Genome-wide SNP analysis reveals an increase in adaptive genetic variation through selective breeding of coral

Kate Quigley [1], Line Bay[1], and Madeleine van Oppen[2]

[1]Australian Institute of Marine Science
[2]University of Melbourne Faculty of Science

May 11, 2020

Abstract

Summer heat waves are the principal global driver of mortality in reef-building corals. Resilience-based genetic management may increase coral heat tolerance, but it is unclear how temperature responses are regulated at a genomic level and thus how corals may adapt to warming naturally or through selective breeding. Here we combine phenotypic, pedigree, and genomic marker data from colonies sourced from a warm reef on the Great Barrier Reef reproductively crossed with conspecific colonies from a cooler reef to produce combinations of warm and cool purebred and hybrid larvae and juveniles. Intra-population breeding created significantly greater genetic diversity across the coral genome and maintained diversity in key regions associated with heat tolerance and fitness. High-density genome-wide scans of single nucleotide polymorphisms (SNPs) identified alleles significantly associated with offspring reared at 27.5°C (87 − 2,224 loci), including loci putatively associated with proteins involved in responses to heat stress (cell membrane formation, metabolism, and immune responses). Underlying genetics explained 43% of PCoA variation in juvenile survival, growth, and bleaching responses at 27.5°C and 31°C between the multilocus genotypes. Genetic marker contribution to total variation in fitness traits (narrow-sense heritability) were high for survival but not for growth and bleaching in juveniles, with heritability of these traits influenced more at 31°C relative to 27.5°C. Using only a limited number of crosses, the mechanistic understanding presented here demonstrates that allele frequencies are affected by one generation of selective breeding, key information for the assessments of genetic intervention feasibility and modelling of reef futures.

Introduction

Coral reefs globally are undergoing significant degradation due to mass coral mortality driven by high sea surface temperatures (Heron et al., 2017; Hughes et al., 2018). Increasingly warm seawater temperatures and losses in coral cover are also leading to changes in fundamental reproductive processes (Baums et al., 2019; Hughes et al., 2019; Shlesinger & Loya, 2019), with major implications for ecosystem resilience and recovery in the aftermath of severe, acute events such as mass coral bleaching. To safeguard the persistence and resilience of coral reefs, assisted evolution interventions to increase the heat tolerance of corals at the early life-history stages as well as in adults are being tested in the laboratory and the field (National Academies of Sciences and Medicine, 2019; Reef Restoration and Adaptation Program: Intervention Technical 2019). These interventions may also protect populations of high importance. Assisted evolution interventions include selective breeding within species, hybridisation between species, and microbial manipulations (Chakravarti, Beltran, & Oppen, 2017; Chan, Peplow, Menéndez, Hoffmann, & van Oppen, 2019; Damjanovic, Blackall, Webster, & Oppen, 2017; Morgans, Hung, Bourne, & Quigley, 2019; Quigley, Bay, & van Oppen, 2019; van Oppen, Oliver, Putnam, & Gates, 2015). The knowledge gained from experimentation of genetic interventions will facilitate optimisation of breeding designs, reduce the incidences and impacts of trait trade-offs and the help to determine the probabilities of success and timeframes associated with genetic gains.

Improved forecasting of future coral reef health based on the scope and rates of recovery is urgently needed.

1

This will rely on estimating rates of acclimatisation and adaptation to predict how populations will respond to a suite of selective pressures. The incorporation of the effects of selection and adaptation to understand population evolutionary trajectories has been applied to the conservation of species like Eltham's butterfly and bighorn sheep (Creech et al., 2017; Roitman et al., 2017). Corals exist close to their thermal limits (reviewed in Drury, 2019) but exhibit a wide range of phenotypes in response to heat stress (Ainsworth et al., 2015; Kenkel et al., 2013), suggestive of a high adaptive capacity predicated on genetic diversity across populations and species (Matz, Treml, Aglyamova, & Bay, 2018). There is evidence that corals have increased their heat tolerance by ~0.5°C in the last decade, which suggests substantial capacity to adapt to increasing temperatures (Sully, Burkepile, Donovan, Hodgson, & van Woesik, 2019). Part of this heat tolerance can be attributed to the diversity and relative abundance of dinoflagellate symbionts (family Symbiodiniaceae) that inhabit coral tissues (Baird, Bhagooli, Ralph, & Takahashi, 2009; reviewed in Quigley, Baker, Coffroth, Willis, & van Oppen, 2018). In particular, changes in symbiont community composition can attribute 1.0-1.5°C of increased tolerance to adults (Berkelmans & van Oppen, 2006) and 26x increased tolerance in juvenile corals relative to warm-warm and cool-warm crosses (Quigley, Randall, van Oppen, & Bay, 2020) and has been found to explain up to 24% of bleaching variability (Mizerek, Baird, & Madin, 2018). The genomic architecture of heat tolerance is often polygenic (Bay & Palumbi, 2015; Jin et al., 2016). Multiple candidate genes of varying effect sizes have been examined, including heat shock proteins and genes involved in immunity (Louis, Bhagooli, Kenkel, Baker, & Dyall, 2017), although some target genes may more broadly be associated with stress tolerance generally and not heat tolerance specifically. Therefore, to understand corals adaptive capacity, comprehensive knowledge is needed concerning the variation in alleles at particular loci and the topology of polymorphic genomic regions. This knowledge can then be used to understand their influence in driving the emergence of different phenotypes that ultimately enable adaptation in populations over ecological and evolutionary time frames.

Evolutionary genetics describes gene frequency changes across populations and over time. The analysis of genetic diversity in wild and artificially-bred populations is central to understanding the rates and potential of adaptation. Shallow and deep whole genome sequencing is now widely available and relatively cost effective, including sequencing whole genome single-nucleotide polymorphisms (SNPs) (Davey et al., 2011; Helyar et al., 2011) and are increasingly illuminating the ecological and evolutionary mechanisms coral use to respond to their environment (Dixon et al., 2015; Fuller et al., 2019; Palumbi, Barshis, Traylor-Knowles, & Bay, 2014). SNPs are particularly valuable for identifying common variants underlying trait variation and can be applied to genetic conservation practices through breeding programs that utilize quantitative genetic principles, breeding values, and narrow and broad-sense heritability (h^2, H^2) (Visscher, Hill, & Wray, 2008). SNPs are intrinsically linked to the h^2 given the proportionality to the product of SNP quantity and effect size (Holland et al., 2019). This marker-assisted selection approach identifies markers ultimately underlying the heritability estimates, although single alleles may explain only a small proportion of the measured heritable variation (Manolio et al., 2009). Specifically, SNPs provide the marker data used as inputs for relatedness models, that when combined with phenotypic data, are used to calculate additive genetic effects and variances and finally, heritability and the potential for selection (O'hara, Cano, Ovaskainen, Teplitsky, & Alho, 2008).

Here we apply high-density genome-wide marker sequencing to samples collected from selectively bred reproductive crosses to elucidate how artificially produced corals may evolve under heat stress. Selective breeding of corals from historically warmer reefs with others sourced from cooler reefs and then exposing juveniles to temperature tolerant symbionts like *Durusdinium trenchii* significantly improved coral fitness when exposed to 31°C (e.g. increased survival, growth and bleaching tolerance (Quigley, Randall, van Oppen, & Bay, 2020)), even upon inspection of a limited number of crosses produced from only five parents. Here we link those traits measured in the five reproductive crosses to their underlying allele frequency changes and genomic contribution (h^2) and identify the putative causative variants separating the five families to explain and potentially link these heritability estimates at both ambient and elevated temperatures to explain the genetic variation in these three quantitative traits.

