Highly Variable and Non-complex Diazotroph Communities in Corals From Ambient and High CO₂ Environments

Laura Geissler*, Valentine Meunier†, Nils Rädecker†, Gabriela Perna, Riccardo Rodolfo-Metalpa, Fanny Houlbrèque and Christian R. Voolstra*

1 Department of Biology, University of Konstanz, Konstanz, Germany; 2 ENTROPIE UMR 9220 (CNRS, IRD, UR, UNC, IFREMER) Institut de Recherche pour le Développement, Nouméa, New Caledonia; 3 Laboratory for Biological Geochemistry, School of Architecture, Civil and Environmental Engineering, Ecole Polytechnique Fédérale de Lausanne (EPFL), Lausanne, Switzerland

The ecological success of corals depends on their association with microalgae and a diverse bacterial assemblage. Ocean acidification (OA), among other stressors, threatens to impair host-microbial metabolic interactions that underlie coral holobiont functioning. Volcanic CO₂ seeps offer a unique opportunity to study the effects of OA in natural reef settings and provide insight into the long-term adaptations under a low pH environment. Here we compared nitrogen-fixing bacteria (diazotrophs) associated with four coral species (Pocillopora damicornis, Galaxea fascicularis, Acropora secale, and Porites rus) collected from CO₂ seeps at Tutum Bay (Papua New Guinea) with those from a nearby ambient CO₂ site using nifH amplicon sequencing to characterize the effects of seawater pH on bacterial communities and nitrogen cycling. Diazotroph communities were of generally low diversity across all coral species and for both sampling sites. Out of a total of 25 identified diazotroph taxa, 14 were associated with P. damicornis, of which 9 were shared across coral species. None of the diazotroph taxa, however, were consistently found across all coral species or across all samples within a species pointing to a high degree of diazotroph community variability. Rather, the majority of sampled colonies were dominated by one or two diazotroph taxa of high relative abundance. Pocillopora damicornis and Galaxea fascicularis that were sampled in both environments showed contrasting community assemblages between sites. In P. damicornis, Gammaproteobacteria and Cyanobacteria were prevalent under ambient pH conditions, while a single member of the family Rhodobacteraceae was present at high relative abundance at the high pH CO₂ site. Conversely, in G. fascicularis diazotroph communities were indifferent between both sites. Diazotroph community changes in response to OA seem thus variable within as well as between host species, potentially arguing for haphazard diazotroph community assembly. This warrants further research into the underlying factors structuring diazotroph community assemblages and their functional role in the coral holobiont.

Keywords: coral holobiont, diazotrophy, nitrogen fixation, climate change, ocean acidification, nifH, next-generation sequencing
INTRODUCTION

Corals are adapted to the oligotrophic conditions of the tropical ocean via symbiotic associations with a variety of microbes, which extend their metabolic repertoire to support the efficient uptake and recycling of limiting nutrients (Bang et al., 2018; Robbins et al., 2019; Pogoreutz et al., 2020). The resulting metaorganism is commonly referred to as the coral holobiont (Knowlton and Rohwer, 2003; Jaspers et al., 2019). Besides endosymbiotic dinoflagellate algae of the family Symbiodiniaceae that transfer photosynthetically fixed carbon to other holobiont constituents in exchange for shelter and inorganic nutrients (Muscatine and Porter, 1977; Lajeunesse et al., 2018), associated microbes include bacteria, archaea, fungi, and viruses that all contribute to coral holobiont physiology and stress resilience (Bourne et al., 2016; Peixoto et al., 2021; Voolstra et al., 2021). Although corals can cover almost up to 100% of their carbon demand via nutrient exchange with their endosymbiotic algae (Falkowski et al., 1984; Muscatine, 1990), biomass synthesis requires the supplementation of translocated carbon with other nutrients (e.g., nitrogen). Consequently, productivity and growth of coral holobionts is argued to be limited by bioavailable nitrogen (Falkowski et al., 1993). Although corals can acquire bioavailable nitrogen through heterotrophic feeding (Grottoli et al., 2006; Houlbreque and Ferrier-Pages, 2009; Meunier et al., 2021a) and the uptake of dissolved organic nitrogen from the surrounding water (Grover et al., 2008; Tilstra et al., 2019; El-Khaled et al., 2020), they also rely on nitrogen-fixing bacteria and archaea (Rädecker et al., 2015). These so-called diazotrophs are capable of reducing N₂ to ammonia (NH₃) via the nitrogenase enzyme and supposedly comprise a significant source of fixed nitrogen to coral holobiont members (Fiore et al., 2010; Lema et al., 2012, 2014; Pogoreutz et al., 2017b). Typically, amplicon sequencing of the nifH gene encoding for a conserved subunit of the nitrogenase enzyme is used for the characterization of diazotroph communities associated with corals (Gaby and Buckley, 2012; Lema et al., 2012; Lesser et al., 2018; Bednarz et al., 2019). These studies indicate that diazotroph communities vary widely between locations and coral host species. Alphaproteobacteria, especially of the orders Rhizobiales and Rhodobacterales, as well as Gammaproteobacteria and Cyanobacteria are commonly found to comprise the largest part of these diazotrophic communities in tropical and temperate corals (Lesser et al., 2004, 2018; Lema et al., 2012; Bednarz et al., 2019).

