**Cladocopium** community divergence in two *Acropora* coral hosts across multiple spatial scales

Sarah W. Davies¹,² | Kelsey N. Moreland² | Drew C. Wham³ | Matt R. Kanke⁴ | Mikhail V. Matz²

¹Department of Biology, Boston University, Boston, MA, USA
²Department of Integrative Biology, The University of Texas at Austin, Austin, TX, USA
³Department of Biology, Pennsylvania State University, University Park, PA, USA
⁴Department of Biomedical Sciences, College of Veterinary Medicine, Cornell University, Ithaca, NY, USA

**Correspondence**
Sarah W. Davies, Department of Biology, 5 Cummington Mall, Boston University, Boston, MA 02215, USA
Email: daviessw@bu.edu

**Funding information**
National Science Foundation, Grant/Award Number: IOS-1258058 and OCE-0928764; National Oceanic and Atmospheric Administration, Grant/Award Number: NA05OAR4301108

**Abstract**
Many broadly-dispersing corals acquire their algal symbionts (Symbiodiniaceae) "horizontally" from their environment upon recruitment. Horizontal transmission could promote coral fitness across diverse environments provided that corals can associate with divergent algae across their range and that these symbionts exhibit reduced dispersal potential. Here we quantified community divergence of *Cladocopium* algal symbionts in two coral host species (*Acropora hyacinthus, Acropora digitifera*) across two spatial scales (reefs on the same island, and between islands) across the Micronesian archipelago using microsatellites. We find that both hosts associated with a variety of multilocus genotypes (MLG) within two genetically distinct *Cladocopium* lineages (C40, C21), confirming that *Acropora* coral hosts associate with a range of *Cladocopium* symbionts across this region. Both C40 and C21 included multiple asexual lineages bearing identical MLGs, many of which spanned host species, reef sites within islands, and even different islands. Both C40 and C21 exhibited moderate host specialization and divergence across islands. In addition, within every island, algal symbiont communities were significantly clustered by both host species and reef site, highlighting that coral-associated *Cladocopium* communities are structured across small spatial scales and within hosts on the same reef. This is in stark contrast to their coral hosts, which never exhibited significant genetic divergence between reefs on the same island. These results support the view that horizontal transmission could improve local fitness for broadly dispersing *Acropora* coral species.

**KEYWORDS**
*Acropora*, asexual lineages, *Cladocopium*, community divergence, coral, horizontal transmission, symbiodiniaceae, symbiosis

**1 | INTRODUCTION**

Many well-known symbioses involve the passing of symbionts from parents to offspring (vertical transmission), fully aligning the evolutionary trajectories of symbiotic partners and typically leading to their deep integration at biochemical and genomic levels (i.e., *Buchnera* in aphids: Nakabachi et al., 2014; Shigenobu & Wilson, 2011). The result of such symbiosis is essentially a novel composite organism, often called the "holobiont", upon which selection can act (Bordenstein & Theis, 2015). In other types of symbioses, the association between partners must be established anew each generation (horizontal transmission), which offers the host's offspring the opportunity to sample a variety of symbiont lineages and select partners that potentially confer some sort of local advantage (Hilario et al., 2011; Schwarz et al., 1999; Usher et al., 2007). In theory, this kind of relationship should generate novel ecological
opportunities for both symbiotic partners through their mixing and matching across environments. For example, association with ecologically specialized algal photobionts can lead to distinct ecological guilds of lichens (Peksa & Skaloud, 2011) or allow a fungal partner to expand its geographic range across a more broad climatic envelope (Fernandez-Mendoza et al., 2011). Similarly, in aphids, association with various horizontally transmitted bacterial symbionts allows these insects to colonize novel host plants across climatic zones (Henry et al., 2013).

Associations with algal symbionts in the family Symbiodiniaceae are obligatory for the majority of shallow water tropical corals since they rely on photosynthetic byproducts from the algae for energy in oligotrophic waters. In turn, the algae benefit from a protected and light-exposed residence as well as inorganic nutrients and CO\textsubscript{2} concentration mechanisms provided by the host (Barott et al., 2015; Muscatine, 1990; Muscatine & Cernichiari, 1969; Trench & Blank, 1987). Coral symbiosis, like many other ecologically important symbioses, is endosymbiotic (occur within cells) and can establish by two fundamentally different modes of transmission: vertical (symbiont inheritance from mother) and horizontal (symbiont from environmental, free-living sources) (Harrison & Wallace, 1990). Vertically transmitting corals guarantee the maintenance of symbiosis in their offspring, however if larvae encounter novel environments, their symbiotic composition may be suboptimal resulting in reduced fitness (Byler et al., 2013; Douglas, 1998; Wilkinson & Sherratt, 2001). During horizontal transmission, aposymbiotic larvae have flexibility in symbiont acquisition and upon arrival to new environments, they can uptake novel symbionts not present in parental populations (Abrego et al., 2009; Ali et al., 2019; Gómez-Cabrera et al., 2008; Little et al., 2004), but availability of symbionts upon arrival is not guaranteed.