Materials and Methods

Spawning, larval rearing, settlement and symbiont infection

Gravid adult corals of the species *Acropora spathulata* were collected from one reef in the far north of the Great Barrier Reef (Tijou, far northern GBR; 13°10'44.0"S, 143°56'54.6"E, permit G16/38488.1) and one reef in the central GBR (Backnumbers, central GBR; 18°30'49.8"S, 147°09'10.7"E, permit G12/35236.1) and brought back to the National Sea Simulator (SeaSim) for coral spawning in November 2017 as outlined in (Quigley et al., 2020). Briefly, on the night of spawning the separated eggs and sperm from three far northern colonies were mixed with those of three central colonies, resulting in larvae from 30 distinct familial crosses that were raised in 27.5°C ultrafiltered, flow-through seawater (see Table 1 for pedigree and parental colony designations). Only five families were available at the time of larval settlement and used to produce recruits (hereafter referred to as juveniles given measurements were taken after 70 days of development). Here we focus on those five familial crosses sequenced at both the larval stage (27.5°C) and the juvenile stage (27.5°C and 31°C). These crosses include three produced from two parents sourced from the warm far northern reef (WW1, WW2, WW3), one cross with a warm dam and cool sire (WC) and one with a cool dam and warm sire (CW). Aposymbiotic juveniles were infected with *Symbiodinium tridacnidorum* ,*Cladocopium goreaui* , and *Durusdinium trenchii* with cultured material obtained from the AIMS Symbiont Culture Facility (further details in Quigley et al., 2020). Full details of the collection, environmental parameters of each location, permits, spawning, larval and juvenile rearing, and symbiosis establishment can be found in Quigley et al., 2020. Briefly, symbiosis establishment in coral juveniles was performed by adding $1x10^5$ ml^{-1}of one of the three cultured symbiont strains separately to each juvenile tank (n = 3 replicate tanks per culture), suspending the flow to allow for contact between corals and symbionts, and repeating one further time until infection was confirmed with microscopy.

Larval sampling and trait measurements in juveniles

Acropora spathulata larvae were sampled into 100% ethanol after rearing at 27.5°C ~2.5 days after fertilization. DNA from tissue samples were extracted using a modified SDS protocol previously used for sequencing of single coral eggs (Quigley, Willis, & Bay, 2017). Juvenile growth (percent change in area), bleaching (percent change in coloration), and survival (alive or dead) were quantified after 70 days of experimental exposure to 27.5°C and 31°C, and a detailed assessment of this physiological data can be found in (Quigley et al., 2020).

Sampling, DNA extraction, library preparation, sequencing

Individual larvae were genotyped (total n = 68) across five families representing three population crosses (n = WW: 34, WC: 15, CW: 19). All samples were sequenced on the Illumina HiSeq2500 and genotyped using proprietary DArT-seq^{TM} technology at Diversity Arrays Technology (Canberra, Australia). Digestion and ligation reactions were carried out with the PstI and HpaII restriction enzymes following (Jaccoud, Peng, Feinstein, & Kilian, 2001; Kilian et al., 2012). DNA purification and library preparation followed DArT proprietary methods (Kilian et al., 2012). As part of the DArT pipeline, raw sequence data was filtered based on Q-scores and filtered against viral and bacterial databases to remove contamination. The restriction site fragments generated were used to call single nucleotide polymorphisms (SNPs) with the KDCompute pipeline and DArTsoft14 algorithm (Diversity Arrays Technology, http://www.kddart.org/kdcompute.html). SNPs were referenced against the *Acropora digitifera* genome (Shinzato et al., 2011) and further quality controlled and filtered to produce a reduced dataset of only high quality loci using both proprietary DART software and the following criteria, which included filtering loci for average repeatability of alleles at the locus ('filter.repavg', >99%), collapsing duplicated sequences ('keep monomorphs'), individuals with missing data ('filter.callrate', >25%), filtering by Minor Allele Frequencies (< 0.02), coverage (< 8 read depth), and filtering loci out of Hardy-Weinberg Equilibrium.

Statistical Analyses

Multivariate statistics

SNP metrics were recalculated in R for post-filtered data using the 'DArTR package' (Gruber, Unmack, Berry, & Georges, 2018). PCoAs were constructed from individual larval SNP genotypes with Gower PCoA

3

ordination and Euclidean distances using the packages 'Adegenet' and 'DArTR' (Gruber et al., 2018; Jombart, 2008). Clusters of genetically related individuals were inferred using Discriminant Analysis of Principal Components (Jombart, Devillard, & Balloux, 2010) constructed with and without the between- and within-group variance components from PCoA outputs. Multiple loading thresholds were explored, including the default value of the third quartile (75%) of x-values (Grünwald, Kamvar, & Everhart, Jombart & Collins, 2015). Diversity metrics were calculated using 'DArTR' following (Sherwin, Chao, Jost, & Smouse, 2017). These analyses are similar to classic population genetic analyses that measure differentiation via F_{st}, in which high allele frequency differentiation suggests population differentiation although they can be fraught with false positives (Rajora, 2019).

Given the influence of close relationships in STRUCTURE analyses (Anderson & Dunham, 2008), DAPC was used as an alternative method to infer admixture (Grünwald et al., 2010). Hence, "true k" was calculated based on Bayesian Information Criteria (BIC, Jombart et al., 2010), in which the largest drop in BIC was at $k = 5$.

Single Nucleotide Polymorphism annotation

SNPs significantly contributing to the PCoA separation between each family cross were identified, translated to nucleotide sequences by searching and aligning sequences to the *Acropora digitifera* genome (v. OIST_v1.1) using custom scripts. All nucleotide matches were downloaded and re-formatted into a searchable BLAST database using the "blastdbcmd" command. Nucleotide matches were blasted to annotated proteins (OIST v.0.9, see reference for full instructions on annotations). Predicted proteins were filtered for only protein matches with E-values <1e-4 were retained (https://blast.ncbi.nlm.nih.gov/Blast.cgi?PROGRAM=

blastx&PAGE_TYPE=BlastSearch&LINK_LOC=blasthome).

Modelling SNP frequency distributions

The distributions of expected Hardy-Weinberg (HW) equilibrium genotype frequencies were simulated using a custom R script (Chivers, 2011) for three population sizes of the SNP profiles from single larval samples (number of larvae in each population, n = 34, 19, 15). A null distribution of allele frequencies in a population that is not evolving follows HW proportions. SNP profiles were constructed from single larval samples for all families as well as bioinformatically pooled by geographic source of parental corals (WW, WC, CW). Parameters included were number of parents, total population sizes, number of genotypes, and number of generations to simulate in randomly mated populations. Expected HW genotype frequencies were statistically compared to the experimentally observed genotype frequencies from the families and a multidimensional, distribution-free Kolmogorov-Smirnov Two-Sample test was used to assess statistical significance of differences between paired continuous two-dimensional distributions using the package 'Peacock.test' (Fasano & Franceschini, 1987).

Narrow-sense heritability

Physiological data on survival, bleaching, and growth (Quigley et al., 2020) were combined with pedigree information (Supplementary Table 1) to calculate narrow-sense heritability estimates for each trait at 27.5°C and 31°C. Relatedness between individual juveniles (sibships and family groups) are calculated from the pedigree input as a random effect. Narrow-sense heritability estimates (h^2) were derived from additive genetic variance calculated from the 'animal model' (Kruuk, 2004) using the package 'MCMCglmm' (Hadfield, 2010), in which the coefficient of relatedness between individual juveniles is parameterized as a random effect. The 'animal model' is a type of specialized mixed-effect model that calculates narrow-sense heritability by parameterizing relatedness between measured individuals using random effects (Kruuk, 2004). Hence, this analysis calculates the relative contribution of genetic and environmental influence on phenotypic variance of each trait (Wilson et al., 2010). This analysis builds upon that performed in Quigley et al., 2020 by: 1) using the underlying physiological data and 2) extending the mixed effects models used in that study that parameterized the influence of population of origin (WW, WC, CW) and parental cross (WW1, WW2, WW3, WC, CW) on survival, growth, and bleaching. Three Symbiodiniaceae strains were used for inoculation

of juveniles. For clarity, heritability estimates for only *Cladocopium* , the punitively principal symbiont of *Acropora spathulata* on the GBR (*sensu* van Oppen, Palstra, Piquet, & Miller, 2001), is presented in the main results and the additional strains in the Supplementary Information.