Given that the coral holobiont typically exists in a nitrogen-limited state (Rädecker et al., 2015, 2021), diazotrophy may help corals adjust to differences in environmental nitrogen availability (Rädecker et al., 2015; Benavides et al., 2017). In support of this, changes in environmental conditions are often met with changes in the composition of microbial communities associated with corals (Jessen et al., 2013; Roder et al., 2015; Ziegler et al., 2019). Shifts in microbial community composition and abundance supposedly change the metabolic repertoire to support the coral holobiont to adapt to changing environmental conditions (Reshef et al., 2006; Ziegler et al., 2017; Voolstra and Ziegler, 2020). Given the potential importance of diazotrophs, changes in diazotroph activity, abundance, and community composition could directly affect the response of the coral holobiont to environmental change (Santos et al., 2014; Pogoreutz et al., 2017a). In this light, ocean acidification (OA), the reduction of the pH in the Earth’s oceans due to increasing absorption of CO₂ from the atmosphere, poses a direct threat to calcifying organisms such as corals (Orr et al., 2005; O’Brien et al., 2016). Reduced carbonate ion concentrations as a result of OA increase energy requirements of calcification in the coral holobiont (Cohen and Holcomb, 2009; Crook et al., 2013; Strahl et al., 2015). The resulting reduction in energy availability in the coral holobiont could thus directly affect the symbiotic interactions within the coral holobiont (McCulloch et al., 2017; D’Olivo et al., 2019).

Currently, our knowledge regarding the effects of OA on nitrogen fixation in coral holobionts is limited. Rädecker et al. (2014) found that gross nitrogen fixation decreases in *Seriatopora hystrix* coral holobionts when exposed to OA conditions for short periods of time. These findings, however, cannot account for any long-term environmental adaptation of coral holobionts under such conditions, including potential shifts in diazotroph community composition. Volcanic CO₂ seeps offer a unique opportunity to study the long-term effects of OA on coral holobionts. Local pCO₂ at some of these sites resemble levels that are expected for global oceans in the coming century (Caldeira, 2005). Coral communities at these low pH seeps show a reduced diversity with massive corals like *Porites* spp. being favored over more structurally complex (e.g., branching or foliose growth forms) species (Fabricius et al., 2011). Corals at these sites have been reported to have higher net photosynthesis rates (Biscéré et al., 2019), suggesting a stimulation of Symbiodiniaceae photosynthetic activity due to enhanced CO₂ availability (Rädecker et al., 2017). In addition, the bacterial microbiomes of some coral species were less diverse (Morrow et al., 2015). However, a more recent study showed that massive *Porites* species showed almost no difference in their associated bacterial communities under high pCO₂ (O’Brien et al., 2018).

In this study, we compared the diazotroph communities of four species of hermatypic corals using a *nifH*-based amplicon sequencing approach to characterize changes in abundance, diversity, and microbiome composition of samples from ambient and high pCO₂ volcanic seep sites in Papua New Guinea. After stringent selection of functional *nifH* phylotypes and samples, we show that in at least one of the species, *Pocillopora damicornis*, different phylogenetic groups of nitrogen-fixing bacteria are present between ambient and high pCO₂ sites, suggesting separate functional roles in the coral holobiont across host species.

MATERIALS AND METHODS

Sample Collection

Sample collection took place during the boreal spring, between May 26th and June 7th 2018, aboard the R/V Alis, as described in Meunier et al. (2021b). At a hydrothermal vent area at Tutum Bay (Ambitle Island, New Ireland Province, Papua New Guinea), samples were collected from one site exposed to high pCO₂.
D Nuniversity, P. rus, and A. secale. Collections took place at the same depths (between 3–8 m) for both sampling sites. Coral fragments were cut with a bone cutter while SCUBA diving and transported in individual zip-lock bags to the laboratory within 30 min of collection. We collected a total of 24 samples from 4 coral species at the high pCO2 site: 7 samples of Acropora secale, 5 of Galaxea fascicularis, 7 of Pocillopora damicornis, and 5 of Porites rus. Similarly, we sampled 27 colonies from 4 coral species at the control site: 10 fragments of A. secale, 5 of G. fascicularis, 9 of P. damicornis, and 3 of P. rus. Thus, a total of 51 samples were collected from both sites. Samples were immersed in DNA/RNA Shield (Zymo, Freiburg, Germany) and shipped on dry ice for DNA extraction, nifH amplification, and sequencing. The P. damicornis samples are the same as those analyzed in Meunier et al. (2021b).

**DNA Collection, nifH Amplification, and Sequencing**

DNA was isolated from 51 coral fragments and 1 negative control (null input sample: to assess for DNA extraction kit contamination) using the Qiagen DNeasy 96 Blood & Tissue kit (Qiagen, Hilden, Germany) with slight modifications to the manufacturer’s instructions. Briefly, coral samples were fragmented with a bone cutter and transferred to Eppendorf tubes with 400 µl of ATL buffer and 20 µl of protease. Samples where then incubated in a thermal mixer (Eppendorf, Hamburg, Germany) for 2 h at 56°C and 600 rpm. Afterward, for each sample, 200 µl of the lysate were transferred into microtubes. To aid the disintegration of coral mucus, tubes were incubated at 56°C for 1 h. DNA extraction then proceeded according to the manufacturer’s instructions. DNA concentrations were quantified using Qubit and the dsDNA High Sensitivity Assay Kit (Invitrogen, Carlsbad, United States). DNA concentrations were quantified using Qubit and the dsDNA High Sensitivity Assay Kit (Invitrogen, Carlsbad, United States). Sequencing was performed at 6 pM with 20% phiX on the Illumina MiSeq platform at 2 × 301 bp paired-end V3 chemistry according to the manufacturer’s specifications. Sequencing data determined in this study are available at NCBI BioProject PRJNA668126 at https://www.ncbi.nlm.nih.gov/bioproject/PRJNA668126.

