

Title:

Deep-sea corals provide new insight into the ecology, evolution, and the role of plastids in widespread apicomplexan symbionts of anthozoans.

Authors:

Samuel A. Vohsen^{1*}, Kaitlin E. Anderson¹, Andrea M. Gade¹, Harald R. Gruber-Vodicka², Richard P. Dannenberg^{1,3}, Eslam O. Osman^{1,4}, Nicole Dubilier², Charles R. Fisher¹, Iliana B. Baums¹

¹Biology Department, Pennsylvania State University, University Park, Pennsylvania, United States of America

²Department of Symbiosis, Max Planck Institute for Marine Microbiology, Bremen, Germany.

³Currenty at Epic, Madison, WI, United States of America

⁴Marine Biology Department, Faculty of Science, Al Azhar University, Cairo, Egypt

Corresponding Author: Samuel A. Vohsen, sav146@psu.edu

Abstract:

Background

Apicomplexans are the causative agents of major human diseases such as malaria and toxoplasmosis. A novel group of apicomplexans, recently named corallicolids, have been detected in corals inhabiting tropical shallow reefs. These apicomplexans may represent a transitional lifestyle between free-living phototrophs and obligate parasites. To shed light on the evolutionary history of apicomplexans and to investigate their ecology in association with corals, we screened scleractinians, antipatharians, alcyonaceans, and zoantharians from shallow, mesophotic, and deep-sea communities. We detected corallicolid plastids using 16S metabarcoding, sequenced the nuclear 18S rRNA gene of corallicolids from selected samples, assembled and annotated the plastid and mitochondrial genomes from a corallicolid that associates with a deep-sea coral, and screened the metagenomes of four coral species for corallicolids.

Results

We detected 23 corallicolid plastotypes that were associated with 14 coral species from three orders and depths down to 1400 m. Individual plastotypes were restricted to coral hosts within a single depth zone and within a single taxonomic order of corals. Some clusters of closely related corallicolids were revealed that associated with closely related coral species. However, the presence of divergent corallicolid lineages that associated with similar coral species and depths suggests that corallicolid/coral relations are flexible over evolutionary timescales and that a large diversity of apicomplexans may remain undiscovered. The corallicolid plastid genome from a deep-sea coral contained four genes involved in chlorophyll biosynthesis: the three genes of the LIPOR complex and *acsF*.

Conclusions

The presence of corallicolid apicomplexans in corals below the photic zone demonstrates that they are not restricted to shallow-water reefs and are more general anthozoan symbionts. The presence of LIPOR genes in the deep-sea corallicolid precludes a role involving photosynthesis and suggests they may be involved in a different function. Thus, these genes may represent another set of genetic tools whose function was adapted from photosynthesis as the ancestors of apicomplexans evolved towards parasitic lifestyles.

Key Words: deep-sea, corals, apicomplexans, LIPOR, symbiont

Background:

Colonial corals at all ocean depths [1] form diverse [2-4] and economically [5] valuable communities that face a variety of threats, including increasing ocean temperatures [6], ocean acidification [7, 8], nutrient runoff [9], disease [6], and oil and gas extraction and production activities [10, 11]. These stressors have prompted the investigation of the role of microbes in the physiology of the coral holobiont [12, 13] and how these microbes may mitigate or exacerbate these stressors [14]. The best-studied coral symbionts are the widespread, eukaryotic, photosynthetic algae in the family Symbiodiniaceae [15]. Recently, two additional alveolate, phototrophic coral symbionts were discovered, which are both close relatives of apicomplexans: *Chromera velia* [16, 17] and *Vitrella brassicaformis* [18]. These organisms attracted attention not only for their potential role as auxiliary phototrophic symbionts, but also for their phylogenetic position basal to apicomplexans. They have been postulated to represent a transitional lifestyle between free-living phototrophs and obligate parasites and may shed light on the evolutionary history of important parasites of humans, such as the apicomplexans *Plasmodium* spp. and *Toxoplasma gondii* [16, 19].

In addition to these phototrophic organisms, evidence of other widespread apicomplexan-related symbionts present in corals is mounting. First, a coccidian apicomplexan, *Gemmocystis cylindrus*, was described by transmission electron microscopy in the mesenterial filaments of multiple Caribbean corals [20]. Later, nuclear encoded 18S rRNA gene sequences of apicomplexan relatives were found in corals in the Caribbean and Great Barrier Reef and named ‘genotype N’ [21, 22]. Genotype N was later shown to cluster alongside coccidians within Apicomplexa [19]. Similarly, 16S rRNA gene sequences from the plastid of apicomplexan relatives appeared in bacterial surveys of corals pantropically [23-25]. These plastid sequences form a highly diverse group that clusters basally to apicomplexans and were thus named Apicomplexan-Related Lineage V (ARL-V) [24].

It has been hypothesized that some or all of these sequences assigned to genotype N and ARL-V arise from the same organism despite the differences in phylogeny between these nuclear and plastid-encoded genes (Fig 1) [22]. Recently, co-localized signals were observed in the mesenterial filaments of *Rhodactis* sp. using probes designed for cytosolic and plastid ribosomes of genotype N and ARL-V, respectively, supporting the hypothesis. These organisms were named ‘corallicolids’ [26]. The localization of signals from ARL-V, genotype-N, and *G. cylindrus* within the tissue of corals suggests that they are endosymbionts, yet their role in the holobiont remains unclear.

In contrast to some apicomplexans like the parasite *Plasmodium* spp., neither genotype N nor ARL-V are associated with any pathology and are commonly found in association with apparently healthy corals [23, 27, 28]. In addition, genotype N and ARL-V occur at high prevalence, having been detected in over 80% of adult colonies in some scleractinian species [27, 28]. Further, genotype N occurs in a high percentage of the planulae of some brooding coral

species suggesting vertical transmission in these species and an association that persists throughout the life cycle of the coral [27, 28].

The role of ARL-V was thought to be phototrophic since it appears basal to apicomplexans and was initially reported to decrease in abundance with depth [23]. However, it was recently shown that the plastid genome of a corallicolid isolated from *Rhodactis* sp. lacks the full genomic repertoire necessary to conduct photosynthesis [26]. The plastid genome did, however, include genes involved in chlorophyll biosynthesis suggesting that corallicolids may produce chlorophyll or a derivative of chlorophyll that interacts with light or is involved in heme regulation [26]. Whether corallicolid apicomplexans are parasites with minimal impact on host fitness, commensals, or mutualists is still unclear.

Most work on corallicolids has focused on shallow-water coral species from reefs in tropical latitudes [19-24, 26-29]. These apicomplexans have only been detected in a few cold-water coral species [25]. However, the majority of known coral species live deeper than 50 m [1], and the communities they form have a much wider range than tropical corals, occurring along most continental margins from the Arctic to the Antarctic [30, 31]. Here, we investigate the occurrence of corallicolids in deep-sea and mesophotic scleractinians, antipatharians, zoantharians, and alcyonaceans from the Gulf of Mexico, ranging in depth from 60 to 2224 m, with comparison to shallow-water corals from the Florida Keys and Curaçao. We also investigate phylogenetic patterns among corallicolids across host phylogeny, geography, and depth using plastid and nuclear marker sequences. Finally, we probe its role in deep-sea corals by analysis of the plastid genome sequence of a corallicolid that associates with the globally distributed deep-sea antipatharian *Leiopathes glaberrima* [Esper 1788] [32-34].

Results:

Plastid and mitochondrial genomes of corallicolids

A 6300 bp mitogenome of apicomplexan origin was assembled from a colony of *Leiopathes glaberrima* collected from a depth of 450 m at site Viosca Knoll (VK) 826. Site names follow the Bureau of Ocean Energy Management's lease block designations which consist of a two-letter region abbreviation followed by 3 digits. Like other apicomplexans, this mitogenome contained multiple rRNA fragments and only three complete genes: *cox1*, *cox3*, and *cytb* [35, 36]. A phylogenetic tree placed this apicomplexan with other corallicolids alongside coccidians using amino acid sequences of these three genes (Fig 2).

From this same metagenomic library, a 45,543 bp plastid genome corresponding to ARL-V was assembled that was composed of two scaffolds. This plastid genome contained four genes involved in chlorophyll biosynthesis that are not present in the apicoplast genome of any other, non-corallicolid, apicomplexan [19, 26]. Three of these genes compose the light independent protochlorophyllide oxidoreductase (LIPOR) complex (*chlL*, *chlN*, *chlB*), and the fourth is aerobic magnesium-protoporphyrin IX monomethyl ester cyclase (*acsF*). A phylogenetic tree of the 5 longest protein-coding plastid genes also showed that the corallicolid in *Leiopathes* clustered alongside coccidian apicomplexans (Fig 2).