Models of h^2 of juvenile survival at both temperatures were run with time as a fixed factor, pedigree and individual as random factors, using binary logistic regression (categorical distribution), 150×10^6 iterations, 0.1% burn-in of total iterations and at multiple thinning levels of 7500 (27.5°C:*Symbiodinium tridacnidorum* , *Cladocopium goreaui* , and*Durusdinium trenchii* , 31°C: *C. goreaui* and *D. trenchii*) or 37,500 (31°C: *S. tridacnidorum*). Models of h^2 of juvenile bleaching and growth at both temperatures were run using 1.5×10^6 iterations, a thinning level of 1,500 and a burn-in of 10% of the total iterations. Fixed and random factors were parameterized in the survival models with a Gaussian distribution.

A non-informative flat prior specification was used, following an inverse gamma distribution (Wilson et al., 2010). Assumptions of chain mixing, normality of posterior distributions and autocorrelation were met. The posterior heritability was calculated by dividing the model variance attributed to relatedness by the sum of additive and residual variance. Given differences in iterations and thinning levels across models, a scaling factor ($x' = x - x_{min}/x_{max} - x_{min}$) was applied to the posterior distributions of each trait using the "rescale" function in the 'plottrix' package (Lemon, 2006) to visualize each heritability estimate equally to ease interpretation. Specifically, each posterior was re-scaled to the minimum ($y = 0.00598$) and maximum ($y = 57.42421$) y-density values for the *S. tridacnidorum* treatment at 31°C.

Results

Interpopulation differentiation

After filtering, 9,031 high quality SNPs remained from the larval samples collected at ambient temperature across 68 individually genotyped larvae from the five families ($n = 7 – 20$ for each). The five families separated in multidimensional space, in which genetic variation explained 43% of the variation between multilocus genotypes from the five families (PC1-PC3: 25.8; 13.4; 4.1%) (Figure 1A). Families that shared the same sire (Tijou 2A) roughly clustered together in Principle Components Analysis (PCoA) space (WW1 and CW with WW2), whereas the other two families were more divergent.

Analysis of population structure using Discriminant Analysis of Principle Components (DAPC) further confirmed this pattern of clustering between WW1 and CW along the first two discriminant functions. The 95% confidence ellipses demonstrate that the dispersion of multilocus genotypes also varied across the families, where WW2 shows the greatest and WW1 shows the least variance. Cluster membership probabilities calculated from DAPC analysis also showed flat distributions in between-group and within-group allele variances, demonstrating that admixture has occurred, especially between WW3, WC, WW1, and to a lesser extent with WW2, and particularly CW.

Allele frequency profiles

The majority of loci were fixed at either 0 or 1. The average distribution of loci in population purebred crosses was significantly different compared to interpopulation crosses, with a greater abundance of loci at intermediate frequencies in the interpopulational families (compared to null distributions; Kolmogorov-Smirnov Two-Sample test,$p = 2.1e-26$ and $1.3e-21$, Figure 2). This suggests that these interpopulation hybrids have significantly greater genetic diversity than within population purebreds. When individual purebred families were analysed, the majority of the on average reduced diversity (e.g. lower abundance of loci at intermediate frequencies) was attributed to crosses WW1 and WW2, whereas WW3 demonstrated a similarly greater abundance of loci at intermediate frequencies as the interpopulational crosses (Supplementary Figure 1). WW1 and WW2 exhibited significantly different distributions compared to the distribution of WC ($p = 5.5e-15$ and $2.4e-4$, respectively) and CW ($p = 1.2e-11$ and $6.7e-4$, respectively) (Figure 2). In contrast, the distribution of loci in the WW3 cross was not significantly different compared to WC ($p =8.7e-2$) or CW ($p = 0.77$). All families differed significantly from expected, modelled HW frequencies when modelled again populations of the same size ($p =$ WW and WC: $1.03e-45$; CW: $2.2e-46$). Observed heterozygosity (H_o),

and alpha Diversity ($^0D_\alpha$ and $^2D_\alpha$ at q = 0 and 2) was highest in the interpopulation crosses (WC, CW) compared to the purebreds, with the exception of WW2, which was also high (Table 1). Alpha Entropy at q = 1 was the highest in WC, WW2 and lowest in WW3.

QTL analysis

SNPs were correlated with the five families reared under ambient temperature. The number of SNPs associated with these five families varied depending on threshold cut-offs ranging from 2,224 SNPs (75% quartile) to 87 SNPs at the stricter threshold value (99% quartile). Twenty-six of the 87 SNPs (30%) could not be assigned annotations within the *A. digitifera* genome, a common issue in coral and symbiont 'omics (Figure 3 inset, *sensu* (Barshis et al., 2013; Peng et al., 2010; Shinzato et al., 2011; Sogin, Putnam, Anderson, & Gates, 2016)).

In the 87 SNPs that contributed to the separation of crosses at ambient temperatures, we identified proteins involved in a range of biological processes, mainly immunity and stress, cell functioning and metabolism, and calcification (Figure 3 inset). Stress related proteins included lysosomal-trafficking regulator-like proteins, CEPU-1-like protein, E3 ubiquitin-protein ligase RNF213-like, spondin, NFX1-type zinc finger-containing protein, NACHT, MAP/microtubule affinity and other processes, NLRC3-like proteins. Proteins involved with collagen production varied across families, as well as 2 dTDP-glucose 4,6-dehydratase-like and two sodium bicarbonate transporter-like proteins. A host of kinase related proteins, including those involved in cell functioning and transcription were detected.

PCoA loadings defined cut-offs for significant SNPs and hence assigned proteins. Homeodomain-interacting proteins, NFX1-type zinc finger-containing proteins and src proteins (Figure 3A), which are broadly associated with transcription, immunity and stress, and cell functioning were broadly positively related to the separation of families WW2/WW1/CW and WW3 and WC along PCoA axis 1 (Figure 1A). Alternatively, differences in cell functioning proteins were more negatively associated with their separation. Proteins contributing to the separation of WW2/WW3, WW1 and CW, and WC along PCoA2 were associated with responses to UV radiation, and MAP proteins (associated with cell functioning and immunity/stress, respectively) whereas multiple proteins associated with immunity and transcription were negatively associated with their separation.

Narrow-sense heritability

Population of origin (WW, WC, CW), family cross identity (WW1, WW2, WW3, WC, CW) and symbiont type significantly influenced juvenile responses to growth, bleaching and survivorship at 27.5°C and 31°C (full summary in Quigley et al., 2020). Specifically, juveniles with at least one parental coral sourced from the northern GBR survived significantly better, grew more, and bleached less at 31°C compared to other crosses, especially if infected with the Symbiodiniaceae symbiont, *D. trenchii* .

Physiological data for juvenile survival, growth, and bleaching were previously presented in Quigley et al., 2020. These raw data were used to calculate Bayesian narrow-sense heritability (h^2) estimates for survival, bleaching, and growth demonstrate that only survival of corals infected with *C. goreaui* was strongly affected by temperature (Figure 4, Supplementary Table 1). Bleaching responses were generally associated with low heritability estimates at both temperatures, although juveniles associated with *C. goreaui* exhibited more moderate heritability estimates. Heritable genotypic variation across families contributed little to differences in growth rates at 27.5 and 31°C (Table 2). With the other three Symbiodiniaceae taxa, h^2 estimates across the three traits varied in juveniles from the five families infected at 27.5 and 31°C (Supplementary Figure 1, Supplementary Table 1).

Discussion

Selective breeding produces increased occurrence of alleles with intermediate frequencies on which natural selection can act

Selective breeding is one genetic intervention strategy that may quickly increase adaptive genetic variation in

corals, as such facilitating adaptation to increasing sea surface temperatures (Chan, Hoffmann, & van Oppen, 2019; van Oppen et al., 2015). This approach has been used across a range of aquaculture and mariculture species to improve a number of commercially important traits. Commercially important species like oysters and mussels share similar life-history attributes with corals, including the production of numerous, potentially low-quality gametes, and high levels of genetic diversity (e.g. R-strategist) (Ellegren & Galtier, 2016). In those systems, only a few generations of selective breeding has resulted in 30% higher growth under elevated pCO_2 conditions compared to wild populations (Parker et al., 2012) and 50% higher growth in redclaw crayfish (Stevenson, Jerry, & Owens, 2013) without significantly eroding genetic diversity compared to wild populations after pooling breeding combinations (O'Connor, Dove, & Knibb, 2016). Selective breeding over only one generation in corals has shown that significant increases in heat tolerance of larvae (Dixon et al., 2015) and juveniles (Quigley et al., 2020) are possible using these methods. Lessons learnt here may therefore provide insights into the mechanisms underpinning the success of these techniques in a wild marine context.