**nifH Amplicon Sequencing Data Analysis Pipeline**

Quality control of MiSeq nifH amplicon paired reads was done using the TaxADivA (TAXonomy Assignment and DIVERSity Assessment) perl wrapper script (Gaby et al., 2018), as detailed in Lesser et al. (2018) with the following settings: -y -k -r 26 -l 29 --keepc4 -g 500. Subsequent to TaxADivA-based read merging, quality control and filtering, samples with <500 sequences were removed from subsequent analyses (Lesser et al., 2018). The remaining 41 samples (including the negative control) were combined into one fasta file and minimum entropy decomposition (MED) was employed at default settings to cluster reads into oligotypes in an unsupervised fashion (Eren et al., 2015). Representative sequences of the resultant 197 final MED nodes were translated to the protein level using FrameBot (Wang et al., 2013). As a reference dataset, the “nifh_prot_ref.fasta” file included with the FrameBot program was used. To separate nifH-like genes belonging to clusters IV and V, from true functional nifH genes belonging to clusters I-III (Raymond et al., 2004), we employed a three-criteria model sensu (Angel et al., 2018). First, after alignment in MAFFT (Katoh et al., 2002), representative sequences for each final MED node were analyzed using a Classification and Regression Tree (CART) model (Frank et al., 2016). Second, we constructed an Evolutionary Placement Algorithm (EPA) Tree (Berger et al., 2011) using the corresponding RaxML implementation (Stamatakis, 2014). Minimum entropy decomposition node representatives were aligned with a reference dataset from the NifMAP pipeline (Angel et al., 2018) that included representatives of all nifH-like gene clusters (“cluster.rep.nifH-chL-bchX.0.9_noDot.fasta”). Query sequences within this combined reference alignment were then placed on the reference tree (“RaxML_bipartitions.MultipleOriginal.tree” also from the NifMAP pipeline). Third, using node representative sequences as queries, a BLASTP (Camacho et al., 2009) search was conducted against the nifH gene database taken from the Zehr lab website (accessed October 2nd 2020). In addition, we looked for matches in the NCBI non-redundant (nr) protein database (excluding uncultured/environmental sample sequences and

1https://www.jzehrlab.com/nifh
Geissler et al.

Coral Diazotroph Communities

automatically generated model annotations) to further confirm the results of the initial BLASTP search. In case of mismatching results but similar percentages of protein identity between the two queries, the result to the curated Zehr nifH database was preferred. In order to determine the cluster designation for the representative sequences of all final MED nodes, a node was considered to be a true nifH-representative (i.e., belonging to clusters I–III) if at least two of the three criteria produced consistent results.

In order to obtain putative species-level annotations and to collapse the MED count matrix for downstream analyses, an OTU (operational taxonomic unit) clustering at the nucleotide level was conducted via CD-HIT-EST (Fu et al., 2012). The clustering cutoff was set to 91.9%, corresponding to the species-level cutoff previously reported (Gaby et al., 2018). Reference sequences of final MED nodes that only differed by their length, i.e., by means of additional or missing nucleotides at the 5′ or 3′ of the sequence, were combined in this clustering step, which led to the removal of 42 redundant MED nodes from an initial total of 197 MED nodes. Lastly, all samples that contained <100 sequences of cluster I–III nifH representatives were removed from subsequent analyses, retaining 25 of the 40 coral samples and 1 negative control. A maximum likelihood phylogenetic tree of all final OTUs was constructed using RAxML from 20 starting trees using the PROTGAMMAAUTO method to determine an optimal substitution model (LG + G in this case). Additionally, 1,000 bootstrap trees were calculated to support the bifurcations of the best ML tree. Trees were visualized using the R package ggtree (Yu et al., 2017).

Statistical analyses within and between sites and species as well as plotting were done in R using the packages phyloseq, ape, and vegan (R Development Core Team, 2015). Pairwise PERMANOVA tests within and between coral host species for a given sampling site were done using the custom R function "pairwise_adonis.R"2 based on a Bray-Curtis dissimilarity matrix. Comparisons of Shannon diversity, inverse Simpson’s Index, and the Chao1 estimator between the two sampling sites over all total of 197 MED nodes. Lastly, all samples that contained <100 sequences of cluster I–III nifH representatives were removed from subsequent analyses, retaining 25 of the 40 coral samples and 1 negative control. A maximum likelihood phylogenetic tree of all final OTUs was constructed using RAxML from 20 starting trees using the PROTGAMMAAUTO method to determine an optimal substitution model (LG + G in this case). Additionally, 1,000 bootstrap trees were calculated to support the bifurcations of the best ML tree. Trees were visualized using the R package ggtree (Yu et al., 2017).

To determine OTUs that were differentially abundant between the control and seep site, linear discriminant analysis (LDA) was conducted in the LEfSe implementation on the Galaxy web application of the Huttenhower lab (‘accessed October 2nd 2020). LEfSe statistical tests for differential abundance were carried out with alpha values for factorial Kruskal-Wallis and subsequent pairwise Wilcoxon tests considered significant at ≤0.05. For the LDA analysis, a threshold score of ≥2.0 for OTUs was considered indicative for a significant difference in the relative abundance between sites. This analysis was done across samples from all species, for samples of P. damicornis, and for samples of G. fascicularis.

