Diversity and flexibility of algal endosymbionts in globally-distributed large benthic foraminifera

Martina Prazeres1,

martina.prazeres@naturalis.nl (Corresponding author)

T. Edward Roberts1,

thomas.roberts@naturalis.nl

Shadrina Fildzah Ramadhani1,

sframadhani94@gmail.com

Steve S. Doo2,3,

steve.doo@leibniz-zmt.de

Christiane Schmidt4,5,

cschmidt@marum.de

Marleen Stuhr2,6,7,

marleen.stuhr@leibniz-zmt.de

Willem Renema1

willem.renema@naturalis.nl

1 Naturalis Biodiversity Center, Leiden, Netherlands

2 Leibniz Centre for Tropical Marine Research (ZMT), Bremen, Germany

3 Department of Biology, California State University, Northridge, United States of America

4 MARUM, University of Bremen, Bremen, Germany

5 University of the Ryukyus, Nishihara, Okinawa, Japan

6 Interuniversity Institute for Marine Sciences (IUI), Eilat, Israel

7 Bar-Ilan University (BIU), Ramat Gan, Israel
Abstract

Background
Revealing the specificity and flexibility of the algal symbiont-host association is fundamental for understanding how species can occupy a diverse range of habitats. Here we assessed the global distribution of the algal symbiont diversity for three shallow-water species of large benthic foraminifera (LBF) of the genus *Amphistegina*. Specifically, we investigated the role of habitat and host identity on the diversity of algal biome.

Results
Here we used next-generation sequencing to identify the algal biome in the species *A. lobifera*, *A. lessonii*, and *A. radiata*, collected from shallow habitats in 16 sites, spanning from the Mediterranean Sea to French Polynesia. Results showed the consistent presence of *Fragilariales* as the main algal taxa associated with all species across sites analysed. However, we uncovered unprecedented diversity of algal phylotypes found in low abundance. We further found high variability in algal biomes both within and between species and sites, indicating a substantial level of flexibility in symbiont associations. The effect of site was significant in all species analysed, and showed that local habitat is the main factor influencing the identity of algal symbionts. Symbiotic associations are also not constrained by the species identity nor the phylogenetic relationship among closely related hosts, suggesting symbiont identity plays a limited role in the evolutionary history of the genus *Amphistegina*.

Conclusions
We found that species of *Amphistegina* form flexible symbiotic relationship with algal taxa, which are primarily shaped by their local habitat. These observations strongly suggest that the capacity of *Amphistegina* species to utilise a diverse array of available symbionts likely underpins the ecological success of these crucial calcifying organisms across their extensive geographic range.

**Key words:** microbiome, phylogeography, symbiosis, photosymbionts, coral reefs, Lessepsian migrants, *Amphistegina.*
Background

Algal symbiosis is at the foundation for coral reef ecosystems. It plays an important role in facilitating organisms to adapt or acclimate to environmental change, and to colonise new habitats [1]. The algal endosymbionts allow hosts to exploit light as an energy source in oligotrophic conditions [2], as symbiont communities can often confer distinct physiological capacities allowing hosts to colonise a specific geographic region or habitat (e.g. [3]). For example, the symbiosis between reef-building corals and dinoflagellates is crucial for the persistence of coral reefs [4], but makes them inherently vulnerable to environmental change. However, the capacity of hosts to utilise a diverse pool of symbionts may alleviate this vulnerability and provide hosts with the capacity to acclimate to ongoing ocean warming (e.g. [5, 6]). Symbiosis is influenced by complex interactions between the host and the local environment [7], which shapes the fitness of the holobiont (i.e. host-symbiont complex) [8].

Of marine species which host algal symbionts, large benthic foraminifera (LBF) are the most common and abundant organisms [9], and are known to have obligatory symbiosis with multiple groups of algae, such as rhodophytes, chlorophytes, diatoms, and dinoflagellates [10]. The ecological advantage of maintaining algal symbiosis is evident in the recurrent emergence of symbiosis in foraminifera over the past 350 My, despite repeated extinction events of symbiotic species [11]. Symbiosis has driven morphological differentiation and speciation of symbiont-bearing species along depth gradients (e.g.[12, 13]), but also ‘horizontally’ across trophic gradients (e.g. [14]). For example, the depth distribution of diatom-bearing species is viewed as indicative of their adaptive potential to a wide light intensity and spectrum [15, 16], whereas other LBF species harbouring symbionts other than diatoms are largely restricted to a small portion of the available light spectrum (reviewed in [17]).
Symbiont diversity is also closely linked to host identity and phylogeny. It has been shown that in diatom-bearing LBF *Amphistegina*, the similarities and differences in lineages of endosymbionts of four closely related species are consistent with what is known of their evolutionary histories [18]. Similarly, dinoflagellate symbionts found in the Caribbean and the Indo-Pacific show phylogenetic divergence, which is consistent with the phylogenetic relationship within their LBF hosts in these two regions [19]. Some species of *Amphistegina* show a stable and persistent algal symbiosis unaffected by water quality gradients [20], while the composition of algal symbionts in diatom-bearing nummulitids changes over depth [21]. The diversity of symbionts might also play a key role in thermal stress tolerances [22], and potentially facilitates geographic range expansion in response to ocean warming [23, 24]. The presence of a consortium of diverse algal species that can be functionally relevant within different environmental conditions may include thermo-tolerant genotypes or species [20, 25]. Similar patterns of changes in algal symbiont consortium can also be found in some species of dinoflagellate-bearing LBF, where mixed infections are common [7, 19]. Yet it remains unclear whether symbiont community is driven by host identity or habitat (including physicochemical conditions), or a combination of both, and to what extent the diversity of associated symbionts allows LBF to respond to changes in environmental conditions and expand their distribution range across shallow platforms worldwide.