Given the obligatory nature of this symbiosis for the host, it is somewhat surprising that in the majority of coral species (~85%), algal symbionts must be acquired by the juvenile coral from its local environment post settlement (Baird et al., 2009; Fadlallah, 1983; Harrison & Wallace, 1990; Hartmann et al., 2017). However, this prevalence of horizontal transmission in coral-algal symbiosis is consistent with a recent meta-analysis on transmission modes in bacteria-eukaryotes. This study demonstrated that horizontal transmission was the dominant transmission mode in marine environments (Russell, 2019). One possible benefit to this horizontal transmission strategy in marine environments is that these aposymbiotic coral larvae can disperse great distances with ocean currents (Davies et al., 2015; Foster et al., 2012; van Oppen et al., 2011; Rippe et al., 2017). Yet, coral larvae can encounter a great variety of reef habitats (Gorospe & Karl, 2011), and therefore conditions on the reef where they eventually settle can be very different from their natal reef (Baird et al., 2007; LaJeunesse et al., 2004). To improve their chance of survival in this novel environment, corals could potentially associate with locally available, and putatively ecologically specialized, algal strains (Byler et al., 2013; Howells et al., 2012; Rowan & Knowlton, 1995).

Indeed, the diversity of algal symbionts in the family Symbiodiniaceae is rich (LaJeunesse et al., 2018) and specific coral-algal associations have been suggested to play pivotal roles in holobiont adaptation to climate change (Berkelmans & van Oppen, 2006; Howells et al., 2012). The genus Cladocopium (formerly clade C Symbiodinium; LaJeunesse et al., 2018) originated and diversified most recently among Symbiodiniaceae, and has achieved the highest diversity of all lineages (Lesser et al., 2013; Pochon & Gates, 2010; Pochon et al., 2006; Thornhill et al., 2014, 2017). This diversity has been associated with functional variation in symbiont thermal performance across reefs (Davies et al., 2018; Howells et al., 2012) as well as with functional differences in gene expression between reef zones (Barfield et al., 2018; Davies et al., 2018), lending support for the potential for reef-specific symbiont communities. In addition, the draft genome of Cladocopium goreau confirmed the divergence of this genus from other Symbiodiniaceae genera and specifically highlight that gene families related to the establishment and maintenance of symbiosis (photosynthesis, host-symbiont interactions, nutrient exchange) were under positive selection (Liu et al., 2018).

However, much less is known about the population biology of Cladocopium spp. algal symbionts, including how their populations are structured in comparison to their coral hosts. Understanding the relative importance of reef environment, coral host, and geographical distance in structuring coral-associated algal symbiont communities is essential to identifying the adaptive capacity of this symbiosis. However, thus far, there are only a handful of population genetic studies of Symbiodiniaceae based on multilocus markers, none of which address all of the above-mentioned potential sources of genetic variation. Here, using microsatellites, we examined the community divergence of Cladocopium spp. algal symbionts hosted by two common, co-occurring species of Acropora corals—A. hyacinthus and A. digitifera—collected from the same reef locations across the Micronesian Pacific (Figure 1a,b). We explore this community divergence across several ecological scales including host species, islands across the Micronesian archipelago, and unique sites within each island. We then discuss these results for the algal symbionts to the previously published population genetic structure of their coral hosts (based on a subset of the exact same coral samples), which demonstrated that both coral species exhibited extensive genetic connectivity and their genetic structure was well explained by the biophysical connectivity between sites (Davies et al., 2015).

2 | MATERIALS AND METHODS

2.1 | Sampling of coral-associated algal symbionts

This study comprised a subset of samples previously analysed for coral host genetics in Davies et al. (2015) (Table 1, Figure 1a). A total of 25 individuals of each coral host species (Acropora hyacinthus and Acropora digitifera) were examined at two reef sites within each of seven islands (Figure 1b). There were two exceptions to this sampling design. First, at Ngulu only one site was visited and only
A. hyacinthus was collected. Second, at Guam, no A. hyacinthus was found on either of the sampled reefs, so only A. digitifera was collected. In total, 13 reef sites were included in this experimental design. All samples were collected between 3–7 m depth, all colonies were >2 m apart, and all samples from both species at a given site were collected at the same approximate GPS coordinates (Table 1).

### 2.2 Laboratory procedures

DNA was isolated following Davies et al. (2013). Microsatellite primers consisted of five previously described Cladocopium-specific loci (previously described as clade C Symbiodinium) (Bay et al., 2009; Wham et al., 2014) and one novel locus mined using MsatCommander (Faircloth, 2008) from nucleotide EST data for Cladocopium lineage C3 (Leggat et al., 2007), for a total of six loci (Table S1). Loci were multiplexed in 20 µl polymerase chain reaction (PCR) mixtures containing 10 ng of template DNA, 0.1 µM of each forward and reverse primers, 0.2 mM dNTP, 1X ExTaq buffer, 0.025 U ExTaq Polymerase (Takara) and 0.0125 U Pfu Polymerase (Promega). Amplifications began at 94°C for 5 min, followed by 35 cycles of 94°C for 40 s, annealing temperature for 120 s, and 72°C for 60 s and a 10 min 72°C extension period. Molecular weights were analysed using the ABI 3130XL capillary sequencer. Data were binned by repeat size and individuals failing to amplify at ≥3 loci were excluded from downstream analyses.