We found that alleles at intermediate frequencies were at greater abundances than expected under HWE in WC and CW crosses relative to the on average distribution of loci in WW crosses, suggesting an overall increase in genetic diversity resulting from selective breeding of corals from different regions of the GBR. Distributions of allele frequencies typically follow that most loci nearly reach fixation at either 0 or 1, where alleles at intermediate frequencies are less common but are important given an increased abundance of intermediate alleles are the raw material for selection (Jombart & Collins, 2015). Hence, genetic material on which selective adaptive processes can operate likely exists at greater abundances in the WC and CW relative to the WW crosses. Hybridisation may lead to increased performance/fitness relative to the parental generation, with this increased performance having been linked to dominant, regulatory, single quantitative trait nucleotides (Jakobson & Jarosz, 2019). Increases in genetic diversity, hybrid vigour and genetic rescue are well-known but inconsistent features of intra- and inter-specific hybridization (Chan, Hoffmann & van Oppen, 2019; Flowers et al., 2019; Hazzouri et al., 2019; Weeks et al., 2011) but has not yet been demonstrated in the selective breeding of corals. For example, it was unknown whether the breeding of divergent populations would cause a decrease in genetic diversity due to processes like genome incompatibility (Hogenboom, 1975). These results support that even in a small number of crosses and contributing parents, genetic diversity improves. Observed heterozygosity and alpha Diversity ($^{0}D_{\alpha}$ and $^{2}D_{\alpha}$ at q = 0 and 2) was highest in the interpopulation crosses (WC, CW). The $^{0}D_{\alpha}$ metric is sensitive to the presence of rare alleles, which is important given breeding across populations may initially introduce novel mutants (Sherwin et al., 2017), suggesting that WC, CW, and WW2 had an increased occurrence of rare variants. $^{1}H_{\alpha}$, also known as the Shannon Information criteria, is informative as a natural measure of evolvability, and was highest in WC and WW2. Finally, preliminary analysis of adult colonies shows Tijou corals had on average higher observed heterozygosity compared to Backnumbers corals (H_{o} = 0.065 vs. 0.0217, unpublished data), suggesting that there could have been a risk of reduced diversity in offspring. Importantly, we did not observe a loss of genetic diversity by crossing these two divergent populations.

Finally, we found that interpopulation selective breeding significantly changed the resulting allele frequencies of offspring compared to modelled HWE distributions. This is perhaps unsurprising given assumptions of random mating, no gene flow, and infinite population sizes were not met. However, this suggests the influence of breeding on the genetic architecture of the resulting F1 generation. Therefore, the interbreeding of even a small number of corals from different reefs across the GBR may result in extensive introgression and therefore accelerate the potential for adaptation to warming, although the production of F2 generations would be needed to confirm this.

We also demonstrate that variants associated with immune responses, growth and cellular operating may be re-arranged during breeding but are maintained within the next generation. This suggests that the important functional diversity (i.e., at stress tolerance genes) associated with focal populations can be maintained in the breeding process. In response to heat stress, corals alter their gene and protein expression patterns, as reflected in changes in their structural lipids, metabolism, and immune responses (Barshis et al., 2013; Sogin et al., 2016). Differences in proteins associated with collagen production and sodium bicarbonate transport were important in differentiating the five families produced in our study. Collagen is important for the

7

production of the extracellular matrix, required for multicellularity and the spatial organization of functional units of cells (Helman et al., 2008) whereas bicarbonate transporters are pivotal for coral calcification and hence growth (Zoccola et al., 2015). Basic cellular functioning also potentially varied through the differences in dTDP-glucose 4,6-dehydratase-like proteins detected and their involvement in the non-oxidative pentose phosphate pathway (Buerger, Wood-Charlson, Weynberg, Willis, & van Oppen, 2016; Yuyama, Watanabe, & Takei, 2011), critical for glucose utilization. Differences between crosses in these foundational processes like cellular organization and biomineralization therefore suggests that even breeding across relatively few individuals has the potential to substantially create distinct genetic combinations.

Protein NLRC3-like, which has been previously implicated in acroporid immune suppression in response to heat stress by acting in Toll-like receptor modification (Zhou et al., 2017), also varied significantly across families. Other proteins involved in immunity and stress were also detected (lysosomal-trafficking regulator-like proteins), and these have also been linked to the mounting of innate immune responses through Toll-like receptor activity in mice (Westphal et al., 2017). Additional immunity related proteins were detected, including CEPU-1-like protein (Spaltmann & Brummendorf, 1996), E3 ubiquitin-protein ligase RNF213-like (Iguchi et al., 2019), NACHT (Hamada et al., 2013), and spondin (Palmer & Traylor-Knowles, 2012). Interestingly, NFX1-type zinc finger-containing protein was found here and downregulated in resistant corals exposed to disease (Polato et al., 2010). Expression of nascent polypeptide-associated complex subunit alpha protein (Bellantuono, Hoegh-Guldberg, & Rodriguez-Lanetty, 2012) was similarly downregulated in resistant corals, suggesting that these proteins provide an important role in protective immune responses.

Genotyping individual aposymbiotic larva from the five families reared under ambient conditions (27.5°C) provides foundational knowledge as to how selective breeding influences underlying genetic architecture. It also sheds light on the underlying molecular origins and mechanisms of heritability, a long-standing goal of quantitative genetics (Jakobson & Jarosz, 2019). Broad and narrow-sense heritability has been quantified for a range of traits, but the underlying mechanisms have rarely been described (Dixon et al., 2015; Dziedzic, Elder, & Meyer, 2017). The majority of alleles fixed at the extremes of allele frequency distributions ("U-shaped" *sensu* Hill, Goddard, & Visscher, 2008) are likely driven by the small number of parents from which the larvae in each cross were derived (5 unique parental colonies used across the families, Supplementary Table 1), but may also suggest selection against heterozygotes during the aquarium rearing period. Inter-population hybrids displayed greater genetic diversity relative to within population purebreds, which may be the result of the varying effects of selection on purebreds versus hybrids in the aquarium environment. U-shaped distributions may arise under strong cases of artificial selection (e.g. aquariums) combined with rare mutations (Hill et al., 2008).

Functional variation associated with selectively bred families

How does genomic variation lead to phenotypic differences between corals? The location and effect size of the SNP difference are important to determining its eventual phenotypic effect, and simplistically, differences in coding vs. non-coding regions are predicted to cause phenotypic effects through a variety of mechanisms (Cavallo & Martin, 2005). Assigning function to SNP differences is challenging given the majority of SNPs detected by association studies are non-coding (Nishizaki & Boyle, 2017), whereas the majority of key changes are coding (Cavallo & Martin, 2005), setting up a situation of difficult detection and classification. Analysis of population structure using DAPC links the genomic patterns seen in the multilocus genotypes with the underlying biological processes quantified in the heritability models. Using this approach, we identified alleles contributing to the separation of these selectively bred families, revealing that breeding of the selected populations targets changes to the immunity and stress responses and growth, likely important processes in survival generally. Assigning differences in SNPs to phenotypic differences between individuals will be key to understanding and increasing thermal tolerance for intervention methods.

Selective breeding influences the level of admixture and population discontinuity

Admixture events affect members of species, populations and individuals differently (Lawson, Van Dorp, & Falush, 2018). We saw this in the extent of admixture and its associated variance across the five families, in

which some families exhibited very little admixture relative to others (especially CW). Although this can be somewhat dependent on the k structure of the model used, both patterns were explored independently using two techniques (DAPC and PCoA) and in conjunction with AIC, confirmed statistically the likelihood of population differentiation. Therefore, this may suggest that the shared ancestry of the colony sourced from Backnumbers2 may be limited between the other Backnumbers and Tijou corals or whose origins were from few or divergent founders (Lawson et al., 2018). This would suggest that adult colony Backnumbers2 is not highly related to other Backnumbers or Tijou colonies. Furthermore, PCoA and DAPC both demonstrated that the five families separate out in multidimensional space given the magnitude of allelic covariance between individuals. Irrespective of population labels, DAPC analysis also recapitulated the number of selectively bred families produced, although interestingly, the purebred families were not assigned to single population clusters but instead retained hybrid clustering structure in which the proportion of ancestry was shared between multiple two to three clusters simultaneously. The discontinuity between populations was also surprising, as demonstrated by the reduced spread of individuals between clusters, especially in WW1, suggestive that selective breeding produces discrete differences in the underlying genomic architecture of F1 individuals, even in populations likely undergoing some underlying level of gene flow (Lukoschek, Riginos, & van Oppen, 2016).