Here, we examined the diversity of algal symbionts (also referred to as algal biome) in *Amphistegina lobifera*, *A. lessonii* and *A. radiata* within shallow habitats across the Mediterranean, Red Sea, Indian and Pacific Oceans. Specifically, we investigate the influence of host identity and local habitat on the composition of their symbiont communities. We address the level of symbiont specificity (i.e. algal taxa that were only found in a given species of *Amphistegina*) between the different host species, and evaluate whether these host-
symbiont associations are partitioned between geographic regions and sites. We further evaluate whether symbiont phylogenies are consistent with the evolutionary histories of host species.

Results

Phylogenetic analysis

Analysis of 18S rRNA of algal biome showed that specimens of *Amphistegina* predominately host diatoms of the order *Fragilariaceae*, which was consistently abundant across sites and species. On average, *Fragilariaceae* represented over 60% of identified diatoms. Diatoms from the order *Toxariales* (class *Mediophyceae*) were found to be abundant within specimens of *A. radiata* collected from Micronesia (47.85 ± 6.65% of the identified taxa), which also showed the lowest relative abundance of *Fragilariaceae* (42.01 ± 8.98%; Fig. 1a). However, *Toxariales* was rare in all specimens of *A. lessonii* and absent in *A. lobifera*. Within *Fragilariaceae*, the most common genus was *Serratifera*, followed by *Nanofrustulum* and *Staurosira* (Fig. 1b). Other groups such as *Bacillariophyceae* were found to represent a substantial proportion of the diatoms in *A. lobifera*, up to an average of 7.20 ± 3.70% in Sicily (Fig. 1a; Supplementary table 1). In contrast, *A. lessonii* was dominated by *Fragilariaceae*, and this order represented nearly 100% of the algal assemblage in Zanzibar and Kimbe Bay (Fig. 1a).

Prevalence of amplicon sequence variants (ASV) was less than 70% among samples and sample groups (Fig. 2). Within taxa, *Fragilariaceae* was the most common order, and *Serratifera* represented a substantial proportion of taxa found across species and sites (Fig. 2). However, most ASVs were found in low relative abundance in only a few samples. Phylogenetic analysis showed that there is no clear congruency between symbiont identity and host relationship (Fig. 3), nonetheless it is worth noting that alpha diversity in *A. lessonii*
is consistently lower than in *A. lobifera* and *A. radiata* (Fig. 4). There was also a significant
difference in alpha diversity among sites and host species (Table 1). The highest average
diversity of ASVs was found in *A. radiata* samples collected from Kimbe Bay, whereas the
lowest diversity was consistently found in *A. lessonii* specimens (Fig. 4; Supplementary table
2). *Amphistegina lessonii* was found to host a few phylotypes of diatoms of the genus
*Nanofrustulum*, in addition to the common genus *Serratifera* (Fig. 3). Additionally, we did
not identify a core algal biome utilising the 80% cut-off across species and sites analysed.

*Multivariate analysis of diversity of diatoms among species and sites*

There was a significant difference between the algal community of *A. lessonii*, *A. lobifera*,
and *A. radiata* among sites (PERMANOVA$_{\text{site x species}}$: $R^2 = 0.199$, Pseudo-$F=8.92$; $p<0.01$;
Table 2). However, group dispersions analysis also revealed that homogeneity of variances
was uneven, indicating that reported $p$-values should be treated with a high degree of caution.
Species provided little overall explanatory value ($R^2 = 0.09$; Pseudo-$F=16.84$; Table 2),
which is reflected in the extensive overlap of algal biomes found in each of the three species
(Fig. 5a). Algal biomes of *A. radiata* were more constrained. Site provided the highest overall
explanatory value ($R^2 = 0.31$; Pseudo-$F=7.58$; Table 2) and was also significant within each
species (*A. lobifera* $R^2 = 0.53$, *A. lessonii* $R^2 = 0.66$, *A. radiata* $R^2 = 0.62$; Table 2). Despite
algal biomes being more distinct between sites than between species, variability within sites
was highly uneven (Figs. 5b-q). Sites contained algal biomes with low variability shared by
all species (i.e. Lord Howe Island; Fig. 5j), low variability with discrete species-specific
communities (i.e. Eilat, Palau, Zanzibar; Figs. 5e,o,z), and high variability regardless of
species identity (i.e. Micronesia, Ningaloo reef; Figs. 5m,n).

**Discussion**
Algal symbiosis plasticity can help host organisms respond to changes in environmental conditions (e.g. [26]), and, as a consequence, allow them to expand their realised niche. The ability to change endosymbiont composition or acquire different types of algal symbionts is therefore an important mechanism by which hosts may survive stress, such as ocean warming. Our results show that symbiont flexibility in Amphistegina may underpin their wide distribution range and adaptation capacity [24, 27]. However, the similarity and differences in endosymbiont associations of the three species of Amphistegina reflect their ecological niche and is determined by the species’ local microhabitat, rather than their evolutionary histories. Instead, algal biomes of symbiotic LBF is shaped by the site at which the individual occurs and is not constrained by species. This association between site and algal biome is not unexpected (e.g. [3]), but our results reveal an unexpectedly high level of variability in the algal communities within sites, especially in A. lobifera. Each site is subject to a unique set of environmental conditions, which dictate the performance of differing algal symbiont types, and consequently the fitness of the host LBF [20, 22]. Although some sites contained a limited range of algal taxa (e.g. French Polynesia and Lord Howe Island; Figs. 5b,j) the majority supported a diverse algal biome (> 10 algal phylotypes).

Our results show that the species within the genus Amphistegina are able to host a wide array of diatoms as symbionts, but each has only a single dominant taxon. We found that the most common and abundant symbionts in Amphistegina belong to the order Fragilariales, confirming and expanding on previous results on localised populations [20] and 2000 + isolations of diatoms in culture (reviewed in [28]). The consistent presence of Fragilariales suggests that Amphistegina likely evolved with a preference for this group of diatoms, although the origin of this association remains unclear. Species of the genus Serratifera were
found in every specimen and at all sites, however the overall diversity of all diatom phylotypes within species and specimens is higher than previously thought (Fig. 3; [21, 29]).