### 2.3 Analyses of allele presence-absence data

Although Symbiodiniaceae in hospite are assumed to be haploid (Santos & Coffroth, 2003), the genus Cladocopium are generally
observed to have two copies of every allele (Thornhill et al., 2014; Wham et al., 2014; Wham & LaJeunesse, 2016). This apparent genome duplication may or may not correspond to a change in chromosome number, or the actual diploid state (Wham & LaJeunesse, 2016), and it has been previously suggested that these lineages should be scored as if they were effectively diploid (i.e., with the expectation of two alleles per locus) to appropriately construct multilocus genotypes (MLGs) from samples (LaJeunesse et al., 2014; Pettay et al., 2015; Thornhill et al., 2014; Wham et al., 2014; Wham & LaJeunesse, 2016). However, ploidy of the Cladocopium samples in our study is unknown, and a single coral could potentially contain several genetically distinct Cladocopium clones. Therefore, data were analysed as “communities of alleles”, i.e., binary (allele presence/absence) values for each sample. This conservative approach recognizes that each multilocus genotype (MLG) could represent multiple genomes from mixed Cladocopium lineages and allowed us to retain all individuals in analyses (569 total: 282 A. digitifera and 287 A. hyacinthus). The drawback of this approach is that it confounds genetic divergence and community divergence in cases of multiple strains per host. However, since multiple-strain infections are rare in Cladocopium (Thornhill et al., 2017), genetic divergence is expected to be the major contributor to our distance measures. Still, we chose to refer to our distances as “Cladocopium community divergence” throughout, to ensure that there is no confusion with true genetic distances.

First, all binary allele data (N = 569 samples: File S1) were converted into a genind object with allele presence/absence data using adegenet 2.0.0 (Jombart et al., 2010) in R (R Development Core Team, 2018). Next, discriminant analysis of principal components (DAPC) was implemented, which classifies samples into user-defined groups based on their coordinates in principle components’ space. Because DAPC does not rely on traditional population genetics models, it is free from Hardy-Weinberg equilibrium and linkage disequilibrium assumptions and thus is considered to be applicable across organisms regardless of their ploidy and genetic recombination rates (Jombart et al., 2010). Here, identification of clusters was achieved using the find.clusters function with a maximum number of 40 clusters. Eighteen principle components (PCs) were maintained and the Bayesian Information Criterion (BIC) indicated that two clusters were optimal in our data. In this initial analysis of all data, samples exhibited strong assignments into two highly-supported clusters - light green and dark green (Figure 1c). These data were therefore split into two subsets, corresponding to these two clusters, for downstream analyses. Only samples assigning to one of the two clusters with a probability > 0.9 were retained, resulting in N = 388 for the light green cluster and N = 172 for the dark green cluster (Table 1, Files S2 and S3).

### 2.4 | Sequencing analysis of Cladocopium psbA\textsubscript{ncr}

To confirm phylogenetic affiliation of the two highly-supported Cladocopium clusters, the non-coding region of the psbA chloroplast gene (psbA\textsubscript{ncr}) was amplified in representative samples. The
psbA\textsuperscript{ncr} region was chosen because of its utility for differentiating species of Symbiodiniaceae (i.e., Lewis et al., 2019). Amplifications were conducted using the primers and settings described by Moore et al. (2003). Amplified products were directly sequenced using the reverse primer. Phylogenetic analysis of psbA\textsuperscript{ncr} reference sequences for C40 (from various scleractinians) and C21 (from Acropora), provided by the LaJeunesse laboratory, was conducted using PAUP Version 4.4a147 (Swofford, 2014) using maximum parsimony. Statistical significance was confirmed via bootstrapping (based on 1,000 replicates). A nexus file (File S4) was used to generate an unrooted phylogenetic tree to demonstrate that the representative samples from the two highly-supported clusters are separated by large differences in sequence divergence. These results indicated that the two major clusters by our genetic data were C40 (sensu LaJeunesse et al., 2004) and C21 (sensu Thornhill et al., 2014), respectively. Further community divergence analysis of these data were completed for each cluster separately and these lineages are referred to as \textit{Cladocopium} C40 and \textit{Cladocopium} C21 throughout the rest of the paper.