Taxon-Specific nifH Quantitative PCR

Patterns of relative abundance based on nifH amplicon sequencing data were confirmed using qPCR (quantitative polymerase chain reaction). For this, OTU_31 was chosen due to its pervasive presence among P. damicornis samples. A specific primer pair for OTU_31 was designed based on its MED representative sequence using Primer-BLAST (Ye et al., 2012). The OTU_31 primer pair (forward: 5′-TTGCGATCCAAAAGTGCAGT-3′; reverse: 5′-GCGGGACTCAACACAGAGAA-3′) amplifies a 157 bp long region of the nifH amplicon determined from the MiSeq sequencing. Primer products were checked on a 1.2% agarose gel for single bands after PCR amplification using the following protocol: 90°C for 10 min, followed by 40 cycles of 95°C, 57°C, and 72°C for 30 s each, and a final extension step of 72°C for 10 min. Upon confirmation of a single amplification product using gel electrophoresis, amplicon gel slices from different samples were cut out, and DNA was extracted and purified using the QIAquick Gel Extraction Kit (Qiagen, Germany) following the manufacturer’s protocol. The purified product was Sanger-sequenced to confirm that the amplified sequence was identical to the OTU_31 representative sequence (Supplementary Figure 1). The 16S rRNA gene amplicon primers 784F and 1061R (Andersson et al., 2008; Bayer et al., 2013) were used to produce a standard for relative abundance calculation. To confirm that primer amplification efficiency for the nifH and 16S primer pairs were similar and close to 100%, a 10-fold serial dilution of DNA from a sample with a known presence of OTU_31 was prepared for the determination of qPCR standard curves (Supplementary Figure 2). All qPCR reactions were run on a qTOWER® 84 using the innuMIX qPCR DSGreen Standard master mix (both Analytik Jena GmbH, Germany) with final primer concentrations of 0.5 µM in a reaction volume of 10 µl with a cycling protocol of 95°C for 10 min, 50 cycles of 95°C for 30 s, 57°C for 30 s, and 72°C for 30 s, followed by 72°C for 10 min, and a subsequent melting curve to assess uniformity of amplification and to confirm the absence of primer dimers. The qPCR samples were run in triplicate in addition to a no-template and an extraction kit negative control for both primer pairs, respectively. Fold change differences were calculated using the ΔΔCT method. For this, the mean ΔCT (CT_{nifH} – CT_{16S}) across all samples was subtracted from the ΔCTs of individuals samples and the calculated ΔΔCTs were expressed as relative fold changes of nifH abundance using 2−ΔΔCT. Pearson’s correlation coefficient was calculated to confirm that qPCR and nifH amplicon sequencing showed similar distribution patterns of OTU_31 in all tested samples.

RESULTS

Overview of nifH Sequencing Effort and Filtering

We collected 51 samples (plus 1 negative control, see Methods) from four different coral species: 7 and 10 samples of Pocillopora

---

2https://github.com/pmartinezarbizu/pairwiseAdonis
3https://huttenhower.sph.harvard.edu/galaxy/
daccominus, 5 and 5 samples of *Galaxea fascicularis*, 7 and 9 samples of *Acropora secale*, and 5 and 3 samples of *Porites rus* from the high and control CO₂ site, respectively. For each species, samples were collected from a high carbon dioxide volcanic seep site with a partial pressure pCO₂ = 1179 µatm (pH 7.73) and a nearby control site with an ambient pCO₂ = 436 µatm (pH 8.12). After amplifying *nifH*-like genes using the IKG3 and DVV primer pair and MiSeq sequencing, a total of 7,143,933 read pairs (on average 137,383.3 read pairs/sample) were available for analysis. After quality filtering and read pair merging, 10 samples were excluded due to low sequence counts. With the remaining 41 samples, 6,240,993 sequences were assessed using MED. Post-MED, a total of 4,531,377 counts remained, of which 637,413 (14.1%) were annotated as functional nitrogenase genes of *nifH* sequences. 3,893,964 (85.9%) were identified as belonging to cluster IV and as functional nitrogenase genes of *nifH* sequences remained, of which 637,413 (14.1%) were annotated as belonging to cluster IV and V nitrogenase gene homolog representatives. On average, coral samples were represented by 110,521 sequences after MED. Of the 40 coral host samples that were considered during MED, 15 were discarded due to having <100 sequences annotated to *nifH* clusters I–III: Consequently, 25 samples remained for downstream analysis: 4 *A. secale*, 11 *P. damicornis*, 7 *P. rus* (Table 1). These 25 samples comprised 2,662,462 sequences in total. On average 23.5% of sequences from all samples could be assigned to *nifH* clusters I–III. Although there were no statistically significant differences in the ratio of functional nitrogenase to *nifH* homolog annotated sequences between species (ANOVA, p = 0.08), *P. damicornis* samples showed the highest presence of cluster I–III to cluster IV–V sequences (84% of assessed sequences) (Figure 1). In addition, samples from the control site showed higher relative numbers of cluster I–III sequences (33% of assessed sequences) in comparison to the seeps site (15% of assessed sequences, t test, p = 0.1). The negative control (no input DNA extraction) comprised 6,984 sequences after MED and showed a higher relative number of functional nitrogenase sequences, with 92.7% of retained sequences annotated as true *nifH* (Table 1 and Supplementary Table 1).

**Phylogenetic nifH Analysis**

Clustering of the 197 determined MED nodes into OTUs using a 91.9% sequence identity threshold in CD-HIT-EST resulted in 88 OTUs, of which 61 were designated as cluster IV *nifH* homologs or cluster V bacteriochlorophyll synthesis

**Table 1** | nifH sequencing overview and alpha diversity measures of 25 coral samples from a high pCO₂ (seep) and an ambient pCO₂ (control) site and 1 negative control, retained after sequence processing and nifH gene filtering.

