Deep connections: Divergence histories with gene flow in mesophotic Agaricia corals

Katharine E. Prata1,2, Cynthia Riginos1, Ryan N. Gutenkunst3, Kelly R. W. Latijnhouwers4, Juan A. Sánchez5, Norbert Englebert1, Kyra B. Hay1, Pim Bongaerts1,2,4

1School of Biological Sciences, The University of Queensland, St Lucia, Queensland, Australia
2California Academy of Sciences, San Francisco, California, USA
3Department of Molecular and Cellular Biology, University of Arizona, Tuscon, Arizona, USA
4Caribbean Research and Management of Biodiversity Foundation, Willemstad, Curacao
5Laboratorio de Biología Molecular Marina (BIOMMAR), Departamento de Ciencias Biológicas, Universidad de los Andes, Bogotá, Colombia

Correspondence
Katharine E. Prata, School of Biological Sciences, The University of Queensland, St Lucia, QLD, Australia.
Email: k.prata@uq.edu.au

Funding information
Australian Research Council, Grant/Award Number: DE160101433; Catlin Group; University of Queensland; Explorers Club

Handling Editor: Loren Rieseberg

Abstract
Largely understudied, mesophotic coral ecosystems lie below shallow reefs (at >30 m depth) and comprise ecologically distinct communities. Brooding reproductive modes appear to predominate among mesophotic-specialist corals and may limit genetic connectivity among populations. Using reduced representation genomic sequencing, we assessed spatial population genetic structure at 50 m depth in an ecologically important mesophotic-specialist species Agaricia grahamae, among locations in the Southern Caribbean. We also tested for hybridisation with the closely related (but depth-generalist) species Agaricia lamarcki, within their sympatric depth zone (50 m). In contrast to our expectations, no spatial genetic structure was detected between the reefs of Curacao and Bonaire (~40 km apart) within A. grahamae. However, cryptic taxa were discovered within both taxonomic species, with those in A. lamarcki (incompletely) partitioned by depth and those in A. grahamae occurring sympatrically (at the same depth). Hybrid analyses and demographic modelling identified contemporary and historical gene flow among cryptic taxa, both within and between A. grahamae and A. lamarcki. These results (1) indicate that spatial connectivity and subsequent replenishment may be possible between islands of moderate geographic distances for A. grahamae, an ecologically important mesophotic species, (2) that cryptic taxa occur in the mesophotic zone and environmental selection along shallow to mesophotic depth gradients may drive divergence in depth-generalists such as A. lamarcki, and (3) highlight that gene flow links taxa within this relatively diverse Caribbean genus.

KEYWORDS
cryptic species, isolation with migration, mesophotic, population genetics, Scleractinia, spatial connectivity
1 | INTRODUCTION

Mesophotic coral reef ecosystems lie below the well-studied shallow coral reefs (at ~30–150 m depth) (Lesser et al., 2009) and represent a substantial proportion of the world’s potential coral reef habitat. These reefs have received considerable recent attention due to their hypothesised role as ecological refuges (Bongaerts, Ridgway, et al., 2010; Bongaerts & Smith, 2019; Glynn, 1996; Semmler et al., 2017). Consequently, vertical genetic boundaries have been assessed within “depth-generalist” hard corals, namely those found in both shallow and mesophotic reefs. In such studies, vertical genetic structure has been commonly observed, although often varying among species and location (Bongaerts et al., 2012; Brazeau et al., 2013; Eckert et al., 2019; Serrano et al., 2014, 2016; Studivan & Voss, 2018; van Oppen et al., 2011), refuting the concept of universal maternal inheritance of symbionts (Baird et al., 2008; Miller & Benzie, 1997; Pocillopora spp., Combosch & Vollmer, 2015; Portes spp., Forsman et al., 2017; Psammocora spp., Stefani et al., 2008, Seriatopora spp., Bongaerts, Riginos, et al., 2010; Stylophora spp., Arrigoni et al., 2016; and octocorals, Prada & Hellberg, 2013; Quattrini et al., 2019). Recently, demographic modelling that is sensitive to detecting low levels of historical gene flow has shown evidence of this between taxa (e.g., Cooke et al., 2020; Ladner & Palumbi, 2012; Favia spp., Carlson & Budd, 2002; Madracis spp., Frade et al., 2010; Platygryra spp., Miller & Benzie, 1997; Pocillopora spp., Combosch & Vollmer, 2015; Portes spp., Forsman et al., 2017; Psammocora spp., Stefani et al., 2008, Seriatopora spp., Bongaerts, Riginos, et al., 2010; Stylophora spp., Arrigoni et al., 2016; and octocorals, Prada & Hellberg, 2013; Quattrini et al., 2019). The examination of co-occurring closely related but genetically distinct groups can provide further insights into evolutionary dynamics of divergence with gene flow (Bird et al., 2012; Nosil, 2008). For example, morphologically cryptic and sympatric coral taxa may represent incipient species where specific habitats delineate taxa (e.g., Carlson & Budd, 2002; Warner et al., 2015) or differences in reproductive timing (e.g., Rosser, 2015). Determining the extent to which scleractinians can exchange alleles intra- or interspecifically can provide insights into possible rates of adaptive evolution such as in response to anthropogenic stresses.

Here, we focus on the genus Agaricia (Order: Scleractinia), which is one of the most speciose genera in the Caribbean and perhaps the most dominant group at mesophotic depths (Loya et al., 2019). Agaricia species have predominately plating morphologies and have been described as both hermaphroditic (Fadlallah, 1983) and gonochoric (Kerr et al., 2013). A brooding reproductive mode has been observed (through larval experiments) for three out of the seven species (A. humilis, A. tenuifolia and A. agaricina; Morse et al., 1988) and has therefore been assumed for the genus. Agaricia species are also presumed to have maternal inheritance of symbionts (Baird et al., 2009) resulting in host-endsymbiont specificity and with most species harbouring a distinct Cladoccopium strain (Bongaerts et al., 2013). Within the Southern Caribbean, Agaricia species segregate by depth with some habitat overlap. Three species occur abundantly at mesophotic depths: depth-generalist, Agaricia lamarcki (most commonly found: ~15–50 m), and depth-specialists A. grahamae (~50–90 m) and A. undata (~60–90 m) (Bongaerts, Frade, et al., 2015; Bongaerts et al., 2013). Genetic structure has been assessed previously in Agaricia species, with horizontal spatial structure described in corals (Arrigoni et al., 2019; Bongaerts et al., 2021; Gómez-Corráles & Prada, 2020; Ladner & Palumbi, 2012; Nakajima et al., 2017; Warner et al., 2015) and both homoplaspy and phenotypic plasticity can lead to the misidentification of genetically distinct taxa (Forssman et al., 2009; Kitahara et al., 2016; Paz-García et al., 2015). Incomplete reproductive barriers between closely related taxonomic species are common (reviewed in Willis et al., 2006), as are successful interspecific laboratory crosses (Willis et al., 1997). Regular hybridisation between species or, reticulate evolution, has long been suspected to be an important aspect of coral evolution (Veron, 1995). Multilocus genetic or genomic approaches are moreover uncovering evidence for historical introgression (Mao et al., 2018) and frequent observations of contemporary admixture are consistent with intermixing of semi-differentiated taxa (e.g., Acropora spp., Ladner & Palumbi, 2012; Favia spp., Carlson & Budd, 2002; Madracis spp., Frade et al., 2010; Platygryra spp., Miller & Benzie, 1997; Pocillopora spp., Combosch & Vollmer, 2015; Portes spp., Forsman et al., 2017; Psammocora spp., Stefani et al., 2008, Seriatopora spp., Bongaerts, Riginos, et al., 2010; Stylophora spp., Arrigoni et al., 2016; and octocorals, Prada & Hellberg, 2013; Quattrini et al., 2019). Recently, demographic modelling that is sensitive to detecting low levels of historical gene flow has shown evidence of this between taxa (e.g., Cooke et al., 2020; Ladner & Palumbi, 2012; Prada & Hellberg, 2020; Rippe et al., 2021). The examination of co-occurring closely related but genetically distinct groups can provide further insights into evolutionary dynamics of divergence with gene flow (Bird et al., 2012; Nosil, 2008). For example, morphologically cryptic and sympatric coral taxa may represent incipient species where specific habitats delineate taxa (e.g., Carlson & Budd, 2002; Warner et al., 2015) or differences in reproductive timing (e.g., Rosser, 2015). Determining the extent to which scleractinians can exchange alleles intra- or interspecifically can provide insights into possible rates of adaptive evolution such as in response to anthropogenic stresses.
found over small to moderate distances in A. agaricites and A. fragi-
lies (<40 km) (Bongaerts et al., 2017; Brazeau et al., 2005) but not in A. lamarcki and A. undata (considering similarly small to moderate