Interestingly, *A. lobifera* and *A. radiata* showed a higher alpha diversity and variability in algal biome, while *A. lessonii* showed a less diverse and more variable symbiont community (Figs. 3, 4), featuring phylotypes of the genera *Serratifera* and *Nanofrustulum*. Despite the low number of specimens collected in our study, *A. radiata* showed the highest alpha diversity among species analysed, and therefore patterns of diversity are unlikely to be an artefact of sampling effort. Previous studies have also shown low levels of diversity in *A. gibbosa*, which is a species restricted to the Atlantic Ocean. This species has shown to be associated with a single sequence type [22] or very low symbiont diversity [18]. However, the limited geographic distribution of *A. gibbosa* and the small spatial extent of the study conducted by Stuhr et al. [22] might have contributed to the reduced number of symbionts found. For instance, the Caribbean Sea is far smaller than the Indo-Pacific Ocean and supports a reduced diversity of LBF and *Amphistegina* [30], which is only represented in this region by the single species *A. gibbosa*. In the Caribbean, the diversity of symbionts available is also likely to be low.

Acquisition of a more diverse array of symbionts may offer a route for hosts to colonise different ecological niches and habitats, as is the case with *A. radiata*. *Amphistegina radiata* tend to be more abundant at greater depths and common within reef rubble [13, 31]. *Amphistegina lobifera* is constrained to shallow habitats [13, 31], but has an exceptional capacity of extend its distributional range [23]. It is also suggested that *A. radiata* belongs to a separate lineage that evolved independently from *A. lessonii* and *A. lobifera* [32]. We found that *A. radiata* specimens host phylotypes belonging to the diatom order *Toxariales* (class
Mediophyceae; Fig. 1), which was also found in very low abundances in A. lessonii. Previous barcoding analysis also showed that A. radiata has preference for diatoms other than Fragilariales [18]. It seems that the preference for Fragilariales or Toxariales (Figs. 3, 5) is consistent with morphological adaptation to light and habitat preference between the two groups (A. lessonii-A. lobifera and A. radiata; [32, 33]. Despite the presence of Toxariales in A. radiata specimens, all Amphistegina specimens analysed in our study predominantly host algae belonging to Fragilariales. As a result, species identity informed only 9% of variation (Table 2). Increasing the number of A. radiata specimens and geographic regions covered will help elucidate whether the presence of Toxariales is consistent across sites or driven by species identity.

Through the assessment of algal biome diversity in these three species of the genus Amphistegina, we demonstrate that symbiont communities are dictated to a certain degree by site but are not constrained by species identity or phylogenetic relationships. While patterns of alpha diversity are partly informed by species identity, levels of flexibility are mostly shaped by site. This confirms the role of the symbiont community as an important interface between the host and the local environment (e.g. [22, 25]). Symbiont communities respond differently to differing conditions, and the high variability within sites reveals that a wide array of symbiont communities remains viable within most sites (Fig. 7). However, the algal biome is more constrained in some sites than others. For example, A. lobifera populations that occur at the edge of their geographic distribution tend to have a highly variable algal biome, with high variability between specimens from the same site (e.g. Greece, Ningaloo Reef, Sicily; Fig. 7d, n, p), whereas in the core of their distribution a consistent algal biome across specimens is more common (e.g. Great Barrier Reef, Indonesia, Kimbe Bay; Fig. 7c, g, i). This allows speculations that novel invading species are at an advantage to increased
environmental tolerance given by a pool of different endosymbionts. In contrast, *A. lessonii*
not only showed lower alpha diversity of algal phylotypes than other *Amphistegina* species,
but also a more conserved algal biome among the sites analysed. As a result, we were unable
to find a universal core algal biome across all *Amphistegina* species and sites analysed, and
only a local-scale species-specific core biome was detected, further supporting the hypothesis
that the composition of the biome is largely driven by site [20]. Ultimately, our results
suggest that local microhabitat, and the environmental factors associated with it, are likely to
impose the strongest influence on both the algal phylotype available, and how hosts acquire
their algal symbionts.

The ability to acquire a wide array of algal taxa (i.e. flexibility) or constraints in algal
acquisition (i.e. specificity) appear to vary according to the taxonomic scale being analysed.
Other diatom-bearing genera are hosts to diatoms of families other than *Fragilariales*. For
example, *Pararotalia calcariformata* primarily hosts *Minutocellus polymorphus* [34], which
belongs to the family *Cymatosirales*. Whereas nummulitids such as *Heterostegina*,
*Cycloclypeus*, and *Nummulites* host diatoms belonging to the family *Thalassionematales*
[21]. A similar pattern is also found in dinoflagellate-bearing species. The majority of
dinoflagellate-bearing genera consistently retain a specific symbiont group [35]. Conversely,
analysis of algal biomes along a natural environmental gradient showed that the
dinoflagellate-bearing *Marginopora vertebralis* has highly flexible symbiosis at species level
[36]. Similar to our results, these different populations of *M. vertebralis* select their algal
symbionts according to their local environment. This means that specificity may be more
prevalent at higher taxonomic levels (i.e., class to family), and increasingly flexible as
taxonomic scale decreases (i.e., genus and species).
Another possibility to consider is that the rare (i.e. less abundant diatom taxa and other algal groups found in low abundance) are used as food by the hosts. Those could be retained in the cytoplasm of the host and show up in the sequences despite being functionally irrelevant for the symbiont pool. It has been demonstrated that *Amphistegina* rely on the photosynthesis for most of its energy requirements [37]. However, *Amphistegina* is known to utilise heterotrophic feeding on algae and bacteria for nutrient acquisition [38]. Species within the algal biome found in low abundance (between 1 and 5%) detected in our study, and in previous culturing studies (reviewed in [39]), could play an important role as associates, but they are likely to be used as food as opposed to as to be primary endosymbionts [18].