Presence of corallicolids in metagenomes

Corallicolid 18S, mitogenomes, and/or plastid genomes were detected in the metagenomes of *L. glaberrima* (9/12), *Callogorgia delta* (3/8), and *Acropora palmata* (9/21) but not in any *Paramuricea* sp. type B3 (n=4) (Fig 3). Corallicolids were detected in multiple colonies of *A. palmata* in all four sampled regions spanning the Caribbean: Florida Keys, US Virgin Islands, Belize, and Curaçao. Corallicolids were only detected in *C. delta* colonies from

lease blocks Mississippi Canyon (MC) 885 (1/2) and Green Canyon (GC) 234 (2/2) and were not detected in any colony from MC751 (n=4). Finally, LIPOR genes and/or *acsF* of corallicolid origin were detected in two deep-sea coral species: *L. glaberrima* (6/12) and *C. delta* (2/8), as well as the shallow-living *A. palmata* (1/21).

16S screening for plastids across a wide diversity of corals

In total, 25 plastotypes were identified based on the 16S rRNA gene and 23 of these corresponded to corallicolid plastids (ARL-V). Corallicolid plastotypes were detected among 14 coral species including scleractinians (five plastotypes, 1 coral species), antipatharians (five plastotypes, 3 coral species), alcyonaceans (thirteen plastotypes, 10 coral species), and between three major depth zones: shallow (0-20 m, six plastotypes 1-6, 3 coral species), mesophotic (60-100 m, eight plastotypes 7-14, 3 coral species), and deep-sea (250-1400 m, nine plastotypes 15-23, 8 coral species) (Fig. 4b). Two plastotypes (24 and 25) clustered outside of ARL-V (Fig S2 in Additional File 1). One was found in two colonies of the mesophotic coral *Swiftia exserta* (plastotype 24), and a second was found in a single deep-sea *Stichopathes* sp. colony (plastotype 25).

Depth and Geographic patterns of plastid markers

Corallicolid plastids were detected down to 1400m but were rare below 700m (Fig 4). All plastotypes were restricted to a single depth zone (shallow, mesophotic or deep). However, some had wide geographic ranges within a depth zone (Fig 4, 5, 6). For instance, plastotypes 7 and 8 were present in *Swiftia exserta* and/or *Muricea pendula* at all seven mesophotic sites spanning 563 km, and plastotype 16 was found in *L. glaberrima* from 6 deep-sea sites spanning 794 km east to west (Fig 5). While some plastotypes had wide ranges, others were not found at every site where their host corals were sampled (Fig 6). For instance, plastotype 19 associated with *C. delta*

colonies at sites GC234 (13 out of 37), MC885 (16/25), and VK826 (2/3), but no plastotype was detected in any *C. delta* colony from MC751 (n=19), GC249 (n=14), or GC290 (n=10). Additionally, plastotype 17 was only found in *L. glaberrima* colonies from sites on the West Florida Slope and was absent from 4 other sites where *L. glaberrima* was sampled and other plastotypes were present in this species (Fig 6).

Host Flexibility and Specificity of plastid markers

Many pairings between coral host and apicomplexan symbiont were flexible. At least 6 plastotypes were detected in multiple coral species (Fig 5 and 6). For instance, plastotypes 7, 8, and 9 were found in both *S. exserta* and *M. pendula* across their ranges. Similarly, plastotype 19 was found in five deep-sea octocorals including *C. delta*, *Acanthogorgia aspera*, *Paragorgia* sp. 1, *Paramuricea* sp., and *Swiftia* sp. Pairings were also flexible from the host perspective since many coral species hosted multiple plastotypes as did many individual colonies. Eight coral species hosted at least 2 plastotypes. Among colonies with apicomplexans, about 47% hosted multiple plastotypes, and some hosted up to four in the same sample. Despite these flexibilities, some potential patterns of specificity were observed. For instance, plastotype 18 was only detected in the genus *Paramuricea* at both MC751 and GC234 despite the presence of other coral species that hosted corallicolids at these sites. Further, every plastotype was restricted to one order of corals.

Phylogenetic patterns among corallicolids using plastid markers

Phylogenetic relationships among corallicolids based on plastid sequences were not strictly governed by coral host phylogeny, depth, or geography (Fig 7). For instance, corallicolid plastid sequences obtained from octocorals did not form a single cluster and were instead dispersed among those obtained from scleractinians. Similarly, sequences from deep-sea corals

180 did not form a single cluster. Further, plastotypes detected in the same coral colony were not
181 always their closest relatives such as plastotypes 16 and 17 which co-occurred within some *L.*
182 *glaberrima* colonies but had divergent sequences. Despite this lack of general patterns, five well-
183 supported clusters included sequences from host coral species restricted to a single order or
184 lower taxonomic level. Two clusters with good bootstrap support (>95%) included only
185 sequences obtained from octocorals. One of these clusters included octocoral hosts from shallow,
186 mesophotic, and deep-sea habitats and included five plastotypes obtained in this study (9, 13, 19,
187 20, and 23). Further, two well supported clusters (>98%) only included sequences recovered
188 from shallow-water scleractinian corals. One of these clusters included four plastotypes found in
189 *A. palmata* (plastotypes 1-4), which formed a well-supported cluster (99%) by themselves.
190 Finally, plastotypes 15 and 16 were obtained from *L. glaberrima* and formed their own well-
191 supported cluster (98%).

192 Phylogenetic patterns among corallicolids using nuclear markers (18S sequences)

193 Thirty-two nuclear-encoded 18S sequences corresponding to genotype N were recovered
194 from eight coral species including one scleractinian, 3 antipatharians, and 4 alcyonaceans. Like
195 the phylogenetic relationships inferred from plastid sequences, those inferred from nuclear
196 sequences were not strictly clustered based on coral host, depth, or geography. Despite this,
197 several well-supported clusters included sequences obtained from a single coral species or order
198 (Fig 7). For instance, sequences obtained from *Leiopathes glaberrima* from Viosca Knoll sites
199 (VK826 and VK906) formed a well-supported cluster (100%) which was distantly related to a
200 second cluster of sequences obtained from *L. glaberrima* from the West Florida Slope (95%).
201 Two other clusters were well supported (>97%) and only included sequences obtained from
202 octocorals. One of these clusters included sequences from *Callogorgia delta* and *Ellisella* sp. and

further contained another well-supported subcluster (98%) of sequences that were obtained from *C. delta* only. The second octocoral cluster included sequences from the mesophotic corals *Ellisella* sp., *Swiftia exserta*, and *Muricea pendula* from all 7 sites. Within this cluster, sequences from *S. exserta* and *M. pendula* clustered with 90% support; however, the relatively small amount of structure within this cluster did not reflect coral host species or site.

Discussion:

The newly discovered apicomplexans named “corallicolids” have been identified in diverse anthozoan hosts, yet their function is unknown. By surveying a broad diversity of anthozoans from a variety of habitats, we have obtained insight into the ecology, evolution, and potential role of this novel group of apicomplexans in corals.

Ecology

This dataset provides some clues regarding the potential niche preferences of corallicolids and their mode of transmission. Here, we expand the known distribution of corallicolids to a depth of at least 1400 m. Many corallicolids were detected in mesophotic and deep-sea coral species demonstrating that corallicolids as a group are general anthozoan associates that are not restricted to corals in tropical shallow-water reefs. However, corallicolids were not as prevalent at sites deeper than about 700 m despite extensive sampling down to 2200 m (Fig 4). The underlying cause of this pattern is not clear. Corallicolids may be limited to a certain depth if they have a function that utilizes light. While corallicolids do have plastids which encode some genes in chlorophyll synthesis, there is no indication of any function that involves light. Alternatively, if corallicolids rely on vector species as other apicomplexans do, their transmission may be limited in deeper water where mobile species exist at lower densities. Another possibility is that corallicolids were not present in the particular coral species that we

sampled below 700m but are present in other corals at those depths since corallicolids were not detected in all coral species at shallower depths. Comparisons to additional detection methods, coral species, and/or ocean basins will be necessary to establish the consistency of this pattern.