distances: <40 km) (Gonzalez-Zapata et al., 2018; Hammerman et al., 2018). However, the spatial genetic structure of a dominant

mesophotic-specialist, A. grahamae, has not yet been determined. This

species shares the same Cladocium strain with A. lamarcki at their sympatric depth zone (50 m), with A. lamarcki predominantly

hosting a different strain at shallower depths (Bongaerts, Carmichael, et al., 2015; Bongaerts et al., 2013). Furthermore, mitochon-
drial markers (atp6, nad5 and cox1-1-rRNA) have been unable to
geneically differentiate A. grahamae and A. lamarcki (Bongaerts,

Frade, et al., 2015; Bongaerts et al., 2013), although this is not

surprising for Anthozoa with slow mutation rates of mtDNA, and

shared haplotypes probably indicates their close-relatedness.

Consequently, there is potential for hybridisation between A. gra-
hamae and A. lamarcki as well as host divergence of shallow and

mesophotic populations within A. lamarcki.

To evaluate horizontal genetic structure and interspecific gene

flow of these ecologically important mesophotic species, we used a

reduced representation genome sequencing approach (nextRAD) on

specimens collected using a manned submersible and deep technical
diving. We tested the following three hypotheses: (1) there is hori-

tzontal genetic structure between populations of mesophotic depth-
specialist scleractinian species, Agaricia grahamae between Curacao and Bonaire (~40 km); (2) gene flow occurs or has occurred between A. grahamae and congener, A. lamarcki within the depth zone they

share (50 m); and (3) depthpartitioning occurs within A. lamarcki be-
tween 15 and 50 m. After initial examination of the genetic data,

we found two sympatrically occurring cryptic taxa each within both
taxonomic species and thus post hoc decided to test the hypothe-
sis that; (4) gene flow occurs between cryptic taxa or has occurred
during their divergence history. To examine spatial genetic structure

and identify hybrids, we used individual-based assignment models

and multivariate analyses. For testing whether and when gene flow

occurred during the divergence of taxa, we used the diffusion ap-

proximations for demographic inference (dadi) (Gutenkunst et al.,

2009) to compare various demographic scenarios.

2 MATERIALS AND METHODS

2.1 Sample collection

Specimens of Agaricia were collected at eight different locations

on the leeward side of the islands of Curacao and Bonaire in the

Southern Caribbean, as part of the “XL Catlin Seaview Survey” carried

out between March and April 2013 (Figure 1, Table S2-1). Samples

were collected using technical SCUBA or the manned submersible

“Curasub” operated by “Substation Curacao”, under permits from

the Curacao Government and the Bonaire Island Council. Specimens of

the focal species Agaricia grahamae (Wells, 1973) were collected at

a sampling depth of 50 m (±2 m), with two additional populations

sampled at 60 m (±2 m) and 80 m (±5 m), whereas specimens of

Agaricia lamarcki (Milne-Edwards & Haime, 1851) were collected at

a subset of four locations within Curacao at a sampling depth of 15 m

(±2 m) or 50 m (±2 m). At one site, “CR60” (Figure 1, Table S2-1),

additional samples were collected from the same colony (tissue con-

nected) to assess for potential chimeras and from closely adjacent
colonies (tissue not connected) to assess for potential clones due to

fission. Morphological classification of the two species followed

the taxonomic features specified by Wells (1973), Veron (2000) and

Humann and DeLoach (2002). An additional 12 A. grahamae samples

from San Andres, Colombia (SA) were collected under permits by the

National Environmental Licensing Authority (ANLA), and added as

goout group samples: three were collected from the upper mesophotic

zone (60–65 m) and nine from the lower mesophotic zone (85 m).

Small fragments of colonies were stored in salt-saturated buffer so-
lution containing 20% DMSO and 0.5 M EDTA, and for a subset of

specimens a skeletal voucher was collected.

2.2 DNA isolation, library

preparation and sequencing

Isolation of genomic DNA from the coral host was carried out as

reported in Bongaerts et al. (2017), using centrifugation steps to

reduce endosymbiont contamination. Symbiodiniaceae were then

FIGURE 1 Sampling locations for collected samples of Agaricia grahamae and A. lamarcki. Samples were collected from Curacao and Bonaire located in the

Southern Carribean, outgroup samples are from San Andres (top left). Samples of A. grahamae were collected from five sites in Curacao and three sites in Bonaire and A. lamarcki were collected in four sites in Curacao: CK, CR, CS and CE (bottom left). Photograph of the two study species (right)
isolated from two *A. grahamae* specimens (to sequence separately as a subsequence reference), using fluorescence-activated cell sorting (BD FACSAria Cell Sorter) at the Queensland Brain Institute. Quality and yield of gDNA were assessed using gel electrophoresis and a Qubit fluorometer, with a subset of samples (*A. grahamae* n = 176; *A. lamarcki* n = 51; Symbiodiniaceae n = 2) selected for downstream sequencing. For *A. lamarcki*, the Symbiodiniaceae ITS2 profile was determined for several of these samples in a previous study (n = 6; Bongaerts, Carmichael et al., 2015), and we screened the profiles of an additional 41 samples using the same ITS2-DGGE method against reference samples from that study. Library preparation was carried out using the nextRAD method (SNPsaurs, LLC), using a 9 bp selective sequence (“GTGTAGAGG”) to amplify loci consistently between samples. Genomic DNA was fragmented and ligated with adapter sequences using Nextera reagent (Illumina, Inc), and sequenced across a total of six HiSeq 2500 (Illumina, Inc) lanes using 100 bp single-end chemistry and following the manufacturer’s recommended protocol. Samples that failed in the initial run (three lanes), were purified using AMPure XP beads to remove potential inhibitors and sequenced again on the additional HiSeq lanes.

### 2.3 Sequence clustering and variant calling

TRIMGALORE v.0.4.5 ([https://github.com/FelixKrueger/TrimGalore](https://github.com/FelixKrueger/TrimGalore)) was used to remove adapters and low-quality ends (Phred below 20) and discarding reads that were less than 30 bp. Read clustering was conducted using the IPYRAD pipeline v.0.7.22 (Eaton & Overcast, 2017) using default settings, excepting: minimum depth statistical/majority = 10, filter for adapters = 1, maximum uncalled bases = 5, maximum heterozygotes = 8, and maximum number of SNPs per locus = 20. Initial filtering, symbiont contamination removal and defining clonal lineages followed Bongaerts et al. (2017) (available through: [https://github.com/pimbongaerts/radseq](https://github.com/pimbongaerts/radseq)), unless otherwise indicated. We used BLASTN to identify and remove any matches to three Symbiodiniaceae databases (RAD isolates from Bongaerts et al. (2017), *Breviolum minutum* (ITS2 type B1) genome (Shoguchi et al., 2013) and *Cladocopium* (ITS2 type C1) genome (Liu et al., 2018). Other potential microbial contamination was removed through a BLASTN search against the NCBI nonredundant nucleotide database, extracting positive matches (max. e-value of $10^{-5}$) that were classified as non-Cnidarian taxa (using the NCBI Taxonomy Database), RAD loci were truncated to 100 bp prior to downstream analysis. Two sequencing duplicates (but from the same library preparation) of each species were included in the data set to assess genotyping error and as a comparison to identify natural clones. The occurrence of any genetically identical individuals (clones) were then evaluated through assessing the distribution of allelic similarities between all pairs of individuals. Pairs of individuals that had 96% similar reads and above were deemed as clonal groups and one representative of each pair was retained. This threshold was chosen due to a combination of the genetic similarity of sequencing replicates (at similarities of 99%), a break in the distribution of pairwise allele similarities, and the maximum similarity (95%) that was observed for individuals occurring at different sites. Two data sets were retained: one removing putative clonal individuals and a second with all individuals.