Heritability

Heritability estimates provide information as to the extent to which phenotypic trait variation has a genetic basis and the potential for adaptive change in these traits (Visscher et al., 2008). SNPs are intrinsically linked to narrow-sense heritability given h^2 is proportional to the product of the number of SNPs and their effect size (Holland et al., 2019). We found a high level of variation in survival, bleaching, and growth attributed to narrow-sense heritability (h^2) when pooled at the level of population cross, which varied in juveniles infected with *C. goreaui* across the two temperatures tested here (see Supplementary Information for further details on *S. tridacnidorum* and *D. trenchii*).

Interestingly, additive genetic variance was only important in influencing survival at 31°C when infected with *C. goreaui* but not at 27.5°C. Hence, it is likely that there is significant selection on survival at high temperatures for *C. goreaui* , consistent with the generally overall poorer performance of these symbionts at both temperatures (*S. tridacnidorum*) or at higher temperatures in coral juveniles (*C. goreaui*) (see Table 6.1 in Quigley, Baker, Coffroth, Willis, & van Oppen, 2018). The shapes of the posterior heritability distributions are informative for generating hypotheses concerning the underlying genomic architecture associated with selective breeding. Bimodal heritability estimates can be driven by single large effect eQTLs (as seen generally for traits with high heritability), whereas multiple loci of small effect size generally result in unimodal distributions (Rudra et al., 2017). The influence of *cis* and *trans* co-regulation of the traits may also influence the underlying distribution (Yang et al., 2014). Bimodal heritability (often with wide credibility intervals) estimates can also be driven by high underlying within-population variability even with large sample sizes (mirrored in the high phenotypic variability between genotypes, Quigley et al., 2020), and have been recorded across a range of disease-related expression of genes (Yang et al., 2014). High heritability at ambient temperatures may also be common (Kronenberg et al., 2019), and potentially indicative of core traits under continual selective pressures. Finally, it is likely that not all the SNPs associated with these three phenotypes have been captured although narrow-sense heritability should incorporate additively all the common SNPs associated (Holland et al., 2019) and it will be important to use these methods under a variety of experimental settings to fully develop models of adaptive selection.

Conclusion

The identification of potential allele frequency differences among familial crosses can link specific genotypes to important fitness traits. Although only a small number of crosses are presented in this pilot study produced from the mixing of five adult colonies, our findings suggest that selective breeding has measurable effects on the genomic diversity of the resulting offspring. This was apparent in the significant differences in genetic variants and phenotypes associated with stress tolerance when exposed to varying temperatures. These results are in line with studies demonstrating the influence of recent immigrants having significant influence

on allele frequency shifts.

Acknowledgements

References

Ainsworth, T. D., Krause, L., Bridge, T., Torda, G., Raina, J.-B., Zakrzewski, M., . . . Smith, C. (2015). The coral core microbiome identifies rare bacterial taxa as ubiquitous endosymbionts. *The ISME Journal* , *9* (10), 2261–2274.

Anderson, E. C., & Dunham, K. K. (2008). The influence of family groups on inferences made with the program Structure. *Molecular Ecology Resources* , *8* (6), 1219–1229.

Baird, A. H., Bhagooli, R., Ralph, P. J., & Takahashi, S. (2009). Coral bleaching: the role of the host. *Trends in Ecology & Evolution* ,*24* (1), 16–20. doi: 10.1016/j.tree.2008.09.005

Barshis, D. J., Ladner, J. T., Oliver, T. a, Seneca, F. O., Traylor-Knowles, N., & Palumbi, S. R. (2013). Genomic basis for coral resilience to climate change. *Proceedings of the National Academy of Sciences of the United States of America* , *110* (4), 1387–1392. doi: 10.1073/pnas.1210224110

Baums, I. B., Baker, A. C., Davies, S. W., Grottoli, A. G., Kenkel, C. D., Kitchen, S. A., . . . Shantz, A. A. (2019). Considerations for maximizing the adaptive potential of restored coral populations in the western Atlantic. *Ecological Applications* , *29* (8), e01978. doi: 10.1002/eap.1978

Bay, R. A., & Palumbi, S. R. (2015). Rapid acclimation ability mediated by transcriptome changes in reef-building corals. *Genome Biology and Evolution* , *7* (6), 1602–1612.

Bellantuono, A. J., Hoegh-Guldberg, O., & Rodriguez-Lanetty, M. (2012). Resistance to thermal stress in corals without changes in symbiont composition. *Proceedings of the Royal Society B-Biological Sciences* , *279* (1731), 1100–1107.

Berkelmans, R., & van Oppen, M. J. H. (2006). The role of zooxanthellae in the thermal tolerance of corals: a 'nugget of hope' for coral reefs in an era of climate change. *Proceedings of the Royal Society B: Biological Sciences* , *273* (1599), 2305–2312. doi: 10.1098/rspb.2006.3567

Buerger, P., Wood-Charlson, E. M., Weynberg, K. D., Willis, B. L., & van Oppen, M. J. H. (2016). CRISPR-Cas defense system and potential prophages in Cyanobacteria associated with the coral Black Band Disease.*Frontiers in Microbiology* , *7* , 2077.

Cavallo, A., & Martin, A. C. R. (2005). Mapping SNPs to protein sequence and structure data. *Bioinformatics* , *21* (8), 1443–1450.

Chakravarti, L. J., Beltran, V. H., & Oppen, M. J. H. (2017). Rapid thermal adaptation in photosymbionts of reef-building corals.*Global Change Biology* .

Chan W.Y., Hoffmann A.A., & van Oppen M. J. H. (2019) Hybridization as a conservation management tool. Conservation Letters 0:e12652Chan, W. Y., Peplow, L. M., Menendez, P., Hoffmann, A. A., & van Oppen, M. J. H. (2019). The roles of age, parentage and environment on bacterial and algal endosymbiont communities in Acropora corals. *Molecular Ecology* , *28* (16), 3830–3843. doi: 10.1111/mec.15187

Chivers, C. (2011). https://gist.github.com/cjbayesian/468725f4bd7d6f3f027d#file-hw_sim-r

Creech, T. G., Epps, C. W., Landguth, E. L., Wehausen, J. D., Crowhurst, R. S., Holton, B., & Monello, R. J. (2017). Simulating the spread of selection-driven genotypes using landscape resistance models for desert bighorn sheep. *PLOS ONE* , *12* (5), 1–26. doi: 10.1371/journal.pone.0176960

Damjanovic, K., Blackall, L. L., Webster, N. S., & Oppen, M. J. H. (2017). The contribution of microbial biotechnology to mitigating coral reef degradation. *Microbial Biotechnology* , *10* (5), 1236–1243.

Davey, J. W., Hohenlohe, P. A., Etter, P. D., Boone, J. Q., Catchen, J. M., & Blaxter, M. L. (2011). Genome-wide genetic marker discovery and genotyping using next-generation sequencing. *Nature Reviews Genetics* , *12* (7), 499.

Dixon, G. B., Davies, S. W., Aglyamova, G. V, Meyer, E., Bay, L. K., & Matz, M. V. (2015). Genomic determinants of coral heat tolerance across latitudes. *Science* , *348* , 1460–1462. Retrieved from http://science.sciencemag.org/content/348/6242/1460.abstract

Drury, C. (2019). Resilience in reef-building corals: The ecological and evolutionary importance of the host response to thermal stress.*Molecular Ecology* .

