orbicella annularis*, *Orbicella faveolata*, and *Acropora cervicornis*, and the sponge *Cymbastela concentrica*. While most members of this group were identified in marine hosts, several species were identified in freshwater organisms, including the ciliate *Euplotes woodruffi* and freshwater cnidarian *Hydra oligactis*. “*Ca. A. rohweri*” shared species-level homology with multiple 16S rRNA sequences from endosymbionts of *Trichoplax adhaerens*, as well as with sequences from both sponge and coral hosts (Fig. 3, colored by percentage BLAST identity). The bootstrap value for this distinct clade of marine endosymbionts was 91.

The high confidence bootstrap values joining this clade together support the assignment of a new genus to these intracellular symbionts of marine invertebrates. We propose the genus name “*Candidatus Aquarickettsia*,” with the draft genome from this study named as “*Candidatus A. rohweri*” in reference to its association with aquatic organisms and Dr. Forest Rohwer, who first described this taxon in corals [10].

Comparative genomics confirms genome content similar to other members of Rickettsiales

Based on analysis of orthologous gene sets using OrthoFinder 2.2.3 (Fig. 4), “*Candidatus A. rohweri*” shares 188 gene clusters with other species of Rickettsiales, 327 gene clusters with only members of “*Ca. Midichloriaceae*” (“*Ca. M. mitochondrii*”, “*Ca. Fokinia solitaria*”, and “*Ca. Jidai-bacter acanthamoeba*”), and encodes two unique gene clusters (genes that share a similar function, but are present only in the genome of “*Ca. A. rohweri*”) as well as 324 unique genes (not assigned to any orthogroup). While 262 of the unique genes were hypotheticals, 14 of the 62 annotated unique genes (Supplementary Table S9) are involved in oxidative phosphorylation. Other unique genes include transporters, FeS cluster assembly proteins (SufB, SufC, SufD, SufE, not commonly found in Rickettsiales), and flagellar proteins FlgF and FlgL. Of the two gene clusters unique to “*Ca. A. rohweri*,” one was comprised 87 genes with homology to transposase IS66. Transposable elements are common in Rickettsiales, and IS66 family transposases are annotated in genomes of *Wolbachia* on NCBI (NC_010981 and NC_010981). The other unique

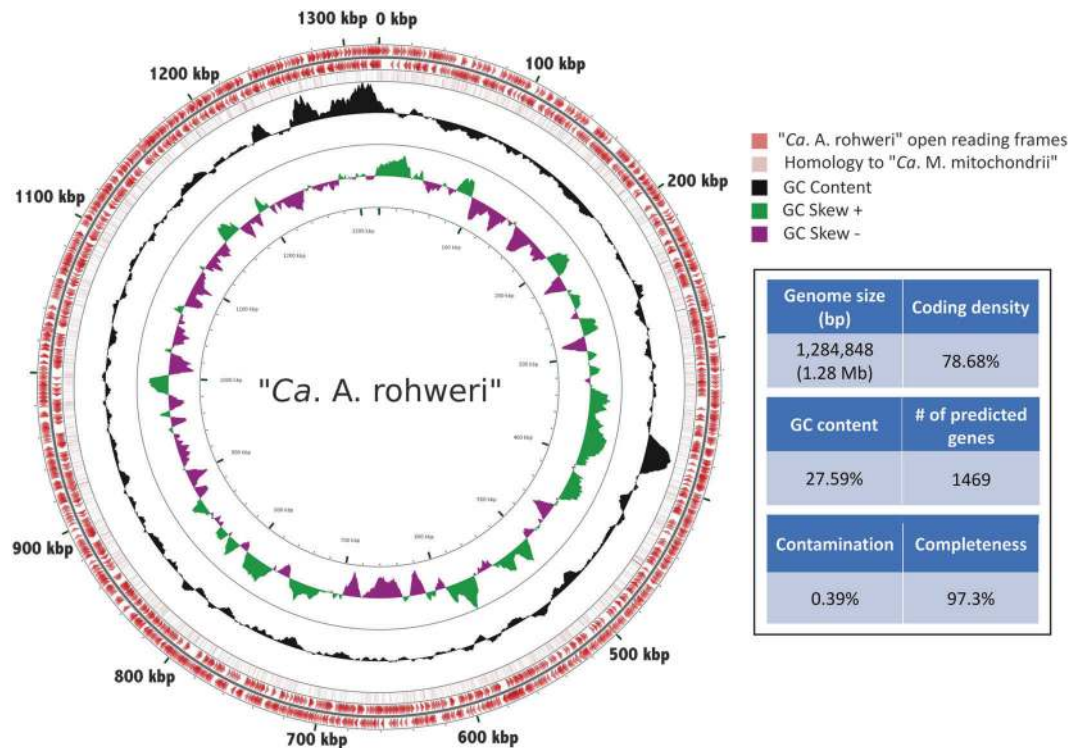


Fig. 2 Circular representation of “*Ca. A. rohweri*” genome (red) with red arrows indicating open-reading frames (ORF), light red bars indicating homology to “*Candidatus Midichloria mitochondrii* IricVA,” black histogram indicating GC content with purple and green

histograms indicating GC+/- skew. GC skew is calculated as $(G-C)/(G+C)$, with positive skew indicating overabundance of G over C and negative skew indicating overabundance of C over G

cluster consists of hypothetical genes and ankyrin repeats. Ankyrin repeats, along with tetratricopeptide repeats (also present in the genome of “*Ca. A. rohweri*”), are eukaryotic-like repeat domains which were observed in the genome of “*Ca. J. acanthamoeba*” and may be involved in symbiont–host interactions [75].

“*Ca. A. rohweri*” is auxotrophic for many essential compounds and relies on the host for metabolic byproducts

Protein prediction and metabolic network analysis revealed 627 KEGG orthology (KO) numbers matching to coding sequences. The KEGG pathways with the greatest numbers of KOs found were: ribosome (47 KOs), oxidative phosphorylation (38), carbon metabolism (34), two-component system (29), biosynthesis of amino acids (24), aminoacyl-tRNA biosynthesis (24), purine metabolism (22), and pyrimidine metabolism (22). Clusters of Orthologous Groups of proteins (COG) [78] functional category analysis mapped 962 genes and indicated that, in comparison with *Pelagibacter ubique* HTCC1062 and *Rhodobacter sphaeroides*, more genes mapped to the categories “Translation” (13% vs. 9% and 4%, respectively), “Replication and Repair” (9% vs. 4% and 4%), “Intracellular Trafficking and Secretion”

(4% vs. 2% and 1%), and “Energy Production and Conversion” (10% vs. 8% and 7%), while fewer genes mapped to the categories pertaining to amino acid, nucleotide, and inorganic ion metabolism and transport (Fig. S7).

Module completion ratio analysis revealed incomplete pathways related to carbon and nitrogen metabolism (Fig. S8) and amino acid biosynthesis (Fig. S9), but the presence of a complete *rvh* Type IV secretion system (Fig. S10). Rickettsiales rely heavily on their hosts for nutrients and metabolic byproducts, and “*Ca. A. rohweri*” exhibits many of the same incomplete metabolic pathways as *Rickettsia* spp. and “*Ca. M. mitochondrii*” (Fig. S11).