The VCF file containing SNP data was filtered using VCFTOOLS v.0.1.16 (Danecek et al., 2011) to have a minor allele count of three due to many singletons and doubletons probably being sequencing or PCR errors and thus have consequences on downstream population genetic analyses (Andrews et al., 2016; Linck & Battey, 2019). To remove sites that were not represented across most individuals, we removed sites that had $>$50% missing data across individuals and used a minimum depth of five per site. Certain individuals with low coverage across sites ($<$50% of sites genotyped) were removed. The removal of individuals with high missing data was conducted before more stringent filtering as these individuals will bias which sites are retained. Lastly, as different missing data filtering thresholds can change observed genetic patterns substantially, we applied four different thresholds for sites: (1) 50% maximum missing data of sites across individuals, (2) 20%, (3) 10% and (4) 5% and these data sets were compared for congruence across analyses. Results from the 20% missing data data set are presented here unless reported otherwise due to congruence of results across the four thresholds.

### 2.4 Statistical analysis

#### 2.4.1 Population structure

To assess population structure of both species and initial evidence for hybridisation and introgression between species, methods that do not use a priori population assumptions were used. Models employed include: model-based multilocus population assignment methods based on maximum likelihood (ADMIXTURE, v.1.3.0 Alexander et al., 2009) and Bayesian (STRUCTURE, v.2.3.4 Pritchard et al., 2000) optimisation criteria as well as principal component analysis (PCA) and likelihood-based genetic clustering (using functions “gPCA”, “snap.clust” and “find.clusters” in the package, ADEGENET v.2.1.3 (Jombart, 2008) in R v.3.6.3 R Core Team, 2020). The consensus of all four methods was used to identify the number of genetic clusters (K) within the data sets and to assign individuals to these clusters. The data sets including clones and without clones were compared for differences. Both “all SNPs” and putatively "neutral" data sets were compared for differences in genetic clustering patterns. The “neutral” data sets were created by removing outlier SNPs found using PCADAPT R package v.4.3.3 (Luu et al., 2017). The PCADAPT method identifies SNPs that exhibit significantly large correlations with certain PC axes relative to the genomic background, based on Mahalanobis distance and corrections using a genomic inflation factor. The calculation of q-values was used to determine which SNPs to retain with a false discovery rate $<$10% calculated by the QVALUE package v.2.18.0 (Storey et al., 2019) in R.

For ADMIXTURE and STRUCTURE analyses, the data sets were randomly trimmed to one SNP per contig to reduce correlations
caused by physical linkage. Replicate data sets (10 replicates) with one random SNP per contig were created for comparison. In ADMIXTURE we ran each data set with a cross-validation of 100 for \( K = 1\)–7. In STRUCTURE we ran each data set with a Burn in of 100,000 and 50,000 MCMC replicates for \( K = 1\)–7. Cross-validation error between runs, log-likelihood ratios, and Evanno’s Best K (Evanno et al., 2005) were evaluated to find the most likely number of clusters for each data set in both analyses. For the PCA, the number of PC axes were deemed appropriate by assessing a drop between eigenvalues (where the slope becomes less steep as the cutoff) as well as qualitatively assessing any structure on each axis iteratively until structure dissipates. For likelihood-based genetic clustering using the snap.clust function, we chose the number of genetic groups (\( K \)) with lowest value for the Akaike information criterion (AIC) and the Bayesian information criterion (BIC) between \( K = 1\)–10 for each data set. We found two cryptic and sympatric genetic groups within each species (referred to as AG1 and AG2 within \( A. grahamae \) and AL1 and AL2 within \( A. lamarcki \) hereafter), these were treated as separate groups in subsequent analyses.

We calculated deviations from the Hardy-Weinberg equilibrium for genetic groups within each taxonomic species using the “summary” function in ADEGENET, then used Bartlett’s t-test for homoscedasticity and a t-test for differences in means between the expected and observed heterozygosity in the package STATS v.3.6.3 (R core team). We calculated population genetic statistics for each species treating each genetic group (e.g., AG1 and AG2) as populations, in the HIERFSTAT package v.0.0.4 (Goudet, 2005) in R for \( F_{st} \) and Weir and Cockerham’s \( F_{ST} \). We tested for significance using Goudet’s G-statistic with 1000 permutations. A \( \chi^2 \) test was used to detect correlations between the frequency of the genetic group in each depth profile for \( A. lamarcki \) using STATS.

### 2.4.2 Hybridisation and introgression

Potential hybrids and individuals with various levels of mixed ancestry were found within in the assignment methods results. We investigated putative hybridisation using NEWHYBRIDS v.1.1 (Anderson & Thompson, 2002). NEWHYBRIDS incorporates the predictable patterns of inheritance seen in mating events rather than using only allele frequencies. We assessed whether individuals had genotypes consistent with any of eight hybrid classes: \( (1 \text{ and } 2) \) pure parental, \( P1 \text{ or } P2, (3 \text{ first-} \text{ F1 or (4 second-generation, F2 hybrids (F1 hybrid offspring), (5 and 6) first generation backcross from F1 into each parent group, b-1 or b-2, or (7 and 8) second-generation backcross into each parent group, b-1-1 or b-2-2. The analysis was run for \( >10,000 \) steps and \( >10,000 \) MCMC with both Jeffery’s and Uniform distributions testing for the eight possible genotype scenarios and each run was repeated five times. The number of steps the analyses was run for was determined by convergence of the parameters. We show results from the 5% missing data sets with neutral and unlinked SNPs due to inability to estimate hybrid classes with higher missing data thresholds.

### 2.4.3 Spatial genetic structure

We assessed the spatial genetic structure of the genetic groups within both \( A. grahamae \) and \( A. lamarcki \) to look for potential spatial and or environmental barriers to gene flow. We applied redundancy analysis (RDA) to assess the relative exploratory power of geographic distance and depth environment to SNP genotype data. RDA performs a multiple linear regression between matrices (Legendre & Legendre, 2012) and is a commonly used technique for assessing the relative contribution of multiple predictors. It is effective for uncovering predictors of population genetic structure (e.g., Forester et al., 2018; Legendre & Fortin, 2010). We utilised the partial RDA model for each taxonomic species: SNPs –latitude + longitude + depth environment (categorical) + condition (genetic group). Because cryptic genetic groups within each species were found, conditioning was used to evaluate the relative contribution of depth and geographic location to SNP variance considering the variance explained by the genetic groups. Missing genetic data was imputed with the most common SNP following Forester et al. (2018). For \( A. grahamae \), we calculated the over-water distance matrix between locations using the function “lc.dist” in the package MARMAP v.1.0.4 (Pante & Simon-Bouhet, 2013) in R. Here, we calculated the least-cost path between our locations which avoids land masses. These distances were then transformed into PCoA coordinate scores for input as exploratory factors into the RDA model. For \( A. lamarcki \), which was only sampled in sites in Curaçao, we used the raw latitudes and longitudes as input into the RDA model because the over-water distance was similar. We used the “rda” function in the package VEGAN v.2.5 (Oksanen et al., 2018) in R. We began with the full models and used permutation tests (1000 permutations) to permute the genotypes randomly and assess the global model, RDA axes and marginal significance using PERMANOVA (“anova.cca” function) (Legendre et al., 2011).