While corallicolids as a group had a wide range, specific plastotypes differed in niche characterized by habitat and varying host coral specificity. For instance, many apicomplexan plastotypes were widespread, but all were restricted to corals within a depth zone. The only exception was plastotype 16 which was present at very high relative abundances in *L. glaberrima* colonies at Viosca Knoll sites but was also detected at low abundances in samples from other depth zones and coral orders (Fig S1 in Additional File 1). We suggest this is likely to have resulted from contamination or sample cross-talk [37, 38] because the only samples which contained plastotype 16, other than those from *L. glaberrima*, were processed on a sequencing run which included 30 *L. glaberrima* samples. Plastotype 16 was present at very low relative abundance in these samples (including 2 of 4 Powersoil kit extraction blanks and 1 of 4 PCR negative controls) and many had fewer reads that passed quality control and likely had low DNA concentrations and were more susceptible to contamination or sample cross-talk. Further, no plastotype 16 was found in any coral except *L. glaberrima* processed on other sequencing runs. Finally, 21 colonies had replicates on both sequencing runs. In 10 colonies, plastotype 16 was detected in the replicate on the first sequencing run but was not detected in any replicate on the second sequencing run. Plastotype 16 was not detected in either the first or second sequencing run in the remaining 11 colonies.

The restriction of all other plastotypes to a single depth zone may, in part, reflect the fact that all coral species sampled were also restricted to a single depth zone. However, some plastotypes were found in divergent alcyonaceans over a 500 km span but were not found in

other alcyonacean species as close as 30 km away in deeper waters. In addition to habitat preference, all corallicolids were limited by host phylogeny. No corallicolid was detected with confidence in coral species which belonged to multiple orders. Some plastotypes seemed restricted to a single coral species or genus, such as plastotype 18, which was only found in *Paramuricea* sp. at two separate sites despite the presence of other octocorals nearby. Others, however, were flexible, such as the corallicolids which were detected in both *Swiftia exserta* and *Muricea pendula* which belong to separate families [39]. The presence of the same plastotypes in these two coral species is unlikely to be due to contamination (see Supplement for explanation). Host specificity of individual corallicolids has not yet been reported for this group. Previous work, however, has only found patterns in the community of plastid sequences in three coral species [26].

The data further suggest that some corallicolids may be horizontally transmitted. First, specific corallicolids were detected in divergent coral species indicating that these corallicolids may disperse between colonies independent of their host corals' larvae. Second, some corallicolids were not detected at all sites where their host coral species were sampled. For example, we failed to detect corallicolids in *C. delta* at MC751 using both 16S surveys and metagenomes despite detecting corallicolids in around 50% of *C. delta* colonies within three other sites. Third, the distribution of some corallicolids did not reflect their host population structure. For instance, plastotype 17 was only found in *Leiopathes glaberrima* on the West Florida Slope, and we failed to detect it in colonies at GC140 and GB299, even though these host populations are connected via gene flow [34]. It is possible that additional corallicolids may be present at sites where they were not detected because they were below our threshold of detection. However, this seems unlikely given the number of colonies screened.

Evolution

As previously reported [26], mitochondrial sequences of corallicolids clustered within Apicomplexa alongside coccidians. However, our analysis of plastid sequences also placed corallicolids within apicomplexans while Kwong et al. placed them basal to apicomplexans. This may be due to our selection of only the 5 longest protein coding genes while Kwong used 19 genes and examined multiple tree building methods. This discrepancy may be resolved as a greater diversity of apicomplexan relatives are analyzed.

Several clades of corallicolids were associated with coral hosts from a single order or lower taxonomic level (Fig 7) which suggests the possibility of coevolution between corallicolids and their coral hosts within these groups. However, the relationships between these clades did not reflect the coral phylogeny, depth zone, or geography. Multiple clusters of corallicolids grouped with other corallicolids that associated with corals from different orders and depth zones. Some corallicolids even infected multiple coral species while corallicolids from divergent clades infected the same coral species and even the same individual. This suggests that corallicolid lineages have transitioned between depth zones and coral order multiple times and thus exhibit flexibility over evolutionary timescales.

The lack of strong large-scale phylogenetic correlations with host, habitat, or geography may be due to the relatively sparse sampling of the tremendous anthozoan diversity and their wide geographic ranges. The inclusion of corallicolids from additional coral species, depths, and locations may reveal undetected phylogenetic patterns. Taxonomic and phylogenetic descriptions of another coral-associated alveolate group, dinoflagellates in the family Symbiodiniaceae, was stymied for decades by similar challenges because the family is composed of multiple speciose genera with varying levels of host specificity [15]. Some species of zooxanthellae show high

host specificity while others are host generalists [40, 41]. Conversely, some corals only associate with one species of zooxanthellae while other corals are more flexible [40].

Several patterns appeared in both nuclear and plastid phylogenies including clusters of sequences from the same or closely related host coral species (Fig 7). For instance, a clade including sequences from both *C. delta* and *Ellisella* sp. appeared in both nuclear and plastid phylogenies (Fig 7, blue boxes). Similarly, a second clade including sequences from *Swiftia exserta* and *Muricea pendula* was present in both phylogenies (Fig 7, gray boxes). However, the nuclear and plastid phylogenies based on single marker genes of corallicolids showed limited phylogenetic congruence. The shared clusters mentioned above did not cluster in the same positions relative to each other when comparing trees (Fig 7). For instance, the 18S sequences obtained from shallow-water antipatharians and the cluster including our sequence from *Acropora palmata* were the most basal sequences yet their associated plastid sequences were placed in more derived positions. These incongruencies may reflect the low support values for deeper level branches and the limited power of these sequences to resolve these relationships.

Definitive alignments between the nuclear and plastid phylogenies are also limited by the difficulty in assigning plastotypes to nuclear sequences because the variation of plastid sequences associated with a single nuclear lineage is unknown and may vary at multiple levels (Fig 1). First, the two copies of 16S may differ within a plastid genome. Second, corallicolids may be heteroplasmic if individual cells harbor plastids which exhibit sequence variation [42]. Third, corallicolids with similar nuclear sequences may harbor divergent plastids. Finally, there may be hidden diversity in nuclear sequences that was not detected and is associated with some of the divergent 16S sequences. For instance, all corallicolid 18S sequences obtained from *S. exserta* and *M. pendula* formed a single cluster yet the plastotypes detected in the same samples

belonged to two divergent clades. Plastotypes 7 and 8 may be associated with the cluster of 18S sequences detected in these corals but plastotype 9, which was present at lower relative abundances, may be associated with an 18S lineage that was not detected.

Role of the symbiont

Contrary to early speculation [29], we conclude that corallicolid plastids are not likely to be involved in photosynthesis. Kwong et al. [26] reached a similar conclusion because the plastid genome of corallicolids did not encode photosystem genes; however, they did not rule out their presence in the nuclear genome. Our survey provides further evidence against photosynthesis since many corallicolids were found in deep-sea coral species at depths below the photic zone. Further, LIPOR genes and *acsF* were confirmed in the plastids of corallicolids that associate with corals at depths greater than 400 m.

While a role involving photosynthesis can be confidently ruled out for corallicolids, we cannot rule out a parasitic lifestyle like that of their apicomplexan relatives. Previous speculation has discounted parasitism since corallicolids were not associated with a recognizable pathology in their coral hosts. However, even the well-established parasitic coccidian apicomplexan, *Toxoplasma gondii*, appears asymptomatic in most human infections. It only results in more severe pathologies in a small percentage of immunocompetent humans [43-46]. Corallicolids may be similar parasites in corals, with a minor impact on host fitness under normal circumstances but increasing in abundance and impact on the host when the corals are stressed. This is consistent with the observation that apicomplexans increased in abundance in corals afflicted with “white plague” disease [47]. Further, the original work that described *Gemmocystis cylindrus* reported patchy bleaching and tissue necrosis in some infected corals [20]. However,

without appropriate controlled experiments, it is difficult to determine the role of these apicomplexans in corals.

In addition to the lifestyle of corallicolids, the function of their LIPOR genes and *acsF* remains uncertain. These genes may still be producing chlorophyll for an unknown purpose. This seems unlikely since Kwong et al. [26] reported no autofluorescence from corallicolids while autofluorescence was observed in co-occurring photosynthetic Symbiodiniaceae in the shallow-water corallimorph they examined. Another possibility is that the LIPOR gene products may have evolved a different function altogether. This gene family has a history of dramatic functional shifts since the genes of the LIPOR complex share an evolutionary origin with *nif* genes which compose a nitrogen-fixing enzyme complex [48]. In corallicolids, the LIPOR complex may perform a different role that uses its Fe-S centers as active sites, such as the synthesis of isoprenoids or lipoic acid. These compounds are synthesized by other enzymes with Fe-S centers in the apicoplasts of apicomplexans and are essential for their survival in their hosts [49].