“*Ca. A. rohweri*” is deficient in most pathways involved with amino acid biosynthesis

The exploration of functional gene categories of “*Ca. A. rohweri*” provided us with an outline of this organism’s potential metabolic function. Due to the absence of most amino acid biosynthesis genes (Fig. 5), this bacterium requires supplementation of every amino acid from the host or algal symbiont except for serine, glycine, glutamate, aspartate, lysine, and threonine. Acroporid corals lack the ability to synthesize the essential amino acids or synthesize cysteine from homocysteine or serine and are thus

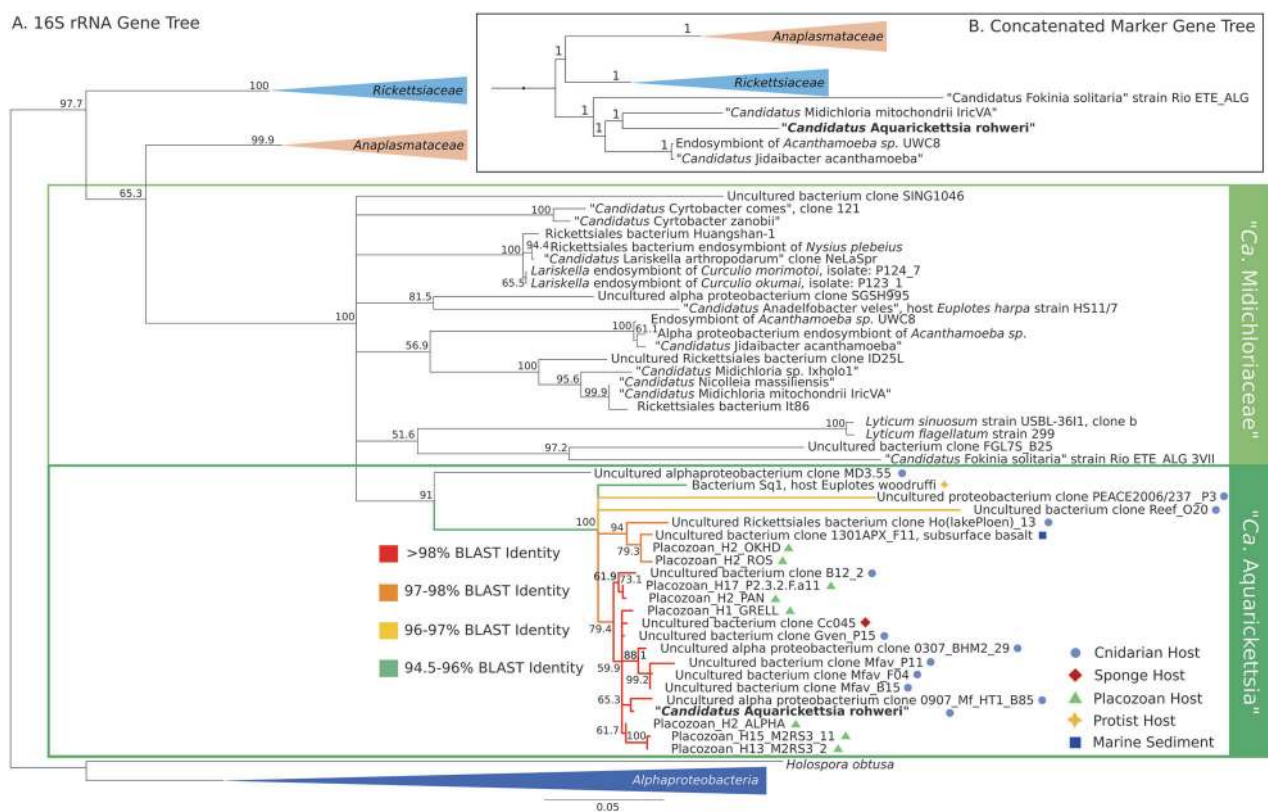


Fig. 3 Consensus maximum likelihood and concatenated marker gene tree of Rickettsiales. **a** Phylogeny generated from bootstrapped data sets using 16S rRNA sequences and PhyML with 1000 bootstrapped replicates. Branches within the proposed genus *Aquarickettsia* are colorized by BLAST identity to the 16S rRNA sequences for “*Ca. a rohweri*.” Bootstrap confidence values listed at nodes, representing the certainty of that node in the phylogenetic tree, where 100 is maximum

certainty. **b** Bayesian phylogeny generated from concatenated amino acid alignment of 92 orthologous, single-copy marker genes for bacteria. Node support was evaluated using posterior probabilities. Branch lengths in both phylogenies correspond to evolutionary distance. Outgroups and Rickettsiales families other than “*Ca. Midichloriaceae*” are collapsed; uncollapsed phylogenetic trees are presented as Supplementary Figs. S5 and S6

dependent on *Symbiodiniaceae* for these compounds [34]. *Symbiodiniaceae* are, however, able to synthesize all essential amino acids with the exception of histidine and lysine [79]. The close proximity of “*Ca. A. rohweri*” to *Symbiodiniaceae* spp. in some FISH imagery in the gastrodermis suggests that amino acids may be acquired directly from the algal symbiont.

Experimental studies in other Rickettsiales (summarized in ref. [80]) have indicated the ability to import amino acids that cannot be synthesized. In addition, while genes (*glyA* and *ltaE*) are present to interconvert Gly to Ser and Thr (Fig. S9), the organism requires at least one of these three amino acids, and studies of *Rickettsia* show the import of Ser and Gly to be required for growth [80]. Lys, Asp, Gln, and Glu can be synthesized from oxaloacetate, but can also be imported. Proton glutamate symport protein (GltP) imports Asp and Glu, ABC transporter GlnHPQ is able to transport Gln in other Rickettsiales but is missing a gene in our existing annotations, and an uncharacterized ABC transporter is predicted to transport

Lys. Another ABC transporter, ArtMPQ, can transport Arg and the proline/betaine symporter can import Pro (ProP, Fig. 5). ProP exists in the genome of “*Ca. A. rohweri*” in multiple sequence variants that may transport other amino acids. Other species of Rickettsiales have been shown to import Met, Ser, and Gly, but these transporters are uncharacterized.

Mechanisms of energy synthesis, transport, and storage in “*Ca. A. rohweri*”

The cell membrane of Rickettsiales spp. has been shown to be permeable to host NAD^+ [80]. Although “*Ca. A. rohweri*” potentially can produce ATP through oxidative phosphorylation coupled to electron transport and driven by Glu oxidation, it also possesses an ATP/ADP symporter (nucleotide translocase; Tlc1). Tlc1 allows for the import of ATP in exchange for ADP, effectively sapping the host of energy (Fig. 5). Rickettsiales spp. utilize this system to compensate for their reduced metabolism, and are thus

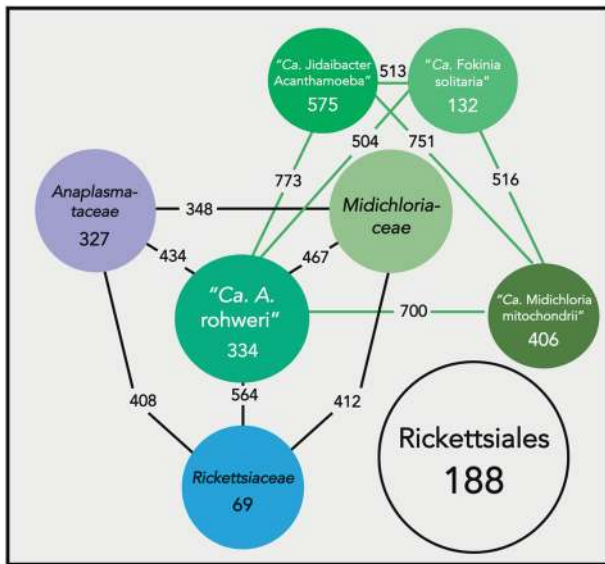


Fig. 4 Orthologous gene sets shared by species of Rickettsiales. Network graph showing orthologous gene sets shared by species of Rickettsiales at the order level (Rickettsiales), family level (*Rickettsiaceae*, *Anaplasmataceae*, *Midichloriaceae*), and genus level (“*Ca. A. rohweri*”, “*Ca. M. mitochondrii*”, “*Ca. J. acanthamoeba*”, and “*Ca. F. solitaria*”). Numbers in black indicate shared gene sets between two compared organisms or families, or gene sets shared by an entire family sampled. Numbers in white indicate individual genes that were unique to a given organism. Orthologous gene families were identified using OrthoFinder

considered energy parasites [16, 75]. Nucleotide transport proteins are essential to the lifestyle of many obligate parasites including *Chlamydia* and *Rickettsia* where they function as the main energy supply, but are also found in free-living bacteria such as Cyanobacteria [81, 82]. *Wolbachia* spp., though host-associated, lack *tlc*. Species of *Rickettsia* encode up to five copies of *Tlc* (*Tlc1* to *Tlc5*); only *Tlc1* functions as an ATP transporter, while *Tlc4* and *Tlc5* import host ribonucleotides [80]. “*Ca. A. rohweri*” encodes only one copy, with 51% amino acid identity to the *Tlc1* gene of “*Ca. J. acanthamoeba*”.