### 2.5 | Analyses of asexual lineages within each cluster

To determine the prevalence of identical asexual lineages within \textit{Cladocopium} C40 and C21, individuals with matching MLGs were investigated. Singleton alleles were removed (12 alleles in C40; 10 alleles in C21). Genotypic identity, the probability that two MLGs sampled without replacement from the data set were identical, was calculated as \( G_i = \sum p_i^2 \), where \( p_i \) is the frequency of the \( i \)th repeated MLG. Genotypic diversity, the probability that two MLGs sampled without replacement from the data set were different, is the complement of genotypic identity: \( G_0 = 1 - G_i \). Hierarchical clustering tree was constructed in R (R Development Core Team, 2018) using \texttt{vegdist(x, binary = T, method="manhattan")} function from the vegan package (Oksanen et al., 2013) and processed with the function \texttt{hclust(x)}. Manhattan distance was chosen for this analysis because it corresponds to the total number of nonshared alleles between two MLGs, which is zero for identical MLGs. Samples sharing identical MLGs were then identified using the function \texttt{cutree(x, h = 0.2)}, which grouped samples with less than one (i.e., zero) nonshared alleles. The probability of chance occurrence of identical MLGs was assessed by a resampling simulation in R. To create simulated MLGs assuming random sorting of alleles, we first created a matrix of allele presence-absences, where rows were samples and columns were alleles using 44 non-singleton alleles in C40 and 46 non-singleton alleles in C21. In each column, 1 marked the presence and 0 marked the absence of an allele. Then, 100,000 simulated MLGs were created by taking a single random draw from each column. In this way, the probability of sampling an allele is equal to its frequency in the total population, and multiple alleles per locus can be sampled since the allele presence-absence matrix did not contain locus information. The probability that \( n \) MLGs in the data set were identical by chance was calculated using the formula \( \left(\frac{a}{n}\right)^{n-1} a \), where \( a \) is the number of times the MLG was observed in the simulation and \( n \) is the number of times it was observed in the actual data. For downstream DAPC analyses, we have created dataframes including only a single representative of each MLG within a site within the same host species (Files S5 and S6 for C40 and C21, respectively).

The geographical distances spanned by MLGs were investigated by calculating a distance matrix from reef site coordinates, in decimal degrees, using the dist function in R. The \texttt{dms2dec} function from Zanolla et al. (2018) was used to convert degrees minutes seconds to decimal degree format. The largest distance was taken for MLGs spanning more than one site. Distances were converted to kilometres using the National Hurricane Center and Central Pacific Hurricane Center’s calculator (https://www.nhc.noaa.gov/gccalc.shtml). Differences in per site genotypic diversity between C40 and C21 and between coral host species were tested using \texttt{t.test()} function in R. Differences in frequencies of repeated MLGs between C40 and C21 was tested using the function \texttt{chisq.test()} based on Monte Carlo simulations with 10,000 replicates (Hope, 1968).

### 2.6 | Within-cluster analyses across coral hosts, islands and sites within islands

To visualize \textit{Cladocopium} community divergence between host species, between islands, and between sites and host species within each island, assignment of samples to genetic clusters with prior grouping of island/host/site was performed in R (R Development Core Team, 2018) using DAPC (Jombart, 2008; Jombart et al., 2010) separately for C40 and C21. Successful reassignment, indicated by a high proportion of samples correctly assigning back to their a priori group, indicates that these user-defined groups are distinct, which in our case implies divergence between in hospite \textit{Cladocopium} communities. Here, data were converted into principal components (PCs) and then \( a \)-scores were computed to determine the optimal number of PCs to retain. \( a \)-scores determine the proportion of successful reassignment corrected for the number of PCs retained and protect against model overfitting (Jombart et al., 2010). Assignment rates, PCs and discriminant functions (DF) retained, and the overall proportion of variance explained by each of the models are included in Table S2. Proportion of successful assignments within each model are shown on all figures.

### 2.7 | Unconstrained \textit{Cladocopium} community analyses

Because DAPC analyses aim to maximize variation between predefined groups, we also visualized all C40 and C21 data independently using a principal coordinate analysis of allele presence-absence data using the \texttt{vegdist(x, method="bray")} function implemented in the vegan package in R (Oksanen et al., 2013). \textit{Cladocopium} community divergences between host species,
FIGURE 2  Repeated MLGs (asexual lineages) in Cladocipium. Fan trees of Cladocipium (a) C40; and (b) C21 MLGs. Host species, A. hyacinthus and A. digitifera, are colour-coded on the inside of the tree and the seven islands in Micronesia are indicated in the ring around the tree. (c) Frequencies of repeated MLG group sizes. C40 has larger repeated MLG groups than C21. (d) Frequencies of repeated MLGs spanning hosts, binned by MLG group size. There is no clear difference in the proportion of host-spanning MLGs between C40 and C21. (e) Greatest geographical distance spanned by a MLG group of a given size. C21 MLGs span considerably larger distances than C40 MLGs. (f, g) Proportions of coral colonies hosting repeated MLGs at each reef site in each host species. Bar colours correspond to host species, where faded bar segments represent unique MLGs and bright bar segments represent identical MLGs. h, A. hyacinthus; d, A. digitifera. Reef site colours correspond to Figure 1 and Table 1 [Colour figure can be viewed at wileyonlinelibrary.com]
islands and host species and sites within islands were then tested with a distance-based PERMANOVA using the vegan::adonis function (method = "bray").