The analysis of our global-scale dataset show that the nature of the symbiont community is primarily shaped by local habitat, and to a lower degree the host’s identity. The ability of *Amphistegina* to acquire a wide range of diatom species might underlie their ubiquitous presence throughout the Indo-Pacific and Red Sea, and most recently their successful invasion of the Mediterranean Sea [40]. We suggest that other LBF species that have similar plasticity in symbiont assemblages are likely to be similarly influenced by local scale factors as opposed to host identity. In addition to symbiont plasticity, *Amphistegina* can shift their life history strategy from asexual division towards an increased dependence on sexual generations, allowing horizontal transmission of symbionts best suited to their environment [41]. Hence, *Amphistegina* populations have the capacity to respond to shifts in environmental conditions and occupy a wide range of habitats, making them well-adapted to cope with ongoing climate change.

**Conclusions**
We show that the three most common shallow-water species of *Amphistegina* are able to establish symbiosis with a wide range of algal endosymbionts, suggesting that while the presence of an algal biome is crucial to the host, the identity of species within the biome is not. Diversity of symbionts is higher in *A. lobifera* and *A. radiata* than in *A. lessonii*, and subtle differences in diversity might be correlated with the evolutionary history of species through niche partitioning. Patterns of diversity can be associated with species identity, in contrast to host-symbiont specificity. We demonstrated that *A. lobifera*, *A. lessonii* and *A. radiata* form flexible associations with their algal endosymbionts, and that site (likely driven by environmental conditions) rather than species identity informs the composition of algal biomes. Further studies should investigate whether flexibility of the host-symbiont system is also prevalent in other LBF species across their distribution range, and whether host species actively select their algal biome.

**Material and methods**

**Study sites and collection of samples**

Live specimens of *Amphistegina lobifera*, *A. lessonii*, and *A. radiata* (Fig. 1) were collected across a broad geographic range spanning the Mediterranean Sea, Red Sea, Indian Ocean, and Pacific Ocean. In total, 16 reef sites were selected (Fig. 2; Table 1). These sites encompass a wide range of environmental conditions and the known distribution of these species [42]. *Amphistegina lobifera* and *A. lessonii* occur in similar habitats and occasionally occupy the same niche. *Amphistegina lobifera* frequently occupy shallow areas (0-12 m), whereas *A. lessonii* can be found as deep as 50 m [13]. Despite this habitat preference, both species show an overlap in distribution that is not generally defined and depends on local geography and environmental conditions of sites. In contrast, *A. radiata* is usually found in deeper areas of the reef from 30 to 90 m [13, 31], occasionally being found in shallow
regions on the reef slope. For this study, samples were collected from shallow areas of the reef slope (<10 m water depth) by snorkelers or SCUBA divers following previously described methods [27]. Briefly, pieces of reef rubble containing the targeted species were collected, scrubbed, and specimens picked out and placed in 96% ethanol or air-dried for further analysis. All specimens were collected from the same rubble sample.

Samples processing and DNA extraction

Between twelve and four specimens per species per site were selected (n = 206, Table 1). In the laboratory, specimens were cleaned with 96% molecular grade ethanol under a stereomicroscope, and individual photos were taken utilising a stacking microscope (Zeiss SteREO Discovery V12). Individuals were subsequently placed in tubes containing 96% molecular grade ethanol for further wash and removal of any contamination on the shell. Individuals were air dried, then placed in individual tubes containing 200 µl of lysis buffer with Proteinase K. DNA extractions were conducted using the QIAamp® DNA Micro kit (Qiagen, Germany) according to manufacturer’s instructions. DNA concentration was measured using the DropSense96 platform (Trinean, Belgium). Total DNA concentration was standardised to 2 ng per µl of DNA across all samples.

Library preparation, next-generation sequencing, and sequences analysis

To test for differences in algal microbiome between reef sites and species, we amplified the hypervariable V4 region of the 18S SSU rRNA [43] utilising the universal 18S primers TAREuk454FWD (5′- CCAGCASCYCGGTATTCC -3′) and TAREukREV3 (5′- ACTTTGCCTTGTAGTYRA -3′) [44] with Nextera™ tags (Illumina, USA). All reactions were performed in 20 µl volumes containing 1x TaqMan™ Environmental Master Mix 2.0 (ThermoFisher, USA), 100 pmol of each primer, and approximately 6 ng of template DNA.
This hypervariable region was amplified utilising the following conditions: initial
denaturation at 95 °C for 10 min, followed by 40 cycles of 30 s denaturation at 95 °C,
annealing at 52 °C for 45 s, and final extension at 72 °C for 1 min, and ended with a final
extension at 72 °C for 5 min. Libraries were visualised by a 1% agarose gel electrophoresis
stained with ethidium bromide. In total, amplification was successful for 175 specimens.
Positive libraries were purified using a magnetic-beads based NucleoMag® NGS clean-up kit
following manufacturer’s instructions manual. Each purified library was then barcoded with
unique Nextera™ labels in a second PCR reaction as follows: initial denaturation at 95 °C for
10 min, followed by 30 cycles of 30 s denaturation at 95 °C, annealing at 52 °C for 45 s, and
final extension at 72 °C for 1 min, and ended with a final extension at 72 °C for 5 min. Size
distribution of libraries were checked using the capillary electrophoresis in QIAxcel (Qiagen,
Germany). Afterwards, libraries were normalised and pooled using the QIAgility system
(Qiagen, Germany). Finally, the quality of the library products was assessed and standardised
on a Bioanalyzer 2100 (Agilent) using a High Sensitivity DNA Chip. Libraries were
sequenced utilising the Illumina MiSeq platform using the 2 × 300 bp paired-end protocol
yielding paired-end reads that overlap almost completely. Sequencing was conducted by
BaseClear (Leiden, Netherlands). Two negative control samples were used to monitor any
contamination during DNA extraction and PCR amplifications, however no quantifiable
DNA was detected for further analysis. A single negative control containing 18.2 Ω MilliQ
H₂O was used during library preparation and sequenced. The obtained .fasta files containing
all amplicon sequences including the negative control sample were deposited to NCBI under
the accession number TBA. Sequence data were processed using the statistical program R
v3.6.1 [45], using the DADA2 workflow described in detail by Callahan et al. [46, 47].
Briefly, forward and reverse sequences lacking adaptors and primer sequences were checked
for quality, trimmed, and filtered to remove low-quality sequence reads. Quality score cut-off
point was determined based on quality of both forward and reverse sequence reads, maintaining the recommended overlap for merging the sequences. The DADA2 method was utilised for barcoding filtering, de-replication, chimeric identification and removal, and merging pair-end reads. DADA2’s error model automatically filters out singletons, removing them before the subsequent sample inference step. Sample inference was performed using the inferred error model. Afterwards, an ASV table was constructed, which is an analogue of the traditional Operational Taxonomic Unit (OTU). A total of 10,425 ASVs was identified after chimera removal. Sequences were then aligned, and OTUs defined at 99% similarity against the curated 18S SILVA v132 database [48]. Any sequences that were not assigned at phylum level were filtered out of the dataset. Phylogenetic tree was constructed using the inferred ASV table without chimeras. A multiple-alignment was performed using the decipher package [49] in R. Subsequently the phylogenetic tree was constructed by first building a neighbour-joining tree, and then using this tree as a starting point to fit a GTR+G+I (Generalised time-reversible with Gamma rate variation) maximum likelihood tree using the phangorn [50] package in R.

