










## SPECIAL ISSUE ARTICLE

# Cryptic diversity and spatial genetic variation in the coral *Acropora tenuis* and its endosymbionts across the Great Barrier Reef

Ambrocio Melvin A. Matias<sup>1,2</sup>  | Iva Popovic<sup>2</sup>  | Joshua A. Thia<sup>3</sup>  | Ira R. Cooke<sup>4</sup>  | Gergely Torda<sup>5</sup>  | Vimoksalehi Lukoschek<sup>5,6</sup>  | Line K. Bay<sup>7</sup>  | Sun W. Kim<sup>2</sup>  | Cynthia Riginos<sup>2</sup> 

<sup>1</sup>Institute of Biology, University of the Philippines Diliman, Quezon City, Philippines

<sup>2</sup>School of Biological Sciences, The University of Queensland, St. Lucia, Queensland, Australia

<sup>3</sup>Bio21 Institute, School of BioSciences, The University of Melbourne, Parkeville, Victoria, Australia

<sup>4</sup>College of Public Health, Medical and Veterinary Sciences, James Cook University, Townsville, Queensland, Australia

<sup>5</sup>ARC Centre of Excellence for Coral Reef Studies, James Cook University, Townsville, Queensland, Australia

<sup>6</sup>Gold Coast University Hospital, QLD Health, Southport, Queensland, Australia

<sup>7</sup>Australian Institute of Marine Science, Townsville, Queensland, Australia

## Correspondence

Cynthia Riginos, School of Biological Sciences, The University of Queensland, St. Lucia, Qld, Australia.  
Email: [c.riginos@uq.edu.au](mailto:c.riginos@uq.edu.au)

## Funding information

Great Barrier Reef Foundation

## Abstract

Genomic studies are uncovering extensive cryptic diversity within reef-building corals, suggesting that evolutionarily and ecologically relevant diversity is highly underestimated in the very organisms that structure coral reefs. Furthermore, endosymbiotic algae within coral host species can confer adaptive responses to environmental stress and may represent additional axes of coral genetic variation that are not constrained by taxonomic divergence of the cnidarian host. Here, we examine genetic variation in a common and widespread, reef-building coral, *Acropora tenuis*, and its associated endosymbiotic algae along the entire expanse of the Great Barrier Reef (GBR). We use SNPs derived from genome-wide sequencing to characterize the cnidarian coral host and organelles from zooxanthellate endosymbionts (genus *Cladocodium*). We discover three distinct and sympatric genetic clusters of coral hosts, whose distributions appear associated with latitude and inshore–offshore reef position. Demographic modelling suggests that the divergence history of the three distinct host taxa ranges from 0.5 to 1.5 million years ago, preceding the GBR's formation, and has been characterized by low-to-moderate ongoing inter-taxon gene flow, consistent with occasional hybridization and introgression typifying coral evolution. Despite this differentiation in the cnidarian host, *A. tenuis* taxa share a common symbiont pool, dominated by the genus *Cladocodium* (Clade C). *Cladocodium* plastid diversity is not strongly associated with host identity but varies with reef location relative to shore: inshore colonies contain lower symbiont diversity on average but have greater differences between colonies as compared with symbiont communities from offshore colonies. Spatial genetic patterns of symbiont communities could reflect local selective pressures maintaining coral holobiont differentiation across an inshore–offshore environmental gradient. The strong influence of environment (but not host identity) on symbiont community

Ambrocio Melvin A. Matias and Iva Popovic joint first authors.

This is an open access article under the terms of the [Creative Commons Attribution](https://creativecommons.org/licenses/by/4.0/) License, which permits use, distribution and reproduction in any medium, provided the original work is properly cited.

© 2022 The Authors. *Evolutionary Applications* published by John Wiley & Sons Ltd.

composition supports the notion that symbiont community composition responds to habitat and may assist in the adaptation of corals to future environmental change.

#### KEYWORDS

cryptic species, demographic history, endosymbiont, gene flow, introgression, population genomics

## 1 | INTRODUCTION

Genetic biodiversity is not always packaged into tidy species groups. Rather, dynamic tensions between divergence and homogenization shape genetic diversity among taxa that interbreed occasionally in spatially and (or) temporally variable manners (Duranton et al., 2018). How and when taxa diverge to create reproductive barriers are central problems of speciation (Coyne & Orr, 2004; Roux et al., 2016). But also, how and when distinct taxa solidify or soften their delineations, and whether divergence inevitably results in speciation, remain open questions in evolutionary biology (Galtier, 2019; Hey & Pinho, 2012; Sousa & Hey, 2013). With the increasing use of genomic approaches, it is evident that genomic regions can vary in their permeability to gene flow; regions of the genome that can resist the homogenizing effect of gene flow can maintain local adaptation and facilitate divergence, especially for high dispersal taxa (Andrew & Rieseberg, 2013; Martin et al., 2013; Riquet et al., 2019). Consequently, cryptic diversity can manifest over spatial scales that are expected to be homogenized by gene flow. Characterizing such cryptic diversity and understanding its spatial distribution are important for developing biodiversity management strategies in our rapidly changing world (Hey et al., 2003).

Corals form the basis of reef ecosystems that harbour exceptional biodiversity (Fisher et al., 2015), yet delineation of coral species remains contentious due to frequent mismatches between morphological and molecular systematics, morphological plasticity, cryptic genetic diversity and porous taxonomic boundaries. Genome-scale investigations provide more robust and better-resolved estimates of genetic relationships than previous approaches due to the analytic depth afforded by thousands or more loci. Recent genetic studies have revealed surprising divisions and gene exchange among coral taxa, provoking re-evaluations of both deep evolutionary relationships (Fukami et al., 2008; Kitahara et al., 2014; Ying et al., 2018) and relationships among closely related taxa. For example, population genomic surveys have uncovered distinct genetic clusters within *Acropora* (Cooke et al., 2020; Fifer et al., 2021; Rose et al., 2018; Underwood et al., 2020), *Agaricia* (Bongaerts et al., 2017; Prata et al., 2022), *Montastraea* (Rippe et al., 2021; Sturm et al., 2020), *Pachyseris* (Bongaerts et al., 2021), *Pocillopora* (Smith et al., 2017; van Oppen et al., 2018) and *Siderastrea* (Rippe et al., 2021), where some earlier microsatellite-based studies had failed to detect these clusters. (For full taxonomic authorities for all species mentioned

throughout this paper, see the World Register of Marine Species: [www.marinespecies.org](http://www.marinespecies.org)). Cryptic differentiation in corals often coincides with environmental gradients, including depth (Bongaerts et al., 2017; Rippe et al., 2021; van Oppen et al., 2018) and inshore–offshore shelf location (Kenkel et al., 2013; Warner et al., 2015), indicating that processes structuring variation can occur on small spatial scales relative to the species' range. There is, thus, an evolutionary tug-of-war between processes that drive divergence versus those that maintain genetic cohesion among many closely related coral taxa. Characterizing the breadth and distribution of coral diversity therefore requires sampling across relevant spatial scales and environments and employing analytical approaches that are sufficiently resolute to capture the evolutionary processes contributing to this diversity (as in Bierne et al., 2013).

The cnidarian host's intimate symbiosis with dinoflagellates from the family Symbiodiniaceae further complicates the delineation of corals into discrete species. Symbiont compositions have profound effects on the response of the coral holobiont (or composite organism) to environmental stress (Cooper et al., 2011; Hoadley et al., 2021; Manzello et al., 2019; Morikawa & Palumbi, 2019; Sampayo et al., 2008). Within coral hosts, symbiont compositions can vary spatiotemporally in response to environmental gradients, such as depth (Bongaerts et al., 2010, 2013, 2015) and shelf location (Cooper et al., 2011; Howells et al., 2009; Sawall et al., 2014; Tonk et al., 2013; Ziegler et al., 2015). Inshore habitats generally experience more extreme and variable conditions relative to offshore habitats, including warmer and more variable temperatures, greater turbidity and greater concentrations of inorganic nutrients (Camp et al., 2020; Manzello et al., 2019; Ziegler et al., 2015). Symbionts from inshore corals may also facilitate greater recovery from bleaching relative to those from offshore corals (Manzello et al., 2019). Patterns of symbiont variation across environmental gradients may represent important sources of variation that confer adaptive responses to environmental stress, comprising an important axis of coral genetic variation that is not necessarily constrained by taxonomic boundaries of the cnidarian host. The decoupling of symbionts and host inheritance is likely to typify many broadcast spawning corals, where coral juveniles acquire their endosymbionts from the local environment (Baird et al., 2009; Madin et al., 2016; but see Quigley et al., 2020). Spatial costructuring between symbionts and hosts is sparsely documented, and this study builds on previous studies that simultaneously consider spatial genetic variation of

coral hosts and symbionts at the colony level (exceptions include: Bongaerts et al., 2021; Cooke et al., 2020; Kenkel et al., 2013; Rose et al., 2018, 2021; Thomas et al., 2022; van Oppen et al., 2018; Warner et al., 2015).

In this study, we simultaneously examine host and symbiont genetic variation of the reef-building coral *Acropora tenuis* (Dana, 1846–1849) across the entire GBR, sampling both latitudinal and inshore–offshore gradients. *A. tenuis* is a common branching coral of the Great Barrier Reef (GBR) and the wider tropical Indo-Pacific, with broadcast spawning reproduction. Previous microsatellite surveys of *A. tenuis* (Lukoschek et al., 2016; Riginos et al., 2019) spanning much of the GBR (1500+ km) found greater genetic diversity in southern GBRs (especially the Swains and Capricorn–Bunker reef groups) compared with central and northern reefs, with subtle signs of possible admixture from the Coral Sea. There was no compelling evidence for cryptic speciation, and symbiont diversity was not examined. In contrast, using low-coverage whole-genome sequencing and sampling across ~350 km of inshore central GBR reefs, Cooke et al. (2020) found substantial differentiation of colonies from one site, Magnetic Island, compared with colonies from the other four sampling sites, with low levels of gene flow between Magnetic Island and other locations. In the same study, symbiont mitochondrial haplotypes were differentiated among all five sampling locations. The three locations most affected by riverine plumes, including Magnetic Island, contained symbiont haplotypes not found in the two more maritime locations, consistent with environmental filtering of symbionts. Thus, the initial genomic survey by Cooke et al. (2020) suggests that GBR *A. tenuis* may comprise more than one distinctive gene pool, that is, multiple species sensu Mallet (1995). Moreover, this prior work shows that geographic differentiation of symbionts can exceed that of the host animal.

Here, we characterize *A. tenuis* colonies across the GBR using genome-wide single nucleotide polymorphisms (SNPs) in the cnidarian hosts and SNPs from symbiont plastids. We aimed to: (1) determine whether *A. tenuis* resolved into two or more cryptic taxa as evidenced by distinct genetic clusters; (2) document spatial extents of such cryptic taxa and their associations with latitude and shelf position; (3) test for gene flow among cryptic taxa using demographic modelling; and (4) characterize patterns of symbiont diversity within and among coral colonies with respect to latitude, shelf position and cryptic cnidarian host taxon identity. Our data suggest that at least three discrete taxa comprise what is currently recognized as *A. tenuis* on the GBR. We demonstrate that diversity in the cnidarian host and its symbionts are constructed by similar environmental processes but with much stronger genetic–environment associations for symbionts than coral hosts. Additionally, we find that structuring of symbiont variation is independent of genetic identity of cnidarian hosts, suggesting that these discrete but cryptic cnidarian taxa share the same, environmentally determined symbiont pool. This study provides the most spatially comprehensive joint survey of genomic variation in cnidarian hosts and their symbiont partners for any coral species on the GBR to date.

## 2 | MATERIALS AND METHODS

### 2.1 | Sample collection, DNA extraction and genomic library preparation

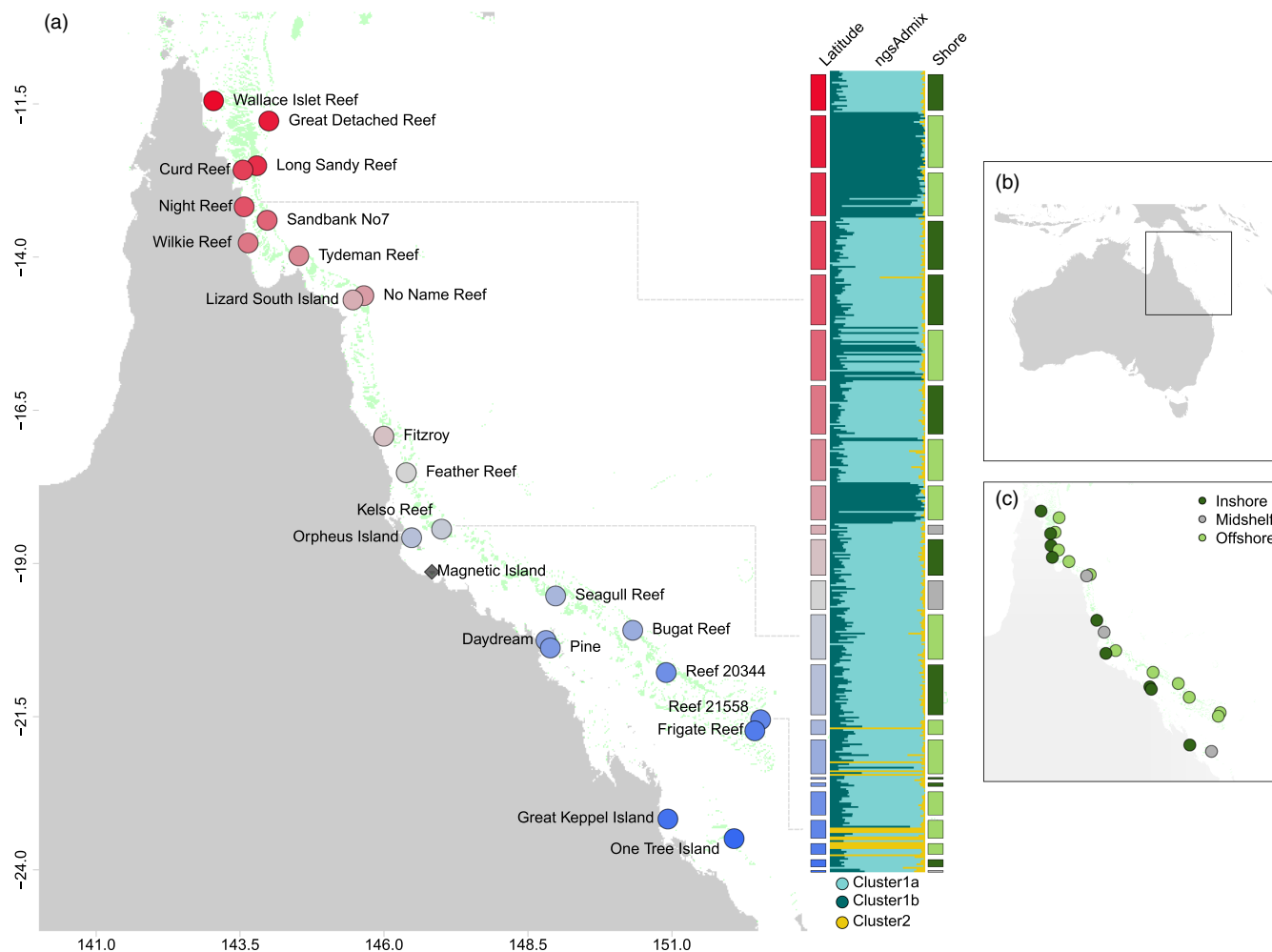
Details for all methods including function calls and settings can be found in the Supplemental Information in Appendix S1. Adult colonies from 26 locations along the span of the GBR were sampled (Figure 1 and Table S1). Throughout this paper, we refer to our study organism as *A. tenuis* despite recent evidence that this species on the GBR may be different from the *A. tenuis* type specimen from Fiji and the appropriate GBR species name may be *Acropora kenti* (T. Bridge pers comm; also see comment in Cooke et al., 2020). In the absence of a formal taxonomic revision of the species, we continue with recent conventions to avoid further confusion. Whole genome libraries were prepared and enriched for a selection of ~21,000 nuclear SNPs randomly distributed across the *A. tenuis* genome using custom bait probes (Arbor Biosciences). Libraries for 698 individual colonies and four technical replicates were sequenced on an Illumina HiSeq HT (125 bp paired end reads) to approximately 5× coverage.