2.8 | Data and code availability

All data and code used for all analyses and figure generation are publicly available at https://github.com/daviessw/Cladocopium_Micronesia.

3 | RESULTS

3.1 | Two clusters of Cladocopium symbionts observed in Micronesian acroporids

Across the two coral host species in Micronesia (Figure 1a,b), two distinct Cladocopium clusters were observed with 98.4% of samples (560/569) strongly assigning to one of the two clusters (Figure 1c). Sequencing of the psbA gene from representative samples from each cluster identified them as Cladocopium C40 and C21 (LaJeunesse et al., 2004; LaJeunesse & Thornhill, 2011; Thornhill et al., 2014) (Figure S1). It is important to note that the possible presence of other background Symbiodiniaceae genera would not affect Cladocopium genotyping results since our microsatellite assays are genus-specific (Bay et al., 2009; Wham et al., 2014). Corals of both Acropora species from Palau and Ngulu were found to almost exclusively host Cladocopium C40 (Figure 1c, dark green bars). C40 was also prevalent in A. digitifera at one reef site on Yap (Goofnuw Channel: GO.2) and was occasionally found in A. digitifera throughout Micronesia (Figure 1c). All other Acropora hosts associated with Cladocopium C21 (Figure 1c, light green bars). Both Cladocopium lineages possessed high allelic diversity, with a total of 44 unique alleles in C40 (N = 127 corals) and 49 unique alleles in C21 (N = 328 corals).

3.2 | Asexual lineages in Cladocopium symbionts

C40 comprised a total of 105 unique MLGs, 22 of which were found more than once (Figure 2a). In C21 there were 309 unique MLGs, 53 of which were found more than once (Figure 2b). Using resampling simulations, we determined that 16 out of 22 repeated MLGs in C40 and all 53 repeated MLGs in C21 were unlikely to occur due to random assortment of microsatellite alleles (p < .001). Six repeated MLGs in C40 were less robustly supported (p-values ranging from 0.0014 to 0.0395), but all still passed the p < .05 significance threshold. We therefore posit that repeated MLGs constitute evidence of identity by descent, i.e., represent lineages descending by asexual reproduction from a common MLG ancestor. Asexual lineage group size was on average 4.05 for C40 and 2.49 for C21, ranging up to 14 in C40 and five in C21. This difference was
significant ($p = .013$, Figure 2c,d). For the whole data set, the genotypic identity level (probability that two randomly sampled MLGs are identical) in C40 was almost 10-fold higher than in C21 (0.0204 versus 0.0024), but this difference was not readily apparent when per-reef measures of genotypic diversity were compared ($p = .13$; Figure S2a). There was also no significant difference in overall genotypic diversity of algal symbionts (of any type) hosted by the two coral species ($p = .9$; Figure S2b). Summaries of proportions of repeated MLGs for each reef site are shown in Figure 2f,g.

Notably, many asexual lineages spanned host species, reef sites, and even islands (Figure 2a,b). Larger group size in C40 compared to C21 did not translate into larger geographic distance spanned by an asexual lineage (Figure 2e). The largest distance spanned by C40 lineages was between Goofnuw Channel (GO.2), Yap and Lighthouse Reef (LH.1), Palau (~578 km), while the largest distance spanned by C21 lineages was between South Tip (ST.1), Yap and Hiroshi Point (HP.2), Kosrae (~3,714 km) (Figure 2e). Proportions of asexual lineages spanning host species also differed between C40 and C21: 36.4% of them spanned host species in C40 (Figure 2a), compared to 20.8% in C21 (Figure 2b).

3.3 Cladocopium community divergence by coral host species, islands, and local reef environments

Unlike MLGs that occurred repeatedly and thereby could be attributed to individual asexual lineages, singleton MLGs could represent
individual symbiont genotypes or mixtures of genotypes hosted by the same coral. Therefore, all MLGs were analysed as "communities of alleles", making no genetic assumptions. Discriminant analysis of principal components (DAPC) strongly differentiated between host species for both Cladocopium C40 and C21 (Table S2, Figure 3a,b), with assignment rates ranging from 0.75 (A. hyacinthus hosting C21) to 0.93 (A. digitifera hosting C40). In addition, unconstrained analyses confirmed that distinctions between host species was significant for both C40 and C21 (PERMANOVA p < .001) (Figures S3a,b). These results confirm that host species play a role in structuring Cladocopium communities across Micronesia. In addition, DAPC demonstrated clustering among islands for each Cladocopium species irrespective of host species: generally high per-island assignment rates were obtained both for C40 (Figure 3c, 0.61–0.91) and C21 (Figure 3d, 0.54–0.88), which were also confirmed using unconstrained analyses for both C40 and C21 (PERMANOVA p < .001, Figures S3c,d). Notably, algal symbionts from Yap consistently showed some of the lowest assignment rates for both C40 (0.61) and C21 (0.63). Another notable fact was that algal symbiont communities from Ngulu and Kosrae were highly distinct, suggesting the possibility of additional Cladocopium lineages (besides C40 and C21) existing there (Figure 3c), which was not further explored here.