Dziedzic, K., Elder, H., & Meyer, E. (2017). Heritable variation in bleaching responses and its functional genomic basis in reef-building corals (*Orbicella faveolata*). *BioRxiv* . Retrieved from http://biorxiv.org/content/early/2017/10/07/185595.abstract

Ellegren, H., & Galtier, N. (2016). Determinants of genetic diversity.*Nature Reviews Genetics* , *17* (7), 422–433.

Fasano, G., & Franceschini, A. (1987). A multidimensional version of the Kolmogorov–Smirnov test. *Monthly Notices of the Royal Astronomical Society* , *225* (1), 155–170.

Flowers, J. M., Hazzouri, K. M., Gros-Balthazard, M., Mo, Z., Koutroumpa, K., Perrakis, A., . . . Aberlenc, F. (2019). Cross-species hybridization and the origin of North African date palms.*Proceedings of the National Academy of Sciences* , *116* (5), 1651–1658.

Fuller, Z. L., Mocellin, V. J. L., Morris, L., Cantin, N., Shepherd, J., Sarre, L., . . . Andolfatto, P. L. (2019). Population genetics of the coral *Acropora millepora* : Towards a genomic predictor of bleaching. *BioRxiv* , 867754.

Gruber, B., Unmack, P. J., Berry, O. F., & Georges, A. (2018). dartr: An r package to facilitate analysis of SNP data generated from reduced representation genome sequencing. *Molecular Ecology Resources* ,*18* (3), 691–699.

Grunwald, N. J., Kamvar, Z. N., & Everhart, S. E. (n.d.).*Discriminant analysis of principal components (DAPC)* .

Hadfield, J. D. (2010). MCMC methods for multi-response generalized linear mixed models: the MCMCglmm R package. *Journal of Statistical Software* , *33* (2), 1–22.

Hamada, M., Shoguchi, E., Shinzato, C., Kawashima, T., Miller, D. J., & Satoh, N. (2013). The complex NOD-like receptor repertoire of the coral*Acropora digitifera* includes novel domain combinations.*Molecular Biology and Evolution* , *30* (1), 167–176.

Hazzouri, K. M., Gros-Balthazard, M., Flowers, J. M., Copetti, D., Lemansour, A., Lebrun, M., . . . Fresquez, Z. A. (2019). Genome-wide association mapping of date palm fruit traits. *Nature Communications* , *10* (1), 1–14.

Helman, Y., Natale, F., Sherrell, R. M., LaVigne, M., Starovoytov, V., Gorbunov, M. Y., & Falkowski, P. G. (2008). Extracellular matrix production and calcium carbonate precipitation by coral cells in vitro.*Proceedings of the National Academy of Sciences* , *105* (1), 54–58.

Helyar, S. J., Hemmer-Hansen, J., Bekkevold, D., Taylor, M. I., Ogden, R., Limborg, M. T., . . . Carvalho, G. R. (2011). Application of SNPs for population genetics of nonmodel organisms: new opportunities and challenges. *Molecular Ecology Resources* , *11* , 123–136.

Heron, S. F., Eakin, C. M., Douvere, F., Anderson, K. D., Day, J., Geiger, E., . . . Obura, D. O. (2017). Impacts of Climate Change on World Heritage Coral Reefs : A First Global Scientifc Assessment.*Paris, UNESCO World Heritage Centre* .

Hill, W. G., Goddard, M. E., & Visscher, P. M. (2008). Data and theory point to mainly additive genetic variance for complex traits. *PLoS Genetics* , *4* (2).

Hogenboom, N. G. (1975). Incompatibility and incongruity: two different mechanisms for the non-functioning of intimate partner relationships.*Proceedings of the Royal Society of London. Series B. Biological Sciences* , *188* (1092), 361–375.

Holland, D., Frei, O., Desikan, R., Fan, C.-C., Shadrin, A. A., Smeland, O. B., . . . Dale, A. M. (2019). Beyond snp heritability: Polygenicity and discoverability of phenotypes estimated with a univariate gaussian mixture model. *BioRxiv* , 133132.

Hughes, T. P., Anderson, K. D., Connolly, S. R., Heron, S. F., Kerry, J. T., Lough, J. M., . . . Wilson, S. K. (2018). Spatial and temporal patterns of mass bleaching of corals in the Anthropocene.*Science* , *359* (6371), 80 LP – 83.

Hughes, T. P., Kerry, J. T., Baird, A. H., Connolly, S. R., Chase, T. J., Dietzel, A., . . . Woods, R. M. (2019). Global warming impairs stock–recruitment dynamics of corals. *Nature* . doi: 10.1038/s41586-019-1081-y

Iguchi, A., Yoshioka, Y., Forsman, Z. H., Knapp, I. S. S., Toonen, R. J., Hongo, Y., . . . Yasuda, N. (2019). RADseq population genomics confirms divergence across closely related species in blue coral (*Heliopora coerulea*). *BMC Evolutionary Biology* ,*19* (1), 187.

In, V.-V., O'Connor, W., Dove, M., & Knibb, W. (2016). Can genetic diversity be maintained across multiple mass selection lines of Sydney rock oyster, *Saccostrea glomerata* despite loss within each?*Aquaculture* , *454* , 210–216.

Jaccoud, D., Peng, K., Feinstein, D., & Kilian, A. (2001). Diversity arrays: a solid state technology for sequence information independent genotyping. *Nucleic Acids Research* , *29* (4), e25–e25.

Jakobson, C. M., & Jarosz, D. F. (2019). Molecular Origins of Complex Heritability in Natural Genotype-to-Phenotype Relationships. *Cell Systems* , *8* (5), 363–379.

Jin, Y. K., Lundgren, P., Lutz, A., Raina, J.-B., Howells, E. J., Paley, A. S., . . . van Oppen, M. J. H. (2016). Genetic markers for antioxidant capacity in a reef-building coral. *Science Advances* ,*2* (5), 1–7. Retrieved from http://advances.sciencemag.org/content/2/5/e1500842.abstract

Jombart, T. (2008). adegenet: a R package for the multivariate analysis of genetic markers. *Bioinformatics* , *24* (11), 1403–1405.

Jombart, T., & Collins, C. (2015). *Analysing genome-wide SNP data using adegenet 2.0. 0* .

Jombart, T., Devillard, S., & Balloux, F. (2010). Discriminant analysis of principal components: a new method for the analysis of genetically structured populations. *BMC Genetics* , *11* (1), 94.

Kenkel, C. D., Goodbody-Gringley, G., Caillaud, D., Davies, S. W., Bartels, E., & Matz, M. V. (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.

Kilian, A., Wenzl, P., Huttner, E., Carling, J., Xia, L., Blois, H., . . . Hopper, C. (2012). Diversity arrays technology: a generic genome profiling technology on open platforms. In *Data production and analysis in population genomics* (pp. 67–89). Springer.

Kronenberg, L., Yates, S., Boer, M. P., Kirchgessner, N., Walter, A., & Hund, A. (2019). Temperature response of wheat affects final height and the timing of key developmental stages under field conditions. *BioRxiv* , 756700.

Kruuk, L. E. B. (2004). Estimating genetic parameters in natural populations using the 'animal model.' *Philosophical Transactions of the Royal Society of London B: Biological Sciences* ,*359* (1446), 873–890.

Lawson, D. J., Van Dorp, L., & Falush, D. (2018). A tutorial on how not to over-interpret STRUCTURE and ADMIXTURE bar plots. *Nature Communications* , *9* (1), 3258.

Lemon, J. (2006). Plotrix: a package in the red light district of R. *R-News* , *6* (4), 8–12.

Louis, Y. D., Bhagooli, R., Kenkel, C. D., Baker, A. C., & Dyall, S. D. (2017). Gene expression biomarkers of heat stress in scleractinian corals: Promises and limitations. *Comparative Biochemistry and Physiology Part C: Toxicology & Pharmacology* , *191* , 63–77.

Lukoschek, V., Riginos, C., & van Oppen, M. J. H. (2016). Congruent patterns of connectivity can inform management for broadcast spawning corals on the Great Barrier Reef. *Molecular Ecology* .

Manolio, T. A., Collins, F. S., Cox, N. J., Goldstein, D. B., Hindorff, L. A., Hunter, D. J., ... Chakravarti, A. (2009). Finding the missing heritability of complex diseases. *Nature* ,*461* (7265), 747–753.