### 2.2 | Processing, filtering and clone identification from genomic data

*FastQC* (<http://www.bioinformatics.bbsrc.ac.uk/projects/fastqc>) and *bbtools* (Bushnell, 2014) were used for initial data processing, and *BWA* (Li & Durbin, 2009) was used to map paired reads against the *A. tenuis* draft genome (*aten\_final\_0.11.fasta*; Cooke et al., 2020). Data conversion, indexing, sorting and annotations were undertaken with *samtools* v.1.10 (Li et al., 2009) and *picard* (<http://broadinstitute.github.io/picard/>).

We used *ANGSD* v0.921 (Korneliussen et al., 2014) to identify polymorphic sites and estimate genotype likelihoods in a probabilistic framework that incorporates statistical uncertainty associated with sequencing errors and missing genotypes in low-coverage genomic datasets (Kim et al., 2011). For low-coverage genomic data, using genotype likelihoods improves the accuracy of population genetic inferences (Warmuth & Ellegren, 2019) and diversity estimates (Korunes & Samuk, 2021) and *ANGSD* enabled analyses are the most accepted for low-coverage population genomic data (Matz, 2018; Therkildsen & Palumbi, 2016). For *A. tenuis*, only sites that included 3+ reads in ≥40% of individuals were retained, and only the top 65% of individuals, ranked by the total number of positions covered by 3+ reads passing minimum quality criteria of mapping quality 30 and base quality 30, were kept.

Potential clones were identified following the approach of Manzello et al. (2019), where this method identifies likely clones by applying a threshold on genetic similarities estimated as identity-by-state (IBS) between individuals. Four pairs of technical sequence replicates allowed us to determine the upper 95% confidence interval for pairwise genetic differences (i.e. 1–IBS) for multilocus



**FIGURE 1** Map of sampling locations across the Great Barrier Reef (GBR) and major genetic clusters of *Acropora tenuis* samples based on *NGSadmix*. (a) Geographic locations of *A. tenuis* samples from 23 reefs across the GBR with each location coloured by their latitude. The *NGSadmix*-derived genetic ancestry proportions and clustering of each of 448 colonies is summarized by the plot (with  $K = 3$ ) at the right side of (a). Bars on the left side of the *NGSadmix* plot indicate sampling locations (coloured as latitude), and bars on the right side show the shore position of the sampling location. Black dashed lines connecting the points on the map to the bars of the *NGSadmix* plot are added for quick reference. (b) Map of Australia illustrating where the sampling locations are situated. (c) Sampling locations coloured based on shore position

genotypes (MLG). We applied this upper limit as a threshold and then pruned sets of individuals falling below this value to a single representative.

### 2.3 | Identification of major host clusters, verification of taxonomy and integration with previous *A. tenuis* studies

Genotype likelihoods were used to infer genetic clusters among our samples with *ANGSD*. The sites used for this analysis required minimum criteria of mapping quality 30, base quality 30 and coverage  $\geq 3$  reads in 85% of individuals. We called major and minor alleles directly from the genotype likelihoods assuming biallelic sites with a likelihood ratio test  $p$ -value  $< 0.000001$ . An individual covariance matrix was extracted using *PCAngsd* (Meisner & Albrechtsen, 2018)

with a 0.05 minor allele frequency (MAF) threshold, and the optimal number of ancestral populations ( $K$  genetic clusters) was determined using the ‘-admix’ option. Principal components analysis (PCA) was performed using R (v4.0.1; R Core Team, 2019). The Bayesian hierarchical clustering method *NGSadmix* (Skotte et al., 2013) was applied to estimate individuals' ancestry proportions assuming 2–6  $K$  genetic clusters (MAF = 0.05). Our data were most consistent with three distinct genetic clusters, also inferred by *PCAngsd*, which we refer to as Clusters 1A, 1B and 2 (see Section 3).

To contextualize the taxonomic relationships of these clusters with other *Acropora* species, we identified a subset of samples that were representative of their genetic clusters (assignment probability  $q \geq 0.85$  obtained through *NGSadmix*, see Supplemental Information in Appendix S1 for more details). We performed an analysis of ultraconserved element (UCE) and exon consensus sequences to confirm that specimens used in this study were closely related to

the specimens used to build the *A. tenuis* reference genome and the samples from Cooke et al. (2020). UCEs and exon loci have broad utility for resolving evolutionary relationships within the Anthozoa (Quattrini et al., 2018) and species-level relationships within the Acroporidae (Cowman et al., 2020). We extracted a subset of our representative samples with the greatest number of mapped reads (Cluster 1A,  $n = 14$ ; Cluster 1B,  $n = 14$ ; and Cluster 2,  $n = 6$ ) and combined them with two previously published samples (from Magnetic Island and Fitzroy Island: Cooke et al., 2020), and aligned them to loci extracted from the *A. tenuis* reference genome and to two outgroup species (*Acropora millepora* and *Acropora digitifera*). To extract UCE and exon sequences from whole-genome data (from Cooke et al., 2020), we performed an in-silico target-capture and locus assembly process (detailed in the Supplemental Information in Appendix S1). We then performed maximum likelihood phylogenetic inference based on these alignments using IQ-Tree (ver. 2.1.2: Minh et al., 2020).

Complementing the UCE/exon phylogenetic analyses, we also undertook phylogenetic analysis using *A. tenuis* *ITS1* sequences. Using all representative coral samples (as above), we extracted read alignments at the *ITS1* region and obtained consensus sequences using *bcftools* v1.3 (Danecek et al., 2021). A neighbour-joining tree of *Acropora ITS1* sequences was constructed by combining our *ITS1* consensus sequences with reported *Acropora ITS1* sequences available in GenBank. Alignment (*msa* v1.20.1 package: Bodenhofer et al., 2015), DNA substitution model evaluation (*phangorn* v2.6.2 package: Schliep, 2011), genetic distance and neighbour-joining tree construction (*ape* v5.4-1 package: Paradis et al., 2004), and tree visualization (*ggtree* package: Yu et al., 2017) were undertaken in R.

Additionally, to better understand spatial distributions of major genetic clusters, we visualized our data together with previously published *A. tenuis* populations from five inshore reefs ( $n = 148$ : Cooke et al., 2020) using genotype likelihoods as inferred by *ANGSD* and *PCA* in *PCAngsd*. Despite exploring a variety of filtering strategies, the combined dataset showed slight but consistent evidence of a study (or batch) effect (see Section 3 for details), and no further analyses were undertaken that included data from Cooke et al. (2020). We also performed a clustering analysis in *STRUCTURE* (Pritchard et al., 2000) for a set of overlapping colonies (Table S1) from this study and genotyped by Lukoschek et al. (2016) with microsatellites using the same *STRUCTURE* run conditions as the original study.

## 2.4 | Population genetic structure and demographic analysis among host genetic clusters

To evaluate genetic structure within clusters and demographic histories between clusters, we focussed on individuals with  $q \geq 0.85$  from *NGSAdmix* (Cluster 1A,  $n = 141$ , Cluster 1B,  $n = 92$  and Cluster 2,  $n = 14$ ), as these individuals exhibit low admixture and therefore are most characteristic of their genetic cluster (Figure 1). We calculated pairwise  $F_{ST}$  between genomic clusters and between reefs within clusters, considering only reefs with at least three sampled

individuals. We generated unfolded 2D-SFS for each comparison and calculated pairwise  $F_{ST}$  for each site and in overlapping 50kb windows (10 kb step size). We calculated pairwise geographic distances from latitude and longitude using the Vicinity formula for ellipsoids that accounts for the Earth's curvature and tested for isolation by distance (IBD) between reefs using Mantel tests with 1000 permutations in the R package *vegan* v2.5-7 (Oksanen et al., 2017).

We performed model-based simulations in *dadi* v1.7.0 (Gutenkunst et al., 2009) to examine population demographic history, and to test for the presence of gene flow between major clusters, by evaluating the likelihood of different demographic models to the joint site frequency spectra (SFS) between clusters. Although discrimination between ancestral and derived alleles may lend greater inferential power to SFS-based methods, unfolding SFS relies on accurate assignment of ancestral and derived allele states. However, deep evolutionary divergences between *A. tenuis* and available outgroup genomes (e.g. 15 million years between *A. tenuis* and *A. digitifera*; Mao et al., 2018 and among the most divergent but undated branches within *Acropora* for *A. tenuis* - *A. hyacinthus*; Cowman et al., 2020) preclude our ability to infer accurate ancestral states: outgroup sequences are likely to represent independently fixed, lineage-specific mutations rather than ancestral alleles or may lack orthologous sequences. We therefore performed all *dadi* analyses using folded SFS to avoid the risk of incorrect assignments and false demographic inferences. Cooke et al. (2020) also used unfolded SFS for *A. tenuis* whole-genome sequences. We used *ANGSD* and the 'realSFS' method to calculate cluster-specific SFS using the quality criteria as stated above.

We used *dadi* to examine four pairwise demographic models: (i) a divergence model (div), which represents strict isolation with no gene flow between clusters; two isolation-with-migration (IM) models with (ii) symmetric and (iii) asymmetric gene flow; and (iv) a secondary contact (SC) model, which entails strict isolation followed by SC with gene flow. To balance sample sizes and improve computational efficiency, we subsampled individuals from clusters 1A ( $n = 15$ ) and 1B ( $n = 13$ ) to match the low total number of individuals in Cluster 2 ( $n = 14$ ) and prioritized individuals with the highest depth of coverage. From Cluster 1A, we used three groups of individuals (each group composed of 15 samples) sampled in the northern (Wilkie Reef,  $n = 9$ ; Sandbank Reef,  $n = 6$ ), central (Fitzroy Reef,  $n = 7$ ; Kelso Reef,  $n = 8$ ) and southern (Bugatti Reef,  $n = 6$ ; Daydream,  $n = 1$ ; Pine,  $n = 2$ ; Reef 20-344,  $n = 4$ ; Seagull Reef,  $n = 2$ ) GBR to represent the large spatial coverage due to detected IBD within this cluster. For each cluster and each group, we selected individuals that were from the same or nearby reef(s). Using these subsets of individuals, we then generated seven different joint-SFS representing seven comparisons between three Cluster 1A datasets (north, central and south) and Cluster 1B (north) and Cluster 2 (south). Our pairwise comparisons were thus made between sympatric corals (northern Cluster 1A with Cluster 1B; and southern Cluster 1A with Cluster 2), and allopatric corals (central and southern Cluster 1A with Cluster 1B; northern and central Cluster 1A with Cluster 2; and Cluster 1B with Cluster 2). All the folded joint-SFS were generated using *ANGSD*



v0.934 following the same settings as described for single population SFS. For each of the seven joint-SFS, we estimated the likelihood of our four different models and estimated their parameters following the optimization of Portik et al. (2017; see Supplemental Information in Appendix S1). Divergence dating followed the same assumptions as detailed by Cooke et al. (2020).