When clustering was performed within islands for C40 (Palau) and C21 (Yap, Chuuk, Pohnpei, Kosrae, Guam), of the two top eigenvalues in DAPC analysis, generally one discriminant function (DF) explained Cladocopium community divergence by host species while the other DF explained differences between reef sites (Figure 4). Unconstrained analyses corroborated this result: Cladocopium communities were always significantly different between coral host species and sites within islands (Figure S4). There was only one instance when DAPC and unconstrained analysis did not show strong support for clustering by sites and host species within island: C21 from Yap (Figures 4b, S4b). This is probably due to unbalanced sampling of sites:symbiont groups for A. digitifera: this species showed high prevalence of C21 relative to C40 across all Yap sites except Goofnuw channel, where C40 was more prevalent (Figure 1c). The strongest separation between host:site groups was observed at Chuuk (Figures 4d, S4d) and at Kosrae (Figures 4e, S4e).

4 | DISCUSSION

4.1 | Acropora corals establish symbiosis with distinct Cladocopium communities

Across the Micronesian Pacific (Figure 1a), both Acropora coral hosts associated with two distinct lineages of Cladocopium (Figure 1c), which were identified as C40 and C21 (Figure S1), with the potential for additional species present (e.g., the highly distinct C21 from Ngulu, Figure 3c). This observation suggests that both coral hosts show flexibility in their symbiotic associations with Cladocopium across their range and within their specific environments (Abrego et al., 2009; Berkelmans & van Oppen, 2006). This association with Cladocopium is consistent with the wealth of previous community composition studies suggesting that Indo-Pacific acroporids are dominated by algal symbionts in this genus (i.e., LaJeunesse et al., 2003, 2004; Thornhill et al., 2014). Initial symbiont infection is probably determined by local availability of symbionts, either free-living or, those that have been recently evacuated from local coral hosts (Thornhill et al., 2017). Diverse infections are made possible by the flexibility of arriving coral recruits (Abrego et al., 2009; Ali et al., 2019; Cumbo et al., 2013; Little et al., 2004). After infection, a winnowing process - competition between symbiont strains modulated both by the host and by the environment - leads to the eventual dominance of a single asexual lineage of symbionts in a single host colony and distinct symbiont communities across coral hosts in a specific habitat (Rowan et al., 1997; Thornhill et al., 2017).

Strict associations of a single coral with a single Symbiodiniaceae asexual lineage have been observed across a variety of coral species and Symbiodiniaceae genera (Baums et al., 2014; Pinzón et al., 2011; Thornhill et al., 2014), however this is not always the case (see Howells et al., 2009, 2013). In our study, it is also important to acknowledge that we only explored community divergence patterns within Cladocopium because we leveraged Cladocopium-specific microsatellite loci (Bay et al., 2009; Wham et al., 2014). This genus is most commonly known to associate with Acropora in this region, which is consistent with our previous ITS2 metabarcoding results on the same coral samples from Palau reefs, which showed that Acropora hosts strictly associated with one of two Cladocopium symbiont haplotypes (Quigley et al., 2014). Here, we tested several samples (N = 4) for community level algal species identification (Figure S1), which confirmed C40 and C21 designations; however these more coarse-grained genus-level analyses were not performed on samples from across the range. Therefore, we are unable to comment on other algal genera known to inhabit corals at background levels (Silverstein et al., 2012; Ziegler et al., 2018).

4.2 | Distinct asexual lineages within Cladocopium C40 and C21

We posit that symbiont MLGs shared between coral colonies represent asexual lineages descending from the same MLG ancestor, because, as we demonstrate through resampling simulations, repeated occurrence of an MLG through random sorting of alleles is highly unlikely. Note that we call groups of shared MLGs “asexual lineages” rather than “clones”, to recognize that their representatives might have accumulated mutations throughout their genomes since their divergence from the common ancestor, despite retaining the ancestral MLG at the six microsatellite loci analysed here. Previous Symbiodiniaceae studies based on microsatellite loci demonstrated that rates of MLGs sharing can differ substantially between Symbiodiniaceae genera, between lineages within a genus, and between regions (Thornhill et al., 2017). For example, work on D. trenchii hosted by Acropora colonies found very low rates of shared MLGs between colonies
(Hoadley et al., 2019), and similarly low rates have been observed in Dusurdinium from Galaxea fascicularis from the South China Sea (Chen et al., 2020). However, Pettay et al. (2011) found that unique Pocillopora hosts frequently associated with the same S. glynni MLG. Here, we find 22 repeated MLGs in C40 and 53 in C21, which account for 51.7% of C40 corals and 34% of C3 corals (Figure 2a,b). Unlike Caribbean Acropora (Baums et al., 2014), all coral hosts analysed here represent distinct genets (i.e., the small proportion of clones detected in Davies et al., 2015 were avoided) and therefore sharing of symbiont MLGs cannot be attributed to clonality of their hosts. While few studies have investigated MLG sharing in Pacific Cladocopium, Howells et al. (2013) found that only 13% of A. millepora from the Great Barrier Reef hosted identical MLGs. However, rates of MLG sharing appear to be different across Cladocopium species. For example, Thornhill et al. (2014) observed that 17% of C3 hosted by Siderastrea siderea, 70% of C7 hosted by Orbicella spp., and 47% of C7a/C12 hosted by Orbicella spp. represented shared MLGs. In light of these data, the prevalence of asexual lineages that we have observed are well within previously published estimates.