Matz, M. V, Treml, E. A., Aglyamova, G. V, & Bay, L. K. (2018). Potential and limits for rapid genetic adaptation to warming in a Great Barrier Reef coral. *PLoS Genetics* , *14* (4), e1007220.

Mizerek, T. L., Baird, A. H., & Madin, J. S. (2018). Species traits as indicators of coral bleaching. *Coral Reefs* , 1–10.

Morgans, C. A., Hung, J. Y., Bourne, D. G., & Quigley, K. M. (2019). Symbiodiniaceae probiotics for use in bleaching recovery. *Restoration Ecology* .

National Academies of Sciences and Medicine, E. (2019). *A research review of interventions to increase the persistence and resilience of coral reefs* . National Academies Press.

Nishizaki, S. S., & Boyle, A. P. (2017). Mining the unknown: assigning function to noncoding single nucleotide polymorphisms. *Trends in Genetics* , *33* (1), 34–45.

O'hara, R. B., Cano, J. M., Ovaskainen, O., Teplitsky, C., & Alho, J. S. (2008). Bayesian approaches in evolutionary quantitative genetics. *Journal of Evolutionary Biology* , *21* (4), 949–957.

Palmer, C. V, & Traylor-Knowles, N. (2012). Towards an integrated network of coral immune mechanisms. *Proceedings of the Royal Society B: Biological Sciences* , *279* (1745), 4106–4114.

Palumbi, S. R., Barshis, D. J., Traylor-Knowles, N., & Bay, R. A. (2014). Mechanisms of reef coral resistance to future climate change. *Science* , *344* (6186), 895–898.

Parker, L. M., Ross, P. M., O'connor, W. A., Borysko, L., Raftos, D. A., & Portner, H. (2012). Adult exposure influences offspring response to ocean acidification in oysters. *Global Change Biology* ,*18* (1), 82–92.

Peng, S., Wang, Y., Wang, L., Chen, W. U., Lu, C., Fang, L., & Chen, C. (2010). Proteomic analysis of symbiosome membranes in cnidaria-dinoflagellate endosymbiosis. *Proteomics* , *10* . doi: 10.1002/pmic.201000126

Polato, N. R., Voolstra, C. R., Schnetzer, J., DeSalvo, M. K., Randall, C. J., Szmant, A. M., ... Baums, I. B. (2010). Location-specific responses to thermal stress in larvae of the reef-building coral *Montastraea faveolata* . *PLoS One* , *5* (6), e11221.

Quigley, K., Willis, B., & Bay, L. (2017). Heritability of the *Symbiodinium* community in vertically-and horizontally-transmitting broadcast spawning corals. *Scientific Reports* , *7* , 8219.

Quigley, K M, Baker, A. C., Coffroth, M. A., Willis, B. L., & van Oppen, M. J. H. (2018). Bleaching Resistance and the Role of Algal Endosymbionts. In *Coral Bleaching* (pp. 111–151). Springer.

Quigley, K M, Randall, C. J., van Oppen, M. J. H., & Bay, L. K. (2020). Assessing the role of historical temperature regime and algal symbionts on the heat tolerance of coral juveniles. *Biology Open* .

Quigley, Kate M, Bay, L. K., & van Oppen, M. J. H. (2019). The active spread of adaptive variation for reef resilience. *Ecology and Evolution* , *Accepted* .

Rajora, O. P. (2019). *Population Genomics: Concepts, Approaches and Applications* . Springer.

Reef Restoraiton and Adaptation Program. Bay LK , Rocker M, Bostrom-Einarsson L, Babcock R, Buerger P, Cleves P, Harrison D, Negri A, Quigley K, Randall CJ, van Oppen MJH Webster N. (2019). A report provided to the Australian Government by the Reef Restoration and Adaptation Program.

Roitman, M., Gardner, M. G., New, T. R., Nguyen, T. T. T., Roycroft, E. J., Sunnucks, P., . . . Harrisson, K. A. (2017). Assessing the scope for genetic rescue of an endangered butterfly: the case of the Eltham copper. *Insect Conservation and Diversity* , *10* (5), 399–414.

Rudra, P., Shi, W. J., Vestal, B., Russell, P. H., Odell, A., Dowell, R. D., . . . Kechris, K. (2017). Model based heritability scores for high-throughput sequencing data. *BMC Bioinformatics* ,*18* (1), 143.

Sherwin, W. B., Chao, A., Jost, L., & Smouse, P. E. (2017). Information theory broadens the spectrum of molecular ecology and evolution.*Trends in Ecology & Evolution* , *32* (12), 948–963.

Shinzato, C., Shoguchi, E., Kawashima, T., Hamada, M., Hisata, K., Tanaka, M., . . . Ikuta, T. (2011). *Using the Acropora digitifera genome to understand coral responses to environmental change* .

Shlesinger, T., & Loya, Y. (2019). Breakdown in spawning synchrony: A silent threat to coral persistence. *Science* , *365* (6457), 1002–1007. doi: 10.1126/science.aax0110

Sogin, E. M., Putnam, H. M., Anderson, P. E., & Gates, R. D. (2016). Metabolomic signatures of increases in temperature and ocean acidification from the reef-building coral, *Pocillopora damicornis* . *Metabolomics* , *12* (4), 71.

Spaltmann, F., & Brummendorf, T. (1996). CEPU-1, a novel immunoglobulin superfamily molecule, is expressed by developing cerebellar Purkinje cells. *Journal of Neuroscience* , *16* (5), 1770–1779.

Stevenson, J., Jerry, D., & Owens, L. (2013). *Redclaw Selective Breeding Project* . Rural Industries Research and Development Corporation.

Sully, S., Burkepile, D. E., Donovan, M. K., Hodgson, G., & van Woesik, R. (2019). A global analysis of coral bleaching over the past two decades. *Nature Communications* , *10* (1), 1264. doi: 10.1038/s41467-019-09238-2

van Oppen, M. J. H., Palstra, F. P., Piquet, A. M.-T., & Miller, D. J. (2001). Patterns of coral–dinoflagellate associations in Acropora: significance of local availability and physiology of *Symbiodinium*strains and host–symbiont selectivity. *Proceedings of the Royal Society of London, Series B: Biological Sciences* , *268* (1478), 1759–1767. doi: 10.1098/rspb.2001.1733

van Oppen, Oliver, J., Putnam, H., & Gates, R. (2015). Building coral reef resilience through assisted evolution. *Proceedings of the National Academy of Sciences* , *112* (8), 2307–2313.

Visscher, P. M., Hill, W. G., & Wray, N. R. (2008). Heritability in the genomics era—concepts and misconceptions. *Nature Reviews Genetics* , *9* (4), 255–266.

Weeks, A. R., Sgro, C. M., Young, A. G., Frankham, R., Mitchell, N. J., Miller, K. A., . . . Hoffmann, A. A. (2011). Assessing the benefits and risks of translocations in changing environments: a genetic perspective. *Evolutionary Applications* , *4* (6), 709–725. doi: 10.1111/j.1752-4571.2011.00192.x

Westphal, A., Cheng, W., Yu, J., Grassl, G., Krautkramer, M., Holst, O., . . . Lee, K.-H. (2017). Lysosomal trafficking regulator Lyst links membrane trafficking to toll-like receptor–mediated inflammatory responses. *Journal of Experimental Medicine* , *214* (1), 227–244.

Wilson, A. J., Reale, D., Clements, M. N., Morrissey, M. M., Postma, E., Walling, C. A., . . . Nussey, D. H. (2010). An ecologist's guide to the animal model. *Journal of Animal Ecology* , *79* (1), 13–26.

Yang, S., Liu, Y., Jiang, N., Chen, J., Leach, L., Luo, Z., & Wang, M. (2014). Genome-wide eQTLs and heritability for gene expression traits in unrelated individuals. *BMC Genomics* , *15* (1), 13.