## 2.5 | Spatial distribution of cnidarian host genetic clusters relative to latitude and shelf position

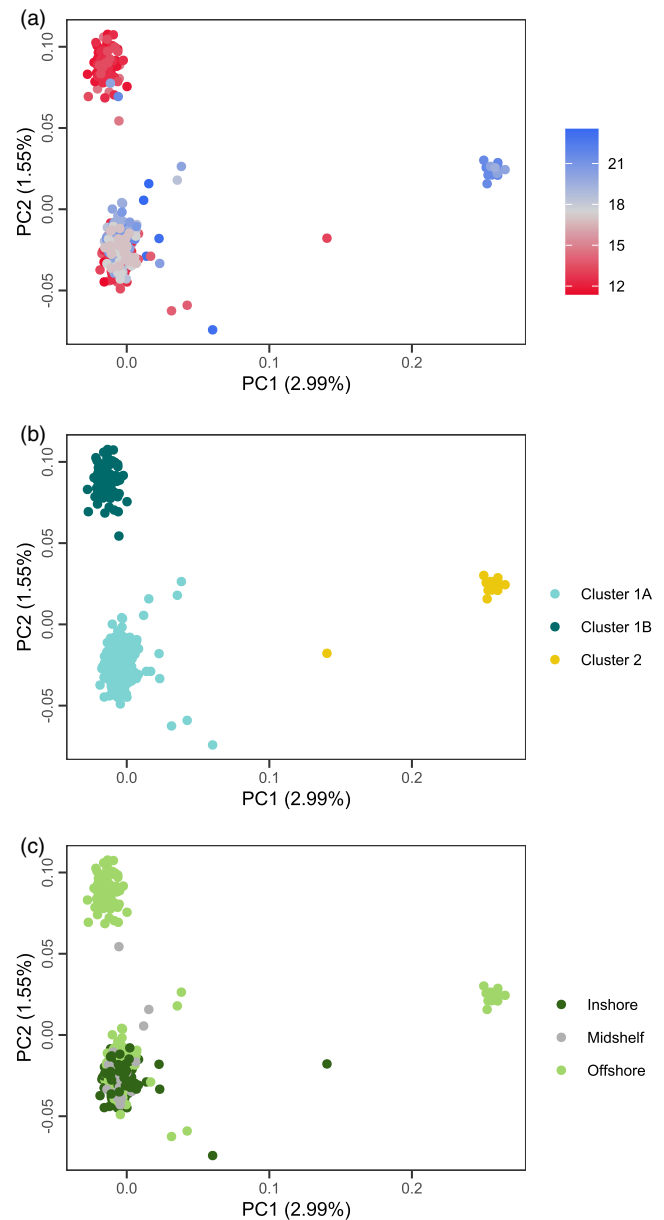
To characterize how genomic variation in coral hosts was distributed across the GBR, we performed a partial redundancy analysis (RDA) to partition the effects of latitude and shelf position on coral host genotypes. We fit the model:

$$Y = \text{Latitude} + \text{Shore} \mid (\text{PC1} + \text{PC2}).$$

Here, our response,  $Y$ , was a matrix of scaled inferred individual allele frequencies ranging from 0 to 1: 0, a homozygote for the major allele; 0.5, a heterozygote; and 1, a homozygote for the minor allele.  $Y$  comprised 448 coral samples and 14,789 SNP loci. The response matrix was conditioned on PC scores of genomic variation (that is, PC1 and PC2; Figure 2) before testing the effects of our focal constraining predictors. The effect of latitude was fit as a continuous predictor (scaled to a mean of 0 and variance of 1), and the effect of shore was fit as a categorical predictor (inshore, midshelf and offshore). The inshore factor was set as the reference factor level for the shore effect. RDA was executed using *vegan* package in R (Oksanen et al., 2017).

## 2.6 | Assaying symbiont diversity using ITS2 and plastid genomes

Our data on symbionts include only short reads obtained from shotgun sequencing of the coral colony, and therefore, typical approaches based on matching longer sequence reads to reference databases were not feasible. Rather, we draw upon strategies for examining allelic variation in pooled sequence data and focus on the community of symbionts within each coral as a pooled sample. We began analyses of symbiont variation by testing whether there were compositional differences of *ITS2* symbiont clades among coral host genetic clusters. A reference set of Symbiodiniaceae *ITS2* sequences was obtained from the SymPortal database (Hume et al., 2019); we used the 'published named' sequence set. We mapped short reads of all coral samples to these symbiont *ITS2* sequences using *bwa mem* and alignments were quality filtered with *samtools* v1.3 for a MAPQ  $\geq 20$ . A custom R script was used to parse SAM files to obtain haplotype matches. A minimum of 3 mapped reads was required as evidence of a symbiont *ITS2* reference in a coral sample. We only observed hits to *ITS2* reference sequences from species of the genus *Cladocopium* with little



**FIGURE 2** Population genomic structure in *Acropora tenuis*. Scatterplots illustrate the first two principal component axes separating *A. tenuis* samples into three major genetic clusters. (a) Samples coloured by latitude south. (b) Samples coloured by genetic cluster, where  $q > 0.50$  was used to determine cluster assignment. (c) Samples coloured by shore position. See respective legends for colour keys

variation in *ITS2* among colonies (see Section 3). Note that we did not attempt to define intragenomic variants as described by Hume et al. (2019); rather, we used the SymPortal reference only to verify that our *ITS2* samples matched *Cladocopium*. Then, we proceeded with downstream analyses from the perspective of within-genus symbiont variation and focussed on sequence diversity in plastid genome of *Cladocopium goreau* (LaJeunesse et al., 2018), where within-colony polymorphism should represent orthologous diversity among strains (without additional paralogous diversity within strains as typified by *ITS2*).

Our novel approach of using plastid genome allele frequencies to characterize symbiont variation is analogous to that well established in pool-seq studies, whereby the goal is to estimate the composition of allele frequencies in a pooled sample and not determine individual genotypes or haplotypes (Schlötterer et al., 2014). Thus, we describe the 'population' diversity of symbionts for each coral colony. Illumina short reads were mapped to the *Cladocopium* plastid genome (Clade C1: Liu et al., 2018) to identify SNP and indel variants, herein referred to as 'symbiont plastid loci'. Details of our custom mapping and filtering procedures are detailed in the Supplementary Methods in Appendix S1. Our final data set of symbiont allele frequencies comprised  $n = 223$  coral samples, with no missing data. The median number of alleles per locus was 2, with a maximum of 17 alleles, and a 75th and 97.5th percentile of 3 and 15.65 alleles. As is the norm with pooled short-read data, we are unable to phase alleles, but can use allele counts (frequencies) to (i) measure symbiont community diversity within each coral colony, and (ii) characterize the distribution of symbiont diversity among different coral taxa or across environmental gradients.

## 2.7 | Spatial distribution of symbiont plastid allele diversity relative to latitude and shelf position

To test how symbiont allelic variation was distributed among coral hosts, we performed two analyses: a generalized linear model (GLM) for the number of symbiont alleles and an RDA. Our GLM of the number of symbiont alleles was used to assess how symbiont diversity varied over latitude, shelf position and coral host. We determined the number of unique alleles per plastid locus for each sampled coral colony:

$$N_A = \text{Counts} + \text{Latitude} + \text{Shore} + \text{Host cluster} + \text{Locus} + \text{Sample}.$$

Here,  $N_A$  was the number of alleles at a locus, counts were the number of mapped reads at a locus, latitude and shore reflect colony position, host cluster was based on PCAs of the host, and locus and sample were fit as random effects to allow a unique intercept for each plastid locus and each coral colony, respectively. All variables were scaled and normalized. The GLM was fit using the *lme4* v1.1-27.1 R package (Bates et al., 2014).

A partial RDA was used to examine symbiont structure among corals and across space. In this analysis, we were interested in the allelic composition of symbionts, that is, which alleles were present and their relative abundance in a coral host.

$$Y = \text{Latitude} + \text{Shore} + \text{Host cluster} | \text{Counts}.$$

Here,  $Y$  was the symbiont allele count matrix, and the remaining variables were as above for the GLM. Both the GLM and the partial RDA were conditioned on read counts to control for the effects of greater sequencing effort at a plastid locus.

## 3 | RESULTS

### 3.1 | Processing, filtering and clone identification from genomic data

We sequenced a total of 702 samples including four sequencing replicates. These samples yielded on average 4,088,885 reads (SD = 3,060,236 reads). After the initial filtering, each sample retained a mean of 3,181,264 reads (SD = 22,850,422). Alignment of the filtered sequences to the reference genome resulted in an average 94% mapping rate (SD = 11%) with ~5% of samples having a mapping rate <73%. The baited regions (including the 500bp regions flanking them, covering a total of 2,172,865 positions) had a mean coverage of 2.01 reads (across all the 702 samples; SD = 2.89), with read depth  $\geq 3$  for 25% of SNPs. The final high-quality data set resulted in 463 individuals and 296,667 loci falling within baited and nonbaited genomic regions (Table S1), where minimum inclusion criteria are detailed in Table S2. Many of the samples were collected in 2010 or earlier and not specifically stored for genomic work; thus, a high rate of individual dropout was expected and was consistent with visual inspection of total DNA visualized on agarose gels indicating substantial DNA degradation for many specimens.

We classified putative clones as individuals showing levels of pairwise genetic distances (1-IBS) below a threshold of 0.17, the upper 95% confidence limit among technical replicates. We identified 11 putative clones as evidenced by shared MLG from eight reefs based on these criteria: Tydeman Reef ( $n = 1$ ), Fitzroy Island ( $n = 2$ ), Kelso Reef ( $n = 2$ ), Orpheus Island ( $n = 1$ ), Seagull Reef ( $n = 1$ ), Bugatti Reef ( $n = 2$ ), Pine Island ( $n = 1$ ) and Reef 20-344 ( $n = 1$ ). After filtering, reducing MLGs to a single representative colony, and removing technical replicates, we retained 60,733 high-confidence loci falling within baited and nonbaited genomic regions (Table S1) across 448 individuals from 23 reefs (Figure 1).

### 3.2 | Identification of major host clusters, verification of taxonomy and integration with previous *A. tenuis* studies

Principal components analysis of 14,789 loci ( $MAF \geq 0.05$ ) in *PCAngsd* revealed three discrete genetic groups (Figure 2) with 1.55% and 2.99% of genetic variation explained on the first and second PCs. We observed strong structuring of coral groups by both latitude and inshore-offshore reef position. PC1 separated two larger genetic groups (Cluster 1A and Cluster 1B) from a small group of individuals (Cluster 2) sampled in the southern GBR (Bugatti Reef, Reef 21-558, Seagull Reef, Frigate Reef). Cluster 1A and 1B were separated on the second PC. Cluster 1A was geographically widespread and represented by individuals sampled from all reefs, except for Great Detached Reef. In contrast, Cluster 1B consisted entirely of individuals sampled from offshore reefs, with almost all individuals originating from northern reefs below 14.7 latitude; two individuals belonging to Cluster 1B are from the southern GBR (Bugatti Reef

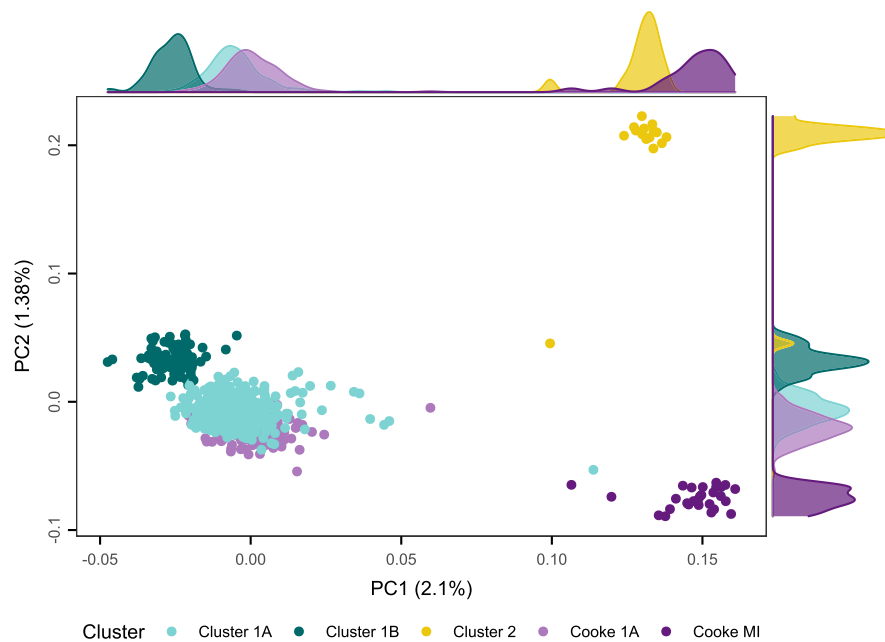
and Reef 21-558), suggesting that this taxon is not restricted to the northern GBR, but likely widespread and sympatric with Cluster 1A. Cluster 2 was restricted to southern offshore reefs, although the group had the fewest samples. *NGSadmix* analyses showed concordant groupings for  $K = 3$  clusters, with nine individuals showing mixed admixture proportions  $>0.30$  (Figure 1). Notably, one individual showed mixed ancestry from divergent Clusters 1A and 2 (Night Reef) and showed an intermediate position on the PCA (Figure 2).

Extracted UCE and exon consensus sequences for phylogenetic analyses yielded 1681 loci with 207,000 variable sites across all samples. A maximum likelihood tree built using IQ-Tree clearly resolved the outgroup species and placed all the samples from this study in a single well resolved (1.0 node support) clade containing the *A. tenuis* reference sequence (Figure S4).

After filtering ( $\leq 5\%$  missing nucleotides) *ITS1* host consensus sequences for additional phylogenetics analysis, we were left with  $n = 46$ ,  $n = 26$  and,  $n = 8$ , for genetic Clusters 1A, 1B and 2, respectively. All *ITS1* consensus sequences from our study were grouped in a single monophyletic clade with the *A. tenuis* sequence AF538489 (Figure S4). An additional *A. tenuis* reference sequence, AF538524, was more closely related to *Acropora cytherea* sequences, indicative of a possible misidentification of this sequence. Analysis of variation in *ITS1* sequences in our samples was unable to resolve Clusters 1A and 1B whereas Cluster 2 formed a monophyletic clade nested within the *A. tenuis* *ITS1* clade, indicating distinctness of this genetic cluster relative to Cluster 1A and 1B, which share *ITS1* variation.

Principal components analysis of 6354 SNPs ( $MAF > 0.5$ ) from our data set combined with five previously published *A. tenuis* populations (Cooke et al., 2020; Figure 3) indicated shared diversity between Cluster 1A and four inshore reefs sampled between Cairns and Townsville (Fitzroy Island, Dunk Island, Pelorus Reef and Pandora Reef). A single individual from our data set sampled from Great Keppel Island in the Capricorn-Bunker Group showed an affinity to the genetically distinct Magnetic Island cluster. Although we found concordant clustering across the two data sets, regardless of filtering parameters employed, there was persistent differentiation between individuals between data sets (see Figure 3, including among individuals sampled from the same reef) indicating some subtle batch effect arising from different attributes of the data, possibly related to genomic library preparation or sequencing platforms (Lou & Therkildsen, 2022). Because of this batch effect, subsequent analyses focussed solely on data from this study.