While “*Ca. A. rohweri*” possess a functional pyruvate dehydrogenase complex (PDC) and tricarboxylic acid (TCA) cycle, there is no evidence of functional glycolysis/gluconeogenesis (Fig. 5). *Rickettsia* can import pyruvate from the host through an uncharacterized transporter, and import UDP-glucose as the main sugar source for lipopolysaccharide and peptidoglycan biosynthesis (Fig. 5). Unlike *Rickettsia* spp., our organism lacks the pathway to synthesize PHB, a storage molecule that retains energy to be used when host energy sources are depleted or unavailable. “*Ca. A. rohweri*” possesses three copies of *EamA*, which transports the cosubstrate *S*-adenosylmethionine (SAM) [80]. The presence of other rickettsial signatures (ProP, GltP, SpoT, and MdlB) [80] confirms the necessity of these

genes for the obligate intracellular lifestyle of these bacteria even in diverged hosts.

“*Candidatus A. rohweri*” can detect, but is incapable of metabolizing nitrogen species

“*Candidatus A. rohweri*” lacks any complete pathways for nitrogen metabolism (Fig. S8), and has only two genes present within this module (*gudB/rocG* glutamate dehydrogenase, *glnA* glutamine synthetase), which are utilized for other purposes. Despite an inability to metabolize nitrogen, “*Ca. A. rohweri*” does possess the NtrY-NtrX two-component system for regulating nitrate metabolism (Fig. 5). NtrY is a transmembrane protein that allows the cell to sense extracellular nitrogen levels, while NtrX is normally involved in regulating *nif*-genes involved with nitrogen fixation [83, 84], which the genome of “*Ca. A. rohweri*” does not encode. The function of this system is poorly understood, but these genes are upregulated by uptake of Pro and Gln from the host cell in other Rickettsiales systems and are likely involved with cell proliferation within the host [83]. Both Rickettsiales and *Pelagibacter* encode the NtrY-NtrX two-component system as well as the EnvZ-OmpR system (Fig. 5) for the control of osmotic stress [17, 85]. The presence of two-component regulatory systems enables these organisms to sense environmental changes and rapidly respond to stimuli by activating or repressing certain genes [84, 85]. These systems have been implicated not only in environmental response but also in the ability of pathogenic bacteria to infect and survive in their hosts, making them advantageous for both intracellular and free-living bacteria [17]. Unlike *Rickettsia* spp. and “*Ca. M. mitochondrii*,” “*Ca. A. rohweri*” possesses the PhoR-PhoB two-component system for detecting inorganic phosphate limitation, which may also play a role in responding to nutrient enrichment and depletion.

Rickettsiales vir homolog T4SS may be involved in host attachment and infection

Due to their role in human disease, the mechanisms of pathogenicity in *Rickettsiaceae* are well described, but homologs in “*Candidatus Midichloriaceae*” are not. All *Rickettsiaceae* genomes sequenced thus far encode a reduced Type IV secretion system (T4SS) known as Rickettsiales vir homolog (*rvh*), which compared with canonical T4SSs lacks a homolog of *virB5*, the gene encoding the minor pilus subunit [15, 86–89]. T4SSs are annotated in the genomes of “*Ca. M. mitochondrii*,” which has one *rvh* T4SS [90], and “*Ca. J. acanthamoeba*,” which has three T4SS clusters, only one of which is related to *rvh* [75].

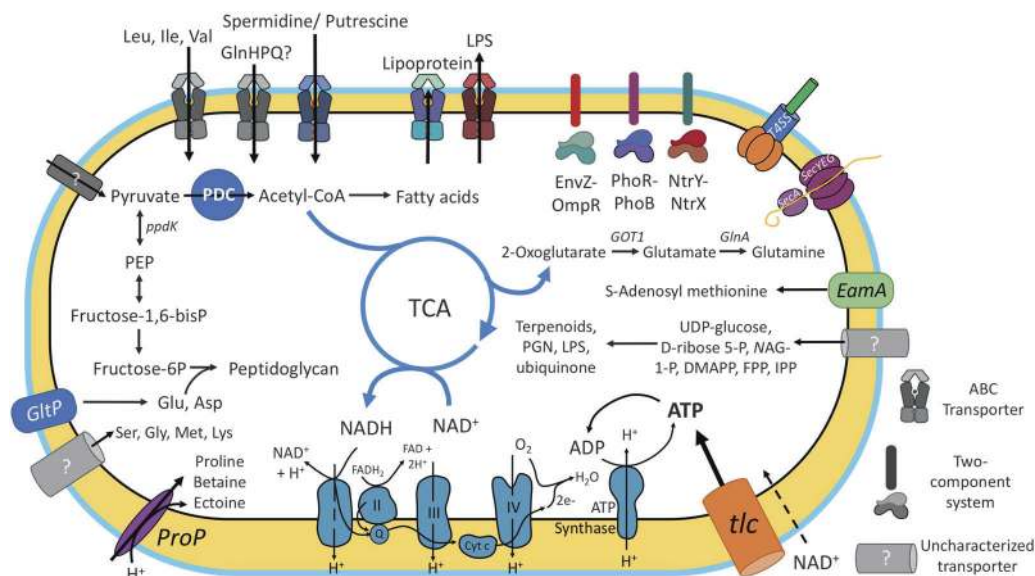


Fig. 5 Metabolic reconstruction of “*Ca. A. rohweri*” illustrating known and uncharacterized transporters and known metabolic pathways. “*Candidatus A. rohweri*” must import many host precursors to synthesize compounds such as terpenoids, peptidoglycan, and fatty acids. While some transporters and symporters have been characterized, others are predicted (gray coloration) based on the inability of the organism to synthesize precursors and evidence of uptake by other species of Rickettsiales. DMAPP dimethylallyl pyrophosphate, EamA rickettsial S-adenosylmethionine (SAM) transporter, EnvZ-OmpR osmoregulatory two-component system, FPP *trans,trans*-farnesyl diphosphate, GlnA glutamine synthetase, GlnHPQ glutamine permease ABC transporter, GltP Glu symporter, GOT1 glutamic-oxaloacetic transaminase 1, IPP isopentenyl diphosphate, LPS

lipopolysaccharide, NAG-1-P *N*-acetylglucosamine-1-P, NtrY-NtrX nitrogen sensing two-component system, Oxidative phosphorylation pathway: complex I NADH:ubiquinone oxidoreductase, complex II succinate:ubiquinone oxidoreductase, complex III ubiquinol:cytochrome *c* oxidoreductase, *cyt c* cytochrome *c*, and complex IV, cytochrome *c* oxidase, PEP pyruvate dehydrogenase complex, PEP phosphoenolpyruvate, PGN peptidoglycan, PhoR-PhoB phosphate and ferric iron sensing two-component system, ppdk pyruvate phosphate dikinase, ProP proline/glycine betaine transporter, SecYEG/SecA bacterial secretion system (Sec), T4SS Type IV secretion system, TCA tricarboxylic acid cycle, *tlc* ADP/ATP translocase gene, UDP-glucose uridine diphosphate glucose

T4SSs are cell envelope-spanning complexes through which bacterial cells secrete or take up macromolecules [87, 91, 92]. While “*Ca. A. rohweri*” has annotations for all necessary components of the *rvh* T4SS (Fig. S10), homology to *rvh* genes from other Rickettsiales varied. Trb1 VirB10 was the most diverged gene, with a maximum of 52.7% amino acid identity (to “*Ca. M. mitochondrii*”), while VirB11 was the most conserved at 76.8% amino acid identity (to “*Ca. J. acanthamoeba*”) (Supplementary Table S12). Nonetheless, “*Ca. A. rohweri*” appears to retain infective capabilities as *Aquarickettsia* spp. were visualized in coral mucocytes (Fig. 1).

In other species of bacteria such as *Agrobacterium tumefaciens* [91, 92], T4SSs are involved in host cell attachment, DNA transfer, and secretion of virulence factors directly into host cells. A similar role has been proposed for *rvh*, and the function of this system has been demonstrated in *Ehrlichia* and *Anaplasma* to secrete effector proteins [93, 94]. Current data suggest that *rvh* T4SS played a pivotal role in the transition from an extracellular lifestyle to an obligate intracellular lifestyle [17, 87], as it is absent from free-living members of Alphaproteobacteria including *Pelagibacter*.