Interestingly, we found that Cladocopium asexual lineages were not only shared across conspecifics on the same reef, but also across different host species, different reefs on the same island, and even between host species on different islands (Figure 2a,b). Given that unique MLGs have been shown to exhibit functional variation both in culture (i.e., S. psygmophilum, Parkinson et al., 2016), and in hospite (Davies et al., 2018; Howells et al., 2012), these results are counterintuitive for several reasons. First, it is difficult to imagine how an asexual lineage can disperse across such distances, which was especially evident in C21 (Figure 2e), given that the majority of symbioses in corals involve horizontal transmission (Baird et al., 2009) and free-living Symbiodiniaceae are expected to have low dispersal potential (reviewed in Thornhill et al., 2017). Secondly, it is surprising that the same asexual lineage would be successful across both host species and across different environments given that coral-associated symbiont distributions have been proposed to correlate with depth (Andras et al., 2011; Kirk et al., 2009), temperature (Baums et al., 2014; Hume et al., 2016; LaJeunesse et al., 2014), PAR (Rowan et al., 1997), and host species (Thornhill et al., 2014, 2017). Another interesting discussion point is that C21 asexual lineages appear to be more broadly distributed across the seascape than C40 (Figure 2e), suggesting that C21 may have higher dispersal potential than C40. If so, this might explain larger group size in C40 compared to C21: since less dispersal implies less mixing of asexual lineages across locations, the symbiont with less dispersal would be more likely to have larger same-MLG groups detected at any given location. An alternative explanation of the difference between MLG group sizes in C40 and C21 is higher variance in the rates of asexual reproduction among C40 genotypes compared to C21 genotypes (Thornhill et al., 2017).

It is important to note that we are probably underestimating the frequencies of identical asexual lineages given the complexities of peak calling in microsatellite analyses and error rates associated with PCR-based analyses of repeated loci. Our results highlight the urgent need for in-depth population genomic studies of Symbiodiniaceae, which would allow for the investigation of evolution within and among asexual lineages, local adaptation, emergence of novel symbiont-host associations, and interactions between all of these aspects. An effective approach for Symbiodiniaceae genomics would be the recently introduced expression exome capture sequencing (EecSeq, Puritz & Lotterhos, 2018), which would provide a cost-efficient solution to the problem of pervasive host DNA contamination. Intensive sampling of hosts associated with Cladocopium across additional host species and sites coupled with sequencing deeper coverage across the genome will undoubtedly shed light on the population biology of these generalist symbionts.

4.3 | Cladocopium C40 and C21 exhibit imperfect host specificity

The majority of reef-building coral species associate with a specific Symbiodiniaceae type, which have traditionally been coarsely defined based on ribosomal and/or chloroplast markers (Fabina et al., 2013; Rodriguez-Lanetty et al., 2004; Thornhill et al., 2014; Weis et al., 2001). Previous Symbiodiniaceae multilocus genotyping studies revealed that each of these symbiont types can harbour within-type diversity, both at genetic and functional levels (Howells et al., 2009, 2012; Santos et al., 2004). Here we observe significant divergence between Cladocopium communities among two different host species in both C40 and C21 across the Micronesian Pacific (Figures 3a,b; S3a,b), and this pattern of host specificity consistently holds between host species on the same reef (Figures 4; S4). Previous work on octocorals similarly observed significant host differentiation among algal symbionts; however, they found that this genetic divergence was driven by different aged cohorts and depth in their system (Andras et al., 2011). Here, host habitat depth or age class is not relevant for the host specificity observed given that specific attention was paid to collecting colonies located at similar depths and of similar size classes. Instead, our data suggest that for both C40 and C21, local association of hosts and symbionts within the same cluster is due to host specificity in Cladocopium (Figure 4; Figure S4), which has been previously proposed in symbionts hosted by Pocillopora in the south Pacific (Magalon et al., 2006). Since our study rigorously sampled two coral host species across several spatial scales, we also detected that this specificity is imperfect: at every location, there were symbionts in one host species that would have been assigned to another coral host based on their MLG (Figures 4, S4). In fact, there were multiple MLGs both within C40 and C21 that were shared across hosts at the same site and across different islands (Figure 2a,b), further highlighting that this host specificity is imperfect. Overall, these patterns suggest that host specialization in Cladocopium is present, however the boundary between hosts appears permeable in A. hyacinthus and A. digitifera across the spatial scale investigated here.
4.4 | Divergent Cladocopium communities within islands