Clustering analysis using previously published microsatellites for the 395 *A. tenuis* individuals shared between both studies did not resolve the three major genomic clusters (Figure S5). Assignment tests in the original publication found notably more geographic delineation (Lukoschek et al., 2016; Figure 5 and Figure S2) than the *STRUCTURE* results presented here. But the original study included many more colonies ( $n = 2014$ ) and may have been able to leverage extra discriminatory power among groups as compared to a reduced data set ( $n = 365$ ). Regardless, microsatellites both here (Figure S5) and in Lukoschek et al. (2016) failed to identify the distinct clusters revealed by thousands of SNPs (Figures 1 and 2).



**FIGURE 3** Population genetic structure of all *Acropora tenuis* samples from this study ( $n = 448$ ) in combination with those from Cooke et al. (2020) ( $n = 148$ ). Scatterplot illustrates the first two principal component axes summarizing genetic structure. Points are coloured with respect to genetic clusters as described in this study (refer to colour scheme in Figures 1 and 2). Cluster 1A (pale teal), Cluster 1B (dark teal) and Cluster 2 (gold) are samples from this study. Cooke 1A (pale purple) are samples from Cooke et al. (2020) that putatively belong to our defined Cluster 1A. Cooke MI (dark purple) are samples from Cooke et al. (2020) that were sampled from Magnetic Island. Density plots on the edges of the x- and y-axes summarize the density distribution of samples from each cluster



### 3.3 | Population genetic structure and demographic analysis among host genetic clusters

Considering only coral samples with  $q \geq 0.85$  as assigned by *NGSAdmix* as representatives of major genomic clusters (Cluster 1A,  $n = 141$ , Cluster 1B,  $n = 92$  and Cluster 2,  $n = 14$ ), genome-wide average  $F_{ST}$  values between Cluster 2 and either Cluster 1A (0.196) or Cluster 1B ( $F_{ST} = 0.212$ ) were substantively greater than between Cluster 1A and 1B ( $F_{ST} = 0.018$ ). Site-specific  $F_{ST}$  values varied across the genome (Figure S6); Cluster 2 was differentially fixed ( $F_{ST} \geq 0.99$ ) at 42 SNPs relative to Cluster 1A and at 59 SNPs relative to Cluster 1B, whereas comparisons between Clusters 1A and 1B did not indicate any fixed sites ( $F_{ST}$  range: 0–0.77). Within cluster differentiation was globally weak (mean weighted  $F_{ST} = 0.02$ ) and varied between 0.01 and 0.06 among all pairwise comparisons. We found a significant association between genetic differentiation and reef distance within Cluster 1 (Mantel test  $R = 0.44$ ,  $p = 0.01$ ), but associations were not significant within Cluster 1B and Cluster 2 with fewer sampled populations (Figure S7).

Genome-wide estimates of nucleotide diversity ( $\pi$ ) were similar among Clusters 1A (mean = 0.024) and 1B (mean = 0.023), but lower for Cluster 2 (mean = 0.018; Figure S8). Similarly, estimates of average Watterson's  $\theta$  per site were comparable between Cluster 1A (mean = 0.031) and Cluster 1B (mean = 0.030), which were relatively higher than Cluster 2 (mean = 0.018; Figure S8).

In investigating gene flow between clusters, we explicitly compared the divergence-only (strict isolation) model with models involving gene flow and fitted them to 2D joint-SFS involving seven cluster pair combinations as described in the methods (see also Figure S10). The divergence model showed very poor fit for all the seven joint-SFS comparisons that we tested, having consistently more negative log-likelihoods (Figure S9). Among models containing gene flow, we were unable to distinguish isolation-with-migration (IM) scenarios (with either symmetric or asymmetric migration) against SC. The log-likelihoods of these models exhibited a large amount of overlap. Furthermore, the estimates of demographic parameters that were shared across these models were similar (Figure S9). For example, estimates of divergence times between clusters (for IMs and SC) ranged from 0.46 to 1.5 million years (MY) with a median of ~0.8 MY after excluding the extremely high estimates obtained from comparisons of northern Cluster 1A and Cluster 1B. The estimates of SC times all occurred within 0.048–0.094 MY after the estimated divergence (except again for the estimate from northern Cluster 1A and Cluster 1B comparison); the estimated brief period of isolation in our SC model relative to the divergence potentially indicates convergence between our SC and IM results. Because of this convergence, we mainly present the results for IM with asymmetric migration (Figure 4). Thus, the most parsimonious conclusion, based on our demographic analyses, is that divergence among clusters in *A. tenuis* has likely occurred in the presence of weak and ongoing gene flow. For IM and SC models, our results suggest relatively high migration between Cluster 1A and Cluster 1B, with an estimated 1.3 to 21.6 individuals moving between clusters per generation. Indeed, actual

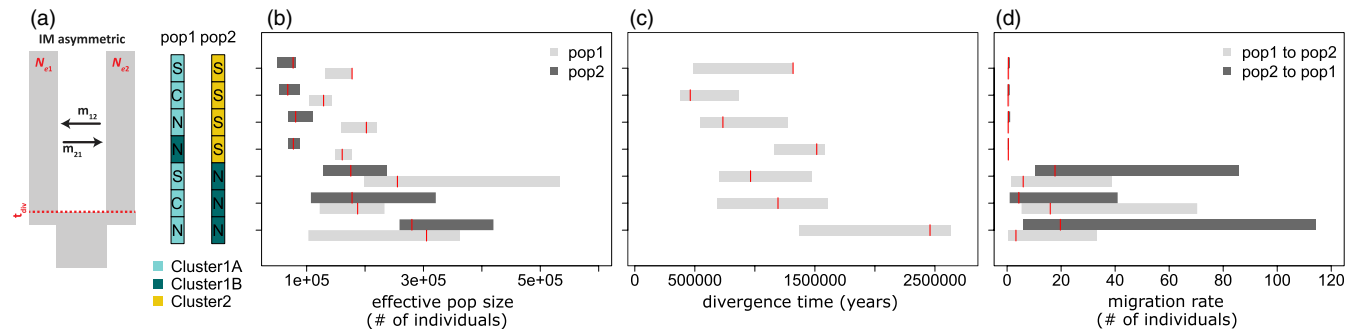
gene flow between Cluster 1A and Cluster 1B could be greater than these estimates as our threshold cut-off of 0.85 excluded many Cluster 1A-like individuals (Figure S3) with secondary affinity to Cluster 1B (Figure 1). While our results indicate ongoing gene flow between Cluster 2 and both Clusters 1A and 1B, the migration rates are low, with an estimated 0.3–0.65 individuals moving between clusters per generation (Figure 4d and Figure S10). Our estimate of  $N_e$  for Cluster 2 is consistently lower than  $N_e$  estimates for Cluster 1A and Cluster 1B (Figure 4b).

### 3.4 | Spatial distribution of cnidarian host genetic clusters relative to latitude and shelf position

The first two PC axes of the partial RDA examining spatial distribution of genotypes across the GBR described 4.7% of the total variation in individual allele frequencies (Figure 5). The full model explained significantly more variation relative to random expectations ( $F_{3,442} = 1.20$ ,  $p = 0.001$ ). The effect of latitude ( $F_{1,442} = 1.48$ ,  $p = 0.001$ ) and shore position ( $F_{2,442} = 1.05$ ,  $p = 0.004$ ) were both significant but small (latitude explained 0.33% of variation in individual allele frequencies, while shelf position explained 0.47%), after partialling out the effect of host genetic cluster. The effect size for spatial variables may be downwardly biased as Clusters 1B and 2 were geographically restricted; therefore, the constrained ordination was driven by samples belonging to the host genetic Cluster 1A, for which we had a large, and well distributed, number of samples. The effects of latitude and shore position on genomic variation were largely uncorrelated, as shown by the essentially perpendicular directions of effect in RDA1/RDA2 dimension space, where RDA1 accounted for 43% of the constrained variation ( $F_{1,442} = 1.52$ ,  $p = 0.001$ ) and RDA2 accounted for 30% of the explainable variation ( $F_{1,442} = 1.10$ ,  $p = 0.006$ ).

### 3.5 | Assaying symbiont diversity using ITS2 and plastid genomes

Mapping coral host short reads to a set of Symbiodiniaceae *ITS2* reference sequences recovered only matches to Clade C (genus *Cladocopium*) (data from 312 colonies retained following filtering). The C40c reference matched 100% of samples, C72k in 0.64% of samples, and C72d in 0.32% of samples, with 4–275 reads mapped (mean = 17.7 per colony) to the *ITS2* locus. Having verified that symbionts were *Cladocopium*, we focussed on plastid sequence variation to describe symbiont diversity within and between coral colonies. Mapping short reads to plastids provided substantial allelic diversity and coverage (data from 223 colonies retained following filtering), with 19 loci (SNPs and indels) from two contigs (contig 1 and contig 7), with 2–17 alleles per locus. The diversity of symbiont plastid alleles within each coral colony varied from a minimum of 1, to a maximum of 8, alleles per locus, including SNPs (diallelic and triallelic) and indels. The average median read depth at a plastid locus was 86.74 reads, with a range of 26–775 median reads.



**FIGURE 4** Parameter estimates for isolation with asymmetric migration for the different pairwise comparisons analysed, where isolation with asymmetric migration models were most consistently among the more likely supported scenarios. (a) Schematic of model and parameters where the coloured boxes indicate the cluster and geographic location of individuals included in the analyses with 'N' being from north, 'C' from central and 'S' for south. The estimates for effective population size (b), divergence time (c) and migration rates (d) were converted from *dadi* parameters following Cooke et al. (2020). Specifically, we used a mutation rate of  $1.86 \times 10^{-8}$  per base per year for the 296,667 SNPs describing the SFS to convert  $\theta$  estimates to  $N_{ref}$ , and a generation time of 5 years was used to convert time estimates to year. In (b) to (d), the bars indicate the 2.5%–97.5% quantile of the bootstrap estimates, while the red horizontal line shows the empirical estimate

### 3.6 | Spatial distribution of symbiont plastid allele diversity relative to latitude and shelf position

Spatial patterns of symbiont plastic diversity showed strong associations with geographic features. The GLM identified a significant effect of latitude ( $F_{1,190} = 12.84$ ,  $p < 0.001$ ), with increasing alleles per locus with increasing latitude ( $\beta = 0.044$ ,  $p = 0.001$ ). There was also a significant effect of shelf position ( $F_{2,196} = 8.32$ ,  $p < 0.001$ ). Relative to inshore reefs, offshore reefs had more alleles per locus ( $\beta = 0.097$ ,  $p = 0.001$ ), whereas midshelf reefs did not significantly differ from inshore reefs ( $\beta = 0.024$ ,  $p = 0.728$ ). The effect of the host genetic cluster ( $F_{2,190} = 0.81$ ,  $p = 0.446$ ) did not significantly predict the number of alleles per colony. Not surprisingly, there was a significant effect of read counts on the number of symbiont alleles per coral colony ( $F_{1,323} = 13.32$ ,  $p < 0.001$ ) and an increase in alleles per locus with increasing read counts ( $\beta = 0.058$ ,  $p < 0.001$ ). In summary, after accounting for variation in read counts, symbiont plastid allele diversity exhibited clinal variation with latitude, and lower diversity in inshore and midshelf reefs.

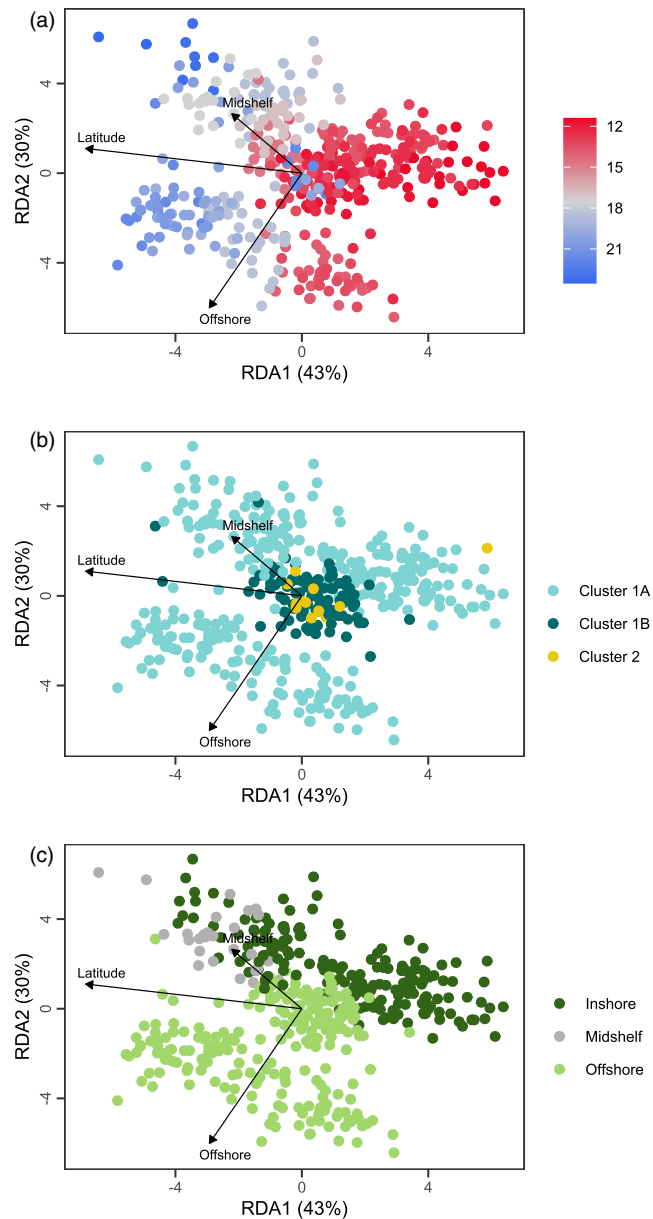
For the partial RDA examining how symbiont plastid allele composition was structured (Figure 6), we observed a significant effect of latitude ( $F_{1,216} = 3.83$ ,  $p = 0.016$ ), which explained 1.4% of variation in symbiont allele composition. The effect of shelf position was also significant ( $F_{2,216} = 24.92$ ,  $p = 0.001$ ) and explained 18.3% of the variation in symbiont allele composition. There was a nonsignificant effect of the host genetic cluster ( $F_{2,216} = 1.68$ ,  $p = 0.094$ ), which explained 1.2% of the variation in symbiont allele composition. The effect of latitude was almost completely perpendicular to the effect of the offshore factor, indicating that the effects of latitude and offshore position had largely uncorrelated effects on symbiont allele composition on the first two RDA dimensions. Only the first dimension (RDA1) was significant ( $F_{1,216} = 49.21$ ,  $p = 0.001$ ), accounting for 86% of the explained variation in symbiont allelic composition. Read counts accounted for 6% of variation in symbiont allele composition.