### 2.4.4 Demographic inference

To explore the possibility of gene flow occurring during the divergence of our taxa, we used the diffusion approximation of demographic inference dadi v.2.1.1 (Gutenkunst et al., 2009) in PYTHON v.3.6. We treated the two clusters found previously within each species as separate populations (within \( A. grahamae \): AG1 and AG2 and within \( A. lamarcki \): AL1 and AL2). The joint allele frequency spectrum (JAFS) was used to examine the likelihood of various demographic scenarios by modelling forward-in-time changes to the JAFS using solutions to the Chapman-Kolmogrov forward equation (i.e., the diffusion approximation). The JAFS is a matrix of bins comprising SNP counts for each combination of haplotype frequencies between the populations. We tested five different divergence scenarios for our 2-population comparisons: (1) divergence with no migration (no mig), (2) divergence with continuous symmetrical migration (sym mig), (3) divergence with continuous asymmetrical migration (asym mig), (4) divergence with ancient symmetrical migration followed by
isolation (anc mig) and (5) divergence in isolation followed by symmetrical migration (sec cont) (see Supporting Information S1-1 for schematics of models). For each model, we incorporated inbreeding due to finding a statistically significant positive $F_{IS}$ within each taxon (Blischak et al., 2020). Parameters for each of these models were fit for the JAFS of each pairwise population comparison, within (i.e., AG1 and AG2) and between the known species (i.e., AG1 and AL1), thus we assessed divergence patterns within six paired groups. The folded JAFS was used due to the ancestral state of each allele being unknown without an available outgroup as a reference. The JAFS were constructed by representative individuals from each population, which had $>0.95$ admixture assignment using neutral, unlinked (one SNP per contig) and 20% missing loci data sets. The "subsample" function in dadi was used to randomly select a subset of haplotypes for analysis and maximise the number of SNPs due to missing data issues and to avoid difficulties with modelling inbreeding when projecting SNP frequencies to smaller sizes. In each paired population comparison, singletons and doubletons were masked because these entries were unreliable due high error rates in sequencing. Model parameters were optimised by simulating a model JAFS and calculating the likelihood of each model fit to our data set JAFS using the Nelder-Mead simplex as the optimising algorithm. We assured convergence by running optimisations until independent runs with the same parameter scores ($\pm 1\%$ difference) and the lowest AIC occurred $\geq$2–3 times. We then compared $\Delta$AIC and log-likelihoods (using the likelihood ratio test when nested) using the most likely replicate of each demographic scenario. The residuals between the simulated JAFS and the data set JAFS were qualitatively assessed for random distributions across the JAFS and forming a normal distribution at zero. To assess goodness-of-fit we used nonparametric bootstrapping to see if the likelihood of parameters from the analysis data set fit the distribution of likelihoods from the bootstraps. We applied the Godambe information matrix (GIM) (Coffman et al., 2016) using bootstraps to calculate the confidence intervals of the parameter estimates.

3 | RESULTS

3.1 | Sequence clustering and variant calling

Reduced representation sequencing was performed on 227 individuals to obtain 686,608 SNPs and 130,890 loci. We identified 41 individuals with $\geq 96\%$ genetic similarity (putative clones), representing 20 clonal groups and one individual of each clonal group was kept for subsequent analyses (Tables S2-1 and S2-2, Figure S1-2). Filtering measures and genetic assignment of species groups obtained four main data sets with a maximum 20% missing data per locus threshold (see Table S2-2 for specific filtering results): 1. All individuals data set including both species (161 individuals, 4306 neutral SNPs), 2a. A. grahamae individuals with San Andrés outgroup included (118 individuals, 3104 SNPs), 2b. A. grahamae individuals with outgroup, outlier individuals and putative hybrid removed (106 individuals, 2725 neutral SNPs) and 3. A. lamarcki (41 individuals 2515 neutral SNPs). The congruent results from clustering analyses agreed with our morphological taxonomic classification of the two species: A. grahamae and A. lamarcki and further analyses were performed on these groups separately. Congruent results also found two distinct genetic clusters within both species that were found sympatrically across sites and depths within each species (named: AG1, AG2, AL1 and AL2).

3.2 | Population structure

Morphologically identified A. grahamae and A. lamarcki formed separate genetic clusters and one individual that was morphologically
classified as A. grahamae had 0.5 mixed ancestry at K = 2 in ADMIXTURE and had intermediate PC1 scores between the two species clusters (Figure 2). This individual was treated as a provisional F1 hybrid. At K = 3, two cryptic clusters within A. grahamae (AG1 and AG2) occurred sympatrically at sites within Curacao, Bonaire and outgroup San Andres, Colombia. At K = 4, A. lamarcki also split into two clusters (AL1 and AL2). Across all methods, K = 4 was deemed most likely for this data set. The same two clusters were found in the subset A. grahamae data set at K = 2 (Figure 3).

At K = 3, three individuals formed a separated cluster and appear as outliers separated from both A. grahamae groups on PC1. AG2 was more commonly found in Bonaire compared to Curacao sites (10 vs. 4) and were more closely related to most individuals collected from San Andres (SA) (7/8 assigned to AG2 at K = 2). The AG2 individuals from San Andres formed a separate cluster at K = 4. Across all population structure methods, K = 3 was deemed most likely. We removed the A. grahamae/A. lamarcki putative hybrid, three outlier individuals and outgroup samples from San Andres from the A. grahamae data set to eliminate confounding elements for Hardy-Weinberg Equilibrium, spatial and hybrid analyses (i.e., data set 2b).

A. lamarcki also had the same two clusters found in the “all individuals” data set that occurred sympatrically at all sites and within both depths. AL1 was more commonly found at 50 m (12/19) and AL2 at 15 m (22/24) (Figure 4, $\chi^2 = 10.96, df = 1, p < .01$). Across all methods for the A. lamarcki data set, K = 2 was deemed most likely. After having subset each of AG1, AG2, AL1 and AL2, into separate data sets, we found no further genetic structure (increasing K did not reveal any more clusters or improve likelihood greatly) across all population structure analyses.

Hardy-Weinberg Equilibrium estimates were calculated for the cryptic taxa and F-statistics were calculated for each species with cryptic taxa treated as populations. Cryptic taxa within each taxonomic species were not in HWE (Table S2-3) with an excess of homozygosity. Inbreeding within populations for each species was high (AG: $F_{IS} = 0.18$ and AL: $F_{IS} = 0.19$) accompanied by substantial population differentiation between cryptic taxa (AG: $F_{ST} = 0.17, p < .01$, AL: $F_{ST} = 0.18, p < .01$).

Regarding the Symbiodiniaceae associated with A. lamarcki, corals predominately associated with the two ITS2 profiles (C3/C3d/C3.N6/C3.N7 and C3/C11/C11.N4/C3.N5) reported in Bongaerts et al. (2013, 2015; although sometimes with an extra unidentified band) and were partitioned between 15 and 50 m depth. However, there was no association observed with the AL1 and AL2 lineages (Figure S1-3), with both profiles occurring in both lineages.

We found groups of colonies with the same genotype (“genets”) at most sites, with 17 genets observed for A. grahamae and three for A. lamarcki populations (Table S2-1, Figure S1-2), and each genet was always restricted to a single site (and depth). Most genets consisted of two colonies (“ramets”), although one genet for A. grahamae and one for A. lamarcki was represented by three ramets. At CR60 we took multiple samples for five A. grahamae colonies as well as four clusters of closely adjacent colonies, and these always represented the same genotype.

### 3.3 | Hybridisation

#### 3.3.1 | Between A. grahamae and A. lamarcki

From the ADMIXTURE results for the data set including both species, 115 individuals assigned to A. grahamae with >0.98 assignment and 42 individuals assigned to A. lamarcki with >0.99 assignment at
3. Of the admixed individuals, three were the outlier individuals identified previously within the *A. grahamae* data set and one putative hybrid (also identified in the PCA results) had 0.5 admixture (Figure 2). The three outlier individuals and outgroup samples from San Andrés were removed to create another data set for input into NEWHYBRIDS (**n** = 149, 5% missing, **SNPs** = 555). In NEWHYBRIDS, all individuals assigned as pure species with 0.99 probability (P1 or P2) apart from the putative F1 hybrid which assigned as a F1 hybrid with a 0.99 probability.

3.3.2 | Between cryptic groups within *A. grahamae*

We also assessed the patterns of admixture between the two *A. grahamae* lineages (AG1 and AG2). ADMIXTURE results presented 55 individuals assigned to AG1 >0.99 assignment and 14 individuals for AG2 at K = 2. The remaining 37 individuals predominately assigned to AG1 (0.8–0.98). We used data set 2b (without outgroup, AL/AG hybrid and outliers) with more stringent filtering for input into NEWHYBRIDS (**n** = 106, 5% missing, **SNPs** = 514). NEWHYBRIDS assigned 86 individuals with predominately P1 assignment (0.62–0.99) (pure parental AG1) and 12 individuals with predominately P2 assignment. The remaining individuals: six individuals (predominately assigned to AG1) and two individuals (predominately assigned to AG2) had higher assignments to hybrid categories (Table S2-4). Most notably, the highest probabilities for these admixed individuals were second generation back-crosses into each parental. Due to the limitations of the NEWHYBRIDS analysis, we were not able to estimate genotype classes of more advanced backcrosses which these individuals may more accurately represent.