Unconventional mechanisms of pathogenicity in “*Ca. A. rohweri*”

The coral-associated “*Ca. A. rohweri*” possesses many genes involved in host signaling that may play a role in recognition and phagocytosis by the host. “*Ca. A. rohweri*” possesses a nearly complete set of genes for flagellar assembly despite its endosymbiotic nature. While “*Ca. J. Acanthamoeba*” and “*Ca. Fokinia solitaria*,” endosymbiotic members of “*Ca. Midichloriaceae*,” possess a full set of genes for flagellar assembly [75, 95], TEM imagery of “*Ca. Fokinia solitaria*” demonstrated the lack of a flagellum despite the presence of necessary genes [95]. Flagellin, encoded by the gene *fljC*, is a known pathogen-associated molecular pattern, recognized by Toll-like receptor 5, a component of innate immune systems of both plants and animals [9]. Upon detection by the host, flagellin causes the release of nuclear factor NF- κ B, a key element of infection response that controls DNA transcription, cytokine production, and cell survival [96]. “*Ca. A. rohweri*” possesses genes to produce and export peptidoglycan and lipopolysaccharides (Fig. 5), also known pathogen-associated molecular patterns [97]. These microbial signals may play

an essential role as they trigger host phagocytosis of “*Ca. A. rohweri*,” after which the parasite is able to escape further attack once inside the mucocyte. Transcriptome analysis has indicated that corals affected by WBD increase expression genes involved in phagocytosis of foreign bacteria [5, 13]. The infection model proposed for “*Ca. J. Acanthamoeba*” and supported by TEM suggests that the parasite is taken up by phagocytosis, releases effector proteins into the host cytoplasm via the T4SS to manipulate host gene expression and signal transduction, and recruits to the host endoplasmic reticulum to escape phagolysosomal degradation [75].

The presence of the *tlc* system in “*Ca. Aquarickettsia*” indicates that these organisms siphon energy from their hosts which, when combined with their dependency on the host for scavenged amino acids and sugars, could rapidly deplete the host of necessary cell-building resources. “*Ca. A. rohweri*” infects coral mucocytes (as shown in Fig. 1), and we suspect that these parasites eventually starve, weaken, and kill them. The coral host must replace its mucocytes in order to maintain defenses against sedimentation and other invading microbes, and the energy expended by the host to replace these cells can lead to chronic stress on the coral. Nutrient enrichment has been shown to increase growth of *Symbiodiniaceae* spp., which may provide “*Ca. A. rohweri*” with more available resources as amino acids, sugars, and lipids are transferred from the algal symbiont to the host. This additional nutrition provided to the host by abundant *Symbiodiniaceae* spp. allows the production of more mucocytes in which “*Ca. A. rohweri*” can proliferate.

Biogeography analysis confirms the association of “*Ca. Aquarickettsia*” with non-bilaterian metazoans worldwide

To survey the global prevalence of this genus and identify new hosts, we used Redbiom to query the publicly available Earth Microbiome Project (EMP) database; in all, our Redbiom analysis represents an evaluation of a total of 173,714 16S V4 Greengenes Illumina 100 nt samples (Fig. 6). “*Ca. A. rohweri*” was prevalent in marine sediments (rocks and sand), seawater, and freshwater, and marine organisms such as corals, sponges, the ctenophore *Mnemiopsis*, and kelp (Fig. 6). The presence of “*Ca. A. rohweri*” in marine sand and sediments suggests either that these habitats provide a potential vector for transmission, represent “*Ca. A. rohweri*” hosted by interstitial organisms, or contain marine invertebrate eDNA [98]. “*Ca. J. acanthamoeba*” is horizontally transferred in aquatic environments [75]; both freshwater and seawater may thus play a role in transmission of “*Ca. Aquarickettsia*.” “*Ca. A. rohweri*” was absent in terrestrial samples, such as dust, groundwater, and “mammal-associated” data sets, including

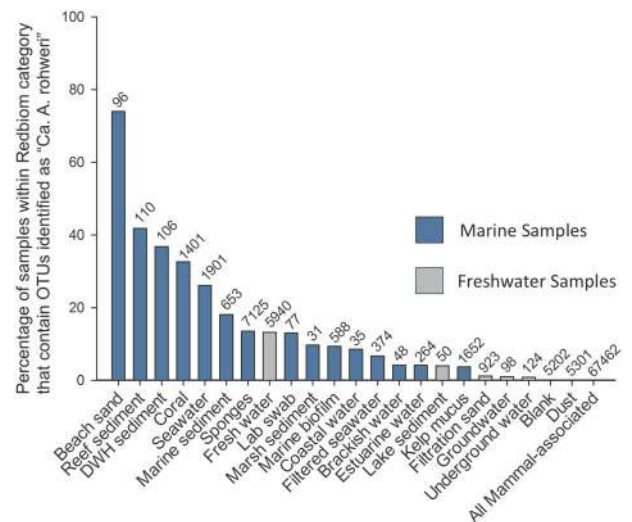


Fig. 6 Prevalence of “*Ca. Aquarickettsia rohweri*” 16S sequence within Earth Microbiome Project samples. Bars represent the percentage of samples within each category that had sequences assigned to OTU 150441 (EMP OTU assignment performed with 97% similarity, closed-reference OTU picking with Greengenes). Numbers above bars indicate the total number of samples in that category catalogued in the EMP Qiita database. Blue bars represent samples from marine or brackish environments, and gray bars represent samples from other environments

those from human medical samples (Fig. 6 gray bars). This analysis was repeated for OTUs within “*Ca. Aquarickettsia*” and compared to other OTUs from “*Ca. Midichloriaceae*” that were not part of the proposed genus (Supplementary Table S13).

Out of the 1401 total coral samples in the EMP database, 457 (32.62%) had OTUs identified as “*Ca. A. rohweri*” (i.e., >97% identity), whereas 494 (38.2%) had OTUs within “*Ca. Aquarickettsia*.” As metadata for coral samples in the EMP database included sampling coordinates and host identity, we had the capability to determine how “*Ca. A. rohweri*” was distributed globally and in which coral species it was identified (Fig. 7). We found that this OTU is present in the Pacific, Atlantic, and Indian Oceans in 51 scleractinian genera, 11 cnidarian outgroups, and a ctenophore, *Mnemiopsis*.

When we mapped the distribution of EMP sponge samples (primarily from the Sponge Microbiome Project [99]) in which “*Ca. A. rohweri*” was identified (Fig. S12), we found an even more global distribution, with samples spanning 76 sponge genera. “*Ca. A. rohweri*” was identified in sponge samples across the world, and is widespread throughout Europe (North Sea, Northern Atlantic, Mediterranean Sea), the Red Sea, the Caribbean, and Australia.

To supplement results from analysis of the EMP database, we queried the SRA database with the full-length 16S rRNA sequence of “*Ca. A. rohweri*” using IMNGS at 97% and 99% similarity thresholds (Supplementary Table S14).