## 4 | DISCUSSION

Here, we present a uniquely rich data set for a common coral species, *A. tenuis*, considering both spatial genomic diversity for the cnidarian host and associated symbiont communities across the entire GBR. We found that the cnidarian host comprised at least three well differentiated but sympatric taxa, with strong evidence for gene flow among them. Our data are consistent with growing evidence that many morphologically defined coral species comprise multiple cryptic co-occurring taxa that are connected over evolutionary time by low-to-moderate levels of ongoing gene flow. Despite this differentiation in the cnidarian host, all identified *A. tenuis* taxa shared a common symbiont pool, dominated by the genus, *Cladocopium*. Rather than segregated by cnidarian host taxa, symbiont allelic variation was largely structured by shelf position, supporting the idea that broadcast spawning corals possess locally acquired and environmentally filtered endosymbiont communities. Habitat matching of coral symbionts might therefore provide an important mechanism for rapid adaptive change in corals.

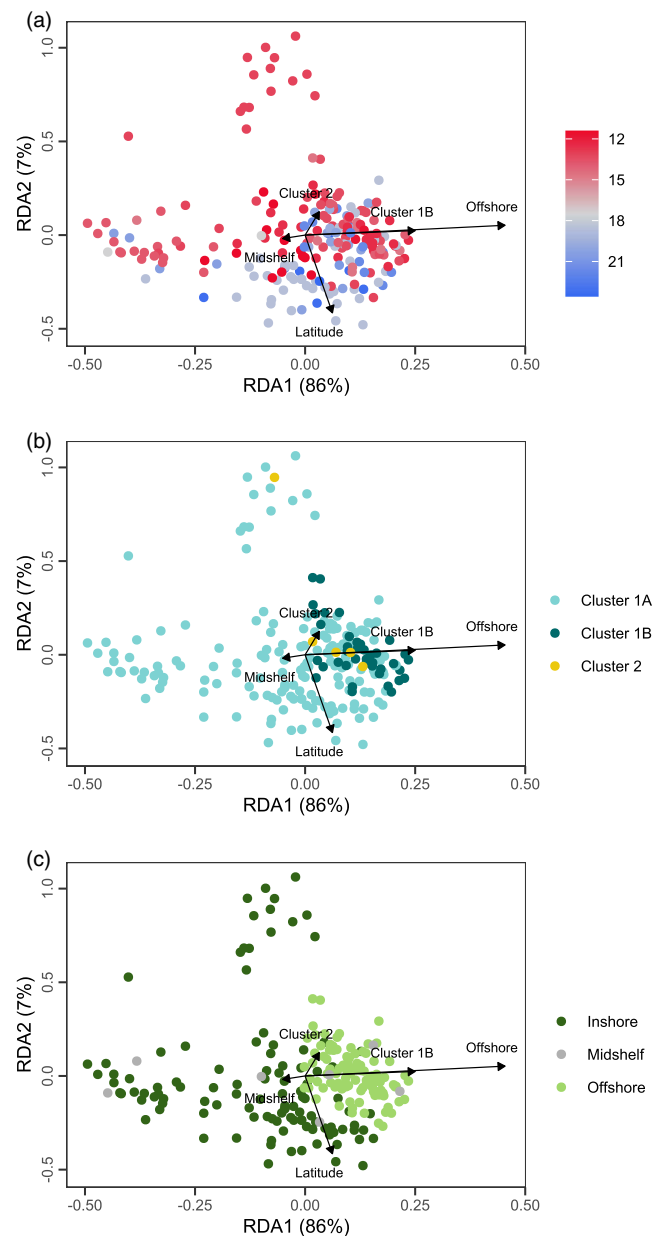
### 4.1 | Cryptic diversity within the *A. tenuis* host

*Acropora tenuis* is common on the GBR and throughout the Indo-Pacific, and among the first corals to be characterized in population genomic investigations (e.g. present study; Cooke et al., 2020; Zayasu et al., 2021), complementing earlier studies based on microsatellites (Lukoschek et al., 2016; Riginos et al., 2019; Underwood, 2009). These new genomic studies all highlight cryptic divisions, largely undetected in studies with fewer loci. Among our samples, we found three distinct genetic groupings with overlapping ranges (Figures 1 and 2), and divergence times exceeding 0.5 MY (Figure 4). Hints of the three clusters were also evidenced in the coral-based *ITS1* phylogeny, where Cluster 2 individuals formed a single monophyletic clade.



**FIGURE 5** Redundancy analysis fitting coral host individual allele frequency estimates as a function of latitude and shore position. Each point represents a coral colony, coloured by (a) latitude, (b) coral host genetic cluster; and (c) shore position; see respective legends for colour keys. Arrows illustrate the relative loadings for eigenvectors corresponding to each RDA dimension. Note, both RDA1 and RDA2 significantly explained variation in the individual allele frequencies of coral hosts. Percentages in parentheses are the amount of explainable variation (0.77% of total variation) captured by each RDA dimension

The *A. tenuis* clusters identified here complement and improve the resolution of previous differentiation based on 10 microsatellite markers (Lukoschek et al., 2016). The prior study identified genetic subdivision between southern offshore and inshore reefs and found two major groupings: (i) a widespread northern and central GBR cluster found predominantly north of 19°S and (ii) a southern offshore cluster. However, some northern reefs (e.g. No Name Reef



**FIGURE 6** Redundancy analysis fitting symbiont allele composition as a function of latitude, coral host genetic cluster, and shore position. Each point represents a coral colony, coloured by (a) latitude, (b) coral host genetic cluster; and (c) shore position; see respective legends for colour keys. Arrows illustrate the relative loadings for eigenvectors corresponding to each RDA dimension, although only RDA1 significantly explained variation in symbiont allele composition. Percentages in parentheses are the amount of explainable variation (28.38% of total variation) captured by each RDA dimension

and Lizard Island) had samples belonging to the southern cluster. This southern-offshore versus north/central division aligns with the separation between our Clusters 1A and 1B that includes samples from No Name Reef and one individual from Lizard Island (Figure 1). Microsatellite variation in these prior studies indicated higher genetic diversity on southern reefs compared with northern reefs, likely representing the sympatric but genetically divergent clusters

identified in the present study. This diversity differential provided much of the signal of north to south gene flow (Riginos et al., 2019) and will need to be re-evaluated by restricting future gene flow estimates to same-cluster individuals. With the present low sample numbers for Cluster 1B and Cluster 2, it is probably premature to make firm conclusions regarding their spatial distributions along the GBR.

By coanalysing genome-wide variation from our study and from Cooke et al. (2020), we can also corroborate and extend these earlier findings for *A. tenuis*. While Cooke et al. (2020) found a distinct taxon restricted to the inshore Magnetic Island on the central GBR, our new results extend the known range for this taxon with a single sample from Great Keppel. Thus, based on the geographically extensive sample pool across both studies, there appear to be at least three genetically distinct taxa (Figures 1 and 2, and maybe four taxa: Figure 3) within what is morphologically defined as *A. tenuis*.

Although field identification of corals is challenging and can lead to cases of misidentification, our analysis of UCEs and *ITS1* sequences confirms that Clusters 1A, 1B and 2 form a monophyletic group that is synonymous with previously reported *A. tenuis* samples (Figure S4). Population genomic data and analyses are well-suited for uncovering clusters of individuals that may have only recently stopped exchanging genes or that continue to be linked by low levels of gene flow. For instance, assignment tests (e.g. *NGSadmix*) leverage genotype compositions (assuming Hardy–Weinberg proportions and linkage equilibria within groups: Pritchard et al., 2000) to identify clusters of individuals that match expectations for unrestricted random mating. Additionally, the demographic modelling approach in *dadi* employs SFS to estimate divergence times while accounting for gene flow (Gutenkunst et al., 2009). Thus, population genomic approaches can recognize recently diverged taxa or locally adapted populations that may represent ecologically relevant diversity. In other words, the fine taxonomic resolution from population genomics (Figures 1, 2 and 4) may point to further subdivisions in *A. tenuis* than current phylogenies suggest.

## 4.2 | Divergence with ongoing gene flow among genetic clusters

Hybridization and interspecific gene flow may be integral aspects of coral evolution (Hobbs et al., 2021; Willis et al., 2006). Yet, few studies have formally tested for gene flow between recent discrete taxa using demographic inference until recently (exceptions include: Cooke et al., 2020; Fifer et al., 2021; Hellberg et al., 2016; Ladner & Palumbi, 2012; Prada & Hellberg, 2020; Prata et al., 2022; Rippe et al., 2021). These historical analyses provide information on the speciation process and timing, including whether divergence occurred in the absence of gene flow (i.e. allopatric speciation) or with ongoing genetic exchange (ecological or sympatric speciation).

Our reconstructions of demographic history for *A. tenuis* show that cryptic taxa diverged prior to initial formation of the GBR and that genomic differences have been maintained despite ongoing gene flow. We found splitting times between 0.46 and 1.5 MY among all genetic clusters, despite higher genome-wide differentiation in Cluster 2 (Figures 2 and 4). If the distinct genetic clusters do differ in distributions or habitat niches, then it is likely that modern GBR distributions reflect taxon filtering (Sommer et al., 2014) rather than adaptation in situ given these divergence times. In all cases, models of strict divergence without gene flow were rejected (Figure S9), whereas comparisons of sympatric and allopatric (non-overlapping) populations provided replicated support for divergence with gene flow between coral taxa (Figure S10). Under the consistently highly supported isolation-migration (IM) models, gene flow estimates were greatest between Clusters 1A and 1B (Figure 4).

Overall, our results corroborate previous estimates for late Pleistocene divergences (0.27–0.6 MY) reported between *A. tenuis* cryptic taxa (Cooke et al., 2020), albeit with older estimates in the present study. The GBR is a geologically young structure that formed in the last 0.5 MY (Chadwick-Furman, 1996; Pandolfi & Kelley, 2008; Webster & Davies, 2003). Cyclical sea level changes during the late Pleistocene and the resulting redistribution of species ranges (Hewitt, 2000), however, likely promoted repeated SCs and periodic gene flow between coral populations. Periodic contact with pulses of historical admixture between diverging taxa may be a common evolutionary history for high dispersal marine species (e.g. Duranton et al., 2018) and could explain why we could not clearly discriminate between SC and IM models in *A. tenuis* (Figure S9). In all of our comparisons, between-cluster migration estimates most likely reflect introgression between taxa that have some level of partial reproductive isolation. Thus, the strength or weakness of gene flow is not an accurate predictor of dispersal but likely reflects the degree of genomic permeability between taxa at different stages of the speciation process (Bierne et al., 2013; Harrison & Larson, 2016). Few fixed differences and regions of elevated differentiation in Cluster 2 comparisons ( $F_{ST} > 0.5$ ; Figure S6) may point to parts of the genome possibly associated with reduced migration. Heterogenous divergence across genomes, however, between taxa may result from a wide range of underlying molecular processes (Cruickshank & Hahn, 2014; Ravinet et al., 2017). Determining whether regions of elevated differentiation are related to differential introgression, the origins of those alleles (e.g. Rose et al., 2021), and whether they help maintain local adaptation in high dispersal *Acropora* corals will be outstanding questions for understanding the processes shaping coral diversity.

## 4.3 | Geographic gradients and associated genetic diversity of hosts and symbionts

In *A. tenuis*, we found that both cnidarian host genotypes and symbiont allelic variation were structured with respect to latitude and shelf position. For the cnidarian host, latitude and shelf position

together explained a very small but significant proportion of variation (0.77%) in individual allele frequencies. This signal was largely driven by Cluster 1A for which we had the most widely distributed samples (Figure 5). Broader sampling of Clusters 1B and 2 might reveal similar patterns, but it is also possible that these clusters are subject to greater environmental forcing than Cluster 1A, leading to restricted distributions (Sommer et al., 2014).

In contrast, genetic differentiation by latitude and shelf position was substantial for symbionts, explaining 19.7% of variation in compositional turnover (Figure 6). *ITS2* sequences showed that all three cnidarian host clusters were dominated by symbionts from the genus *Cladocopium* (100%). All coral colonies contained the C40c strain, and <1% contained C72k and C72d. Yet, by recovering plastid sequences from colony-level shotgun sequencing, we could further characterize allelic variation and test the effects of the environment on symbiont structure. Our novel method of describing symbiont diversity based on plastid derived alleles appears to yield more information than strain identity based on *ITS2* sequences. Because genomic sequencing is typically obtained from samples of the coral holobiont, our approach demonstrates the great potential of using cocaptured symbiont genome sequences to study variation variations in these endosymbiotic bacteria alongside their cnidarian hosts.

Allelic diversity of *Cladocopium* plastids varied by shelf position, and to a lesser extent by latitude, both in terms of alpha (number of alleles within colonies) and beta (change in allelic composition between colonies) diversity. These distinct aspects of spatial diversity were evidenced in the GLM results that focussed on spatial patterns of within-colony diversity. Colonies from northern and further offshore sites had more plastid alleles per locus, compared with colonies from southern, inshore and midshelf locations. Our RDA of symbiont allelic composition showed substantial turnover along latitudinal and shelf gradients (Figure 6). Symbiont plastid diversity or differentiation showed no relationship to host genetic cluster in either GLM or RDA analyses. While coral-symbiont associations may show strong evolutionary specificity (Prada et al., 2014), our observations support the notion that the specificity and maintenance of coral-*Symbiodinium* symbiosis can also be influenced by environmental gradients (Tonk et al., 2013) more so than related host taxa.