Whatever its function, we hypothesize that the LIPOR enzyme in corallicolids is sensitive to oxygen concentrations and may be related to a low oxygen niche for several reasons: First, oxygen impedes the LIPOR enzyme complex in other organisms by binding its Fe-S centers [50-52]. Second, corallicolids were detected in the mesenteries in the gastric cavity of polyps which can be a hypoxic to anoxic environment in corals [53]. Third, some of the closest relatives to corallicolids based on 18S sequences were recovered from anoxic marine sediment from Greenland [19]. Fourth, certain life stages of other apicomplexans appear to be adapted to a low oxygen niche by relying primarily on fermentation [54, 55]. Finally, some enzymes that conduct essential functions of plastids in other apicomplexans, such as isoprenoid biosynthesis, are

impeded by oxygen [56, 57]. Altogether, this suggests that corallicolids may occupy a microoxic niche in corals and share many characteristics with parasitic apicomplexans.

Conclusions:

Here, we provide the first foray into a systematic survey of a novel anthozoan associate: corallicolids. We show that corallicolids are more widespread and diverse than previously thought by acquiring and analyzing samples from underrepresented taxa and habitats. We found that corallicolids form a diverse, apicomplexan clade with members that appear to be limited to coral hosts within a taxonomic order and within broad depth zones. We detected several clades of corallicolids that associated with closely related coral hosts. However, some corallicolid lineages show signs of transitioning between coral hosts and habitats over evolutionary time scales. Finally, the presence of LIPOR genes encoded in the plastids of apicomplexans associated with deep-sea corals suggests that the plastid's role is not photosynthesis and is perhaps related to a low oxygen environment within coral polyps. These genes may represent another toolkit that has been repurposed from a role in photosynthesis to a new role as the ancestor of apicomplexans evolved towards a parasitic lifestyle.

Methods:

Collections

Four hundred and twenty-one coral colonies from forty-two coral morphospecies including scleractinians (hard corals), antipatharians (black corals), zoantharians, and alcyonaceans were sampled from 33 sites in the Gulf of Mexico, Curaçao, and the Florida Keys ranging in depth from 2-2224 m. Thirty-two of those morphospecies were collected from 21 deep-sea sites in the Northern Gulf of Mexico ranging in depth from 249-2224 m and spanning 897 km east to west (Figure 5, Table S1 in Additional File 1). The deep-sea collection sites are

named for the Bureau of Ocean Energy Management lease block in which they are found, for example MC751 is lease-block 751 in the Mississippi Canyon lease area. Five coral species were collected from seven mesophotic sites ranging in depth from 60-100 m. The mesophotic and deep-sea corals were collected on 8 research cruises between 2009 and 2017 with remotely-operated vehicles (ROVs) onboard the vessels NOAA *Ronald H Brown* (2009 and 2010, ROV Jason II), R/V *Atlantis* (2014, HOV Alvin), E/V *Nautilus* (2015, ROV Hercules), DSV *Ocean Inspector* (2016, ROV Global Explorer), MSV *Ocean Intervention II* (2017, ROV Global Explorer), DSV *Ocean Project* (2017, ROV Comanche), and NOAA *Nancy Foster* (2017, ROV Odysseus). Each ROV dive was restricted to one site and thus depth changed little within a dive (10-20m). Branches of mesophotic and deep-sea corals were sampled using specially designed coral cutters on the manipulator arm of the ROV and stored in separate compartments of a temperature-insulated biobox or separated within quivers using specially designed rubber stoppers. Upon recovery of the ROV, corals were transferred to cold seawater (<10°C), and subsamples were flash frozen or stored in 90% ethanol within four hours. Frozen samples were stored at -80°C and ethanol samples at -20°C before DNA extraction. We attempted to flash freeze all samples but since liquid nitrogen was limited on some cruises, we preserved some samples in 90% ethanol instead including *Muricea pendula* from Diaphus Bank and McGrail Bank, *Ellisella* sp. from Diaphus Bank, *Swiftia exserta* from Geyer Bank, *Callogorgia delta* from MC751, MC885, and GC234 in 2015 but not other years, *Leiopathes glaberrima* from sites on the West Florida slope, and a single *Swiftia pallida* colony from GC852.

Three coral species were collected from shallow-water sites. Six colonies of *Acropora palmata* in total were sampled from four sites in the Florida Keys in November 2014 (Table S1 in Additional File 1). Coral fragments were removed using a hammer and chisel, snap frozen at

409 the surface, and stored at -80°C until extraction. Two *Antipathes* sp. and 10 *Stichopathes* sp.
410 colonies were collected from Director's Bay in Curaçao in March 2017 from 10-20 m depth.
411 Coral fragments were removed using bone cutters and kept in individual Ziploc bags with
412 seawater for transportation to CARMABI Research Station where they were preserved in 90%
413 ethanol and stored at -20°C until DNA extraction.

414 Metagenomes

415 To construct metagenomic libraries, DNA was extracted from 12 *Leiopathes glaberrima*
416 colonies using a Powersoil DNA extraction kit (Qiagen, Hilden, Germany) and from 8
417 *Callogorgia delta* and 4 *Paramuricea* sp. B3 colonies using a DNA/RNA Allprep kit (Qiagen,
418 Hilden, Germany) with RLT+ lysis buffer with β -mercaptoethanol. DNA/RNA Allprep kits were
419 used on these species because it produced extracts with higher DNA concentrations and less
420 degradation for these species. Whole tissue including skeleton was homogenized using bead-
421 beating (Powersoil kit) or a tissue homogenizer (Allprep kit). All library preparation was
422 conducted at the Max Planck Genome Center in Cologne, Germany and sequenced on an
423 Illumina HiSeq2500 platform with 100 bp paired-end reads for all *L. glaberrima* colonies and
424 150 bp paired-end reads for both *C. delta* and *P.* sp. B3 metagenomes. Four of these *Leiopathes*
425 *glaberrima* libraries were subsequently sequenced deeper using 250 bp paired-end reads.

426 Twenty-one publicly available *Acropora palmata* metagenomes were used under the
427 sequence read archive accession numbers SRR7235979- SRR7235980, SRR7235982,
428 SRR7235983, SRR7235985- SRR7235988, SRR7236001, SRR7236003, SRR7236007-
429 SRR7236012, SRR7236015, and SRR7236017- SRR7236020. These samples originate from
430 four distinct areas spanning the entire range of this species in the Caribbean: Florida Keys (n=6),
431 US Virgin Islands (n=5), Belize (n=5), and Curaçao (n=5) [58].

Metagenomes were quality filtered by trimming reads to a quality score of greater than 2 using bbdduk (ver 37.52) and screened for ARL-V using phyloFlash (ver3.3) [59]. The *Leiopathes glaberrima* library with the highest ARL-V coverage was assembled using Megahit (ver1.0.5) [60] under default parameters and a k-max of 141. A single circular contig of 6300 bp was identified as the mitochondrial genome based on its sequence similarity to coccidian apicomplexan mitochondrial genes (NCBI BLASTn) [61]. A draft plastid assembly was constructed using a contig containing the ARL-V 16S rRNA gene and two other contigs linked in the De Bruijn assembly graph using Bandage ver0.8.1 [62]. These three contigs were combined into one scaffold manually because they had significant overlap. A final putative plastid contig was identified by coverage and GC content and contained sequences similar to apicoplasts (NCBI BLASTn). The mitogenome and draft plastid genome were annotated using GeSeq within the Chlorobox web server [63] using tRNAscan-SE [64]. The *Eimeria tenella* mitogenome (HQ702484) and a corallicolid mitogenome and plastid genome (MH320093, MH324845) were used as references.

All metagenomes were screened for the presence of corallicolids with bbmap (Bushnell, B. Available from: <https://sourceforge.net/projects/bbmap/>) by mapping reads to the corallicolid 18S, mitogenome, and plastid genome from *Leiopathes glaberrima*. Corallicolid mitogenomes and plastid genomes were considered detected in a library if reads mapped with greater than 90% identity and 96% coverage for 250 bp and 150 bp libraries and 98% identity and coverage for 100 bp libraries. Apicomplexan nuclear 18S was considered detected if reads mapped with greater than 95% identity and 96% coverage for 150 bp and 250 bp reads. No reads from any 100 bp library could be confidently classified as corallicolid 18S.