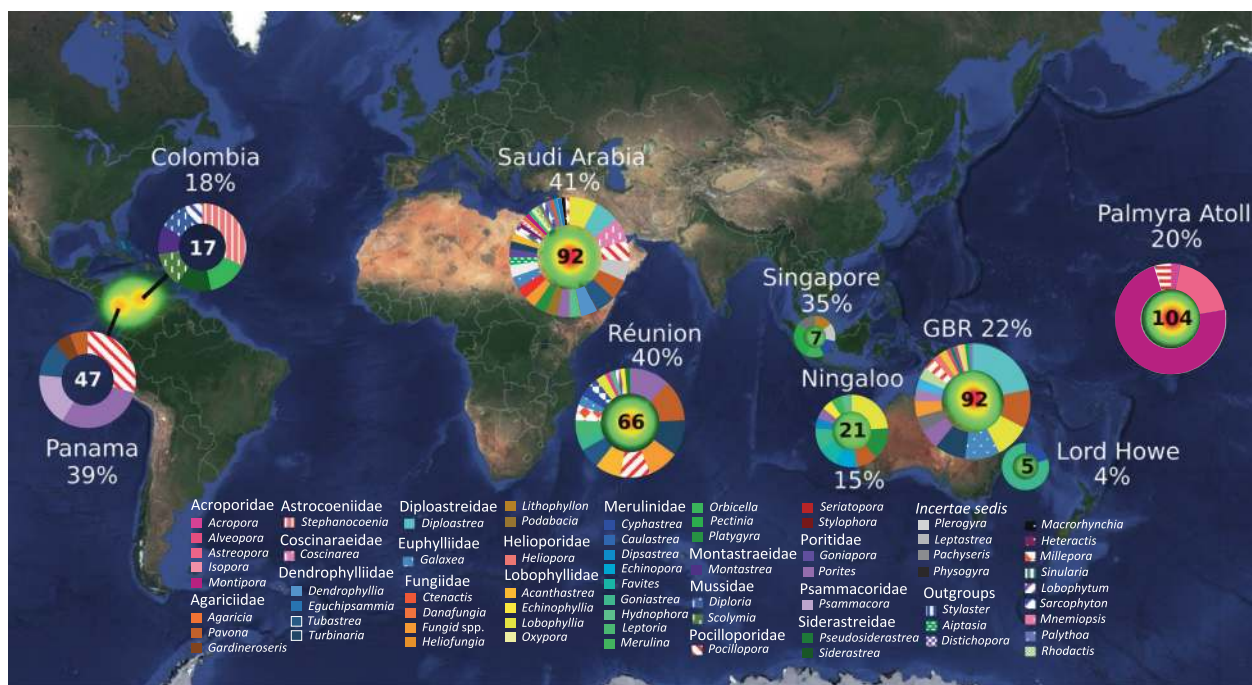


Fig. 7 Map of the distribution of “*Ca. A. rohweri*” in coral and cnidarian outgroups. Coral samples included in map were from the Global Coral Microbiome Project (EMP Study ID 10895) and Palmyra Atoll Corallimorph and Bleaching Surveys (EMP Study ID 10798), a total of 451 samples. Ring charts show the number of samples (numeral inside ring) and the different coral host genera (exterior colored bars)

While metadata for the SRA database is less detailed than the EMP database, these results indicated that sequences identified as “*Ca. A. rohweri*” were found exclusively in marine metagenomes, seven of which were coral metagenomes. Although there were 12,665 marine metagenomes (including seawater, marine sediment, and plankton metagenomic samples) and 1855 coral metagenomes queried, “*Ca. A. rohweri*” was only found in 21 samples. This is likely due to our use of the full-length 16S rRNA sequence for this analysis, which is more informative than an 100-bp fragment.

Together these data suggest that the proposed genus “*Ca. Aquarickettsia*” broadly associates with corals and with many members of the non-bilaterian metazoan phyla (Placozoa, Porifera, Cnidaria, and Ctenophora), as well as the even more ancient protists. This broad host range is not uncommon in Rickettsiales: many species of this family (notably, *Wolbachia* spp.) are able to switch hosts, infecting both vertebrate and invertebrate hosts while still maintaining virulence [75, 100].

“*Candidatus Aquarickettsia*” may drive disease events under nutrient-replete conditions

We characterized the genome and metabolic capabilities of “*Ca. A. rohweri*,” a novel bacterium in a new genus of

in which “*Ca. A. rohweri*” was present. Percentages indicate the number of coral samples from each geographic location in which “*Ca. A. rohweri*” was identified. “*Ca. A. rohweri*” has also been identified in samples from the Caribbean which were sequenced using a different protocol and therefore not included in this analysis

Rickettsiales that is found globally distributed in corals and sponges and also associates with protists and placozoans. As its proposed name suggests, the genus “*Ca. Aquarickettsia*” is primarily associated with marine-dwelling organisms. Based on the metabolic capabilities of “*Ca. A. rohweri*,” we postulate that this group of parasitic organisms plays a role in coral disease development. Previously, we reported evidence that inorganic nutrient exposure increased “*Ca. A. rohweri*” populations, reduced coral growth [3], and increased host tissue loss and mortality [2]. Based on this evidence and the presence of nutrient-sensing genes, we hypothesize that exposure to nitrate and ammonium leads to a transition in the role of “*Ca. A. rohweri*” from commensal to parasitic in coral hosts (Fig. 8). We hypothesize that “*Ca. A. rohweri*” detects nutrient enrichment through NtrY-NtrX, triggering other functions such as cell growth and division or the upregulation of transporters to acquire more metabolites from the host. As nutrient enrichment also allows the coral host and algal symbiont to produce more N-rich amino acids and proteins, surplus resources could easily be scavenged by the parasite. Last, we hypothesize that in nutrient-replete conditions, the presence of surplus nitrogen metabolites encourages the proliferation of this parasite to levels that may overwhelm host metabolic capabilities (Fig. 8).

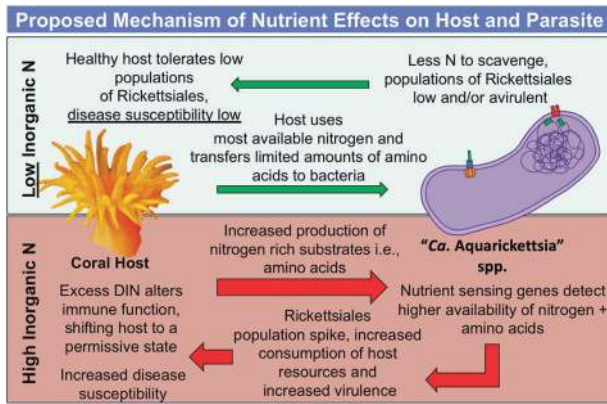


Fig. 8 Model of predicted interactions between species of "*Candidatus Aquarickettsia*" with invertebrate host and how interactions are altered by nutrient input

Rickettsiales have been implicated in coral disease for more than 25 years. However, their presence in seemingly healthy corals has led to doubt that they are the primary pathogen responsible for white-band disease. We show here that "*Ca. A. rohweri*" is globally associated with many coral hosts, and possesses the genomic capacity to parasitize the coral holobiont for amino acids and ATP. We also previously showed that in response to nutrient enrichment, these parasites proliferate and reduce coral growth. Thus through shifts in its population dynamics "*Ca. A. rohweri*" may leave the host more susceptible to other opportunistic pathogens leading to disease. Alternatively, "*Ca. A. rohweri*" may serve as a primary pathogen, although this still remains speculative. As nutrient pollution increasingly affects reefs, we suspect that parasites within the proposed genus will proliferate and may contribute to heightened vulnerability to disease and mortality.

Acknowledgements We would like to thank Drs. Steve Giovannoni and Luis Bolanos for their insight into this paper. Further we thank all of the people involved in the Global Coral Microbiome Project, primarily Drs. Monica Medina, Joe Pollock, and Jesse Zaneveld. This work was funded by a National Science Foundation DOB grant #1442306 to RVT. JGK acknowledges funding through the National Science Foundation Graduate Research Fellowship Program (No. 2018260750). GW acknowledges funding through the LMU Munich's Institutional Strategy LMUexcellent within the framework of the German Excellence Initiative and the European Union's Horizon 2020 Marie Skłodowska-Curie Innovative Training Network IGNITE (No. 764840). GW also acknowledges the Leibniz Supercomputing Center of the Bavarian Academy of Sciences and Humanities (www.lrz.de) for providing access to supercomputing infrastructure.

Funding This work was funded by: a Dimensions of Biodiversity NSF grant (#1442306) to RVT, NSF Ocean Sciences grants to DEB and RVT (#1130786) and BRS (#1056980), and NSF Graduate Fellowships to both JGK (#1840998) and ES (#1106401).

Author contributions The permits FKNMS-2013-073 and FKNMS-2014-081 were obtained from the Florida Keys National Marine

Sanctuary. RVT, DEB, BRS, ECS, and JGK designed the research; ECS and AAS conducted the field research; JGK, SMR, and RM performed the genomic and phylogenetic analysis; ECP performed the histology; ME and GW generated placozoan sequence data; KS performed the FISH; JGK and RVT wrote the paper and all authors contributed to revisions.

Compliance with ethical standards

Conflict of interest The authors declare that they have no conflict of interest.

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

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/>.