Within each island and sympatric host species, all Cladocopium pairwise comparisons exhibit high assignment rates back to their a priori groups (Figure 4), which demonstrates significant community divergence between closely located reef sites (Figure S4). It is tempting to speculate that Cladocopium community divergence among individual reefs might be due not only to dispersal limitation, but also to spatially varying selection, implying environmental specialization (i.e., local adaptation) in the symbionts. However, these islands are remote and understudied and therefore we cannot provide further support for this claim as we did not measure environmental parameters and did not assess symbionts’ fitness across environments. Among factors that might contribute to genetic subdivision among reefs irrespective of distance is high variation in reproductive success among Cladocopium asexual lineages on a local scale, which would elevate divergence due to spatial discordance of short-term allele frequency fluctuations (Thornhill et al., 2017). Yet, previous work has demonstrated that other Cladocopium symbiont populations have exhibited classic signals of local adaptation (Howells et al., 2012), and therefore reef sites investigated here offer an excellent study system for investigating the fine-scale local adaptation potential of Cladocopium. If these algal symbionts are indeed locally adapted, this would ensure that horizontally transmitting coral hosts increase their local fitness by associating with local symbionts. To confirm this hypothesis, future work is required to experimentally demonstrate that these symbionts are achieving their maximum fitness in their local reef environment (Kawecki & Ebert, 2004).

4.5 | Cladocopium communities are more spatially structured than their coral hosts

With our conservative approach to analysis of our symbiont genetic data we cannot directly compare the divergence of symbiont communities to the previously published genetic structure of their coral hosts (Davies et al., 2015). Still, we can compare these results qualitatively. For symbiont communities hosted by the same coral species, we consistently find significant divergence between different sites within the same island (Figure 4; Figure S4). In contrast, no significant within-island genetic divergence was ever detected for either host species, using the exact same coral samples (Davies et al., 2015). This indicates that C40 and C21 algal symbiont communities are more spatially structured that their coral hosts across the same spatial scale.

Strong community divergence in Cladocopium was not surprising given the prevailing view of their life cycle. It involves symbiotic existence in sedentary hosts alternating with a short-term free-living form that largely exists in the benthos. The opportunity for Cladocopium dispersal by ocean currents is therefore limited, and the primary role of the free-living stage is to invade novel hosts (Fitt et al., 1981; Fitt & Trench, 1983; Littman et al., 2008; Magalon et al., 2006; Yacobovitch et al., 2004). Our data support this hypothesis with the observation of significant clustering between all pairs of sampled sites within islands in both C40 and C21 lineages (Figures 4, S4), which was never observed in the coral host (Davies et al., 2015). Overall our data support the prevailing view that Symbiodiniaceae dispersal is limited, especially relative to their coral hosts, across the seascape. Still, the fact that several asexual lineages spanned reef sites and even islands highlights the potential for occasional long-range dispersal in Cladocopium, especially in C21 (Figure 2e).

ACKNOWLEDGEMENTS

Thanks to field assistants Carly Kenkel, Tim Keitt, Irina Yakushenok and David Stump. Nida Khawaja Rahman contributed to molecular work and James Derry assisted with data management. We acknowledge the Federated States of Micronesia Department of Resources and Development (#11-27-09-01, #11-27-09-02), Guam permit authorities, and the Republic of Palau Bureau of Marine Resources and Koror State Government (Marine Research permit: #09-201; Marine Resource Export Certification #RE-09-23) for all collection permits and CITES export permits. Todd LaJeunesse helped classify Cladocopium lineages and provided useful advice on results and the biology of Cladocopium. We are grateful to Ulrich Mueller, Dan Thornhill, Emily Howells, and the many reviewers who have offered their critical feedback through several sets of revisions throughout the publication process. This study was supported by the Coral Reef Conservation program of the National Oceanic and Atmospheric Administration (NA05OAR4301108) to M.V.M. DCW was funded in part by Pennsylvania State University and grants from the National Science Foundation (OCE-0928764 and IOS-1258058).

AUTHOR CONTRIBUTIONS

S.W.D., and M.V.M. conceived, designed and coordinated the study and drafted the manuscript. S.W.D., M.R.K., and M.V.M. collected coral samples. S.W.D. carried out molecular laboratory work, participated in data analysis, and carried out statistical analyses; D.C.W., M.R.K., and K.M. participated in data analysis, statistical analyses and interpretation; All authors gave their final approval for publication.

DATA AVAILABILITY STATEMENT

All data are available in Files S1–S6 and all data and code used for all analyses and figure generation are publicly available at https://github.com/daviessw/Cladocopium_Micronesia.

ORCID

Sarah W. Davies https://orcid.org/0000-0002-1620-2278
Kelsey N. Moreland https://orcid.org/0000-0002-3571-0876
Drew C. Wham https://orcid.org/0000-0002-1946-9015
Mikhail V. Matz http://orcid.org/0000-0001-5453-9819