16S rRNA metabarcoding

All coral samples were screened for the presence of ARL-V by 16S rRNA metabarcoding. DNA was extracted from deep-sea corals, mesophotic corals, and corals from Curaçao using Powersoil DNA extraction kits. Four extraction blanks were processed alongside coral samples using Powersoil DNA extraction kits. Some samples were processed using other extraction kits. DNeasy kits (Qiagen, Hilden, Germany) were used for *Acropora palmata* samples (n=36) because these samples were originally used in a separate project [65] and four colonies of *Callogorgia americana* were processed using DNA/RNA Allprep kits since these were older samples and we wanted to increase the chance of obtaining high quality DNA extracts. For all samples, whole coral tissue including skeleton was homogenized using bead-beating for Powersoil DNA extraction kits or a tissue homogenizer for DNeasy and DNA/RNA Allprep kits. To compare extraction procedures, five colonies each of *Paramuricea biscaya*, *Callogorgia delta*, *Swiftia exserta*, and *Muricea pendula* were extracted using both a DNA/RNA Allprep kit using RLT+ lysis buffer with β -mercaptoethanol and a Powersoil DNA extraction kit. Finally, replicate extractions using the same kit were performed on some colonies allowing comparisons of the detection rate of apicomplexan plastids. Six replicate extractions were performed on all 6 *Acropora palmata* colonies. Also, 2-4 replicate extractions were performed on each of 6 colonies of *Callogorgia delta*.

The V1-V2 region of the 16S rRNA gene was amplified using the general bacterial primers, 27F and 355R, which were appended with Illumina CS1 and CS2 Fluidigm adapter sequences [66]. Polymerase chain reactions of 20 μ L were conducted using 0.1 U/ μ L Gotaq (Promega, Madison, WI), 1X Gotaq buffer (Promega, Madison, WI), 0.25 mM dNTPs (Bioline, Alvinston, Canada), 2.5 mM MgCl₂, and 0.25 μ M of each primer. Initial denaturation lasted 5

min at 95°C. This was followed by 30 cycles of 30 sec at 95°, 1 min at 51°C, and 1 min at 72°C. Final extension lasted 7 min at 72°C. Successful amplification was checked on a 1% agarose gel. The DNA services facility at the University of Illinois Chicago prepared libraries and sequenced all samples on two separate runs on an Illumina MiSeq platform. Four negative PCR controls were processed alongside other samples using sterile water instead of DNA extract and sequenced on the first run.

Amplicon libraries of 16S genes were analyzed using QIIME2 version 2017.11 [67] using default parameters unless otherwise stated. Paired reads were joined using vsearch and quality filtered using ‘quality-filter q-score-joined’. Chimeras were detected and removed, sub-operational taxonomic units (sOTU's) were constructed, and amplicons were trimmed to 300 bp using ‘deblur denoise-16S’. These sOTU’s represent a lower taxonomic level than OTU’s based on clustering and sequence similarity thresholds. These sOTU’s were classified as plastid sequences if their representative sequence had greater than 90% similarity to the plastid 16S rRNA gene assembled from the *L. glaberrima* metagenomes. These are hereafter referred to as plastotypes.

All plastotypes were considered present in a sample if at least 1 read was detected. However, plastotype 16 was only considered present in samples on the first sequencing run if it composed 15% or higher relative abundance. This is because plastotype 16 from *Leiopathes glaberrima* likely contaminated other samples since many other samples on the first sequencing run contained plastotype 16 at very low abundance including negative controls. See discussion for further explanation.

498 18S rRNA amplicon sequencing

499 An 860 bp section of the nuclear-encoded 18S rRNA gene was sequenced from selected
500 samples using a PCR primer pair designed to amplify apicomplexan 18S in the presence of coral
501 and *Symbiodinium* spp. [18N-F2 5'-TAGGAATCTAAACCTCTTCCA-3'; 18N-R1 5'-
502 CAGGAACAAGGGTTCCTCCGACC-3'] [21]. PCR was performed in 50 uL reactions using 3.0
503 mM MgCl₂, 0.5 µM of each primer, 0.5 mM dNTPs, 0.15 U/uL taq polymerase (Bioline,
504 Alvinston, Canada), and 1X NH₄ buffer (Bioline, Alvinston, Canada). Thermocycler conditions
505 followed the touchdown program designed by Kirk et al. [27, 28]. Amplification was verified
506 using 2% agarose gels in 1X TAE buffer, and bands at approximately 850 bp were excised. DNA
507 was extracted from gels using a gel extraction kit (Qiagen, Hilden, Germany), then Sanger
508 sequenced at the Penn State Genomics Core facility. 18S sequences were analyzed, trimmed, and
509 assembled using default parameters in CodonCode aligner (ver 3.7.1).

510 Phylogenetic trees

511 Our phylogenetic analyses included publicly available sequences from an additional 16
512 species of shallow-water scleractinians and 2 species of shallow-water octocorals. The 18S and
513 16S sequences were aligned in Geneious 11.0.3 using a MUSCLE alignment (ver 3.8.425) [68].
514 Maximum likelihood trees were constructed using iqtree (ver1.6.10) [69], model selection was
515 performed using ModelFinder [70], and 1000 bootstrap replicates were performed using
516 UFboot2 [71]. A Tamura-Nei model [72] and a Hasegawa, Kishino, and Yano model [73] were
517 selected as base substitution models for 18S sequences and 16S sequences, respectively. Both
518 models utilized empirical base frequencies and a 2-category free rate model for rate
519 heterogeneity across sites.

Trees from concatenated single gene alignments were constructed to identify the phylogenetic position of the corallicolid mitogenome and plastid genome using three genes (*cytb*, *cox1*, and *cox3*) and 5 genes (*clpC*, *rpoB*, *rpoC1*, *rpoC2*, and *tufA*), respectively. Amino acid sequences were aligned and concatenated in Geneious, an amino acid substitution model test was performed in MEGA X, and a maximum likelihood tree was constructed using the Le Gascuel model with a discrete Gamma distribution with five parameters in MEGA X.

Declarations

Ethics approval and consent to participate

Sampling deep-sea corals in the Gulf of Mexico does not require permits.

Acknowledgement letters were acquired for all research cruises from NOAA following the Magnuson-Stevens Fishery Conservation and Management Act. Collections in the Florida Keys were approved under permit FKNMS-2014-148-A2 and survey permit FKNMS-2010-130-A1 from the Florida Keys National Marine Sanctuary. Collections in Curaçao were approved under CITES permit numbers 16US784243/9 and AN001.

Consent for publication

Not applicable.

Availability of data and material

The datasets generated by this study were deposited in Genbank under accession numbers MK996254-MK996278 (plastotype sequences), MN022428-MN022459 (apicomplexan 18S sequences), MN078134 (mitochondrial genome), **MN586853 and MN586854** (plastid genome). Raw metagenomic and amplicon sequence data generated in this study is available on the National Center for Biotechnology Information (NCBI) Sequence Read Archive (SRA) under

BioProject numbers PRJNA565265 and PRJNA574146 and BioSample IDs SAMN12824622 – 4713, SAMN12824722 – 4964, SAMN12825090 - 5207, SAMN12825209 – 5269, and SAMN12856788 – SAMN12856811 <https://www.ncbi.nlm.nih.gov/sra>. These data are also publicly available through the Gulf of Mexico Research Initiative Information & Data Cooperative (GRIIDC) at <https://data.gulfresearchinitiative.org> (doi:<10.7266/RYMQTDQ9> and R4.x268.000:0126). The genomic sequence data from *Acropora palmata* that was analyzed in this study is available in the NCBI SRA database under accession numbers SRR7235979 - 80, 82 - 83, 85 - 88, SRR7236001 - 03, 07 - 12, 15, and 17 - 20.

Competing interests

The authors declare that they have no competing interests.

Funding

This is contribution number 549 from the Ecosystem Impacts of Oil and Gas Inputs to the Gulf (ECOGIG) consortium. This research was made possible by a grant from The Gulf of Mexico Research Initiative. This grant provided the funding for sample collection and 16S data analysis. Additional funding sources include the Max Planck Society which funded all metagenome sequencing generated by this study; and the National Oceanic and Atmospheric Administration (NOAA) Deep-sea Coral Research and Technology Program's Southeast Deep Coral Initiative (SEDCI); the NOAA RESTORE Science Program under award NA17NOS4510096; National Science Foundation grants OCE-1516763, OCE-1537959 and BIO-OCE-1220478; and Bureau of Ocean Energy Management (BOEM) contracts 1435-01-05-CT-39187 and M08PC20038 to TDI-Brooks which funded additional sample collection. The funders had no role in study design, analysis, decision to publish, or preparation of the manuscript.

Authors' contributions

SAV wrote the manuscript. SAV, IBB, and CRF designed the study. IBB and CRF assisted with writing and analysis. SAV, RPD, KEA, AMG, and EOO collected data and performed analyses. HGV and ND advised and assisted with data analyses. ND, IBB, and CRF acquired funding. All authors read and edited the manuscript

Acknowledgements

We would like to thank the crews and pilots of all vessels and ROVs, Dr. Erik Cordes and Dr. Santiago Herrera who served as chief scientists for cruises that made this work possible, and Dr. Peter Etnoyer for providing *Leiopathes glaberrima* samples from the West Florida Slope. We would also like to thank everyone else who helped at sea and in the field: Dr. Fanny Girard, Alexis Winnig, Alaina Weinheimer, Steve Auscavitch, Dr. Carlos Gomez, Dr. Cherisse Du Preez, Dr. Amanda Glazier, Dr. Jill Bourque, Dr. Amanda Demopoulos, Alicia Yang, Colin Bashaw, Jesse Mee. We would also like to thank Meghann Devlin-Durante for help with extractions and training undergraduates. Finally, we would like to thank Justin Wheeler for his illustration and Miles Saunders for adapting the bathymetric map.