References

1. Apprill A. Marine animal microbiomes: toward understanding host-microbiome interactions in a changing ocean. *Front Mar Sci.* 2017;4:222.
2. Zaneveld JR, Burkepile DE, Shantz AA, Pritchard CE, McMinds R, Payet JP, et al. Overfishing and nutrient pollution interact with temperature to disrupt coral reefs down to microbial scales. *Nat Commun.* 2016;7:11833.
3. Shaver EC, Shantz AA, McMinds R, Burkepile DE, Vega Thurber RL, Silliman BR. Effects of predation and nutrient enrichment on the success and microbiome of a foundational coral. *Ecology.* 2017;98:830–9.
4. Thurber RV, Willner-Hall D, Rodriguez-Mueller B, Desnues C, Edwards RA, Angly F, et al. Metagenomic analysis of stressed coral holobionts. *Environ Microbiol.* 2009;11:2148–63.
5. Libro S, Kaluziak ST, Vollmer SV. RNA-seq profiles of immune related genes in the staghorn coral *Acropora cervicornis* infected with white band disease. *PLoS ONE.* 2013;8:e81821.
6. Gignoux-Wolfsohn SA, Vollmer SV. Identification of candidate coral pathogens on white band disease-infected staghorn coral. *PLoS ONE.* 2015;10:e0134416.
7. Gignoux-Wolfsohn SA, Marks CJ, Vollmer SV. White band disease transmission in the threatened coral, *Acropora cervicornis*. *Sci Rep.* 2012;2:srep00804.
8. Aronson RB, Precht WF. White-band disease and the changing face of Caribbean coral reefs. The ecology and etiology of newly emerging marine diseases. Dordrecht: Springer; 2001. p. 25–38.
9. Mydlarz LD, Jones LE, Harvell CD. Innate immunity, environmental drivers, and disease ecology of marine and freshwater invertebrates. *Annu Rev Ecol, Evol, Syst.* 2006;37:251–88.
10. Casas V, Kline DI, Wegley L, Yu Y, Breitbart M, Rohwer F. Widespread association of a Rickettsiales-like bacterium with reef-building corals. *Environ Microbiol.* 2004;6:1137–48.

11. Godoy-Vitorino F, Ruiz-Diaz CP, Rivera-Seda A, Ramírez-Lugo JS, Toledo-Hernández C. The microbial biosphere of the coral *Acropora cervicornis* in Northeastern Puerto Rico. *Peer J*. 2017;5:e3717.
12. Miller MW, Lohr KE, Cameron CM, Williams DE, Peters EC. Disease dynamics and potential mitigation among restored and wild staghorn coral, *Acropora cervicornis*. *Peer J*. 2014;2:e541.
13. Di Lauro S. Time-series evaluation of suspect rickettsiales-like bacteria presence in *Acropora cervicornis* off of Broward County from Years 2001–2012. Master's thesis. Nova Southeastern University. Retrieved from NSUWorks, 2015. https://nsuworks.nova.edu/occ_stuetd/379.
14. Kang Y-J, Diao X-N, Zhao G-Y, Chen M-H, Xiong Y, Shi M, et al. Extensive diversity of Rickettsiales bacteria in two species of ticks from China and the evolution of the Rickettsiales. *BMC Evol Biol*. 2014;14:167.
15. Fournier P-E, El Karkouri K, Leroy Q, Robert C, Giumelli B, Renesto P, et al. Analysis of the *Rickettsia africae* genome reveals that virulence acquisition in *Rickettsia* species may be explained by genome reduction. *BMC Genom*. 2009;10:166.
16. Gillespie JJ, Joardar V, Williams KP, Driscoll T, Hostetler JB, Nordberg E, et al. A *Rickettsia* genome overrun by mobile genetic elements provides insight into the acquisition of genes characteristic of an obligate intracellular lifestyle. *J Bacteriol*. 2012;194:376–94.
17. Audia JP. Rickettsial physiology and metabolism in the face of reductive evolution. In: Palmer GH, Azad AF, editors. *Intracellular Pathogens II: Rickettsiales*. Intracellular pathogens II: Rickettsiales. ASM Press, Washington, DC: American Society of Microbiology; 2012. p. 221–42.
18. Viklund J, Ettema TJG, Andersson SGE. Independent genome reduction and phylogenetic reclassification of the Oceanic SAR11 Clade. *Mol Biol Evol*. 2012;29:599–615.
19. Darby AC, Cho N-H, Fuxelius H-H, Westberg J, Andersson SGE. Intracellular pathogens go extreme: genome evolution in the Rickettsiales. *Trends Genet*. 2007;23:511–20.
20. Ferla MP, Thrash JC, Giovannoni SJ, Patrick WM. New rRNA gene-based phylogenies of the alphaproteobacteria provide perspective on major groups, mitochondrial ancestry and phylogenetic instability. *PLoS One*. 2013;8:e83383.
21. Merhej V, El Karkouri K, Raoult D. Whole genome-based phylogenetic analysis of Rickettsiae. *Clin Microbiol Infect*. 2009;15:336–7.
22. Szokoli F, Castelli M, Sabaneyeva E, Schrollhammer M, Krensek S, Doak TG, et al. Disentangling the taxonomy of Rickettsiales and description of two novel symbionts (“*Candidatus Bealeia paramacronuclearis*” and “*Candidatus Fokinia cryptica*”) sharing the cytoplasm of the ciliate protist *Paramecium biaurelia*. *Appl Environ Microbiol*. 2016;82:7236–47.
23. Guo W-P, Tian J-H, Lin X-D, Ni X-B, Chen X-P, Liao Y, et al. Extensive genetic diversity of Rickettsiales bacteria in multiple mosquito species. *Sci Rep*. 2016;6:38770.
24. Bazzocchi C, Mariconti M, Sasser D, Rinaldi L, Martin E, Cringoli G, et al. Molecular and serological evidence for the circulation of the tick symbiont *Midichloria* (Rickettsiales: Midichloriaceae) in different mammalian species. *Parasit Vectors*. 2013;6:350.
25. Epis S, Sasser D, Beninati T, Lo N, Beati L, Piesman J, et al. *Midichloria mitochondrii* is widespread in hard ticks (Ixodidae) and resides in the mitochondria of phylogenetically diverse species. *Parasitology*. 2008;135:485–94.
26. Montagna M, Sasser D, Epis S, Bazzocchi C, Vannini C, Lo N, et al. “*Candidatus Midichloriaceae*” fam. nov. (Rickettsiales), an ecologically widespread clade of intracellular Alphaproteobacteria. *Appl Environ Microbiol*. 2013;79:3241–8.
27. Price KL, Peters EC. *Histological techniques for corals*. 2018. <https://www.amazon.com/Histological-Techniques-Corals-Kathy-Price-ebook/dp/B07L1DSCYZ>.
28. Sharp KH, Eam B, Faulkner DJ, Haygood MG. vertical transmission of diverse microbes in the tropical sponge *Corticium* sp. *Appl Environ Microbiol*. 2007;73:622–9.
29. Amann RI, Binder BJ, Olson RJ, Chisholm SW, Devereux R, Stahl DA. Combination of 16S rRNA-targeted oligonucleotide probes with flow cytometry for analyzing mixed microbial populations. *Appl Environ Microbiol*. 1990;56:1919–25.
30. Greuter D, Loy A, Horn M, Rattei T. probeBase—an online resource for rRNA-targeted oligonucleotide probes and primers: new features 2016. *Nucleic Acids Res*. 2016;44:D586–D589.
31. Behrens S, Lösekann T, Pett-Ridge J, Weber PK, Ng W-O, Stevenson BS, et al. Linking microbial phylogeny to metabolic activity at the single-cell level by using enhanced element labeling-catalyzed reporter deposition fluorescence in situ hybridization (EL-FISH) and NanoSIMS. *Appl Environ Microbiol*. 2008;74:3143–50.
32. Zhang J, Kobert K, Flouri T, Stamatakis A. PEAR: a fast and accurate Illumina Paired-End reAd mergeR. *Bioinformatics*. 2014;30:614–20.
33. Schneider V, Church D. *Genome Reference Consortium. The NCBI Handbook* [Internet]. 2nd edition. Bethesda (MD): National Center for Biotechnology Information (US), 2013.
34. Shinzato C, Shoguchi E, Kawashima T, Hamada M, Hisata K, Tanaka M, et al. Using the *Acropora digitifera* genome to understand coral responses to environmental change. *Nature*. 2011;476:320–3.
35. Aranda M, Li Y, Liew YJ, Baumgarten S, Simakov O, Wilson MC, et al. Genomes of coral dinoflagellate symbionts highlight evolutionary adaptations conducive to a symbiotic lifestyle. *Sci Rep*. 2016;6:39734.
36. Langmead B, Salzberg SL. Fast gapped-read alignment with Bowtie 2. *Nat Meth*. 2012;9:357–9.
37. Peng Y, Leung HCM, Yiu SM, Chin FYL. IDBA-UD: a de novo assembler for single-cell and metagenomic sequencing data with highly uneven depth. *Bioinformatics*. 2012;28:1420–8.
38. Bankevich A, Nurk S, Antipov D, Gurevich AA, Dvorkin M, Kulikov AS, et al. SPAdes: a new genome assembly algorithm and its applications to single-cell sequencing. *J Comput Biol*. 2012;19:455–77.
39. Wu Y-W, Tang Y-H, Tringe SG, Simmons BA, Singer SW. MaxBin: an automated binning method to recover individual genomes from metagenomes using an expectation-maximization algorithm. *Microbiome*. 2014;2:26.
40. Camacho C, Coulouris G, Avagyan V, Ma N, Papadopoulos J, Bealer K, et al. BLAST+: architecture and applications. *BMC Bioinforma*. 2009;10:421.
41. O’Leary NA, Wright MW, Brister JR, Ciuffo S, Haddad D, McVeigh R, et al. Reference sequence (RefSeq) database at NCBI: current status, taxonomic expansion, and functional annotation. *Nucleic Acids Res*. 2016;44:D733–745.
42. Driscoll T, Gillespie JJ, Nordberg EK, Azad AF, Sobral BW. Bacterial DNA sifted from the *Trichoplax adhaerens* (Animalia: Placozoa) genome project reveals a putative rickettsial endosymbiont. *Genome Biol Evol*. 2013;5:621–45.
43. Boetzer M, Henkel CV, Jansen HJ, Butler D, Pirovano W. Scaffolding pre-assembled contigs using SSPACE. *Bioinformatics*. 2011;27:578–9.
44. Parks DH, Imelfort M, Skennerton CT, Hugenholtz P, Tyson GW. CheckM: assessing the quality of microbial genomes recovered from isolates, single cells, and metagenomes. *Genome Res*. 2015;25:1043–55.
45. Syberg-Olsen M, Husnik F. Pseudofinder, GitHub repository: <https://github.com/filip-husnik/pseudo-finder/>. 2018.