Statistical analyses

Statistical analyses and graphical representations were performed in R v.3.6.1 [45]. Differences in algal biomes associated with A. lessonii, A. lobifera, and A. radiata specimens collected from different reef sites were analysed using the packages phyloseq [51], vegan [52], and microbiome [53]. For this purpose, only ASVs classified as Ochrophyta (i.e. diatoms, but also includes phaeophyte) at Phylum level were retain in our dataset for further analyses. Negative control sample was also removed from the final dataset. As a result, we ended up with a total of 3,632 ASVs classified as ‘Ochrophyta’. Prior to the analyses, relative abundance was calculated, and only ASVs present in at least 1% summed across all samples
were retained to minimise the influence of rare and incidental ASVs. The final dataset contained a total of 892 ASVs. The phylogenetic tree was subsequently visualised using the online Interactive Tree of Life v5.3 [54, 55] utilising this dataset.

We calculated diversity indices such as Chao 1 [56], which we used to project estimates of taxonomic richness within each specimen (i.e., alpha diversity), and Simpson index that combines evenness and richness of a given specimen [57]. Indices were calculated using ASVs. We compared significant differences in diversity indices among species and sites by performing a rank sum Kruskal-Wallis test. Prevalence, which is the percentage of specimens where a given ASVs is detected, was calculated only for ASVs classified as ‘Fragilariales’.

Differences within and between species and sites were analysed through a two-way Permutation Multivariate ANOVA (PERMANOVA) using weighted-UniFrac resemblance matrices [58] to account for presence/absence, but also abundance of ASVs between samples. ‘Site’ and ‘Species’ were employed as fixed factors. PERMANOVA outcomes were based on 1,000 permutations using Type I Sums of Squares, and permutation of residuals under reduced model. PERMANOVA was performed using the function adonis2 in the vegan package. Homogeneity of multivariate dispersions was confirmed for the fixed factors ‘Site’ and ‘Species’ using the permutational test betadisper in the package vegan, to confirm that PERMANOVA results were not due to differences in group dispersions, but due to differences in algal community. An unconstrained Principal Coordination Analysis (PCoA) was used as a visual representation of the compositional differences among algal community associated with Amphistegina populations from different collection sites, using the weighted-UniFrac distance matrix. Finally, to identify the stable, consistent algal taxa present in Amphistegina specimens collected from different sites, a core algal biome was defined as
ASVs present in at least 80% of the samples, and analysis conducted using the function `core`
in the package `microbiome`.