K = 2. Of the admixed individuals, three were the outlier individuals identified previously within the *A. grahamae* data set and one putative hybrid (also identified in the PCA results) had 0.5 admixture (Figure 2). The three outlier individuals and outgroup samples from San Andrés were removed to create another data set for input into NEWHYBRIDS (**n** = 149, 5% missing, **SNPs** = 555). In NEWHYBRIDS, all individuals assigned as pure species with 0.99 probability (P1 or P2) apart from the putative F1 hybrid which assigned as a F1 hybrid with a 0.99 probability.

3.3.3 | Between cryptic groups within *A. lamarcki*

Within the *A. lamarcki* data set, ADMIXTURE found 12 individuals assigned to AL1 >0.99 and 26 assigned to AL2 >0.99 and three individuals were admixed. Two of these admixed individuals were assigned to AL1 with ~0.7 and 0.78 assignment and one with 0.98 to AL2. A more stringent filtering data set was used for input to NEWHYBRIDS (**n** = 41, 5% missing, **SNPs** = 622). In NEWHYBRIDS, two individuals presented higher assignment to hybrid classes than pure parentals: one *A. lamarcki* individual collected from CE50 had a 0.99 probability of being a F2 (offspring of two F1s), and another from CR50 had mixed assignments, 0.34 for pure parental AL1, 0.41 for a first generation back-cross into AL1 and 0.25 for second generation back-cross into AL1. The second individual, potentially representing a further back cross into AL1.

3.4 | Spatial genetic structure

No strong correlations between genotype or depth, latitude, and longitude for either species were uncovered with redundancy analyses. For *A. grahamae*, although the global model was significant (**R**² = 0.04, **R**² adj = 0.002, **F** = 1.05, **p** = .005) it only explained 4% of the variance. None of the canonical axes were significant in explaining variation in SNPs (i.e., RDA1 = 4; **p** > .1) and PC1 explained more variation than RDA1 (0.015 > 0.011). This was further substantiated through removal of each term, apart from Condition (Clusters), improving the AIC of the model. For *A. lamarcki*, the full model was globally non-significant (**R**² = 0.07, **R**² adj = 0.0002, **F** = 1.00, **p** = .483) as was each reduced model. Thus, neither horizontal (by space) nor
vertical (by depth) locations were substantial predictors of genetic variation within cryptic taxa in both taxonomic species.

3.5 | Demographic inference

The demographic modelling results from *dadi* consistently found that models including migration (gene flow) had higher likelihoods than those with no migration (Tables 1 and 2, Figure 5 and Figure S1-4). In the "no migration" models, divergence time approached the lower bound parameter limit. Thus, gene flow probably occurred during the divergence of the two species (*Agaricia grahameae* and *Agaricia lamarcki*) as well as during the divergence of the two cryptic taxa found within each species (AG1, AG2, AL1 and AL2). The haplotype frequency patterns within the JAFS were more comparable to the divergence with migration scenario than the no migration scenario, particularly in the shared low frequency bins (Figure 5), as shared alleles in such bins are expected to be elevated from migration. Divergence time (T) in the isolation models was always less than migration models due to attempts to optimise these frequency bins (T1, Figure 5). For all pairwise comparisons the symmetrical and asymmetrical continuous migration models showed similar results, with the asymmetrical migration models having only minimal differences in migration rates. Thus, only the continuous symmetrical migration models are reported here. In the ancient migration model optimisations between all population-pairs, T2 (the second epoch where populations diverged in isolation) approached the lower the parameter limit (~0) and thus equating this model (ancient migration) to the continuous symmetrical migration scenario. The continuous symmetrical migration and secondary contact models had equal likelihood (<2 AIC) for within taxonomic species comparisons (AG1 vs. AG2 and AL1 vs. AL2, Table 1). For both AG groups vs. AL1, secondary contact had higher likelihood >2 AIC than continuous migration model (Table 2). But for AG groups versus AL2, the secondary contact model had equal likelihood with continuous migration. Migration rates estimated from *dadi* were larger between more closely related pairs than the more distant pairs, suggesting that more divergent taxa are likely to experience less gene flow (Tables 1 and 2). Parameter uncertainties were often larger than 0.5 of the parameters and thus are not shown. These were also not converted into real time units because of the unreliability of parameter estimates due to a low number of SNPs used as well as having unreliable mutation rate and generation time estimates. The main purpose of this analysis was to detect whether gene flow occurred over the divergence history between all taxa, and thus conversion into real time units was not necessary for the aims of this study.

### Table 1
Demographic modelling performed between genetic groups within *Agaricia grahameae* (AG1 and AG2) and *Agaricia lamarcki* (AL1 and AL2) support histories of divergence with gene flow. Maximum likelihood estimates for each demographic scenario and each parameter is scaled by $\theta$

| Model | LogL | AIC | $\chi^2$ | $\theta^a$ | $N_1^b$ | $N_2^b$ | $F_1^c$ | $F_2^c$ | $m^d$ | $T1^e$ | $T2^e$ |
|-------|------|-----|----------|-----------|---------|---------|---------|---------|-------|--------|--------|
| Between *A. grahameae* genetic groups | | | | | | | | | | | |
| No mig | -558.54 | 1127.09 | 1181.93 | 239.50 | 2.12 | 150 | 0.00 | 0.00 | — | 0.41 | — |
| Sym mig | -499.09 | **1010.18** | 591.89 | 141.06 | 2.98 | 1.40 | 0.00 | 0.40 | 1.05 | 1.19 | — |
| Anc mig | -499.15 | 1012.30 | 593.28 | 140.80 | 2.98 | 1.40 | 0.00 | 0.40 | 1.05 | 1.19 | 0.00 |
| Sec cont | -497.26 | **1008.52** | 536.81 | 148.19 | 2.91 | 1.43 | 0.00 | 0.46 | 1.27 | 0.65 | 0.38 |
| Between *A. lamarcki* genetic groups | | | | | | | | | | | |
| No mig | -551.26 | 1112.52 | 467.64 | 264.47 | 7.58 | 1.98 | 0.03 | 0.00 | — | 0.43 | — |
| Sym mig | -516.33 | **1044.66** | 385.34 | 200.97 | 26.9 | 1.57 | 0.46 | 0.30 | 0.49 | 1.01 | — |
| Anc mig | -516.27 | 1046.54 | 385.56 | 200.26 | 26.9 | 1.56 | 0.46 | 0.24 | 0.51 | 1.01 | 0.01 |
| Sec cont | -515.94 | **1045.88** | 380.58 | 207.65 | 25.0 | 1.67 | 0.46 | 0.39 | 0.64 | 0.60 | 0.29 |

$a^\theta = 4N_{ref}^d$.  
$^bN_1 = \text{the resulting population size change from } N_{ref} \text{ to population 1. }N_2 = \text{the resulting population size change from } N_{ref} \text{ to population 2.}$  
$^cF_1 \text{ and } F_2 = \text{the inbreeding coefficients (F) of population 1 and 2.}$  
$^dmd = \text{the symmetrical migration rate between population 1 and 2, in } 2N_{ref} \text{ generations.}$  
$^eT1 = \text{time since divergence to present for one epoch models and time since divergence to } T2 \text{ in two epoch model. }T2 = \text{time since } T1 \text{ to present.}$  
$^fUnits in } 2N_{ref} \text{ generations.}$  
$^gModels with the highest likelihood in bold.$
was found within Curaçao for the depth-generalist congener, *A. lamarcki* (Figure 4). Instead, we uncovered two sympatrically occurring and yet distinct cryptic taxa within each taxonomic species (*Figures 2–4*) that were incompletely depth-partitioned in *A. lamarcki* but with no detectable depth or geographic segregation in *A. grahamae*. These cryptic taxa appear to be connected by historical and contemporary gene flow as indicated by the presence of backcross individuals and through demographic analyses (*Table 1, Figure 5, Figure S1-4 and Table S2-3*). The divergence history of *A. grahamae* and *A. lamarcki* was also characterised by interspecific gene flow (*Table 2, Figure 5* and Figure S1-4), suggesting past semi-permeable boundaries between these species.