46. Seemann T. Prokka: rapid prokaryotic genome annotation. *Bioinformatics*. 2014;30:2068–9.
47. Moriya Y, Itoh M, Okuda S, Yoshizawa AC, Kanehisa M. KAAS: an automatic genome annotation and pathway reconstruction server. *Nucleic Acids Res*. 2007;35:W182–W185.
48. Kanehisa M, Goto S. KEGG: Kyoto encyclopedia of genes and genomes. *Nucleic Acids Res*. 2000;28:27–30.
49. Wang Y, Coleman-Derr D, Chen G, Gu YQ. OrthoVenn: a web server for genome wide comparison and annotation of orthologous clusters across multiple species. *Nucleic Acids Res*. 2015;43:W78–W84.
50. CGView Server: a comparative genomics tool for circular genomes | *Nucleic Acids Research* | Oxford Academic. https://academic.oup.com/nar/article/36/suppl_2/W181/2505750. Accessed 22 Aug 2018.
51. Darling ACE, Mau B, Blattner FR, Perna NT. Mauve: multiple alignment of conserved genomic sequence with rearrangements. *Genome Res*. 2004;14:1394–403.
52. Hyatt D, Chen G-L, LoCascio PF, Land ML, Larimer FW, Hauser LJ. Prodigal: prokaryotic gene recognition and translation initiation site identification. *BMC Bioinforma*. 2010;11:119.
53. Huerta-Cepas J, Forslund K, Coelho LP, Szklarczyk D, Jensen LJ, von Mering C, et al. fast genome-wide functional annotation through orthology assignment by eggNOG-Mapper. *Mol Biol Evol*. 2017;34:2115–22.
54. Takami H, Taniguchi T, Arai W, Takemoto K, Moriya Y, Goto S. An automated system for evaluation of the potential functionome: MAPLE version 2.1.0. *DNA Res*. 2016;23:467–75.
55. Srivastava M, Begovic E, Chapman J, Putnam NH, Hellsten U, Kawashima T, et al. The *Trichoplax* genome and the nature of placozoans. *Nature*. 2008;454:955–60.
56. Kamm K, Osigus H-J, Stadler PF, DeSalle R, Schierwater B. *Trichoplax* genomes reveal profound admixture and suggest stable wild populations without bisexual reproduction. *Sci Rep*. 2018;8:11168.
57. Eitel M, Francis WR, Varoqueaux F, Daraspe J, Osigus H-J, Krebs S, et al. Comparative genomics and the nature of placozoan species. *PLoS Biol*. 2018;16:e2005359.
58. Eitel M, Schierwater B. The phylogeography of the Placozoa suggests a taxon-rich phylum in tropical and subtropical waters. *Mol Ecol*. 2010;19:2315–27.
59. DeSantis TZ, Hugenholtz P, Larsen N, Rojas M, Brodie EL, Keller K, et al. Greengenes, a chimera-checked 16S rRNA gene database and workbench compatible with ARB. *Appl Environ Microbiol*. 2006;72:5069–72.
60. Nawrocki EP, Eddy SR. Infernal 1.1: 100-fold faster RNA homology searches. *Bioinformatics*. 2013;29:2933–5.
61. Guindon S, Dufayard J-F, Lefort V, Anisimova M, Hordijk W, Gascuel O. New algorithms and methods to estimate maximum-likelihood phylogenies: assessing the performance of PhyML 3.0. *Syst Biol*. 2010;59:307–21.
62. Darriba D, Taboada GL, Doallo R, Posada D. jModelTest 2: more models, new heuristics and parallel computing. *Nat Meth*. 2012;9:772–772.
63. Kearse M, Moir R, Wilson A, Stones-Havas S, Cheung M, Sturrock S, et al. Geneious basic: an integrated and extendable desktop software platform for the organization and analysis of sequence data. *Bioinformatics*. 2012;28:1647–9.
64. Na S-I, Kim YO, Yoon S-H, Ha S, Baek I, Chun J. UBCG: up-to-date bacterial core gene set and pipeline for phylogenomic tree reconstruction. *J Microbiol*. 2018;56:280–5.
65. Mistry J, Finn RD, Eddy SR, Bateman A, Punta M. Challenges in homology search: HMMER3 and convergent evolution of coiled-coil regions. *Nucleic Acids Res*. 2013;41:e121–e121.
66. Katoh K, Standley DM. MAFFT multiple sequence alignment software version 7: improvements in performance and usability. *Mol Biol Evol*. 2013;30:772–80.
67. Dayhoff MO, Schwartz RM, Orcutt BC. A model of evolutionary change in proteins. In: Dayhoff, MO, editor. Atlas of protein sequence and structure. suppl. 3. National Biomedical Research Foundation, Washington, D.C. 1978; pp. 345–352.
68. Lartillot N, Lepage T, Blanquart S. PhyloBayes 3: a Bayesian software package for phylogenetic reconstruction and molecular dating. *Bioinformatics*. 2009;25:2286–8.
69. Pisani D, Pett W, Dohrmann M, Feuda R, Rota-Stabelli O, Philippe H, et al. Genomic data do not support comb jellies as the sister group to all other animals. *PNAS*. 2015;112:15402–7.
70. Feuda R, Dohrmann M, Pett W, Philippe H, Rota-Stabelli O, Lartillot N, et al. Improved modeling of compositional heterogeneity supports sponges as sister to all other animals. *Curr Biol*. 2017;27:3864–3870.e4.
71. Philippe H, Brinkmann H, Copley RR, Moroz LL, Nakano H, Poustka AJ, et al. Acoelomorph flatworms are deuterostomes related to *Xenoturbella*. *Nature*. 2011;470:255–8.
72. Thompson LR, Sanders JG, McDonald D, Amir A, Ladau J, Locey KJ, et al. A communal catalogue reveals Earth's multi-scale microbial diversity. *Nature*. 2017;551:457–63.
73. Lagkourados I, Joseph D, Kapfhammer M, Giritli S, Horn M, Haller D, et al. IMNGS: a comprehensive open resource of processed 16S rRNA microbial profiles for ecology and diversity studies. *Sci Rep*. 2016;6:33721.
74. Bowers RM, Kyrpidis NC, Stepanauskas R, Harmon-Smith M, Doud D, Reddy TBK, et al. Minimum information about a single amplified genome (MISAG) and a metagenome-assembled genome (MIMAG) of bacteria and archaea. *Nat Biotechnol*. 2017;35:725–31.
75. Schulz F, Martijn J, Wascher F, Lagkourados I, Kostanjšek R, Ettema TJG, et al. A Rickettsiales symbiont of amoebae with ancient features. *Environ Microbiol*. 2016;18:2326–42.
76. Matsuura Y, Kikuchi Y, Meng XY, Koga R, Fukatsu T. Novel clade of alphaproteobacterial endosymbionts associated with stinkbugs and other arthropods. *Appl Environ Microbiol*. 2012;78:4149–56.
77. Senra MVX, Dias RJP, Castelli M, Silva-Neto ID, Verni F, Soares CAG, et al. A house for two—double bacterial infection in *Euplotes woodruffi* S_{q1} (Ciliophora, Euplotia) sampled in Southeastern Brazil. *Micro Ecol*. 2016;71:505–17.
78. Tatusov RL, Galperin MY, Natale DA, Koonin EV. The COG database: a tool for genome-scale analysis of protein functions and evolution. *Nucleic Acids Res*. 2000;28:33–36.
79. Lin S, Cheng S, Song B, Zhong X, Lin X, Li W, et al. The symbiodinium kawagutii genome illuminates dinoflagellate gene expression and coral symbiosis. *Science*. 2015;350:691–4.
80. Driscoll TP, Verhoeve VI, Guillotte ML, Lehman SS, Rennoll SA, Beier-Sexton M, et al. Wholly Rickettsia! Reconstructed metabolic profile of the quintessential bacterial parasite of eukaryotic cells. *mBio*. 2017;8:e00859–17.
81. Greub G, Raoult D. History of the ADP/ATP-translocase-encoding gene, a parasitism gene transferred from a Chlamydiales ancestor to plants 1 billion years ago. *Appl Environ Microbiol*. 2003;69:5530–5.
82. Major P, Embley TM, Williams TA. Phylogenetic diversity of NTT nucleotide transport proteins in free-living and parasitic bacteria and eukaryotes. *Genome Biol Evol*. 2017;9:480–7.
83. Cheng Z, Lin M, Rikihisa Y. *Ehrlichia chaffeensis* proliferation begins with NtrY/NtrX and PutA/GlnA upregulation and CtrA degradation induced by proline and glutamine uptake. *mBio*. 2014;5:e02141–14.