References

1. Baker, A.C., Flexibility and specificity in coral-algal symbiosis: diversity, ecology, and biogeography of Symbiodinium. Annual Review of Ecology, Evolution, and Systematics, 2003. 34(1): p. 661-689.
2. Norris, R.D., Symbiosis as an evolutionary innovation in the radiation of Paleocene planktic foraminifera. Paleobiology, 1996. 22(4): p. 461-480.
3. Bongaerts, P., et al., Sharing the slope: depth partitioning of agaricid corals and associated Symbiodinium across shallow and mesophotic habitats (2-60 m) on a Caribbean reef. Bmc Evolutionary Biology, 2013. 13.
4. Muscatine, L. and J.W. Porter, Reef corals - Mutualistic symbioses adapted to nutrient-poor environments. Bioscience, 1977. 27(7): p. 454-460.
5. Chakravarti, L.J. and M.J.H. van Oppen, Experimental evolution in coral photosymbionts as a tool to increase thermal tolerance. Frontiers in Marine Science, 2018. 5: p. 227.
6. Howells, E.J., et al., Coral thermal tolerance shaped by local adaptation of photosymbionts. Nature Climate Change, 2012. 2(2): p. 116-120.
7. Fay, S.A., M.X. Weber, and J.H. Lipps, The distribution of Symbiodinium diversity within individual host foraminifera. Coral Reefs, 2009. 28(3): p. 717-726.
8. Mieog, J.C., et al., The roles and interactions of symbiont, host and environment in defining coral fitness. Plos One, 2009. 4(7).
9. Langer, M.R., M.T. Silk, and J.H. Lipps, Global ocean carbonate and carbon dioxide production; the role of reef Foraminifera. The Journal of Foraminiferal Research, 1997. 27(4): p. 271-277.
10. Leutenegger, S., Symbiosis in benthic Foraminifera - Specificity and host adaptations. Journal of Foraminiferal Research, 1984. 14(1): p. 16-35.
11. Boudagher-Fadel, M.K., Evolution and geological significance of larger benthic foraminifera. Developments in palaeontology and Stratigraphy, ed. P.B. Wignall. Vol. 21. 2008, Amsterdam, The Netherlands: Elsevier.
12. Lee, J.J. and P. Hallock, Algal Symbiosis as the Driving Force in the Evolution of Larger Foraminifera. Annals of the New York Academy of Sciences, 1987. 503(1): p. 330-347.
13. Hohenegger, J., Depth coenoclines and environmental considerations of western pacific larger foraminifera. Journal of Foraminiferal Research, 2004. 34(1): p. 9-33.
14. Hallock, P., Diversification in algal symbiont-bearing foraminifera: A response to oligotrophy? Revue de Paléobiologie. 1988. 2.
15. Hallock, P., Light dependence in Amphistegina. Journal of Foraminiferal Research, 1981. 11(1): p. 40.
16. Renema, W., Terrestrial influence as a key driver of spatial variability in large benthic foraminiferal assemblage composition in the Central Indo-Pacific. Earth-Science Reviews, 2018. 177: p. 514-544.
17. Prazeres, M. and W. Renema, Evolutionary significance of the microbial assemblages of large benthic Foraminifera. Biological Reviews, 2019. 94(3): p. 828-848.
18. Barnes, K.H., Diversity and distribution of diatom endosymbionts in Amphistegina spp. (Foraminifera) based on molecular and morphological Techniques, in College of Marine Science. 2016, University of South Florida: St. Petersburg, FL.
19. Pochon, X., et al., One-year survey of a single Micronesian reef reveals extraordinarily rich diversity of Symbiodinium types in sortitid foraminifera. Coral Reefs, 2007. 26(4): p. 867-882.
20. Prazeres, M., et al., Symbiosis and microbiome flexibility in calcifying benthic foraminifera of the Great Barrier Reef. Microbiome, 2017. 5: p. 38.
21. Holzmann, M., C. Berney, and J. Hohenegger, Molecular identification of diatom endosymbionts in nummulitid Foraminifera. Symbiosis, 2006. 42(2): p. 93-101.
22. Stuhr, M., et al., Variable thermal stress tolerance of the reef-associated symbiont-bearing foraminifera Amphistegina linked to differences in symbiont type. Coral Reefs, 2018. 37(3): p. 811-824.
23. Langer, M.R., et al., Climate-driven range extension of Amphistegina (Protista, Foraminiferida): Models of current and predicted future ranges. Plos One, 2013. 8(2): p. e54443.
24. Weinmann, A.E., et al., Traveling through time: The past, present and future biogeographic range of the invasive foraminifera Amphistegina spp. in the Mediterranean Sea. Marine Micropaleontology, 2013. 105: p. 30-39.
25. Schmidt, C., et al., Diverse internal symbiont community in the endosymbiotic foraminifera Pararotalia calcariformata: Implications for symbiont shuffling under thermal stress. Frontiers in Microbiology, 2018.
26. Silverstein, R.N., A.M.S. Correa, and A.C. Baker, Specificity is rarely absolute in coral-algal symbiosis: implications for coral response to climate change. Proceedings of the Royal Society B-Biological Sciences, 2012. 279(1738): p. 2609-2618.
27. Prazeres, M., S. Utbicke, and J.M. Pandolfi. Influence of local habitat on the physiological responses of large benthic foraminifera to temperature and nutrient stress. Scientific Reports, 2016. 6: p. 21936.
28. Lee, J.J., Diatoms as endosymbionts, in The diatom world. 2011. Springer. p. 437-464.
29. Lee, J.J., et al., Identification and Distribution of Endosymbiotic Diatoms in Larger Foraminifera. Micropaleontology, 1989. 35(4): p. 353-366.
30. Langer, M.R. and L. Hottinger, Biogeography of selected "larger" foraminifera. Micropaleontology, 2000. 46: p. 105-126.
31. Prazeres, M., T.E. Roberts, and J.M. Pandolfi, Shifts in species abundance of large benthic foraminifera Amphistegina: the possible effects of Tropical Cyclone Ita. Coral Reefs, 2017. 36(1): p. 305-309.
32. Larsen, A.R., Phylogenetic and paleobiogeographical trends in the foraminiferal genus Amphistegina. Revista Espanola de Micropaleontolgia, 1978. 10(2): p. 217-243.
33. Larsen, A.R., Studies of recent Amphistegina, taxonomy and some ecological aspects. Israel Journal of Earth Sciences, 1976. 25(1): p. 1-26.
34. Schmidt, C., et al., Recent invasion of the symbiont-bearing foraminifera Pararotalia into the Eastern Mediterranean facilitated by the ongoing warming trend. Plos One, 2015. 10(8): p. e0132917.
35. Garcia-Cuetos, L., X. Pochon, and J. Pawlowski, Molecular evidence for host–symbiont specificity in sortid Foraminifera. Protist, 2005. 156(4): p. 399-412.
36. Momigliano, P. and S. Uthicke, Symbiosis in a giant protist (Marginopora vertebralis, Soritinae): flexibility in symbiotic partnerships along a natural temperature gradient. Marine Ecology Progress Series, 2013. 491: p. 33-46.
37. Lee, J., et al., Nutritional and related experiments on laboratory maintenance of three species of symbiont-bearing, large foraminifera. Marine Biology, 1991. 109(3): p. 417-425.
38. Ter Kuile, B. and J. Erez, The size and function of the internal inorganic carbon pool of the foraminifer Amphistegina lobifera. Marine Biology, 1988. 99(4): p. 481-487.
39. Lee, J.J., Algal symbiosis in larger foraminifera. Symbiosis, 2006. 42(2): p. 63-75.
40. Caruso, A. and C. Cosentino, The first colonization of the genus Amphistegina and other exotic benthic foraminifera of the Pelagian Islands and south-eastern Sicily (central Mediterranean Sea). Marine Micropaleontology, 2014. 111: p. 38-52.
41. Hallock, P., Algal symbiosis - a mathematical analysis. Marine Biology, 1981. 62(4): p. 249-255.
42. Forderer, M., D. Rodder, and M.R. Langer, Patterns of species richness and the center of diversity in modern Indo-Pacific larger foraminifera. Scientific Reports, 2018. 8: p. 8189.
43. Hadziavdic, K., et al., Characterization of the 18S rRNA gene for designing universal eukaryote specific primers. Plos One, 2014. 9(2): p. e87624.
44. Stoeck, T., et al., Multiple marker parallel tag environmental DNA sequencing reveals a highly complex eukaryotic community in marine anoxic water. Molecular Ecology, 2010. 19: p. 21-31.
45. Team, R.C., R: A language and environment for statistical computing. . 2018, R Foundation for Statistical Computing: Vienna, Austria.
46. Callahan, B.J., et al., DADA2: High-resolution sample inference from Illumina amplicon data. Nature Methods, 2016. 13(7): p. 581-583.
47. Callahan, B.J., et al., Bioconductor workflow for microbiome data analysis: from raw reads to community analyses [version 2; referees: 3 approved]. F1000 Research, 2016. 5.
48. Quast, C., et al., The SILVA ribosomal RNA gene database project: improved data processing and web-based tools. Nucleic Acids Research, 2013. 41(D1): p. D590-D596.
49. Wright, E.S., DECIPHER: harnessing local sequence context to improve protein multiple sequence alignment. Bmc Bioinformatics, 2015. 16.
50. Schliep, K.P., phangorn: phylogenetic analysis in R. Bioinformatics, 2011. 27(4): p. 592-593.
51. McMurdie, P.J. and S. Holmes, phyloseq: An R Package for Reproducible Interactive Analysis and Graphics of Microbiome Census Data. Plos One, 2013. 8(4).
52. Oksanen, J., et al., vegan: Community Ecology Package. R package version 2.5-1. 2017.
53. Lahti, L., et al., Tools for microbiome analysis in R. Version 1.1.10012. 2017.
54. Ciccarelli, F.D., et al., Toward automatic reconstruction of a highly resolved tree of life. Science, 2006. 311(5765): p. 1283-1287.
55. Letunic, I. and P. Bork, Interactive Tree Of Life (iTOL) v4: recent updates and new developments. Nucleic Acids Research, 2019. 47(W1): p. W256-W259.
56. Chao, A., *Nonparametric estimation of the number of classes in a population*. Scandinavian Journal of Statistics, 1984: p. 265-270.