### 4.1 Lack of genetic structuring between reefs and islands

Contrary to our expectations of spatial genetic structure over small distances, genetic homogeneity was found for *A. grahamae* between Curaçao and Bonaire (<40 km) as evidenced by the RDA (see Section 3) and genetic structure results (*Figure 3*). Both cryptic taxa found within *A. grahamae* also occurred at San Andrés (>1000 km away) and thus, it is unlikely that the two cryptic groups within *A. grahamae* were recently allopatrically formed. The congener, *A. lamarcki* was assayed only in Curaçao but also showed no genetic differentiation between the sampled reefs (*RDA results, Figure 4*). Genetic homogeneity within either island (Curaçao and Bonaire) may not be too surprising because reef communities are fairly continuous along leeward sides of the islands which could facilitate stepping-stone gene flow and occasional long-distance dispersal. However, Curaçao and Bonaire are separated by deep oceanic water (>500 m depth) with a north-west bound surface current and a west-east subsurface counter current (*Andrade, 2003*), presenting a physical dispersal barrier. In contrast to our results (and consistent with limited dispersal), shallow occurring congeners *Agaricia agaricites* and *A. fragilis* were found to have localised structure (*Bongaerts et al., 2017; Brazeau et al., 2005*). Our results do, however, match those described for mesophotic occurring *A. lamarcki* and *A. undata* (*Gonzalez-Zapata et al., 2018; Hammerman et al., 2018*) where populations did not exhibit local

---

**TABLE 2** Demographic modelling performed using six cross-species comparisons (*A. grahamae*: AG1 and AG2 and *A. lamarcki*: AL1 and AL2) support histories of divergence with gene flow. Maximum likelihood estimates for each demographic scenario and each parameter is scaled by $\theta^b$

| Model | LogL | AIC | $\chi^2$ | $\theta$ | $N_1$ | $N_2$ | $F_1$ | $F_2$ | $m^a$ | $T1^a$ | $T2^a$ |
|-------|------|-----|----------|---------|-------|-------|-------|-------|-------|-------|-------|
| Between species (1) AG1 and AL1 | | | | | | | | | | | |
| No mig | -500.02 | 1010.04 | 615.34 | 246.23 | 0.84 | 42.5 | 0.00 | 0.01 | --- | 0.84 | --- |
| Sym mig | -432.32 | 876.64 | 330.27 | 94.41 | 2.09 | 2.88 | 0.00 | 0.02 | 0.10 | 3.21 | --- |
| Anc mig | -432.32 | 878.64 | 330.03 | 92.97 | 2.13 | 2.92 | 0.00 | 0.00 | 0.10 | 3.28 | 0.00 |
| Sec cont | -426.92 | 867.84 | 307.21 | 201.16 | 0.87 | 76.8 | 0.15 | 0.36 | 0.27 | 1.29 | 0.17 |
| Between species (2) AG1 and AL2 | | | | | | | | | | | |
| No mig | -670.25 | 1350.50 | 1031.49 | 212.76 | 0.78 | 72.2 | 0.00 | 0.00 | --- | 0.68 | --- |
| Sym mig | -578.08 | 1168.16 | 687.91 | 67.10 | 2.40 | 3.61 | 0.00 | 0.00 | 0.09 | 3.64 | --- |
| Anc mig | -578.15 | 1170.30 | 687.81 | 64.93 | 2.47 | 3.72 | 0.00 | 0.01 | 0.09 | 3.80 | 0.00 |
| Sec cont | -576.35 | 1166.70 | 698.43 | 103.01 | 1.61 | 2.47 | 0.01 | 0.00 | 0.15 | 1.19 | 0.69 |
| Between species (3) AG2 and AL1 | | | | | | | | | | | |
| No mig | -277.28 | 564.56 | 229.23 | 174.10 | 0.77 | 42.6 | 0.02 | 0.01 | --- | 0.96 | --- |
| Sym mig | -261.40 | 534.80 | 187.05 | 95.77 | 1.33 | 2.26 | 0.41 | 0.53 | 0.12 | 2.00 | --- |
| Anc mig | -261.37 | 536.74 | 187.54 | 74.17 | 1.69 | 2.87 | 0.33 | 0.52 | 0.11 | 2.91 | 0.04 |
| Sec cont | -259.32 | 532.64 | 179.72 | 149.32 | 0.83 | 87.9 | 0.36 | 0.62 | 0.17 | 1.19 | 0.22 |
| Between species (4) AG2 and AL2 | | | | | | | | | | | |
| No mig | -401.24 | 812.48 | 479.69 | 159.48 | 0.69 | 41.8 | 0.02 | 0.00 | --- | 0.79 | --- |
| Sym mig | -376.24 | 764.48 | 369.51 | 85.90 | 1.34 | 2.55 | 0.54 | 0.11 | 0.11 | 1.85 | --- |
| Anc mig | -375.68 | 765.36 | 389.22 | 67.73 | 1.65 | 3.16 | 0.51 | 0.01 | 0.11 | 2.65 | 0.06 |
| Sec cont | -375.64 | 765.28 | 398.10 | 92.70 | 1.26 | 2.42 | 0.55 | 0.16 | 0.13 | 0.87 | 0.72 |

*a* $\theta = 4N_{ref}\mu$.

*b* $N_1$, the resulting population size change from $N_{ref}$ to population 1. $N_2$, the resulting population size change from $N_{ref}$ to population 2.

$c$ $F_1$ and $F_2$, the inbreeding coefficients ($F$) of population 1 and 2.

$m^a$, the symmetrical migration rate between population 1 and 2, in 2Nref generations.

$T1^a$, time since divergence to present for one epoch models and time since divergence to T2 in two epoch model. T2 = time since T1 to present. Units in 2Nref generations.

*Models with the highest likelihood in bold.*
genetic structure (over similarly short distances: <40 km). Genetic subdivision in A. grahamae does emerge at larger spatial scales (AG2 individuals in San Andrés, Figure 3a). It is conceivable that reproductive strategies within Agaricia species are more variable than assumed and that the dispersal abilities within A. lamarcki, A. grahamae and A. undata are potentially greater than that of A. agaricites and A. fragilis. Taken together, these findings across studies highlight variability in spatial genetic structuring that can occur even among congeners. A comparative study on the spatial genetic structure at the same localities of Agaricia spp. would provide greater resolution to these hypotheses as well as studies on reproduction in A. lamarcki and A. grahamae.

The occurrence of the clone groups at most sites within A. grahamae and two of the sites within A. lamarcki are not surprising results as fission is common in A. lamarcki and in the congener A. agaricites (Hughes & Jackson, 1985). Their growth form (follaceous/plating/encrusting) makes Agaricia spp. vulnerable to partial mortality from overgrowth by algae and sedimentation, which then results in subsequent fission where one colony becomes two or more colonies with separated tissue (Hughes & Jackson, 1980). At CR60, we confirmed this process by identifying clones from fission.

4.2 | Distinct depth distributions for A. lamarcki but not A. grahamae taxa

Depth constitutes a strong environmental gradient for scleractinian corals, as it modulates light and other environmental conditions upon which corals depend (Bongaerts, Carmichael, et al., 2015). Although no depth-partitioning was observed for A. grahamae taxa between the upper (50 m) and lower (80–85 m) mesophotic zone at one site comparison in Curaçao and one in San Andrés, we did find a difference in the relative abundance of the cryptic taxa by depth in A. lamarcki (15 vs. 50 m, Figure 4, $\chi^2 = 10.96, df = 1, p < .01$). Given the exponential decay of irradiance with depth, the most extreme differences appear in first few metres thus comparisons between 15 m and the mesophotic are considerably greater than between differences within mesophotic zones (e.g., between 50 and 85 m depth). The observed depth-differentiation within A. lamarcki and not A. grahamae follows these light transitions, although environmental factors such as nutrient availability, salinity, flow environment and temperature also vary between shallow and mesophotic depth zones (Bongaerts, Frade, et al., 2015; Dollar, 1982; Lesser et al., 2009) and could be contributing factors. Similar to the present study, a survey of A. lamarcki using ddRAD (Hammerman et al., 2018) also found two sympatric cryptic genetic groups but within Puerto Rico (sampled mostly between 10 and 20 m depth). At the population level, depth differentiation has been observed in Bermuda for the congener A. fragilis, where populations at 12 and 40 m were strongly differentiated (Bongaerts et al., 2017). Future surveys that can measure differences in microhabitats within the same depth profile may be informative for determining more precise environmental niches distinguishing cryptic taxa rather than depth alone.