589
590
591
592
593
594
595
596

References:

- 598 1. Roberts JM, Wheeler AJ, Freiwald A. Reefs of the deep: The biology and geology of
599 cold-water coral ecosystems. *Science* 2006;312(5773):543-547.
- 600 2. Idjadi JA, Edmunds PJ. Scleractinian corals as facilitators for other invertebrates on a
601 Caribbean reef. *Mar Ecol-Prog Ser* 2006;319:117-127.
- 602 3. Öhman MC, Rajasuriya A. Relationships between habitat structure and fish communities
603 on coral and sandstone reefs. *Environ Biol Fishes* 1998;53(1):19-31.
- 604 4. Buhl-Mortensen L, Mortensen PB. Distribution and diversity of species associated with
605 deep-sea gorgonian corals off Atlantic Canada. In: *Cold-Water Corals and Ecosystems*.
606 Edited by Freiwald A, Roberts JM. Berlin, Heidelberg: Springer Berlin Heidelberg; 2005:
607 849-879.
- 608 5. Cesar H, Burke L, Pet-Soede L. The Economics of Worldwide Coral Reef Degradation;
609 2003.
- 610 6. Bruno JF, Selig ER, Casey KS, Page CA, Willis BL, Harvell CD, Sweatman H, Melendy
611 AM. Thermal stress and coral cover as drivers of coral disease outbreaks. *PLoS Biol*
612 2007;5(6):1220-1227.
- 613 7. Hoegh-Guldberg O, Mumby PJ, Hooten AJ, Steneck RS, Greenfield P, Gomez E, Harvell
614 CD, Sale PF, Edwards AJ, Caldeira K *et al.* Coral reefs under rapid climate change and
615 ocean acidification. *Science* 2007;318(5857):1737-1742.
- 616 8. Perez FF, Fontela M, Garcia-Ibanez MI, Mercier H, Velo A, Lherminier P, Zunino P, de
617 la Paz M, Alonso-Perez F, Guallart EE *et al.* Meridional overturning circulation conveys
618 fast acidification to the deep Atlantic Ocean. *Nature* 2018;554(7693):515-518.
- 619 9. Thurber RLV, Burkepile DE, Fuchs C, Shantz AA, McMinds R, Zaneveld JR. Chronic
620 nutrient enrichment increases prevalence and severity of coral disease and bleaching.
621 *Glob Change Biol* 2014;20(2):544-554.

10. Guzmán HM, Jackson JBC, Weil E. Short-term ecological consequences of a major oil-spill on Panamanian subtidal reef corals. *Coral Reefs* 1991;10(1):1-12.
11. White HK, Hsing PY, Cho W, Shank TM, Cordes EE, Quattrini AM, Nelson RK, Camilli R, Demopoulos AWJ, German CR *et al.* Impact of the Deepwater Horizon oil spill on a deep-water coral community in the Gulf of Mexico. *Proc Natl Acad Sci U S A* 2012;109(50):20303-20308.
12. Thurber RLV, Willner-Hall D, Rodriguez-Mueller B, Desnues C, Edwards RA, Angly F, Dinsdale E, Kelly L, Rohwer F. Metagenomic analysis of stressed coral holobionts. *Environ Microbiol* 2009;11(8):2148-2163.
13. Bourne DG, Morrow KM, Webster NS. Insights into the Coral Microbiome: Underpinning the Health and Resilience of Reef Ecosystems. In: *Annual Review of Microbiology, Vol 70*. Edited by Gottesman S, vol. 70. Palo Alto: Annual Reviews; 2016: 317-340.
14. Rosado PM, Leite DCA, Duarte GAS, Chaloub RM, Jospin G, da Rocha UN, Saraiva JP, Dini-Andreote F, Eisen JA, Bourne DG *et al.* Marine probiotics: increasing coral resistance to bleaching through microbiome manipulation. *Isme J* 2019;13(4):921-936.
15. LaJeunesse TC, Parkinson JE, Gabrielson PW, Jeong HJ, Reimer JD, Voolstra CR, Santos SR. Systematic Revision of Symbiodiniaceae Highlights the Antiquity and Diversity of Coral Endosymbionts. *Curr Biol* 2018;28(16):2570-2580.
16. Moore RB, Oborník M, Janouškovec J, Chrudimsky T, Vancova M, Green DH, Wright SW, Davies NW, Bolch CJS, Heimann K *et al.* A photosynthetic alveolate closely related to apicomplexan parasites. *Nature* 2008;451(7181):959-963.
17. Cumbo VR, Baird AH, Moore RB, Negri AP, Neilan BA, Salih A, van Oppen MJH, Wang Y, Marquis CP. *Chromera velia* is Endosymbiotic in Larvae of the Reef Corals *Acropora digitifera* and *A. tenuis*. *Protist* 2013;164(2):237-244.
18. Oborník M, Modrý D, Lukeš M, Černotíková-Stříbrná E, Cihlár J, Tesařová M, Kotabova E, Vancova M, Prasil O, Lukeš J. Morphology, Ultrastructure and Life Cycle of *Vitrella brassicaformis* n. sp., n. gen., a Novel Chromerid from the Great Barrier Reef. *Protist* 2012;163(2):306-323.
19. Janouškovec J, Tikhonenkov DV, Burki F, Howe AT, Kolisko M, Mylnikov AP, Keeling PJ. Factors mediating plastid dependency and the origins of parasitism in apicomplexans and their close relatives. *Proc Natl Acad Sci U S A* 2015;112(33):10200-10207.
20. Upton SJ, Peters EC. A new and unusual species of Coccidium (Apicomplexa, Agamococcidiorida) from Caribbean scleractinian corals. *J Invertebr Pathol* 1986;47(2):184-193.
21. Toller WW, Rowan R, Knowlton N. Genetic evidence for a protozoan (phylum Apicomplexa) associated with corals of the *Montastraea annularis* species complex. *Coral Reefs* 2002;21(2):143-146.
22. Šlapeta J, Linares MC. Combined Amplicon Pyrosequencing Assays Reveal Presence of the Apicomplexan “type-N” (cf. *Gemmocystis cylindrus*) and *Chromera velia* on the Great Barrier Reef, Australia. *PLoS One* 2013;8(9):e76095.
23. Klaus JS, Janse I, Heikoop JM, Sanford RA, Fouke BW. Coral microbial communities, zooxanthellae and mucus along gradients of seawater depth and coastal pollution. *Environ Microbiol* 2007;9(5):1291-1305.