<table>
<thead>
<tr>
<th>Coral species</th>
<th>Sample ID</th>
<th>Site</th>
<th>No. of read pairs</th>
<th>No. of sequences post-TaxADivA</th>
<th>No. of sequences post-MED</th>
<th>Nitrogenase cluster IV/V sequences</th>
<th>Nitrogenase cluster I-III sequences</th>
<th>Chao1</th>
<th>Shannon</th>
<th>Inverse Simpson</th>
</tr>
</thead>
<tbody>
<tr>
<td><em>P. damicornis</em></td>
<td>pd_ctrl_3</td>
<td>Control</td>
<td>226521</td>
<td>170583</td>
<td>134379</td>
<td>21022</td>
<td>113357</td>
<td>3</td>
<td>0.5245</td>
<td>1.5013</td>
</tr>
<tr>
<td><em>P. damicornis</em></td>
<td>pd_ctrl_2</td>
<td>Control</td>
<td>192805</td>
<td>161964</td>
<td>132272</td>
<td>35816</td>
<td>96456</td>
<td>2</td>
<td>0.6763</td>
<td>1.9052</td>
</tr>
<tr>
<td><em>P. damicornis</em></td>
<td>pd_ctrl_6</td>
<td>Control</td>
<td>177427</td>
<td>164185</td>
<td>137557</td>
<td>52232</td>
<td>85325</td>
<td>5</td>
<td>0.1371</td>
<td>1.0631</td>
</tr>
<tr>
<td><em>P. damicornis</em></td>
<td>pd_ctrl_4</td>
<td>Control</td>
<td>126702</td>
<td>122496</td>
<td>101913</td>
<td>98432</td>
<td>3481</td>
<td>8</td>
<td>0.6101</td>
<td>1.7047</td>
</tr>
<tr>
<td><em>P. damicornis</em></td>
<td>pd_ctrl_9</td>
<td>Control</td>
<td>97087</td>
<td>67859</td>
<td>52139</td>
<td>27028</td>
<td>25111</td>
<td>8</td>
<td>0.1681</td>
<td>1.0826</td>
</tr>
<tr>
<td><em>P. damicornis</em></td>
<td>pd_sseep_1</td>
<td>Vents</td>
<td>278757</td>
<td>251094</td>
<td>202800</td>
<td>198787</td>
<td>4013</td>
<td>1</td>
<td>0.0000</td>
<td>1.0000</td>
</tr>
<tr>
<td><em>P. damicornis</em></td>
<td>pd_sseep_6</td>
<td>Vents</td>
<td>202409</td>
<td>196132</td>
<td>147096</td>
<td>30286</td>
<td>118610</td>
<td>2</td>
<td>0.0116</td>
<td>1.0031</td>
</tr>
<tr>
<td><em>P. damicornis</em></td>
<td>pd_sseep_4</td>
<td>Vents</td>
<td>189562</td>
<td>142959</td>
<td>109931</td>
<td>69134</td>
<td>40797</td>
<td>4</td>
<td>0.4475</td>
<td>1.2802</td>
</tr>
<tr>
<td><em>P. damicornis</em></td>
<td>pd_sseep_7</td>
<td>Vents</td>
<td>187035</td>
<td>61244</td>
<td>40504</td>
<td>35096</td>
<td>5408</td>
<td>4</td>
<td>0.3209</td>
<td>1.1700</td>
</tr>
<tr>
<td><em>P. damicornis</em></td>
<td>pd_sseep_3</td>
<td>Vents</td>
<td>168818</td>
<td>165233</td>
<td>134449</td>
<td>133531</td>
<td>918</td>
<td>4</td>
<td>0.0170</td>
<td>1.0044</td>
</tr>
<tr>
<td><em>P. damicornis</em></td>
<td>pd_sseep_2</td>
<td>Vents</td>
<td>116080</td>
<td>90641</td>
<td>71056</td>
<td>54087</td>
<td>16969</td>
<td>3</td>
<td>0.9379</td>
<td>2.2315</td>
</tr>
<tr>
<td><em>G. fascicularis</em></td>
<td>gf_ctrl_5</td>
<td>Control</td>
<td>182794</td>
<td>179950</td>
<td>123061</td>
<td>115752</td>
<td>7309</td>
<td>2</td>
<td>0.6326</td>
<td>1.7880</td>
</tr>
<tr>
<td><em>G. fascicularis</em></td>
<td>gf_ctrl_2</td>
<td>Control</td>
<td>142758</td>
<td>141263</td>
<td>25014</td>
<td>18780</td>
<td>6234</td>
<td>1</td>
<td>0.0000</td>
<td>1.0000</td>
</tr>
<tr>
<td><em>G. fascicularis</em></td>
<td>gf_ctrl_1</td>
<td>Control</td>
<td>122450</td>
<td>121656</td>
<td>101913</td>
<td>98432</td>
<td>3481</td>
<td>8</td>
<td>0.6101</td>
<td>1.7047</td>
</tr>
<tr>
<td><em>G. fascicularis</em></td>
<td>gf_ctl_3</td>
<td>Control</td>
<td>105776</td>
<td>104144</td>
<td>32459</td>
<td>31882</td>
<td>577</td>
<td>6</td>
<td>0.2781</td>
<td>1.1118</td>
</tr>
<tr>
<td><em>G. fascicularis</em></td>
<td>gf_sseep_1</td>
<td>Vents</td>
<td>206846</td>
<td>200857</td>
<td>111622</td>
<td>97436</td>
<td>14186</td>
<td>4</td>
<td>0.0015</td>
<td>1.0003</td>
</tr>
<tr>
<td><em>G. fascicularis</em></td>
<td>gf_sseep_3</td>
<td>Vents</td>
<td>131994</td>
<td>126121</td>
<td>28944</td>
<td>27838</td>
<td>1106</td>
<td>4</td>
<td>0.4392</td>
<td>1.2871</td>
</tr>
<tr>
<td><em>G. fascicularis</em></td>
<td>gf_sseep_5</td>
<td>Vents</td>
<td>106418</td>
<td>105541</td>
<td>36400</td>
<td>36151</td>
<td>249</td>
<td>2</td>
<td>0.0984</td>
<td>1.0410</td>
</tr>
<tr>
<td><em>A. secale</em></td>
<td>as_sseep_6</td>
<td>Vents</td>
<td>270583</td>
<td>261795</td>
<td>220754</td>
<td>220322</td>
<td>432</td>
<td>2</td>
<td>0.0164</td>
<td>1.0046</td>
</tr>
<tr>
<td><em>A. secale</em></td>
<td>as_sseep_4</td>
<td>Vents</td>
<td>174004</td>
<td>172218</td>
<td>149228</td>
<td>148229</td>
<td>999</td>
<td>2</td>
<td>0.3652</td>
<td>1.2656</td>
</tr>
<tr>
<td><em>A. secale</em></td>
<td>as_sseep_5</td>
<td>Vents</td>
<td>162866</td>
<td>156486</td>
<td>105834</td>
<td>101522</td>
<td>4312</td>
<td>2</td>
<td>0.0022</td>
<td>1.0005</td>
</tr>
<tr>
<td><em>A. secale</em></td>
<td>as_sseep_1</td>
<td>Vents</td>
<td>144400</td>
<td>135139</td>
<td>104969</td>
<td>92836</td>
<td>12133</td>
<td>3</td>
<td>1.0976</td>
<td>2.9942</td>
</tr>
<tr>
<td><em>P. rus</em></td>
<td>pr_ctrl_1</td>
<td>Control</td>
<td>228821</td>
<td>225378</td>
<td>186338</td>
<td>174427</td>
<td>11911</td>
<td>4</td>
<td>1.2593</td>
<td>3.2371</td>
</tr>
<tr>
<td><em>P. rus</em></td>
<td>pr_ctrl_2</td>
<td>Control</td>
<td>169340</td>
<td>163304</td>
<td>129391</td>
<td>122758</td>
<td>6633</td>
<td>4</td>
<td>0.9426</td>
<td>2.2207</td>
</tr>
<tr>
<td><em>P. rus</em></td>
<td>pr_ctrl_3</td>
<td>Control</td>
<td>140653</td>
<td>137012</td>
<td>79020</td>
<td>62788</td>
<td>16232</td>
<td>2</td>
<td>0.3873</td>
<td>1.2955</td>
</tr>
<tr>
<td>Neg. Control</td>
<td>NA</td>
<td>NA</td>
<td>151253</td>
<td>16974</td>
<td>6984</td>
<td>509</td>
<td>6475</td>
<td>NA</td>
<td>NA</td>
<td>NA</td>
</tr>
</tbody>
</table>