In contrast to cnidarian host genotypes, we found almost complete depth partitioning of the symbiont profile within A. lamarcki (Figure S1-3) which further supports the potential for
depth-associated divergence. However, the symbiont profiles did not correspond directly to the different host taxa as expected due to algal symbiont specificity patterns seen within the genus, brooding taxa within the Caribbean (Bongaerts et al., 2015a) and other taxa (Prada et al., 2014; Thornhill et al., 2014). The symbiont profile found in individuals of A. lamarcki at mesophotic depths was the same profile shared by A. grahamae (Figure S1-3, Bongaerts et al., 2013; Bongaerts, Carmichael, et al., 2015). Thus, this could indicate an ancestral association with this symbiont, more recent acquisition through host hybridisation, or potentially horizontal symbiont acquisition of this depth-specific endosymbiont (Quigley et al., 2018), which is indeed shown in found coral taxa (Bongaerts, Carmichael, et al., 2015; Rowan & Knowlton, 1995). Another explanation for this disparity may be incomplete host-symbiont lineage sorting. Host divergence (between A. lamarcki taxa) could predate symbiont sorting due to retention in the maternal line. Further experimental work should investigate symbiont inheritance patterns and use modern methods of symbiont genotyping (amplicon sequencing) to better determine the relationship between host genotype and algal symbiont within these taxa.

4.3 Gene flow between cryptic taxa and species

Cryptic taxa are commonly found within genetic studies of scleractinians, yet until recently their divergence history has not been examined in detail in either shallow or mesophotic environments (but see Prada & Hellberg, 2020; Rippe et al., 2021). Often, cryptic coral taxa are closely related and sympatrically occurring and thus have the potential to interbreed. Here, we find evidence of two genetically distinct taxa within both species (Figures 2–4) and backcross individuals between taxa (see Section 3, Table S2-4). An F1 hybrid was also found between A. grahamae and the “deep” A. lamarcki (Figure 3) which could suggest a continued low rate of recent interbreeding between species. Finally, demographic models that included gene flow during divergence (i.e., continuous migration, ancient migration, and secondary contact) consistently had higher likelihoods than the null models of no migration for both between cryptic taxa within and between species (Figure 5, Figure S1-4, Tables 1 and 2).

Although results from dadi gave the greatest support to the two-epoch models of divergence in isolation followed by secondary contacts (Tables 1 and 2), if population expansions or bottleneck occurred, divergence time can be overestimated and results can favour the secondary contact model (Momigliano et al., 2021). As RAD data sets with de novo assemblies are prone to high error rates in low frequency SNP calling (Andrews et al., 2016), we did not include singletons and doubletons in the analyses and thus were unable to accurately model population size changes and differentiate between the different migration scenarios (i.e., ancient migration vs. continuous migration vs. secondary contact). Thus, our inferences about timing and relative model support (among models including migration) should be interpreted cautiously. Nonetheless, models with no migration were consistently rejected and therefore we can confidently conclude that limited migration between distinct taxa has occurred between both cryptic taxa within and between species. Additionally, the relative amount of estimated gene flow scaled with divergence time, where migration was less between A. lamarcki and A. grahamae as compared to migration between less diverged taxa within each species. These results are consistent with hypothesis that genetic permeability scales with divergence time (Roux et al., 2016).

We were not able to confidently resolve whether the very early stages of divergence occurred with or without gene flow. However, our results and those of other studies (e.g., Madracis spp., Frade et al., 2010; Eunica felxuosa, Pocillopora damicornis, Prada & Hellberg, 2020; Agaricia fragilis, Bongaerts et al., 2017; Prada & Hellberg, 2020 and Montastraea cavernosa and Siderastrea siderea, Rippe et al., 2021) confirm that low levels of gene flow connect such closely related taxa, and yet gene flow is insufficient to homogenise them. Without physical barriers between cryptic taxa, exogenous and/or endogenous barriers must maintain divergence. The occurrence of exogenous selection has been shown in many famous examples to maintain species barriers through disruptive selection in face of homogenising gene flow (e.g., sticklebacks, Dean et al., 2019; Darwin’s finches, Han et al., 2017; Heliconious butterflies, Merot et al., 2017; and cichlids, Poelstra et al., 2018). Among scleractinarians there is considerable circumstantial evidence for depth-associated environmental attributes (including light, nutrient availability, temperature, water flow, etc.) providing strong exogenous selection that could contribute to maintaining divergence despite gene flow (Bongaerts et al., 2011, 2017; Carlon & Budd, 2002; Gorospe & Karl, 2015; Prada & Hellberg, 2013; Serrano et al., 2014; van Oppen et al., 2011). Consistently, depth partitioning is observed throughout the Agaricia genus, with each taxonomic species inhabiting a distinct depth profile yet each remaining sympatric at the edge of their depth range (Bongaerts et al., 2013). Depth-associated factors appear to play a role in divergence-with-gene flow within A. lamarcki taxa, between A. grahamae and A. lamarcki but not within A. grahamae taxa.

Endogenous barriers may also be maintaining the cohesiveness of cryptic taxa and indeed the replicated pattern of genome-wide divergence (not just divergence at selected outlier loci) implicates isolating mechanisms that may not be solely limited to environmental factors. Pre-zygotic isolation is most likely to occur through either gametic incompatibilities or temporal and spatial isolation in spawning in organisms where one or both gametes spawn, such as in corals and other marine invertebrates. In the Orbicella coral genus, more closely related species experienced temporal differences in spawning time albeit with gametic compatibilities (in crossed experiments). More distantly related species had overlapping times with gametic incompatibilities (Knowlton et al., 1997; Levitan et al., 2004), thus demonstrating the interactions and development of reproductive barriers in corals. On the other hand, in Indo-Pacific closely-related Acropora species spawning times often overlap and interspecific gamete compatibility can be high, although in the presence of conspecific sperm the number of hybrid offspring produced is low (Willis et al., 2006). Post-zygotic
isolation studies through experimental work in crosses of Acropora species have found equal or higher fitness F1 hybrids (compared to parents) in parental habitats as well as greater fitness in F1 hybrids than parents in higher temperature or environmentally variable habitats (Chan et al., 2019; Willis et al., 2006), but such fitness estimates for F2s and backcrosses remain unclear. Due to practical difficulties, there are few spawning and larval crossing experiments in many coral species (especially for brooders), therefore most endogenous pre- or post-zygotic barries have not been fully tested. The common occurrence of sympatric yet incompletely reproductively isolated genetic taxa within scleractinians is perplexing and suggests that common features contribute to this widespread phenomenon.

5 | CONCLUSIONS

Here, we have shown that there is previously undescribed cryptic genetic diversity within Agaricia and that low levels of gene flow among taxa have characterised their divergence history. Thus, species boundaries within Agaricia appear to be semi-permeable. Intertaxon gene flow stresses the importance of considering multiple closely related taxa in population genomic assessments, since erroneous conclusions could be drawn regarding assignment to spatial population structure and species identity by purely relying on morphology to group individuals into taxa. Importantly introgression is likely to be common in scleractinian corals (Mao et al., 2018) and single species studies often ignore this important source of genetic diversity that could be adaptive in times of environmental change.