57. Simpson, E.H., *Measurement of diversity*. Nature, 1949. 163(4148): p. 688-688.

58. Lozupone, C.A., et al., *Quantitative and qualitative beta diversity measures lead to different insights into factors that structure microbial communities*. Applied and Environmental Microbiology, 2007. 73(5): p. 1576-1585.

59. Locarnini, R.A., et al., *World Ocean Atlas 2013 temperature*, in *National Centers for environmental information NOAA Atlas NESDIS 73*, AMTENAN, Editor. 2013, NOAA: USA. p. 40.
Declarations

Ethics approval and consent to participate
Not applicable

Consent for publication
Not applicable

Availability of data and material
Electronic supplementary material is available as supplementary material accompanying the manuscript. Sequences are deposited in NCBI under the accession number TBA.

Authors’ contribution
MP designed the study. MP, SD, CS, MS, WR, and TER contributed samples. MP and SFR conducted all laboratory analysis. MP and TER analysed the data. MP led the writing of the manuscript with contribution from all authors.

Acknowledgements
We would like to thank Dr. Sonia Rowley for assisting with samples collection in Hawai’i, which was conducted under the Special Activity permit #2019-22 issue to MP. Collection in Mauritius were conducted in partnership with Ms. Rohinee Bhiwajee and Mr. Arnoud Nicolas from the Mauritius Oceanography Institute, under the permission of the Department for Continental Shelf, Maritime Zones Administration and Exploration of the Republic of Mauritius. Samples from Indonesia were collected under the collection permit issue to WR.
Prof. Pamela Hallock for kindly providing samples from the Maldives. We acknowledge Amatzia Genin at the Interuniversity Institute of Marine Science (IUI) and Claire E. Reymond for supporting sampling in Eilat, and thank Gita R. Narayan for sampling efforts in Zanzibar, where foraminifera were collected under the Revolutionary Government of Zanzibar 434 (S.M.Z.) Research Permit No. 0671. Samples from Okinawa, Japan were collected during a short-term fellowship financed through the Japan Society of the Promotion of Science to CS in the laboratory of Prof. Kazuhiko Fujita, from the University of the Ryukyus, who provided field work support.

Funding
This research was financially supported by the Netherlands Organisation for Scientific Research for providing funding through the Innovational Research Incentives Schemes Veni to MP, and the Treub-Maastchappij for providing funding to MP to conduct field work in the Republic of Mauritius.

Competing interests
The authors declare that they have no competing interests.
**Figure captions**

**Figure 1.** Relative abundance of algal taxa across different species and sites. (a) Relative abundance of algal groups classifies as ‘Ochrophyta’. (b) Relative abundance of genera in the Fragilariales.

**Figure 2.** Total abundance (log_{10}-transformed) plotted against prevalence of Fragilariales genera in all samples in *Amphistegina lessonii*, *A. lobifera*, and *A. radiata*.

**Figure 3.** Phylogenetic tree of algal biome community associated with *Amphistegina lessonii*, *A. lobifera* and *A. radiata* across all sites. Dendrogram represents the 892 algal taxa with a relative abundance of at least 1% summed across all samples. Bars represent relative abundance of each phylotype (i.e. ASVs) identified in *A. lessonii* (red), *A. lobifera* (green), and *A. radiata* (blue).

**Figure 4.** Observed and estimated richness (Chao 1 and Simpson) of ASVs classified as ‘Ochrophyta’ per individual sample.

**Figure 5.** Differences in algal biome in *A. lobifera*, *A. lessonii*, and *A. radiate* collected from different sites. Two-dimensional plots utilising weighted UniFrac-distance matrix showing the Principal Coordinates Analysis representing (a) all individuals coded by species, and (b-q) algal communities within each site, coded by species.
Figure 6. Specimens of (A) *Amphistegina lessonii*, (B) *A. lobifera*, and (C) *A. radiata* collected from the same habitat in Kimbe Bay, Papua New Guinea. Scale bars represent 1 mm in A and B, and 2 mm in C. Note that specimens were preserved in 96% ethanol, and therefore symbiont pigment colour shown here does not represent natural coloration.