Only samples with at least ≥100 sequences annotated to *nifH* clusters I–III were retained for downstream analyses.
FIGURE 1 | \textit{nifH} sequencing overview. Percentage of retained sequences after quality control and minimum entropy decomposition (MED) that were annotated as true nitrogenase genes belonging to \textit{nifH} Clusters I–III (black) relative to the total amount of post-QC retained sequences. Only the 25 samples with \( \geq 100 \) true \textit{nifH} sequences are shown and were kept for all subsequent analyses.

Diazotroph Community Composition of Coral Hosts From High \( p\text{CO}_2 \) and Ambient \( p\text{CO}_2 \) Sites

Based on the retained 25 OTUs, the highest number of distinct diazotrophs was found among samples of \textit{P. damicornis} (14 OTUs), followed by \textit{G. fascicularis} (12 OTUs), \textit{P. rus} (8 OTUs), and \textit{A. secale} (5 OTUs). Generally, there was little overlap in diazotroph community composition between coral species (Figure 3A). For each coral species, at least 25% of the OTUs were not observed in any other species, and none of the 25 OTUs was present in all four host species. Apart from the differences between coral species, diazotroph assemblage between colonies of the same species was also largely variant (Figure 3B). We could not detect any OTU that was consistently present across all samples of a given coral species, site, or host species-site combination with one exception. OTU_05 (Gammaproteobacteria) was found in all but one control site sample of \textit{G. fascicularis} (Figure 3C). Further, we found three OTUs that were shared between three coral host species: OTU_31 (Cyanobacteria) was present in \textit{P. damicornis}, \textit{G. fascicularis}, and \textit{A. secale}, while OTU_18 and OTU_23 (both Gammaproteobacteria) were present in \textit{P. damicornis}, \textit{G. fascicularis}, and \textit{P. rus}. Thus, we found little similarity in diazotroph abundance between coral species as well as between colonies of a given coral species. A principal
coordinate analysis of the Bray-Curtis dissimilarity matrix of relative OTU abundances between samples did not reveal a pattern of distinction between the seep and control site (Figure 3B; PERMANOVA: $F = 0.88$, $p = 0.53$). Rather, coral host species explained more of the variation between samples than site (coral species $R^2 = 0.22$; site $R^2 = 0.03$), with $G. \textit{fascicularis}$ and $P. \textit{rus}$ samples clustering closer together. On the other hand, samples of $P. \textit{damicornis}$ and $A. \textit{secale}$ appeared more spread out on the ordination. PERMANOVA analysis further supported this observation ($F = 1.93$, $p = 0.0041$) and pairwise comparisons indicated that the difference stems from the difference between $G. \textit{fascicularis}$ and $P. \textit{damicornis}$ samples.
Coral samples were associated with an average of only 3.08 ± 0.26 OTUs. In general, each sample was dominated by one or two highly abundant OTUs (Figure 3C). The OTU with the highest relative abundance in every sample made up on average 82.22 ± 4.02% of $\text{nifH}$ sequences. Alpha diversity was generally very low among all samples (Table 1). There were no differences between samples or species from the control and seep site when looking at Shannon diversity (Wilcoxon = 45.5, $p = 0.081$), Inverse Simpson’s index (Wilcoxon = 44.5, $p = 0.073$), and Chao1 estimator (Wilcoxon = 54.5, $p = 0.195$) for all host species (Supplementary Figure 3), or when only samples

(Supplementary Table 2; $F = 2.84$, FDR-adjusted $p = 0.03$) as well as between $G.\ fascicularis$ and $A.\ secale$ ($F = 2.38$, FDR-adjusted $p = 0.03$). There was no difference between sets of samples when site and coral host species were considered as factors ($F = 0.75$, $p = 0.70$).