84. Pawlowski K, Klosse U, Bruijn FJde. Characterization of a novel *Azorhizobium caulinodans* ORS571 two-component regulatory system, NtrY/NtrX, involved in nitrogen fixation and metabolism. *Molec Gen Genet*. 1991;231:124–38.
85. Smith DP, Thrash JC, Nicora CD, Lipton MS, Burnum-Johnson KE, Carini P, et al. Proteomic and transcriptomic analyses of ‘*Candidatus Pelagibacter ubique*’ describe the first PII-independent response to nitrogen limitation in a free-living Alphaproteobacterium. *mBio*. 2013;4:e00133–12-e00133-12.
86. Cho N-H, Kim H-R, Lee J-H, Kim S-Y, Kim J, Cha S, et al. The *Orientia tsutsugamushi* genome reveals massive proliferation of conjugative type IV secretion system and host–cell interaction genes. *PNAS*. 2007;104:7981–6.
87. Gillespie JJ, Brayton KA, Williams KP, Quevedo Diaz MA, Brown WC, Azad AF, et al. Phylogenomics reveals a diverse Rickettsiales type IV secretion system. *Infect Immun*. 2010;78:1809–23.
88. Gouin E, Gantelet H, Egile C, Lasa I, Ohayon H, Villiers V, et al. A comparative study of the actin-based motilities of the pathogenic bacteria *Listeria monocytogenes*, *Shigella flexneri* and *Rickettsia conorii*. *J Cell Sci*. 1999;112:1697–708.
89. Felsheim RF, Kurti TJ, Munderloh UG. Genome sequence of the endosymbiont *Rickettsia peacockii* and comparison with virulent *Rickettsia rickettsii*: identification of virulence factors. *PLoS One*. 2009;4:e8361.
90. Sasser D, Lo N, Epis S, D’Auria G, Montagna M, Comandatore F, et al. Phylogenomic evidence for the presence of a flagellum and *cbb3* oxidase in the free-living mitochondrial ancestor. *Mol Biol Evol*. 2011;28:3285–96.
91. Cascales E, Christie PJ. The versatile bacterial type IV secretion systems. *Nat Rev Micro*. 2003;1:137–49.
92. Alvarez-Martinez CE, Christie PJ. Biological diversity of prokaryotic Type IV secretion systems. *Microbiol Mol Biol Rev*. 2009;73:775–808.
93. Lockwood S, Voth DE, Brayton KA, Beare PA, Brown WC, Heinzen RA, et al. Identification of *Anaplasma marginale* Type IV secretion system effector proteins. *PLoS One*. 2011;6:e27724.
94. Liu H, Bao W, Lin M, Niu H, Rikihisa Y. Ehrlichia type IV secretion effector ECH0825 is translocated to mitochondria and curbs ROS and apoptosis by upregulating host MnSOD. *Cell Microbiol*. 2012;14:1037–50.
95. Szokoli F, Sabaneyeva E, Castelli M, Krenke S, Schrollhammer M, Soares CAG, et al. “*Candidatus Fokinia solitaria*”, a Novel “Stand-Alone” symbiotic lineage of Midichloriaceae (Rickettsiales). *PLoS ONE*. 2016;11:e0145743.
96. Bosch TCG. Cnidarian-microbe interactions and the origin of innate immunity in metazoans. *Annu Rev Microbiol*. 2013;67:499–518.
97. Anderson DA, Walz ME, Weil E, Tonellato P, Smith MC. RNA-Seq of the Caribbean reef-building coral *Orbicella faveolata* (Scleractinia-Merulinidae) under bleaching and disease stress expands models of coral innate immunity. *PeerJ*. 2016;4:e1616. <https://doi.org/10.7717/peerj.1616>
98. Stat M, Huggett MJ, Bernasconi R, DiBattista JD, Berry TE, Newman SJ, et al. Ecosystem biomonitoring with eDNA: metabarcoding across the tree of life in a tropical marine environment. *Sci Rep*. 2017;7:12240.
99. Moitinho-Silva L, Nielsen S, Amir A, Gonzalez A, Ackermann GL, Cerrano C, et al. The sponge microbiome project. *Giga-science*. 2017;6:1–7.
100. Clec’h WL, Braquart-Varnier C, Raimond M, Ferdy J-B, Bouchon D, Sicard M. High virulence of *Wolbachia* after host switching: when autophagy hurts. *PLoS Pathog*. 2012;8:e1002844.

Affiliations

J. Grace Klinges¹ · Stephanie M. Rosales^{1,2,3} · Ryan McMinds^{1,4} · Elizabeth C. Shaver⁵ · Andrew A. Shantz^{6,7} · Esther C. Peters⁸ · Michael Eitel⁹ · Gert Wörheide^{9,10,11} · Koty H. Sharp¹² · Deron E. Burkepile⁶ · Brian R. Silliman⁵ · Rebecca L. Vega Thurber¹

¹ Department of Microbiology, Oregon State University, Corvallis, OR 97331, USA

² Cooperative Institute for Marine and Atmospheric Studies, University of Miami, Miami, FL 33149, USA

³ Atlantic Oceanographic and Meteorological Laboratory, National Oceanographic and Atmospheric Administration, Miami, FL 33149, USA

⁴ Center of Modeling, Simulation and Interactions, Université Côte d’Azur, Nice, France

⁵ Division of Marine Science and Conservation, Nicholas School of the Environment, Duke University, Beaufort, NC 28516, USA

⁶ Department of Ecology, Evolution and Marine Biology, University of California, Santa Barbara, Santa Barbara, CA 93106-9610, USA

⁷ Department of Biology, Pennsylvania State University, University Park, State College, PA 16802, USA

⁸ Department of Environmental Science and Policy, George Mason University, Fairfax, VA 22030-4444, USA

⁹ Department of Earth and Environmental Sciences, Paleontology and Geobiology, Ludwig-Maximilians-Universität München, Munich, Germany

¹⁰ GeoBio-Center, Ludwig-Maximilians-Universität München, Munich, Germany

¹¹ Staatliche Naturwissenschaftliche Sammlungen Bayerns (SNSB) – Bayerische Staatssammlung für Paläontologie und Geologie, Munich, Germany

¹² Department of Biology, Marine Biology & Environmental Science, Roger Williams University, Bristol, RI 02809, USA