24. Janouškovec J, Horák A, Barott KL, Rohwer FL, Keeling PJ. Global analysis of plastid diversity reveals apicomplexan-related lineages in coral reefs. *Curr Biol* 2012;22(13):R518-R519.
25. Mathur V, del Campo J, Kolisko M, Keeling PJ. Global diversity and distribution of close relatives of apicomplexan parasites. *Environ Microbiol* 2018;20(8):2824-2833.
26. Kwong WK, del Campo J, Mathur V, Vermeij MJA, Keeling PJ. A widespread coral-infecting apicomplexan with chlorophyll biosynthesis genes. *Nature* 2019;568(7750):103-107.
27. Kirk NL, Ritson-Williams R, Coffroth MA, Miller MW, Fogarty ND, Santos SR. Tracking Transmission of Apicomplexan Symbionts in Diverse Caribbean Corals. *PLoS One* 2013;8(11):10.
28. Kirk NL, Thornhill DJ, Kemp DW, Fitt WK, Santos SR. Ubiquitous associations and a peak fall prevalence between apicomplexan symbionts and reef corals in Florida and the Bahamas. *Coral Reefs* 2013;32(3):847-858.
29. Janouškovec J, Horak A, Barott KL, Rohwer FL, Keeling PJ. Environmental distribution of coral-associated relatives of apicomplexan parasites. *Isme J* 2013;7(2):444-447.
30. Post AL, O'Brien PE, Beaman RJ, Riddle MJ, De Santis L. Physical controls on deep water coral communities on the George V Land slope, East Antarctica. *Antarct Sci* 2010;22(4):371-378.
31. Yesson C, Taylor ML, Tittensor DP, Davies AJ, Guinotte J, Baco A, Black J, Hall-Spencer JM, Rogers AD. Global habitat suitability of cold-water octocorals. *J Biogeogr* 2012;39(7):1278-1292.
32. Bo M, Cerrano C, Canese S, Salvati E, Angiolillo M, Santangelo G, Bavestrello G. The coral assemblages of an off-shore deep Mediterranean rocky bank (NW Sicily, Italy). *Mar Ecol-Evol Persp* 2014;35(3):332-342.
33. Deidun A, Andaloro F, Bavestrello G, Canese S, Consoli P, Micallef A, Romeo T, Bo M. First characterisation of a *Leiopathes glaberrima* (Cnidaria: Anthozoa: Antipatharia) forest in Maltese exploited fishing grounds. *Ital J Zoolog* 2015;82(2):271-280.
34. Ruiz-Ramos DV, Saunders M, Fisher CR, Baums IB. Home Bodies and Wanderers: Sympatric Lineages of the Deep-Sea Black Coral *Leiopathes glaberrima*. *PLoS One* 2015;10(10):19.
35. Hikosaka K, Watanabe Y, Tsuji N, Kita K, Kishine H, Arisue N, Palacpac NMQ, Kawazu S, Sawai H, Horii T *et al.* Divergence of the Mitochondrial Genome Structure in the Apicomplexan Parasites, *Babesia* and *Theileria*. *Mol Biol Evol* 2010;27(5):1107-1116.
36. Liu GH, Tian SQ, Cui P, Fang SF, Wang CR, Zhu XQ. The complete mitochondrial genomes of five *Eimeria* species infecting domestic rabbits. *Exp Parasitol* 2015;159:67-71.
37. Mitra A, Skrzypczak M, Ginalska K, Rowicka M. Strategies for Achieving High Sequencing Accuracy for Low Diversity Samples and Avoiding Sample Bleeding Using Illumina Platform. *PLoS One* 2015;10(4):e0120520.
38. Wright ES, Vetsigian KH. Quality filtering of Illumina index reads mitigates sample cross-talk. *BMC Genomics* 2016;17(1):876.
39. Wirshing HH, Messing CG, Douady CJ, Reed J, Stanhope MJ, Shivji MS. Molecular evidence for multiple lineages in the gorgonian family Plexauridae (Anthozoa : Octocorallia). *Mar Biol* 2005;147(2):497-508.

- 712 40. Lewis AM, Chan AN, LaJeunesse TC. New Species of Closely Related Endosymbiotic
713 Dinoflagellates in the Greater Caribbean have Niches Corresponding to Host Coral
714 Phylogeny. *J Eukaryot Microbiol* 2019;66(3):469-482.
- 715 41. Thornhill DJ, Lewis AM, Wham DC, LaJeunesse TC. Host-specialist lineages dominate
716 the adaptive radiation of reef coral endosymbionts. *Evolution* 2014;68(2):352-367.
- 717 42. C. William Birky J. The Inheritance of Genes in Mitochondria and Chloroplasts: Laws,
718 Mechanisms, and Models. *Annual Review of Genetics* 2001;35(1):125-148.
- 719 43. Furtado JM, Smith JR, Belfort R, Jr., Gattley D, Winthrop KL. Toxoplasmosis: a global
720 threat. *Journal of global infectious diseases* 2011;3(3):281-284.
- 721 44. Jones JL, Holland GN. Short Report: Annual Burden of Ocular Toxoplasmosis in the
722 United States. *Am J Trop Med Hyg* 2010;82(3):464-465.
- 723 45. Jones JL, Kruszon-Moran D, Sanders-Lewis K, Wilson M. *Toxoplasma gondii* infection
724 in the united states, 1999-2004, decline from the prior decade. *Am J Trop Med Hyg*
725 2007;77(3):405-410.
- 726 46. Montoya JG, Liesenfeld O. Toxoplasmosis. *Lancet (London, England)*
727 2004;363(9425):1965-1976.
- 728 47. Garcia GD, Gregoracci GB, Santos ED, Meirelles PM, Silva GGZ, Edwards R, Sawabe
729 T, Gotoh K, Nakamura S, Iida T *et al.* Metagenomic Analysis of Healthy and White
730 Plague-Affected *Mussismilia braziliensis* Corals. *Microb Ecol* 2013;65(4):1076-1086.
- 731 48. Fujita Y. Protochlorophyllide reduction: A key step in the greening of plants. *Plant Cell*
732 *Physiol* 1996;37(4):411-421.
- 733 49. van Dooren GG, Hapuarachchi SV. The Dark Side of the Chloroplast: Biogenesis,
734 Metabolism and Membrane Biology of the Apicoplast. In: *Secondary Endosymbioses*.
735 Edited by Hirakawa Y, vol. 84. London: Academic Press Ltd-Elsevier Science Ltd; 2017:
736 145-185.
- 737 50. Fujita Y, Bauer CE. Reconstitution of light-independent protochlorophyllide reductase
738 from purified Bch1 and BchN-BchB subunits - In vitro confirmation of nitrogenase-like
739 features of a bacteriochlorophyll biosynthesis enzyme. *J Biol Chem* 2000;275(31):23583-
740 23588.
- 741 51. Yamazaki S, Nomata J, Fujita Y. Differential operation of dual protochlorophyllide
742 reductases for chlorophyll biosynthesis in response to environmental oxygen levels in the
743 cyanobacterium *Leptolyngbya boryana*. *Plant Physiol* 2006;142(3):911-922.
- 744 52. Muraki N, Nomata J, Ebata K, Mizoguchi T, Shiba T, Tamiaki H, Kurisu G, Fujita Y. X-
745 ray crystal structure of the light-independent protochlorophyllide reductase. *Nature*
746 2010;465(7294):110-114.
- 747 53. Agostini S, Suzuki Y, Higuchi T, Casareto BE, Yoshinaga K, Nakano Y, Fujimura H.
748 Biological and chemical characteristics of the coral gastric cavity. *Coral Reefs*
749 2012;31(1):147-156.
- 750 54. MacRae JI, Dixon MWA, Dearnley MK, Chua HH, Chambers JM, Kenny S, Bottova I,
751 Tilley L, McConville MJ. Mitochondrial metabolism of sexual and asexual blood stages
752 of the malaria parasite *Plasmodium falciparum*. *BMC Biol* 2013;11:10.
- 753 55. Scheibel LW, Pflaum WK. Carbohydrate metabolism in *Plasmodium knowlesi*.
754 *Comparative Biochemistry and Physiology* 1970;37(4):543-553.
- 755 56. Röhrich RC, Englert N, Troschke K, Reichenberg A, Hintz M, Seeber F, Balconi E,
756 Aliverti A, Zanetti G, Kohler U *et al.* Reconstitution of an apicoplast-localised electron