FIGURE 3 | Diazotroph community composition of coral hosts from a high $p$CO$_2$ and an ambient $p$CO$_2$ site (A) Venn diagram showing number of observed OTUs and their distribution across the four examined coral species. (B) Principal coordinate analysis of Bray-Curtis distance matrix for all 25 coral samples with ≥100 sequences from the 25 cluster I-III $\text{nifH}$ representative OTUs. First two principal coordinates are shown, explained variance in parentheses. (C) Relative abundances of OTUs per sample. OTUs are sorted according to their phylogenetic relationship as indicated by the dendrogram to the right. Class level phylogenetic association is shown in different dot colors. Red highlighted OTUs are differentially abundant between $P.\ damicornis$ samples from the control and seeps site according to LEfSe analyses. OTU_31 (highlighted in green) was used for relative gene copy abundance estimates using qPCR.
FIGURE 4 | Validation of nifH amplicon sequencing results via qPCR based on OTU_31. OTU_31 annotates as a Cyanobacterium with broad prevalence across *P. damicornis* samples. The qPCR fold-change in *nifH* gene copy numbers was calculated relative to the 16S rRNA amplicon and shows high correlation to the relative sequence abundance determined from the *nifH* amplicon sequencing dataset across the 11 samples of *P. damicornis* (Pearson correlation coefficient $r = 0.87$, $p = 4 \times 10^{-4}$).

Testing for differential relative abundance of OTUs (LefSe analysis) between the high $p$CO$_2$ and ambient site identified no OTUs that were indicative for either the seep or control site across all four coral host species. When only *P. damicornis* samples were considered, however, there were clear differences in the distribution of diazotroph taxa between samples from both sites. At the class level, Gammaproteobacteria were more prevalent at the ambient site, while Alphaproteobacteria, specifically OTU_46 (closest matching to family Rhodobacteraceae), were more prevalent at the high $p$CO$_2$ site (LDA effect size = 5.41, $p = 0.013$). For control samples, OTU_03 (family Vibrionaceae) was identified as differentially abundant (LDA effect size = 5.13, $p = 0.036$), with Gammaproteobacteria as a whole being discriminant (LDA effect size = 5.40, $p = 0.015$). By contrast, in *G. fascicularis*, none of the OTUs exhibited a significant differential abundance between the high $p$CO$_2$ and control site.

**DISCUSSION**

The functioning of the coral-algae symbiosis depends on the low, yet stable availability of nitrogen in the holobiont (Rädecker et al., 2015). The present study examined diazotroph,
i.e., nitrogen-fixing communities, of four coral species from an ambient and a high pCO₂ site in Papua New Guinea using amplicon sequencing of the \textit{nifH} subunit of the nitrogenase gene to disentangle how host-identity and environmental conditions shape diazotroph assemblages in stony corals.

### High False Positive Rates in \textit{nifH}

#### Sequencing Pose Unique Challenges

When Analyzing Diazotroph Assemblages

To analyze \textit{nifH} amplicon sequencing data, we employed a pipeline based on (Lesser et al., 2018) and modified using methods from Angel et al. (2018) to confidently assign sequences to \textit{nifH} clusters. Amplification of the \textit{nifH} gene is commonly achieved through the use of highly degenerate primers at high concentrations using a high number of PCR cycles. The consequential amplification and sequencing of homologous sequences that do not represent a functional nitrogenase subunit is inevitable. Thus, it is important for downstream analyses to identify and remove such false positives. Employment of three different methods to identify such sequences of \textit{nifH} cluster IV and V (\textit{nifH} homologs without known function and bacteriochlorophyll synthesis genes) (Raymond et al., 2004), we found that only 14.1\% of sequences could be considered a functional nitrogenase gene in the present dataset. This is in line with what (Lesser et al., 2018) and (Angel et al., 2018) observed using the same primers. Notably, we found that the amount of functional \textit{nifH} sequences was highly variable across individual samples, which varied widely in their overall read depths as well as the read depths of cluster I–III \textit{nifH} sequences.

After removal of all cluster IV/V homologs of \textit{nifH} from the initial 88 OTUs, 27 of them were identified as belonging to true, functional nitrogenase genes. Most of these nitrogenase sequences belonged to cluster I Mo-Fe nitrogenases, while cluster II and III were only represented by one OTU each. These clusters I–III \textit{nifH} OTUs covered a wide range of phylogenetic backgrounds, with the commonly found groups of diazotroph associates of corals being represented by Cyanobacteria as well as Alpha- and Gamma proteobacteria. Notably, samples in this study contained two cluster I \textit{nifH} OTUs belonging to the Rhizobiales order, closely related to \textit{Bradyrhizobium} sp. and \textit{Rhodopseudomonas oryzae}, respectively. Rhizobiales are commonly reported in diazotroph assemblages of corals (Santos et al., 2014; Lesser et al., 2018; Bednarz et al., 2019), sometimes comprising a large relative abundance of up to 70\% (Lema et al., 2012). In this study, two Rhizobiales OTUs of the family Bradyrhizobiaceae were identified across samples, but they also amplified in the negative DNA extraction control (no sample input) where they were the most abundant. This finding cautions the high abundance and importance of Rhizobiales in coral diazotroph communities proposed previously and highlights the importance of implementing negative controls, as noted previously (Laurence et al., 2014; Salter et al., 2014; Weyrich et al., 2019). Following the above, we excluded these two OTUs (OTU_70, OTU_77) from consideration in downstream analysis.