Our observations that *A. tenuis* symbionts show strong environmental filtering and are not associated with host structure corroborate findings from Cooke et al. (2020). In their study, *Cladocopium* also was the dominant genus, and *Cladocopium* mitochondrial haplotypes were associated with exposure to riverine plumes, whereas cnidarian host variation was not correlated with freshwater inputs. Those colonies most heavily exposed to riverine plumes shared a single *Cladocopium* mitochondrial haplotype, consistent with our finding of lower symbiont diversity for inshore *A. tenuis* colonies. Lower diversity has also been observed among inshore soft corals on the GBR (Howells et al., 2009). It is possible that the inshore-offshore habitat gradient represents a ubiquitous environmental filter for symbionts across coral species. Additionally, our data from *A. tenuis* suggest that whilst conditions faced by inshore colonies

decreases local symbiont diversity (i.e. reduced alpha diversity), greater environmental variance might promote variation among inshore locations (i.e. increased beta diversity).

Our study focussed on symbiont diversity in *A. tenuis* at the scale of the GBR. However, symbiont strains can exhibit significant turnover across very short physical distances, for example, between leeward and windward sides of coral islands (Howells et al., 2013), and symbiont composition often exhibits significant turnover with respect to depth (Bongaerts et al., 2010; Prada et al., 2014; van Oppen et al., 2018; Warner et al., 2015). Local variation in symbiont composition and diversity has been also linked to environmental adaptation (such as, Rose et al., 2021). Therefore, joint considerations of both coral host and symbiont spatial genetic patterns can help identify the relevant environmental factors and spatial scales of differentiation that are likely to reflect adaptation to local conditions. Understanding the congruence or mismatch in spatial genetic patterns between coral host and symbionts is particularly important today as they provide insights into how adaptive variation is spatially configured, thereby providing information for spatial planning (Riginos & Beger, 2022), restoration (Baums et al., 2019) and assisted management (van Oppen et al., 2017) in light of climate change and other anthropogenic stressors.

#### 4.4 | Challenges in combining genomic data sets

In principle, combining genomic data sets from different studies offers incredible potential to increase spatial replication or investigate temporal shifts in population genetic patterns. Given that taxonomic relationships among corals are sometimes contentious and field identification to the species level can be challenging, it is especially important (in our opinion) that investigators try to build explicit links between new and previously published DNA sequence-based data sets. Throughout this investigation, we have endeavoured to place our new results in the context of previous work on *A. tenuis*, including downsampling our data to allow phylogenetic analyses (Figure S4), and connecting to previous microsatellite results (Figure S5).

Genomic data sets from different studies, however, can thwart amalgamation. Here, we were particularly interested in merging our *A. tenuis* samples with those from a low-coverage whole-genome study by Cooke et al. (2020). Coanalysing these data required a balance between aggressive filtering versus retaining sufficient loci to extract subtle patterns. Despite our best efforts, weak study (batch) effects emerged, differentiating samples from this study and those from Cooke et al. (2020) (see slight separation between individuals from both studies: Figure 4). Batch effects are common in population genomic investigations (Tom et al., 2017) and might be particularly problematic in comparisons between studies using low-coverage genomic sequencing (Lou & Therikildsen, 2022). This challenge in unifying results across studies is worrisome as the value of individual studies is diminished if they cannot be contextualized against earlier results. Fine-tuning quality metrics and masking genomic regions subject to batch effects or resequencing previously published



samples alongside new data sets might provide solutions for identifying and resolving batch effects (Tom et al., 2017). As more studies move to whole-genome sequencing, methods may also evolve to resolve this issue. We are confident that the overall filtering of our *A. tenuis* samples is sufficient to identify the major genetic clusters in this study and those shared with Cooke et al. (2020). Nonetheless, we did not pursue additional analyses on a merged data set because of suspected batch effects.

#### 4.5 | Cryptic taxonomic diversity and implications for coral conservation

Delineating the natural groupings of frequently interbreeding individuals provides critical information for conservation and management (Hey et al., 2003). Although species delimitations are seemingly easier to understand, they may belie the true scope of ecologically and evolutionarily relevant diversity. Our findings add to a growing number of studies that highlight how current taxonomic hypotheses (Cowman et al., 2020) underestimate diversity in reef-building corals, especially within the diverse Acroporid clade. These results indicate that inferences based on species distributions—such as range size, thermal tolerance, habitat usage and other niche attributes—are questionable without first documenting the range extents and niches of the distinct cryptic taxa. For instance, commonly used conservation planning schemes based on optimizing species representation necessitate knowing which species live where (reviewed by, Riginos & Beger, 2022). Without better documentation on the distributions of cryptic coral species, we lack the information to accurately map species ranges and therefore to formulate conservation plans seeking to preserve coral biodiversity.

Conservation planning for future reefs is likely to prioritize coral species that can withstand high temperatures (Mumby et al., 2010). But cryptic species can differ in thermal physiology. Extensive studies of cryptic and sympatric *A. hyacinthus*, for example, highlight extensive differences in physiological responses (reviewed by Thomas et al., 2018) and identify specific genomic regions that may contribute to the cryptic taxa's distinctive bleaching susceptibility (Rose et al., 2021). Similar results are suggested for *Pocillopora* (Burgess et al., 2021) and *Porites* (Boulay et al., 2014), where cryptic taxa differ in survival through bleaching events. Again, uncertainty regarding cohesive genetic species can lead to erroneous biological inferences (Gomez-Corrales & Prada, 2020) and therefore potentially inappropriate conservation recommendations.

Genomic tools are starting to provide windows into the complex ecological and evolutionary dynamics of coral holobionts in relation to geography and environment. The common finding that low levels of gene flow connect divergent taxa, however, may provide opportunities for adaptation and could contribute to the future well-being of coral reefs. Several examples in diverse taxa point to intraspecific gene flow as an important aspect of local adaptation (Huerta-Sánchez et al., 2015; Oziolor et al., 2019), and the geographically widespread nature of gene exchange (Figure 4) in *A. tenuis* suggests

that universally adaptive alleles should cross taxon boundaries. In *A. tenuis*, zooxanthellate endosymbionts appear to easily traverse host taxa, providing yet another distinct aspect of adaptive inter-change, likely allowing coral populations to respond to environmental stressors on ecological timescales. These emerging findings may guide future management strategies that aim to preserve ecological and evolutionary functions including adaptation capacity.

#### ACKNOWLEDGEMENTS

This work was supported by the Great Barrier Reef Foundation's Resilient Reefs Successfully Adapting to Climate Change Program, funded principally by the Australian Government. During the final writing and analysis phase, Popovic was supported by the Reef Restoration and Adaptation Program, funded by the partnership between the Australian Government's Reef Trust and the Great Barrier Reef Foundation, to Riginos. We thank CX Cheng and K Dougan for sharing their draft version of the *C. goreau* plastid genome.

#### CONFLICT OF INTEREST

The authors have no conflict of interest to declare.

#### DATA AVAILABILITY STATEMENT

Raw data can be found on NCBI's Short Read Archive, submission SUB11600324, where geographical metadata are deposited on GEOME (<http://n2t.net/ark:/21547/EFV2>). VCF files and supporting scripts are on Data Dryad (48096232EVA-2022-007-OA.R1).

#### DATA ACCESSIBILITY STATEMENT

Genetic data: Raw sequence reads are deposited in the SRA (BioProject PRJNA849642) and individual genotype data are available on Data Dryad (48096232EVA-2022-007-OA.R1). Sample metadata can be found in GEOME (<http://n2t.net/ark:/21547/EFV2>; including georeferences in decimal degrees and year of sampling event).

#### BENEFIT-SHARING STATEMENT

This research collaboration involves collaborators from Australian institutions where all collaborators are included as co-authors. The results of research have been shared with the provider communities and the broader scientific community (see above), and the research addresses a priority concern, in this case the conservation of corals. Benefits from this research accrue from the sharing of our data and results on public databases as described above.

#### ORCID

Ambrocio Melvin A. Matias  <https://orcid.org/0000-0002-1329-4285>

Iva Popovic  <https://orcid.org/0000-0001-6582-4236>

Joshua A. Thia  <https://orcid.org/0000-0001-9084-0959>

Ira R. Cooke  <https://orcid.org/0000-0001-6520-1397>

Gergely Torda  <https://orcid.org/0000-0002-4585-3045>

Vimoksalehi Lukoschek  <https://orcid.org/0000-0002-0268-3808>

Line K. Bay  <https://orcid.org/0000-0002-9760-2977>

Sun W. Kim  <https://orcid.org/0000-0002-2429-3245>

Cynthia Riginos  <https://orcid.org/0000-0002-5485-4197>

## REFERENCES

- Andrew, R. L., & Rieseberg, L. H. (2013). Divergence is focused on few genomic regions early in speciation: Incipient speciation of sunflower ecotypes. *Evolution*, 67(9), 2468–2482. <https://doi.org/10.1111/evo.12106>
- Baird, A. H., Guest, J. R., & Willis, B. L. (2009). Systematic and biogeographical patterns in the reproductive biology of scleractinian corals. *Annual Review of Ecology, Evolution, and Systematics*, 40(1), 551–571. <https://doi.org/10.1146/annurev.ecolsys.110308.120220>
- Bates, D., Mächler, M., Bolker, B., & Walker, S. (2014). Fitting linear mixed-effects models using lme4. *arXiv.org*, 1303.3750.
- Baums, I. B., Baker, A. C., Davies, S. W., Grottoli, A. G., Kenkel, C. D., Kitchen, S. A., Kuffner, I. B., Lajeunesse, T. C., Matz, M. V., Miller, M. W., Parkinson, J. E., & Shanz, A. A. (2019). Considerations for maximizing the adaptive potential of restored coral populations in the western Atlantic. *Ecological Applications*, 29(8), e01978. <https://doi.org/10.1002/eap.1978>
- Bierne, N., Gagnaire, P.-A., & David, P. (2013). The geography of introgression in a patchy environment and the thorn in the side of ecological speciation. *Current Zoology*, 59(1), 72–86. <https://doi.org/10.1093/czoolo/59.1.72>
- Bodenhofer, U., Bonatesta, E., Horejš-Kainrath, C., & Hochreiter, S. (2015). msa: An R package for multiple sequence alignment. *Bioinformatics*, 31(24), 3997–3999.
- Bongaerts, P., Carmichael, M., Hay, K. B., Tonk, L., Frade, P. R., & Hoegh-Guldberg, O. (2015). Prevalent endosymbiont zonation shapes the depth distributions of scleractinian coral species. *Royal Society Open Science*, 2(2), 140297. <https://doi.org/10.1098/rsos.140297>
- Bongaerts, P., Cooke, I. R., Ying, H., Wels, D., denHaan, S., Hernandez-Agreda, A., Brunner, C. A., Dove, S., Englebert, N., Eyal, G., Forêt, S., Grinblat, M., Hay, K. B., Harii, S., Hayward, D. C., Lin, Y., Mihaljević, M., Moya, A., Muir, P., ... Hoegh-Guldberg, O. (2021). Morphological stasis masks ecologically divergent coral species on tropical reefs. *Current Biology*, 31(11), 2286–2298.e8. <https://doi.org/10.1016/j.cub.2021.03.028>
- Bongaerts, P., Frade, P. R., Ogier, J. J., Hay, K. B., vanBleijswijk, J., Englebert, N., Vermeij, M. J., Bak, R. P., Visser, P. M., & Hoegh-Guldberg, O. (2013). Sharing the slope: Depth partitioning of agariciid corals and associated *Symbiodinium* across shallow and mesophotic habitats (2–60 m) on a Caribbean reef. *BMC Evolutionary Biology*, 13, 205. <https://doi.org/10.1186/1471-2148-13-205>
- Bongaerts, P., Riginos, C., Brunner, R., Englebert, N., Smith, S. R., & Hoegh-Guldberg, O. (2017). Deep reefs are not universal refuges: Reseeding potential varies among coral species. *Science Advances*, 3(2), e1602373. <https://doi.org/10.1126/sciadv.1602373>
- Bongaerts, P., Riginos, C., Ridgway, T., Sampayo, E. M., vanOppen, M. J. H., Englebert, N., Vermeulen, F., & Hoegh-Guldberg, O. (2010). Genetic divergence across habitats in the widespread coral *Seriatopora hystrix* and its associated *Symbiodinium*. *PLoS One*, 5(5), e10871. <https://doi.org/10.1371/journal.pone.0010871>
- Boulay, J. N., Hellberg, M. E., Cortes, J., & Baums, I. B. (2014). Unrecognized coral species diversity masks differences in functional ecology. *Proceedings of the Royal Society B*, 281(1776), 20131580. <https://doi.org/10.1098/rspb.2013.1580>
- Burgess, S. C., Johnston, E. C., Wyatt, A. S., Leichter, J. J., & Edmunds, P. J. (2021). Response diversity in corals: Hidden differences in bleaching mortality among cryptic *Pocillopora* species. *Ecology*, 102(6), e03324.
- Bushnell, B. (2014). BMAP: A fast, accurate, splice-aware aligner. Report Number: LBNL-7065E. Lawrence Berkeley National Laboratory, Berkeley, CA. <https://sourceforge.net/projects/bmap/>
- Camp, E. F., Suggett, D. J., Pogoreutz, C., Nitschke, M. R., Houlbrequé, F., Hume, B. C. C., Gardner, S. G., Zampighi, M., Rodolfo-Metalpa, R., & Voolstra, C. R. (2020). Corals exhibit distinct patterns of microbial re-organisation to thrive in an extreme inshore environment. *Coral Reefs*, 39(3), 701–716. <https://doi.org/10.1007/s00338-019-01889-3>
- Chadwick-Furman, N. (1996). Reef coral diversity and global change. *Global Change Biology*, 2, 559–568.
- Cooke, I., Ying, H., Forêt, S., Bongaerts, P., Strugnell, J., Simakov, O., Zhang, J., Field, M. A., Rodriguez-Lanetty, M., Bell, S. C., Bourne, D. G., van Oppen, M. J. H., Ragan, M. A., & Miller, D. J. (2020). Signatures of selection in the coral holobiont reveal complex adaptations to inshore environments driven by Holocene climate change. *Science Advances*, 6, eabc6318. <https://doi.org/10.1101/2020.02.25.951905>
- Cooper, T. F., Berkelmans, R., Ulstrup, K. E., Weeks, S., Radford, B., Jones, A. M., Doyle, J., Canto, M., O'Leary, R. A., & vanOppen, M. (2011). Environmental factors controlling the distribution of *Symbiodinium* harboured by the coral *Acropora millepora* on the Great Barrier Reef. *PLoS One*, 6(10), e25536. <https://doi.org/10.1371/journal.pone.0025536>
- Cowman, P. F., Quattrini, A. M., Bridge, T. C. L., Watkins-Colwell, G. J., Fadli, N., Grinblat, M., Roberts, T. E., McFadden, C., Miller, D. J., & Baird, A. H. (2020). An enhanced target-enrichment bait set for *Hexacorallia* provides phylogenomic resolution of the staghorn corals (Acroporidae) and close relatives. *Molecular Phylogenetics and Evolution*, 153, 106944. <https://doi.org/10.1016/j.ympev.2020.106944>
- Coyne, J. A., & Orr, H. A. (2004). *Speciation* (Vol. 37). Sinauer Associates.
- Cruikshank, T. E., & Hahn, M. W. (2014). Reanalysis suggests that genomic islands of speciation are due to reduced diversity, not reduced gene flow. *Molecular Ecology*, 23(13), 3133–3157. <https://doi.org/10.1111/mec.12796>
- Dana, J. D. (1846–1849). Zoophytes. United States Exploring Expedition during the years 1838–1842. Lea and Blanchard, Philadelphia. In (Vol. 7, pp. 1–740, 761 pls. [1846: 1841–1120, 1709–1720; 1848: 1121–1708, 1721–1740; 1849: atlas pls. 1841–1861]).
- Danecek, P., Bonfield, J. K., Liddle, J., Marshall, J., Ohan, V., Pollard, M. O., Whitwham, A., Keane, T., McCarthy, S. A., Davies, R. M., & Li, H. (2021). Twelve years of SAMtools and BCftools. *GigaScience*, 10(2), giab008.
- Duranton, M., Allal, F., Fraïsse, C., Bierne, N., Bonhomme, F., & Gagnaire, P.-A. (2018). The origin and remodeling of genomic islands of differentiation in the European sea bass. *Nature Communications*, 9(1), 1–11. <https://doi.org/10.1038/s41467-018-04963-6>
- Fifer, J. E., Yasuda, N., Yamakita, T., Bove, C. B., & Davies, S. W. (2021). Genetic divergence and range expansion in a western North Pacific coral. *Science of the Total Environment*, 813, 152423. <https://doi.org/10.1016/j.scitotenv.2021.152423>
- Fisher, R., O'Leary, R. A., Low-Choy, S., Mengersen, K., Knowlton, N., Brainard, R. E., & Caley, M. J. (2015). Species richness on coral reefs and the pursuit of convergent global estimates. *Current Biology*, 25(4), 500–505. <https://doi.org/10.1016/j.cub.2014.12.022>
- Fukami, H., Chen, C. A., Budd, A. F., Collins, A., Wallace, C., Chuang, Y.-Y., Chen, C., Dai, C. F., Iwao, K., Sheppard, C., & Knowlton, N. (2008). Mitochondrial and nuclear genes suggest that stony corals are monophyletic but most families of stony corals are not (Order Scleractinia, Class Anthozoa, Phylum Cnidaria). *PLoS One*, 3(9), e3222. <https://doi.org/10.1371/journal.pone.0003222>
- Galtier, N. (2019). Delineating species in the speciation continuum: A proposal. *Evolutionary Applications*, 12(4), 657–663. <https://doi.org/10.1111/eva.12748>
- Gomez-Corrales, M., & Prada, C. (2020). Cryptic lineages respond differently to coral bleaching. *Molecular Ecology*, 29, 4265–4273. <https://doi.org/10.1111/mec.15631>
- Gutenkunst, R. N., Hernandez, R. D., Williamson, S. H., & Bustamante, C. D. (2009). Inferring the joint demographic history of multiple