The lack of genetic structuring of mesophotic coral populations among reefs and islands indicates that mesophotic-specialist species may be more horizontally connected than we anticipated, which is important regarding their enhanced ability to recover from localised disturbances through replenishment from neighbouring reefs and islands. Pairing genetic data with hydrodynamic modelling approaches may help to resolve patterns of gene flow for regions of concern. The spatial genetic structure disparity between members of the genus may reflect undescribed differences in reproductive biology, and thus traits may be more variable and diverse within genera. This justifies the need to investigate species where localised dispersal has been assumed, especially those understudied in the mesophotic zone. This study draws attention to the lack of knowledge in the ecology and diversity of mesophotic corals. In light of recent reports of disturbances to mesophotic reefs, inferences about localised dispersal have been made, understanding mesophotic coral connectivity should be a priority in future studies.

ACKNOWLEDGEMENTS

We thank Mark Vermeij, Pedro Frade and Margaux Carmichael for their support in the field. We thank the CARMABI Research Station, Substation Curacao, the CIEE Bonaire Research Station, and STINAPA Bonaire for logistical support. In Colombia, we would like to thank CORALINA (Convenios No. 13, 2014 and No. 21, 2015) and Fanny L. Gonzalez-Zapata, Nacor Balaños and Luisa F. Dueñas for assisting in collections and analyses. We would also like to thank Jane Hughes for generously offering access to her molecular laboratory. This project was supported by the XL Catlin Seaview Survey (funded by the XL Catlin Group in partnership with Underwater Earth and The University of Queensland), “The Explorers Club - Eddie Bauer Grant for Expeditions”, and an Australian Research Council Discovery Early Career Researcher Award (DE160101433). Open access publishing facilitated by The University of Queensland, as part of the Wiley - The University of Queensland agreement via the Council of Australian University Librarians.

AUTHOR CONTRIBUTIONS

Pim Bongaerts designed and planned the study. Pim Bongaerts, Kyra B. Hay and Norbert Englebert conducted fieldwork in Curacao and Bonaire. Juan A. Sánchez provided samples from San Andrés (Colombia). Kelly R. W. Latijnhouwers conducted all laboratory work. Katharine E. Prata analysed data and wrote the manuscript. Pim Bongaerts and Cynthia Riginos contributed to writing the manuscript and provided guidance on all analyses. Ryan N. Gutenkunst provided advice on demographic analyses. All authors contributed to the manuscript edits.

DATA AVAILABILITY STATEMENT

Raw NextRAD sequences have been stored on NCBI BioProject Accession PRJN801008. Metadata with sequences are stored on Genomic Observatories Metadata Database (GEOME) https://n2t.net/ark:/21547/EAd2. Scripts from the demographic modelling using dadi are stored on https://https://github.com/kepra3/kp_dadi as an electronic notebook.

ORCID

Katharine E. Prata https://orcid.org/0000-0001-6679-0066
Cynthia Riginos https://orcid.org/0000-0002-5485-4197
Juan A. Sánchez https://orcid.org/0000-0001-7149-8369

REFERENCES

Alexander, D. H., Novembre, J., & Lange, K. (2009). Fast model-based estimation of ancestry in unrelated individuals. Genome Research, 19(9), 1655–1664. https://doi.org/10.1101/gr.094052.109
Anderson, E. C., & Thompson, E. A. (2002). A model-based method for identifying species hybrids using multilocus genetic data. Genetics, 160(3), 1217–1229. https://doi.org/10.1093/genetics/160.3.1217
Andrade, C. A. (2003). Evidence for an eastward flow along the Central and South American Caribbean Coast. Journal of Geophysical Research, 108(C6), 3185. https://doi.org/10.1029/2002JC001549
Andrews, K. R., Good, J. M., Miller, M. R., Luikart, G., & Hohenlohe, P. A. (2016). Harnessing the power of RADseq for ecological and evolutionary genomics. Nature Reviews Genetics, 17(2), 81–92. https://doi.org/10.1038/nrg.2015.28
Arrigon, R., Benzioni, F., Terraneo, T. I., Caragnano, A., & Berumen, M. L. (2016). Recent origin and semi-permeable species boundaries in the scleractinian coral genus Stylophora from the Red Sea.
Storey, J. D., Bass, A. J., Dabney, A., & Robinson, D. (2019). qvalue: Q-value estimation for false discovery rate control. (2.18.0) [R package]. http://github.com/jdstorey/qvalue

Studivan, M. S., & Voss, J. D. (2018). Population connectivity among shallow and mesophotic Montastraea cavernosa corals in the Gulf of Mexico identifies potential for refugia. Coral Reefs, 37(4), 1183–1196. https://doi.org/10.1007/s00338-018-1733-7

Tay, Y. C., Noreen, A. M. E., Suharsono, Chou, L. M., & Todd, P. A. (2015). Genetic connectivity of the broadcast spawning reef coral Platgyra sinensis on impacted reefs, and the description of new microsatellite markers. Coral Reefs, 34(1), 301–311. https://doi.org/10.1007/s00338-014-1206-6

Thornhill, D. J., Lewis, A. M., Wham, D. C., & LaJeunesse, T. C. (2014). Host-specialist lineages dominate the adaptive radiation of reef coral endosymbionts: Adaptive radiation of symbiotic dinoflagellates. Evolution, 68(2), 352–367. https://doi.org/10.1111/evo.12270

Underwood, J. N., Smith, L. D., van Oppen, M. J. H., & Gilmour, J. P. (2006). Multiple scales of genetic connectivity in a brooding coral on isolated reefs following catastrophic bleaching: Genetic connectivity in a brooding coral. Molecular Ecology, 16(4), 771–784. https://doi.org/10.1111/j.1365-294X.2006.03187.x

van Oppen, M. J. H., Bongaerts, P., Underwood, J. N., Peplow, L. M., & Cooper, T. F. (2011). The role of deep reefs in shallow reef recovery: An assessment of vertical connectivity in a brooding coral from west and east Australia: Vertical connectivity in a brooding coral. Molecular Ecology, 20(8), 1647–1660. https://doi.org/10.1111/j.1365-294X.2011.05050.x

van Oppen, M. J. H., Lukoschek, V., Berkelmans, R., Peplow, L. M., & Jones, A. M. (2015). A population genetic assessment of coral recovery on highly disturbed reefs of the Keppel Island archipelago in the southern Great Barrier Reef. PeerJ, 3, e1092. https://doi.org/10.7717/peerj.1092

Veron, J. E. N. (1995). Corals in space and time: The biogeography and evolution of the Scleractinia. Cornell University Press.

Veron, J. E. N. (2000). Corals of the world (Issue C/593.6 V4).

Warner, P. A., van Oppen, M. J. H., & Willis, B. L. (2015). Unexpected cryptic species diversity in the widespread coral Seriatopora hystrix masks spatial-genetic patterns of connectivity. Molecular Ecology, 24(12), 2993–3008. https://doi.org/10.1111/mec.13225

Warner, P. A., Willis, B. L., & van Oppen, M. J. H. (2016). Sperm dispersal distances estimated by parentage analysis in a brooding scleractinian coral. Molecular Ecology, 25(6), 1398–1415. https://doi.org/10.1111/mec.13553

Wells, J. W. (1973). New and old scleractinian corals from Jamaica. Bulletin of Marine Science, 43.

Willis, B. L., Babcock, R. C., Harrison, P. L., & Wallace, C. C. (1997). Experimental hybridization and breeding incompatibilities within the mating systems of mass spawning reef corals. Coral Reefs, 16, 553–565. https://doi.org/10.1007/s003380050242

Willis, B. L., van Oppen, M. J. H., Miller, D. J., Vollmer, S. V., & Ayre, D. J. (2006). The role of hybridization in the evolution of reef corals. Annual Review of Ecology, Evolution, and Systematics, 37(1), 489–517. https://doi.org/10.1146/annurev.ecolsys.37.091305.110136

How to cite this article: Prata, K. E., Riginos, C., Gutenkunst, R. N., Latijnhouwers, K. R. W., Sánchez, J. A., Englebert, N., Hay, K. B., & Bongaerts, P. (2022). Deep connections: Divergence histories with gene flow in mesophotic Agaricia corals. Molecular Ecology, 31, 2511–2527. https://doi.org/10.1111/mec.16391

SUPPORTING INFORMATION

Additional supporting information may be found in the online version of the article at the publisher’s website.