### Coral-Associated Diazotroph Communities Are Simple and Non-Conserved Within and Between Host Species

Previous studies suggest a mutualistic relationship between corals and their associated nitrogen-fixing bacteria that could provide adaptive potential to the holobiont under a changing environment (Lesser et al., 2007; Cardini et al., 2014; Benavides et al., 2016; Voolstra and Ziegler, 2020). The functional importance of diazotrophs could thus affect the coral holobiont in two opposing ways: (i) diazotroph community assemblages are flexible, thereby allowing to dynamically change in response to changing environmental conditions (Roder et al., 2015; Ziegler et al., 2017; Voolstra and Ziegler, 2020), or (ii) the dependence of the holobiont on fixed nitrogen results in a near obligate association with a low level of flexibility under changing environmental conditions.

We observed that most coral samples harbored only two to five distinct diazotrophs, and each sample was dominated by one or two taxa. The overall validity of the patterns observed in the amplicon sequencing results was confirmed by the high correlation between qPCR-based cycle threshold (C\textsubscript{T}) fold-change difference and relative abundance in the amplicon sequencing data of OTU_31. Thus, considering that we identified on average only one or two OTUs of high relative abundance in most samples, coral-associated diazotroph communities may be even less diverse than reported in previous studies (Lema et al., 2012, 2014; Lesser et al., 2018; Bednarz et al., 2019), at least for the coral species in this study.

Considering all samples analyzed, coral host species were predominantly associated with Alphaproteobacteria, as well as Cyanobacteria (Figure 3C). This is largely in line with previous studies of \textit{nifH}-based analyses in corals that found pre-dominance of Alphaproteobacteria (Zhang et al., 2016; Lesser et al., 2018), Alpha-, Delta-, Gammaproteobacteria as well as Cyanobacteria and Bacteroidetes (Bednarz et al., 2019), or Alphaproteobacteria (majoritively Rhizobiales, but see above) (Lema et al., 2014). Notably, Deltaproteobacteria, although found in other studies (Zhang et al., 2016; Bednarz et al., 2019), were not found in samples from this study.

In \textit{P. damicornis}, diazotroph communities exhibited prevalent association with Gammaproteobacteria at the control site in comparison to samples from the high pCO₂ site that were predominantly associated with an Alphaproteobacterium of the Rhodobacteraceae family (OTU_46, Figure 3C). This resembles findings by Yang et al. (2015) that observed a shifting dominance of Alphaproteobacteria to Gammaproteobacteria in \textit{P. damicornis} in a coast to open sea gradient. Bacteria of the family Rhodobacteraceae are commonly associated with stress, injury, or disease in corals (Sekar et al., 2008; Sunagawa et al., 2009; Roder et al., 2014a,b).

The absence of diazotroph diversity or conserved coral species-specific assemblages are striking and warrant further
investigation regarding the underlying reasons and the broad validity of these observations. From the current data, it seems that diazotrophs may be opportunistically acquired (i.e., diazotroph assemblage is haphazard) from the surrounding environment, based on their functional contribution. As such, it may be unlikely to find a pattern of phyllosymbiosis and cophylogeny as demonstrated previously for coral-associated bacteria (Pollock et al., 2018). This may provide one explanation for the de facto absence of conserved diazotroph community patterns. Considering the notion that corals fix on average less nitrogen than other benthic reef organisms (Cardini et al., 2014), diazotrophy of coral-associated microbes may not be as essential for holobiont functioning as previously thought.

Although volcanic CO₂ seeps offer a unique opportunity to examine the consequences of OA in a natural setting, they offer only a steady-state view of microbiome composition under high pCO₂ conditions and do not allow teasing apart the adaptive potential underlying microbiome assemblage differences. Further studies will be required to complement nifH sequencing results with physiological measurements to disentangle the contribution and implications of diazotroph microbiome composition for holobiont functioning under ambient and high pCO₂ conditions. Ideally, isolating and culturing approaches of differentially abundant OTUs combined with genome sequencing will clarify a potential functional transition from heterotrophic Gammaproteobacteria to autotrophic Alphaproteobacteria under OA, as observed in *P. damicornis*. In addition, *in situ* transplantation experiments of coral fragments between high and ambient pCO₂ sites offer a promising approach to study diazotroph assemblage dynamics.

**DATA AVAILABILITY STATEMENT**

All sequencing data determined in this study are available at NCBI under BioProject PRJNA668126 (https://www.ncbi.nlm.nih.gov/bioproject/PRJNA668126).

**REFERENCES**


**AUTHOR CONTRIBUTIONS**

RR-M, FH, and CRV conceived and designed the study. LG, VM, NR, and GP generated data. LG, VM, and CRV analyzed data. RR-M, FH, and CRV contributed tools, reagents, and supervision. LG, NR, and CRV wrote the manuscript. LG, VM, NR, GP, RR-M, FH, and CRV revised and approved the manuscript. All authors contributed to the article and approved the submitted version.

**FUNDING**

This study was funded to RR-M by the French National Research Agency (ANR, project CARIOLA no. ANR15CE02-0006–01, 2015), by Fonds Pacifique (project AMBITLE no. 1598, 2016), by the Flotte Océanographique Française for using the research vessel Alis and to FH by LabEx-Coral (FLAMENCO project). VM was the beneficiary of a Ph.D. grant from LabEx-Coral (MACADAM project).

**ACKNOWLEDGMENTS**

We would like to kindly thank the KAUST BioScience Core Laboratory (BCL) for assistance with sequencing. We are grateful to the local populations for access to their reef, and to the National Research Institute, the Milne Bay Provincial Research Committee, the New Ireland Provincial Administration, and the Conservation and Environment Protection Authority of Papua New Guinea and the University of Papua New Guinea for permits. Thanks to the crew of the R/V Alis and of the diving boat Chertan.

**SUPPLEMENTARY MATERIAL**

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fmars.2021.754682/full#supplementary-material


**Conflict of Interest:** The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

**Publisher's Note:** All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.