- transfer pathway involved in the isoprenoid biosynthesis of *Plasmodium falciparum*. *FEBS Lett* 2005;579(28):6433-6438.
57. Seemann M, Bui BTS, Wolff M, Tritsch D, Campos N, Boronat A, Marquet A, Rohmer M. Isoprenoid biosynthesis through the methylerythritol phosphate pathway: The (E)-4-hydroxy-3-methylbut-2-enyl diphosphate synthase (GcpE) is a 4Fe-4S protein. *Angew Chem-Int Edit* 2002;41(22):4337-4339.
 58. Kitchen SA, Ratan A, Bedoya-Reina OC, Burhans R, Fogarty ND, Miller W, Baums IB. Genomic Variants Among Threatened *Acropora* Corals. *G3: Genes|Genomes|Genetics* 2019;9(5):1633-1646.
 59. Gruber-Vodicka HR, Seah BKB, Pruesse E. phyloFlash – Rapid SSU rRNA profiling and targeted assembly from metagenomes. *bioRxiv* 2019:521922.
 60. Li DH, Liu CM, Luo RB, Sadakane K, Lam TW. MEGAHIT: an ultra-fast single-node solution for large and complex metagenomics assembly via succinct de Bruijn graph. *Bioinformatics (Oxford, England)* 2015;31(10):1674-1676.
 61. Altschul SF, Gish W, Miller W, Myers EW, Lipman DJ. Basic local alignment search tool. *Journal of Molecular Biology* 1990;215(3):403-410.
 62. Wick RR, Schultz MB, Zobel J, Holt KE. Bandage: interactive visualization of de novo genome assemblies. *Bioinformatics (Oxford, England)* 2015;31(20):3350-3352.
 63. Tillich M, Lehwarck P, Pellizzer T, Ulbricht-Jones ES, Fischer A, Bock R, Greiner S. GeSeq - versatile and accurate annotation of organelle genomes. *Nucleic Acids Res* 2017;45(W1):W6-W11.
 64. Lowe TM, Eddy SR. tRNAscan-SE: A program for improved detection of transfer RNA genes in genomic sequence. *Nucleic Acids Res* 1997;25(5):955-964.
 65. Durante MK, Baums IB, Williams DE, Vohsen S, Kemp DW. What drives phenotypic divergence among coral clonemates of *Acropora palmata*? *Mol Ecol* 2019;28(13):3208-3224.
 66. Rodriguez-Lanetty M, Granados-Cifuentes C, Barberan A, Bellantuono AJ, Bastidas C. Ecological Inferences from a deep screening of the Complex Bacterial Consortia associated with the coral, *Porites astreoides*. *Mol Ecol* 2013;22(16):4349-4362.
 67. Bolyen E, Rideout JR, Dillon MR, Bokulich NA, Abnet C, Al-Ghalith GA, Alexander H, Alm EJ, Arumugam M, Asnicar F *et al.* QIIME 2: Reproducible, interactive, scalable, and extensible microbiome data science. *PeerJ Preprints* 2018;6:e27295v27292.
 68. Edgar RC. MUSCLE: multiple sequence alignment with high accuracy and high throughput. *Nucleic Acids Res* 2004;32(5):1792-1797.
 69. Nguyen L-T, Schmidt HA, von Haeseler A, Minh BQ. IQ-TREE: A Fast and Effective Stochastic Algorithm for Estimating Maximum-Likelihood Phylogenies. *Mol Biol Evol* 2014;32(1):268-274.
 70. Kalyaanamoorthy S, Minh BQ, Wong TKF, von Haeseler A, Jermiin LS. ModelFinder: fast model selection for accurate phylogenetic estimates. *Nature Methods* 2017;14:587.
 71. Hoang DT, Chernomor O, von Haeseler A, Minh BQ, Vinh LS. UFBoot2: Improving the Ultrafast Bootstrap Approximation. *Mol Biol Evol* 2017;35(2):518-522.
 72. Tamura K, Nei M. Estimation of the number of nucleotide substitutions in the control region of mitochondrial DNA in humans and chimpanzees. *Mol Biol Evol* 1993;10(3):512-526.
 73. Hasegawa M, Kishino H, Yano T-a. Dating of the human-ape splitting by a molecular clock of mitochondrial DNA. *Journal of Molecular Evolution* 1985;22(2):160-174.

74. Kayal E, Bentlage B, Pankey MS, Ohdera AH, Medina M, Plachetzki DC, Collins AG, Ryan JF. Phylogenomics provides a robust topology of the major cnidarian lineages and insights on the origins of key organismal traits. *BMC Evol Biol* 2018;18:18.
75. Zapata F, Goetz FE, Smith SA, Howison M, Siebert S, Church SH, Sanders SM, Ames CL, McFadden CS, France SC *et al.* Phylogenomic Analyses Support Traditional Relationships within Cnidaria. *PLoS One* 2015;10(10):e0139068.

Fig 1. The genomic compartments of corallicolid apicomplexans, the possible diversity in plastid and nuclear compositions, and the phylogeny of host and corallicolid genomic compartments. - PS denotes the loss of photosynthesis. Cnidarian phylogeny based on [74, 75]. Art by Justin Wheeler.

Fig 2. Concatenated tree of (a) mitochondrial genes: *cytb*, *cox1*, and *cox3* and (b) plastid genes: *clpC*, *rpoB*, *rpoC1*, *rpoC2*, and *tufA*. Sequences generated in this study are highlighted in yellow.

Fig 3. Annotation of the mitogenome (a) and plastid genome (b) of a corallicolid from *Leiopathes glaberrima*. Radial lines separate two scaffolds of the plastid genome. (c) Detection of corallicolids in the metagenomes of various coral species using mitogenomes, plastid genomes, and selected genes. Each column represents a separate coral colony. US VI = U.S. Virgin Islands.

Fig 4. Detection of apicomplexan plastids using 16S rRNA by site across a wide depth range (a) and the depth ranges of each corallicolid plastotype (b). Numbers above the boxes in (a) represent the number of coral species sampled at each site while the numbers underneath denote the total number of colonies sampled. Numbers above boxes in (b) are plastotype identifications while the numbers underneath denote the number of sites where each plastotype was detected.

Fig 5. Map of coral sampling locations where corallicolid plastids were detected using 16S sequences in the northern Gulf of Mexico (a), Florida Keys (b), and Curaçao (c). Coral species that hosted corallicolids are listed for each site using abbreviations described in Table S2 in Additional File 1. Numbers in parentheses next to species abbreviations denote the number of colonies screened. Each 16S plastotype is denoted by a letter and unique color and/or pattern. Numbers in the tables denote the number of colonies of each species in which each plastotype was detected. Bathymetric map of the Gulf of Mexico is available from U.S. Bureau of Ocean Energy Management Gulf of Mexico Deepwater Bathymetry Grid. *For colonies that were sampled more than once, a plastotype was considered detected if it was found in at least one sample from that colony.

852

853 Fig 6. Site and host patterns in the relative abundances of plastotypes. Each bar represents the
854 relative abundances of each plastotype among all plastid sequences found in an individual coral
855 colony.

856

857 Fig 7. Phylogenetic relationships among corallicolids using both nuclear (a) and plastid (b)
858 markers. Sequences are labeled by host coral species and are designated shallow-water (S),
859 mesophotic (M), or deep-sea (D). Corallicolid plastotypes based on 16S are labeled 1-23.
860 Sequences obtained in this study are highlighted in yellow. Circles represent nodes with greater
861 than 95% ultrafast bootstrap support. Alongside each 18S sequence, the relative abundances of
862 all plastotypes obtained from the same sample are displayed in barplots. For each sample, the
863 total relative abundance of apicomplexan plastotypes among all 16S sequences is denoted on the
864 left of each bar. Lines connect each bar to the 16S plastotypes present within it. Outgroups are
865 not shown. See Fig S2 in Additional File 1 for outgroups.

866

867 **Additional Files**

868 Additional File 1

869 File format: pdf

870 Title: Supplemental Information

871 Description: Additional text, figures, and tables with sampling locations, species abbreviations,
872 description of the potential contamination of plastotype 16, and a comparison of extraction
873 procedures on apicomplexan detection.

874

875 Additional File 2

876 File format: Excel Workbook

877 Title: 16S rRNA amplicon screening data

878 Description: This dataset includes the read counts for all plastotype sequences for all samples.

879

880 Additional File 3

881 File format: Excel Workbook

882 Title: Metagenomic screening data

883 Description: This dataset includes all mapping results of screening the metagenomes for

884 apicomplexans.

885

886 Additional File 4

887 File format: FASTA

888 Title: Plastotype sequences

889 Description: These are the plastotype sequences that were uploaded to NCBI's GenBank. This

890 file is for the reviewers since this data will become publicly available on NCBI when published.

891 A reviewer link for this submission could not be obtained.

892

893 Additional File 5

894 File format: FASTA

895 Title: Apicomplexan (genotype N) 18S rRNA sequences

896 Description: These are the apicomplexan 18S rRNA sequences that were uploaded to NCBI's

897 GenBank. This file is for the reviewers since this data will become publicly available on NCBI

898 when published. A reviewer link for this submission could not be obtained.

899

900 Additional File 6

901 File format: GenBank (text file)

902 Title: The mitochondrial genome of the Apicomplexan associated with *Leiopathes glaberrima*

903 Description: This is the mitogenome of the Apicomplexan that associates with *Leiopathes*

904 *glaberrima*. This is the GenBank file that includes the DNA sequence, the CDS, and amino acid

905 translations of all genes. This file is for the reviewers since this data will become publicly

906 available on NCBI when published. A reviewer link for this submission could not be obtained.

907

908 Additional File 7

909 File format: FASTA

910 Title: The plastid genome of the Apicomplexan associated with *Leiopathes glaberrima*

911 Description: This is the DNA sequence of the plastid genome of the Apicomplexan that
912 associates with *Leiopathes glaberrima*. This file is for the reviewers since this data will become
913 publicly available on NCBI when published. A reviewer link for this submission could not be
914 obtained.

915

916 Additional File 8

917 File format: FASTA

918 Title: The coding regions of the plastid genome of the Apicomplexan associated with *Leiopathes*
919 *glaberrima*

920 Description: This file contains the amino acid sequences of the coding regions of the plastid
921 genome of the Apicomplexan that associates with *Leiopathes glaberrima*. This file is for the
922 reviewers since this data will become publicly available on NCBI when published. A reviewer
923 link for this submission could not be obtained.

924