- populations from multidimensional SNP frequency data. *PLoS Genetics*, 5(10), e1000695. <https://doi.org/10.1371/journal.pgen.1000695>
- Harrison, R. G., & Larson, E. L. (2016). Heterogeneous genome divergence, differential introgression, and the origin and structure of hybrid zones. *Molecular Ecology*, 25(11), 2454–2466. <https://doi.org/10.1111/mec.13582>
- Hellberg, M. E., Prada, C., Tan, M. H., Forsman, Z. H., & Baums, I. B. (2016). Getting a grip at the edge: Recolonization and introgression in eastern Pacific *Porites* corals. *Journal of Biogeography*, 43(11), 2147–2159. <https://doi.org/10.1111/jbi.12792>
- Hewitt, G. (2000). The genetic legacy of the Quaternary ice ages. *Nature*, 405(6789), 907–913.
- Hey, J., & Pinho, C. (2012). Population genetics and objectivity in species diagnosis. *Evolution*, 66(5), 1413–1429. <https://doi.org/10.1111/j.1558-5646.2011.01542.x>
- Hey, J., Waples, R. S., Arnold, M. L., Butlin, R. K., & Harrison, R. G. (2003). Understanding and confronting species uncertainty in biology and conservation. *Trends in Ecology and Evolution*, 18(11), 597–603. <https://doi.org/10.1016/j.tree.2003.08.014>
- Hoadley, K. D., Pettay, D. T., Lewis, A., Wham, D., Grasso, C., Smith, R., Kemp, D. W., LaJeunesse, T., & Warner, M. E. (2021). Different functional traits among closely related algal symbionts dictate stress endurance for vital Indo-Pacific reef-building corals. *Global Change Biology*, 27(20), 5295–5309.
- Hobbs, J.-P. A., Richards, Z. T., Popovic, I., Lei, C., Staedle, T. M., Montanari, S. R., & DiBattista, J. D. (2022). Hybridisation and the evolution of coral reef biodiversity. *Coral Reefs*, 41, 535–549. <https://doi.org/10.1007/s00338-021-02193-9>
- Howells, E. J., vanOppen, M. J. H., & Willis, B. L. (2009). High genetic differentiation and cross-shelf patterns of genetic diversity among Great Barrier Reef populations of *Symbiodinium*. *Coral Reefs*, 28(1), 215–225. <https://doi.org/10.1007/s00338-008-0450-z>
- Howells, E. J., Willis, B. L., Bay, L. K., & vanOppen, M. J. H. (2013). Spatial and temporal genetic structure of *Symbiodinium* populations within a common reef-building coral on the Great Barrier Reef. *Molecular Ecology*, 22(14), 3693–3708. <https://doi.org/10.1111/mec.12342>
- Huerta-Sánchez, E., Jin, X., Asan, Bianba, Z., Peter, B. M., Vinckenbosch, N., Liang, Y., Yi, X., He, M., Somel, M., Ni, P., Wang, B., Ou, X., Huasang, Luosang, J., Cuo, Z. X., Li, K., Gao, G., Yin, Y., ... Nielsen, R. (2014). Altitude adaptation in Tibetans caused by introgression of Denisovan-like DNA. *Nature*, 512(7513), 194–197. <https://doi.org/10.1038/nature13408>
- Hume, B. C. C., Smith, E. G., Ziegler, M., Warrington, H. J. M., Burt, J. A., LaJeunesse, T. C., Wiedenmann, J., & Voolstra, C. R. (2019). SymPortal: A novel analytical framework and platform for coral algal symbiont next-generation sequencing ITS2 profiling. *Molecular Ecology Resources*, 19(4), 1063–1080. <https://doi.org/10.1111/1755-0998.13004>
- Kenkel, C., Goodbody-Gringley, G., Caillaud, D., Davies, S., Bartels, E., & Matz, M. (2013). Evidence for a host role in thermotolerance divergence between populations of the mustard hill coral (*Porites astreoides*) from different reef environments. *Molecular Ecology*, 22(16), 4335–4348.
- Kim, S. Y., Lohmueller, K. E., Albrechtsen, A., Li, Y., Korneliusson, T., Tian, G., Grarup, N., Jiang, T., Andersen, G., Witte, D., Jorgensen, T., Hansen, T., Pedersen, O., Wang, J. & Nielsen, R. (2011). Estimation of allele frequency and association mapping using next-generation sequencing data. *BMC Bioinformatics*, 12(1), 1–16.
- Kitahara, M. V., Lin, M.-F., Forêt, S., Huttley, G., Miller, D. J., & Chen, C. A. (2014). The “naked coral” hypothesis revisited – Evidence for and against scleractinian monophyly. *PLoS One*, 9, e94774.
- Korneliusson, T. S., Albrechtsen, A., & Nielsen, R. (2014). ANGSD: Analysis of next generation sequencing data. *BMC Bioinformatics*, 15(1), 443. <https://doi.org/10.1186/s12859-014-0356-4>
- Korunes, K. L., & Samuk, K. (2021). pixy: Unbiased estimation of nucleotide diversity and divergence in the presence of missing data. *Molecular Ecology Resources*, 21(4), 1359–1368. <https://doi.org/10.1111/1755-0998.13326>
- Ladner, J. T., & Palumbi, S. R. (2012). Extensive sympatry, cryptic diversity and introgression throughout the geographic distribution of two coral species complexes. *Molecular Ecology*, 21(9), 2224–2238. <https://doi.org/10.1111/j.1365-294X.2012.05528.x>
- LaJeunesse, T. C., Parkinson, J. E., Gabrielson, P. W., Jeong, H. J., Reimer, J. D., Voolstra, C. R., & Santos, S. R. (2018). Systematic revision of Symbiodiniaceae highlights the antiquity and diversity of coral endosymbionts. *Current Biology*, 28(16), 2570–2580. e6. <https://doi.org/10.1016/j.cub.2018.07.008>
- Li, H., & Durbin, R. (2009). Fast and accurate short read alignment with Burrows-Wheeler transform. *Bioinformatics*, 25(14), 1754–1760.
- Li, H., Handsaker, B., Wysoker, A., Fennell, T., Ruan, J., Homer, N., Marth, G., Abecasis, G., Durbin, R., & 1000 Genome Project Data Processing Subgroup. (2009). The sequence alignment/map format and SAMtools. *Bioinformatics*, 25(16), 2078–2079.
- Liu, H., Stephens, T. G., González-Pech, R. A., Beltran, V. H., Lapeyre, B., Bongaerts, P., Cooke, I., Aranda, M., Bourne, D. G., Forêt, S., Miller, D. J., vanOppen, M., Voolstra, C. R., Ragan, M. A., & Chan, C. X. (2018). Symbiodinium genomes reveal adaptive evolution of functions related to coral-dinoflagellate symbiosis. *Communications Biology*, 1, 95. <https://doi.org/10.1038/s42003-018-0098-3>
- Lou, R. N., & Therikildsen, N. O. (2022). Batch effects in population genomic studies with low-coverage whole genome sequencing data: Causes, detection, and mitigation. *Molecular Ecology Resources*, 22, 1678–1692. <https://doi.org/10.1111/1755-0998.13559>
- Lukoschek, V., Riginos, C., & vanOppen, M. J. H. (2016). Congruent patterns of connectivity can inform management for broadcast spawning corals on the Great Barrier Reef. *Molecular Ecology*, 25(13), 3065–3080. <https://doi.org/10.1111/mec.13649>
- Madin, J. S., Hoogenboom, M. O., Connolly, S. R., Darling, E. S., Falster, D. S., Huang, D., Keith, S. A., Mizerek, T., Pandolfi, J. M., Putnam, H. M., & Baird, A. H. (2016). A trait-based approach to advance coral reef science. *Trends in Ecology and Evolution*, 31(6), 419–428. <https://doi.org/10.1016/j.tree.2016.02.012>
- Mallet, J. (1995). A species definition for the Modern Synthesis. *Trends in Ecology and Evolution*, 10, 294–299.
- Manzello, D. P., Matz, M. V., Enochs, I. C., Valentino, L., Carlton, R. D., Kolodziej, G., Serrano, X., Towle, E. K., & Jankulak, M. (2019). Role of host genetics and heat-tolerant algal symbionts in sustaining populations of the endangered coral *Orbicella faveolata* in the Florida Keys with ocean warming. *Global Change Biology*, 25(3), 1016–1031. <https://doi.org/10.1111/gcb.14545>
- Mao, Y., Economo, E. P., & Satoh, N. (2018). The roles of introgression and climate change in the rise to dominance of *Acropora* corals. *Current Biology*, 28(21), 3373–3382.e3375. <https://doi.org/10.1016/j.cub.2018.08.061>
- Martin, S. H., Dasmahapatra, K. K., Nadeau, N. J., Salazar, C., Walters, J. R., Simpson, F., Blaxter, M., Manica, A., Mallet, J., & Jiggins, C. D. (2013). Genome-wide evidence for speciation with gene flow in *Heliconius* butterflies. *Genome Research*, 23(11), 1817–1828. <https://doi.org/10.1101/gr.159426.113>
- Matz, M. V. (2018). Fantastic beasts and how to sequence them: Ecological genomics for obscure model organisms. *Trends in Genetics*, 34(2), 121–132. <https://doi.org/10.1016/j.tig.2017.11.002>
- Meisner, J., & Albrechtsen, A. (2018). Inferring population structure and admixture proportions in low-depth NGS data. *Genetics*, 210(2), 719–731. <https://doi.org/10.1534/genetics.118.301336>
- Minh, B. Q., Schmidt, H. A., Chernomor, O., Schrempf, D., Woodhams, M. D., vonHaeseler, A., & Lanfear, R. (2020). IQ-TREE 2: New models and efficient methods for phylogenetic inference in the genomic era. *Molecular Biology and Evolution*, 37(5), 1530–1534. <https://doi.org/10.1093/molbev/msaa015>