Figure 7. Sampling sites across the Mediterranean (Sicily and Greece), Red Sea (Eilat), Indian Ocean (Maldives, Mauritius, Zanzibar, and Ningaloo Reef), and Pacific Ocean (Indonesia, Papua New Guinea, Okinawa, Palau, Micronesia, Hawai’i, Great Barrier Reef, and Lord Howe Island). Red, green and blue circles represent collection sites for *Amphistegina lessonii*, *A. lobifera*, and *A. radiata*, respectively. Background colour represents mean annual sea surface temperature extracted from the World Ocean Atlas 2013 [59].
Table 1. Kruskal-Wallis rank sum test results of diversity indices among sites and species.

| Term      | Sites |          |          |          |          |          |
|-----------|-------|----------|----------|----------|----------|----------|
|           | $\chi^2$ | df | p-value | $\chi^2$ | df | p-value |
| Observed  | 64.213 | 15 | <0.01   | 35.973 | 2 | <0.01   |
| Chao 1    | 63.947 | 15 | <0.01   | 36.195 | 2 | <0.01   |
| Simpson   | 65.27  | 15 | <0.01   | 39.504 | 2 | <0.01   |
Table 2. Two-way Permutation ANOVA results for weighed UniFrac-distance matrix of algal community associated with specimens of *A. lessonii*, *A. lobifera* and *A. radiata* collected across a wide distribution range, analysed together and individually. Results are based on 1000 permutations.

| Term                        | df  | SS     | R-squared | Pseudo-F | P-value |
|-----------------------------|-----|--------|-----------|----------|---------|
| **A. lessonii x A. lobifera x A. radiata** |     |        |           |          |         |
| Species                     | 2   | 5.054  | 0.093     | 16.85    | <0.01   |
| Site                        | 15  | 17.061 | 0.315     | 7.58     | <0.01   |
| Species*Site                | 8   | 10.792 | 0.199     | 8.99     | <0.01   |
| Residual                    | 142 | 21.306 | 0.393     |          |         |
| Total                       | 167 | 54.213 | 1.000     |          |         |
| **A. lessonii**             |     |        |           |          |         |
| Site                        | 8   | 8.906  | 0.662     | 10.76    | <0.01   |
| Residuals                   | 44  | 4.551  | 0.338     |          |         |
| Total                       | 52  | 13.456 | 1.000     |          |         |
| **A. lobifera**             |     |        |           |          |         |
| Site                        | 14  | 18.128 | 0.531     | 7.35     | <0.01   |
| Residuals                   | 91  | 16.040 | 0.469     |          |         |
| Total                       | 105 | 34.168 | 1.000     |          |         |
| **A. radiata**              |     |        |           |          |         |
| Site                        | 1   | 0.914  | 0.622     | 11.56    | <0.01   |
| Residuals                   | 7   | 0.553  | 0.378     |          |         |
| Total                       | 8   | 1.467  | 1.000     |          |         |
| Collection site                  | Lat; long                  | Depth (m) | A. lobifera | A. lessonii | A. radiata |
|---------------------------------|----------------------------|-----------|-------------|-------------|------------|
| Sicily, Italy                   | 36.74470; 15.11820         | 1         |             |             | X          |
| Vravona, Greece                 | 37.9218111; 24.0141889     | 1         |             |             | X          |
| Maldives                        | 1.92499; 73.39966          | 8         | X           |             |            |
| Mauritius                       | -20.28666; 57.36098        | 2         | X           |             |            |
| Okinawa, Japan                  | 26.65182; 127.85624        | 0.1       | X           |             |            |
| Eilat, Israel                   | 29.5023; 34.918            | 2         | X           | X           | X          |
| Zanzibar                        | -6.145603; 39.12445        | 2         | X           | X           |            |
| Ningaloo Reef, Australia        | -23.15007; 113.75268       | 5         | X           | X           |            |
| Makassar, Indonesia             | -4.71898; 119.25418        | 5         | X           | X           |            |
| Kimbe Bay, Papua New            | -5.42119; 150.09434        | 3         | X           | X           | X          |
| Guinea                          |                           |           |             |             |            |
| Pohnpei, Micronesia             | 6.758169; 157.91721        | 5         | X           | X           | X          |
| Palau                           | 7.30573; 134.50250         | 3         | X           | X           |            |
| Hawai’i, USA                    | 21.64144; -157.91791       | 2         |             | X           |            |
| Great Barrier Reef, Australia   | -14.68383; 145.47186       | 8         | X           | X           |            |
| Lord Howe Island, Australia     | -31.51960; 159.05620       | 5         | X           | X           |            |
| Mo’orea, French Polynesia       | -17.47583; 149.82222       | 8         | X           |             |            |
Figure 1

Algal groups
- Asterioclodon
- Asterionella
- Bacillariophyceae
- Dictyotales
- Ectocarpales
- Fragilariales
- Halothrix
- Mediophyceae
- Other unidentified taxa
- Sarcinochrysidales
- Schizochladia
- Sphacelaria
- Uncultured eukaryotes

Fragilariales genera
- Licmophora
- Nanofrustulum
- Serratifera
- Staurosira

Collection sites
Figure 2

Fragilariales genera

- Lichmophora
- Nanofrustulum
- Serratifera
- Staurosira

Total abundance (log10)

Prevalence
Figure 3
Figure 4

Alpha diversity measure

Observed

Chao 1

Simpson

Collection sites

Species

A. tessonii

A. lobifera

A. radiata
Figure 5

(a) PCoA (13.6%) Species
- A. lessonii
- A. lobifera
- A. radiata

(b) French Polynesia
(c) Great Barrier Reef
(d) Greece
(e) Eilat

(f) Hawai’i
(g) Indonesia
(h) Okinawa
(i) Kimbe Bay

(j) Lord Howe Island
(k) Maldives
(l) Mauritius
(m) Micronesia

(n) Ningaloo Reef
(o) Palau
(p) Sicily
(q) Zanzibar

PCoA (25.4%)
Figure 6