- Morikawa, M. K., & Palumbi, S. R. (2019). Using naturally occurring climate resilient corals to construct bleaching-resistant nurseries. *Proceedings of the National Academy of Sciences of the United States of America*, 116(21), 10586–10591. <https://doi.org/10.1073/pnas.1721415116>
- Mumby, P. J., Elliott, I. A., Eakin, C. M., Skirving, W., Paris, C. B., Edwards, H. J., Enríquez, S., Iglesias-Prieto, R., Cherubin, L. M., & Stevens, J. R. (2011). Reserve design for uncertain responses of coral reefs to climate change. *Ecology Letters*, 14(2), 132–140. <https://doi.org/10.1111/j.1461-0248.2010.01562.x>
- Oksanen, J., Blanchet, F. G., Friendly, M., Kindt, R., Legendre, P., McGinn, D., Minchin, P. R., O'Hara, R. B., Simpson, G. L., Solymos, P., Stevens, M. H. H., Szocs, E., & Wagner, H. (2017). *vegan: Community ecology package*. R package version 2.4-2. <https://CRAN.R-project.org/package=vegan>
- Oziolor, E. M., Reid, N. M., Yair, S., Lee, K. M., VerPloeg, S. G., Bruns, P. C., Shaw, J. R., Whitehead, A., & Matson, C. W. (2019). Adaptive introgression enables evolutionary rescue from extreme environmental pollution. *Science*, 364(6439), 455–457. <https://doi.org/10.1126/science.aav4155>
- Pandolfi, J., & Kelley, R. (2008). *The Great Barrier Reef in time and space: Geology and palaeobiology* (Vol. 3). CSIRO Publishing.
- Paradis, E., Claude, J., & Strimmer, K. (2004). APE: Analyses of phylogenetics and evolution in R language. *Bioinformatics*, 20(2), 289–290.
- Portik, D. M., Leaché, A. D., Rivera, D., Barej, M. F., Burger, M., Hirschfeld, M., Rödel, M. O., Blackburn, D. C., & Fujita, M. K. (2017). Evaluating mechanisms of diversification in a Guineo-Congolian tropical forest frog using demographic model selection. *Molecular Ecology*, 26(19), 5245–5263. <https://doi.org/10.1111/mec.14266>
- Prada, C., & Hellberg, M. E. (2020). Speciation-by-depth on coral reefs: Sympatric divergence with gene flow or cryptic transient isolation? *Journal of Evolutionary Biology*, 34, 128–137. <https://doi.org/10.1111/jeb.13731>
- Prada, C., McIlroy, S. E., Beltrán, D. M., Valint, D. J., Ford, S. A., Hellberg, M. E., & Coffroth, M. A. (2014). Cryptic diversity hides host and habitat specialization in a gorgonian-algal symbiosis. *Molecular Ecology*, 23(13), 3330–3340. <https://doi.org/10.1111/mec.12808>
- Prata, K. E., Riginos, C., Gutenkunst, R. N., Latijnhouwers, K. R., Sánchez, J. A., Englebert, N., Hay, K. B., & Bongaerts, P. (2022). Deep connections: Divergence histories with gene flow in mesophotic *Agaricia* corals. *Molecular Ecology*, 9, 2511–2527.
- Pritchard, J. K., Stephens, M., & Donnelly, P. (2000). Inference of population structure using multilocus genotype data. *Genetics*, 155, 945–959.
- Quattrini, A. M., Faircloth, B. C., Dueñas, L. F., Bridge, T. C. L., Brugler, M. R., Calixto-Botía, I. F., DeLeo, D., Forêt, S., Herrera, S., Lee, S. M. Y., Miller, D. J., Prada, C., Rádis-Baptista, G., Ramírez-Portilla, C., Sánchez, J. A., Rodríguez, E., & McFadden, C. (2018). Universal target-enrichment baits for anthozoan (Cnidaria) phylogenomics: New approaches to long-standing problems. *Molecular Ecology Resources*, 18, 281–295. <https://doi.org/10.1111/1755-0998.12736>
- Quigley, K. M., Alvarez Roa, C., Torda, G., Bourne, D. G., & Willis, B. L. (2020). Co-dynamics of Symbiodiniaceae and bacterial populations during the first year of symbiosis with *Acropora tenuis* juveniles. *MicrobiologyOpen*, 9(2), e959.
- R Core Team. (2019). *R: A language and environment for statistical computing*. R Foundation for Statistical Computing. <https://www.R-project.org/>
- Ravinet, M., Faria, R., Butlin, R. K., Galindo, J., Bierne, N., Rafajlović, M., Noor, M. A. F., Mehlig, B., & Westram, A. M. (2017). Interpreting the genomic landscape of speciation: A road map for finding barriers to gene flow. *Journal of Evolutionary Biology*, 30(8), 1450–1477. <https://doi.org/10.1111/jeb.13047>
- Riginos, C., & Beger, M. (2022). Incorporating genetic measures of connectivity and adaptation in marine spatial planning for corals. In M. Van Oppen & M. Aranda (Eds.), *Coral reef conservation and restoration in the 'omics' age*. Springer.
- Riginos, C., Hock, K., Matias, A. M., Mumby, P. J., Oppen, M. J. H., & Lukoschek, V. (2019). Asymmetric dispersal is a critical element of concordance between biophysical dispersal models and spatial genetic structure in Great Barrier Reef corals. *Diversity and Distributions*, 25(11), 1684–1696. <https://doi.org/10.1111/ddi.12969>
- Rippe, J. P., Dixon, G., Fuller, Z. L., Liao, Y., & Matz, M. (2021). Environmental specialization and cryptic genetic divergence in two massive coral species from the Florida Keys Reef Tract. *Molecular Ecology*, 30, 3468–3484.
- Riquet, F., Liautard-Haag, C., Woodall, L., Bouza, C., Louisy, P., Hamer, B., Otero-Ferrer, F., Aublanc, P., Béduneau, V., Briard, O., el Ayari, T., Hochscheid, S., Belkhir, K., Arnaud-Haond, S., Gagnaire, P. A., & Bierne, N. (2019). Parallel pattern of differentiation at a genomic island shared between clinal and mosaic hybrid zones in a complex of cryptic seahorse lineages. *Evolution*, 73(4), 817–835. <https://doi.org/10.1111/evo.13696>
- Rose, N. H., Bay, R. A., Morikawa, M. K., & Palumbi, S. R. (2018). Polygenic evolution drives species divergence and climate adaptation in corals. *Evolution*, 72(1), 82–94. <https://doi.org/10.1111/evo.13385>
- Rose, N. H., Bay, R. A., Morikawa, M. K., Thomas, L., Sheets, E. A., & Palumbi, S. R. (2021). Genomic analysis of distinct bleaching tolerances among cryptic coral species. *Proceedings of the Royal Society B*, 288(1960), 20210678.
- Roux, C., Fraïsse, C., Romiguier, J., Anciaux, Y., Galtier, N., & Bierne, N. (2016). Shedding light on the grey zone of speciation along a continuum of genomic divergence. *PLoS Biology*, 14(12), e2000234. <https://doi.org/10.1371/journal.pbio.2000234>
- Sampayo, E. M., Ridgway, T., Bongaerts, P., & Hoegh-Guldberg, O. (2008). Bleaching susceptibility and mortality of corals are determined by fine-scale differences in symbiont type. *Proceedings of the National Academy of Sciences of the United States of America*, 105(30), 10444–10449. <https://doi.org/10.1073/pnas.0708049105>
- Sawall, Y., Al-Sofyani, A., Banguera-Hinestroza, E., & Voolstra, C. R. (2014). Spatio-temporal analyses of *Symbiodinium* physiology of the coral *Pocillopora verrucosa* along large-scale nutrient and temperature gradients in the Red Sea. *PLoS One*, 9(8), e103179. <https://doi.org/10.1371/journal.pone.0103179>
- Schliep, K. P. (2011). phangorn: Phylogenetic analysis in R. *Bioinformatics*, 27(4), 592–593.
- Schlötterer, C., Tobler, R., Kofler, R., & Nolte, V. (2014). Sequencing pools of individuals – Mining genome-wide polymorphism data without big funding. *Nature*, 15(11), 749–763. <https://doi.org/10.1038/nrg3803>
- Skotte, L., Korneliussen, T. S., & Albrechtsen, A. (2013). Estimating individual admixture proportions from next generation sequencing data. *Genetics*, 195(3), 693–702.
- Smith, H., Epstein, H., & Torda, G. (2017). The molecular basis of differential morphology and bleaching thresholds in two morphs of the coral *Pocillopora acuta*. *Scientific Reports*, 7(1), 10066. <https://doi.org/10.1038/s41598-017-10560-2>
- Sommer, B., Harrison, P. L., Beger, M., & Pandolfi, J. M. (2014). Trait-mediated environmental filtering drives assembly at biogeographic transition zones. *Ecology*, 95(4), 1000–1009. <https://doi.org/10.1890/13-1445.1/pdf>
- Sousa, V. C., & Hey, J. (2013). Understanding the origin of species with genome-scale data: Modelling gene flow. *Nature*, 14(6), 404–414. <https://doi.org/10.1038/nrg3446>
- Sturm, A. B., Eckert, R. J., Mendez, J. G., Gonzalez-Diaz, P., & Voss, J. D. (2020). Population genetic structure of the great star coral, *Montastraea cavernosa*, across the Cuban archipelago with

- comparisons between microsatellite and SNP markers. *Scientific Reports*, 10(1), 15432. <https://doi.org/10.1038/s41598-020-72112-5>
- Therkildsen, N. O., & Palumbi, S. R. (2016). Practical low-coverage genome-wide sequencing of hundreds of individually barcoded samples for population and evolutionary genomics in nonmodel species. *Molecular Ecology Resources*, 17, 194–208. <https://doi.org/10.1111/1755-0998.12593>
- Thomas, L., Rose, N. H., Bay, R. A., López, E. H., Morikawa, M. K., Ruiz-Jones, L., & Palumbi, S. R. (2018). Mechanisms of thermal tolerance in reef-building corals across a fine-grained environmental mosaic: Lessons from Ofu, American Samoa. *Frontiers in Marine Science*, 4, 434. <https://doi.org/10.3389/fmars.2017.00434>
- Thomas, L., Underwood, J. N., Rose, N. H., Fuller, Z. L., Richards, Z. T., Dugal, L., Grimaldi, C., Cooke, I. R., Palumbi, S. R., & Gilmour, J. P. (2022). Mechanisms of ecological divergence with gene flow in a reef-building coral on an isolated atoll in Western Australia. *Science Advances*, 8, eabl9185.
- Tom, J. A., Reeder, J., Forrest, W. F., Graham, R. R., Hunkapiller, J., Behrens, T. W., & Bhargale, T. R. (2017). Identifying and mitigating batch effects in whole genome sequencing data. *BMC Bioinformatics*, 18(1), 351. <https://doi.org/10.1186/s12859-017-1756-z>
- Tonk, L., Sampayo, E. M., Weeks, S., Magno-Canto, M., & Hoegh-Guldberg, O. (2013). Host-specific interactions with environmental factors shape the distribution of *Symbiodinium* across the Great Barrier Reef. *PLoS One*, 8(7), e68533. <https://doi.org/10.1371/journal.pone.0068533.s004>
- Underwood, J. N. (2009). Genetic diversity and divergence among coastal and offshore reefs in a hard coral depend on geographic discontinuity and oceanic currents. *Evolutionary Applications*, 2(2), 222–233. <https://doi.org/10.1111/j.1752-4571.2008.00065.x>
- Underwood, J. N., Richards, Z., Berry, O., Oades, D., Howard, A., & Gilmour, J. P. (2020). Extreme seascape drives local recruitment and genetic divergence in brooding and spawning corals in remote north-west Australia. *Evolutionary Applications*, 13, 2404–2421. <https://doi.org/10.1111/eva.13033>
- van Oppen, M. J. H., Bongaerts, P., Frade, P. R., Peplow, L. M., Boyd, S. E., Nim, H. T., & Bay, L. K. (2018). Adaptation to reef habitats through selection on the coral animal and its associated microbiome. *Molecular Ecology*, 27(14), 2956–2971. <https://doi.org/10.1111/mec.14763>
- van Oppen, M. J. H., Gates, R. D., Blackall, L. L., Cantin, N., Chakravarti, L. J., Chan, W. Y., Cormick, C., Crean, A., Damjanovic, K., Epstein, H., Harrison, P. L., Jones, T. A., Miller, M., Pears, R. J., Peplow, L. M., Raftos, D. A., Schaffelke, B., Stewart, K., Torda, G., ... Putnam, H. M. (2017). Shifting paradigms in restoration of the world's coral reefs. *Global Change Biology*, 23(9), 3437–3448. <https://doi.org/10.1111/gcb.13647>
- Warmuth, V. M., & Ellegren, H. (2019). Genotype-free estimation of allele frequencies reduces bias and improves demographic inference from RADSeq data. *Molecular Ecology Resources*, 19(3), 586–596. <https://doi.org/10.1111/1755-0998.12990>
- Warner, P. A., vanOppen, 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, 2993–3008. <https://doi.org/10.1111/mec.13225>
- Webster, J. M., & Davies, P. J. (2003). Coral variation in two deep drill cores: Significance for the Pleistocene development of the Great Barrier Reef. *Sedimentary Geology*, 159, 61–80.
- Willis, B. L., vanOppen, 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>
- Ying, H., Cooke, I., Sprungala, S., Wang, W., Hayward, D. C., Tang, Y., Huttley, G., Ball, E. E., Forêt, S., & Miller, D. J. (2018). Comparative genomics reveals the distinct evolutionary trajectories of the robust and complex coral lineages. *Genome Biology*, 19, 175.
- Yu, G., Smith, D. K., Zhu, H., Guan, Y., & Lam, T. T. Y. (2017). ggtree: An R package for visualization and annotation of phylogenetic trees with their covariates and other associated data. *Methods in Ecology and Evolution*, 8(1), 28–36.
- Zayasu, Y., Takeuchi, T., Nagata, T., Kanai, M., Fujie, M., Kawamitsu, M., Chinen, W., Shinzato, C., & Satoh, N. (2021). Genome-wide SNP genotyping reveals hidden population structure of an acroporid species at a subtropical coral Island: Implications for coral restoration. *Aquatic Conservation: Marine and Freshwater Ecosystems*, 31(9), 2429–2439. <https://doi.org/10.1002/aqc.3626>
- Ziegler, M., Roder, C., Büchel, C., & Voolstra, C. R. (2015). Niche acclimatization in Red Sea corals is dependent on flexibility of host-symbiont association. *Marine Ecology Progress Series*, 533, 149–161. <https://doi.org/10.3354/meps11365>

## SUPPORTING INFORMATION

Additional supporting information can be found online in the Supporting Information section at the end of this article.

**How to cite this article:** Matias, A. M. A., Popovic, I., Thia, J. A., Cooke, I. R., Torda, G., Lukoschek, V., Bay, L. K., Kim, S. W., & Riginos, C. (2022). Cryptic diversity and spatial genetic variation in the coral *Acropora tenuis* and its endosymbionts across the Great Barrier Reef. *Evolutionary Applications*, 00, 1–18. <https://doi.org/10.1111/eva.13435>