Yuyama, I., Watanabe, T., & Takei, Y. (2011). Profiling differential gene expression of symbiotic and aposymbiotic corals using a high coverage gene expression profiling (HiCEP) analysis. *Marine Biotechnology (New York, N.Y.)* , *13* (1), 32–40. doi: 10.1007/s10126-010-9265-3

Zhou, Z., Wu, Y., Zhang, C., Li, C., Chen, G., Yu, X., . . . Huang, B. (2017). Suppression of NF-ϰB signal pathway by NLRC3-like protein in stony coral Acropora aculeus under heat stress. *Fish & Shellfish Immunology* , *67* , 322–330.

Zoccola, D., Ganot, P., Bertucci, A., Caminiti-Segonds, N., Techer, N., Voolstra, C. R., . . . Casey, J. R. (2015). Bicarbonate transporters in corals point towards a key step in the evolution of cnidarian calcification. *Scientific Reports* , *5* , 9983.

Data Accessibility

Sequencing data is available through NCBI SRA accession X and X.

Author contributions

KMQ, LB, MvO designed the research; KMQ performed the lab work, analysed the data, wrote the manuscript; All authors contributed reagents and critically reviewed the manuscript.

Conflict of Interest Statement

The authors declare no competing or financial interests.

Table 1. Genomic diversity metrics

Cross	H_o	$^0\Delta_\alpha \pm \Sigma\Delta$	$^2\Delta_\alpha \pm \Sigma\Delta$	$^1H_\alpha \pm \Sigma\Delta$
WW1	0.158	1.29±0.45	1.20±0.34	0.17±0.27
WW2	0.219	1.30±0.46	1.22±0.36	0.18±0.28
WW3	0.189	1.28±0.45	1.20±0.36	0.16±0.27
WC	0.223	1.31±0.46	1.23±0.37	0.18±0.28
CW	0.244	1.30±0.46	1.22±0.37	0.17±0.28

H_o = Average observed heterozygosity

$^0D_\alpha \pm SD$ = Alpha within-population Diversity at q = 0

$^2D_\alpha \pm SD$ = Alpha within-population Diversity at q = 2

$^1H_\alpha \pm SD$ = Alpha within-population Entropy at q = 1

Table 2. Mean narrow-sense heritability (h^2) measurements for coral juveniles infected with *C. goreaui*.

	h^2 (mean ± 95% Bayesian Credibility Intervals)	h^2 (mean ± 95% Bayesian Credibility Intervals)
	27.5°C	31°C
Survival	0.0007 (0.00 - 0.88)	0.93 (0.73 - 0.98)

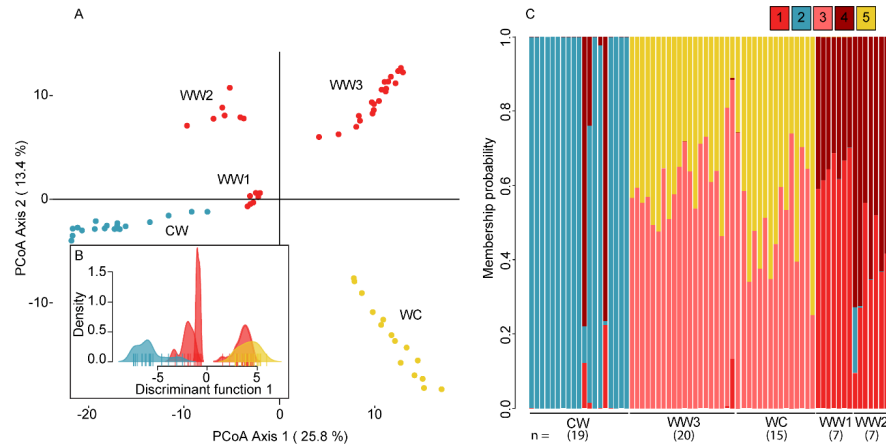|	h² (mean ± 95% Bayesian Credibility Intervals)	h² (mean ± 95% Bayesian Credibility Intervals)
Bleaching	0.13 (0.02 - 0.34)	0.15 (0.03 - 0.34)
Growth	0.01 (0.00- 0.76)	0.01 (0.00 - 0.67)

Figure 1. Genetic differential among population crosses. Single Nucleotide Polymorphism (SNP) cluster analysis for 68 individually genotyped *Acropora spathulata* larvae produced from five reproductive crosses (red: WW1, WW2, WW3), warm-cool (yellow: WC), cool-warm (blue: CW). **(A)** Principal Coordinates Analysis and **(B)** Discriminant Analysis of Principal Components (DAPC) of genetic structure across the SNP dataset. Point colors correspond to five genetic crosses. **(C)** Admixture inference using DAPC at k = 5. Each vertical bar represents an individually genotyped larva, with colors indicative of the membership probability for that population into each of the assigned genetic clusters.
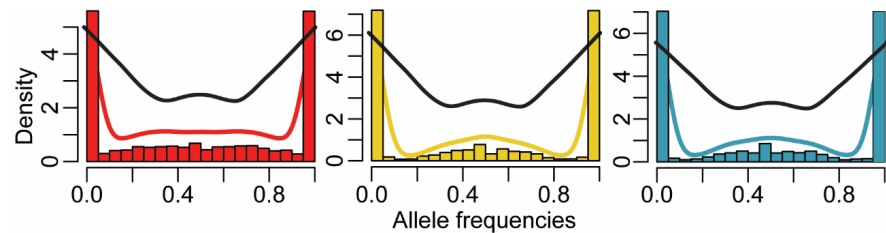

Figure 2. Selective breeding across populations creates signficanlty greater genetic variation across the coral genome. Distribution of allele frequencies for purebred and hybrid offspring groups consisting of warm-warm crosses (red: WW), warm-cool (yellow: WC yellow), cool-warm (blue: CW) for individually genotyped larvae reared at 27.5°C. Black lines correspond to modelled theoretically expected Hardy-Weinberg distributions of allele frequencies simulated for the three population sizes (n = WW: 34, WC: 15, CW: 19). Theoretical versus expected allele frequencies differed signficantly within population crosses (Kolmogorov-Smirnov Two-Sample test, WW: p = 1.0e-45, WC: 2.2e-46 and CW: 1.0e-45) and between purebred and hybrid crosses (WW versus WC: 2.1e-26, WW versus CW: 1.3e-21) but not between hybrids (WC versus CW: 0.34).
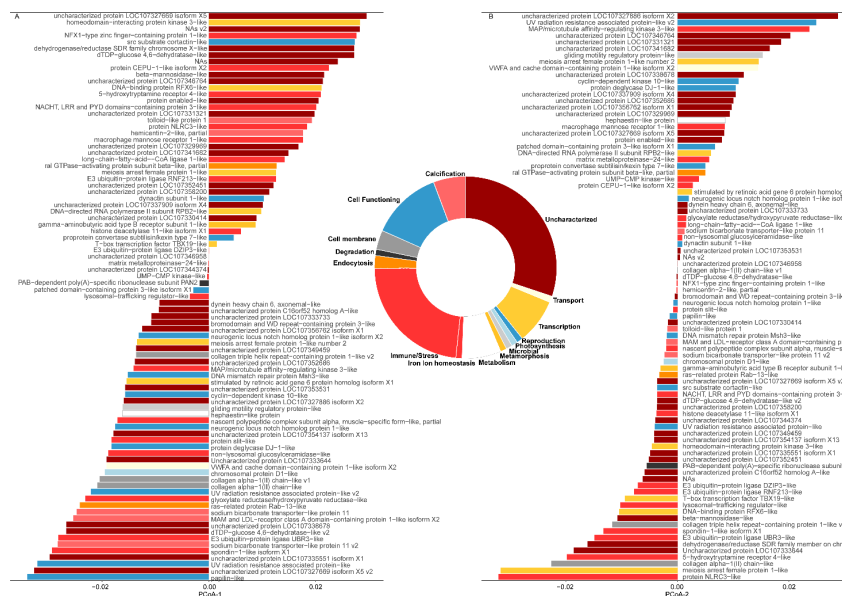
16

Figure 3. Variation in functional proteins as a result of selective breeding of corals. (A) Relative contribution of each protein to PCoA loadings (PC1 variation). Significantly different annotated proteins across the five reproductive crosses in PCoA-1 (A) and PCoA-2 (B). Proportion of proteins per functional category of proteins (inset). Colors correspond to the functional categories of proteins.
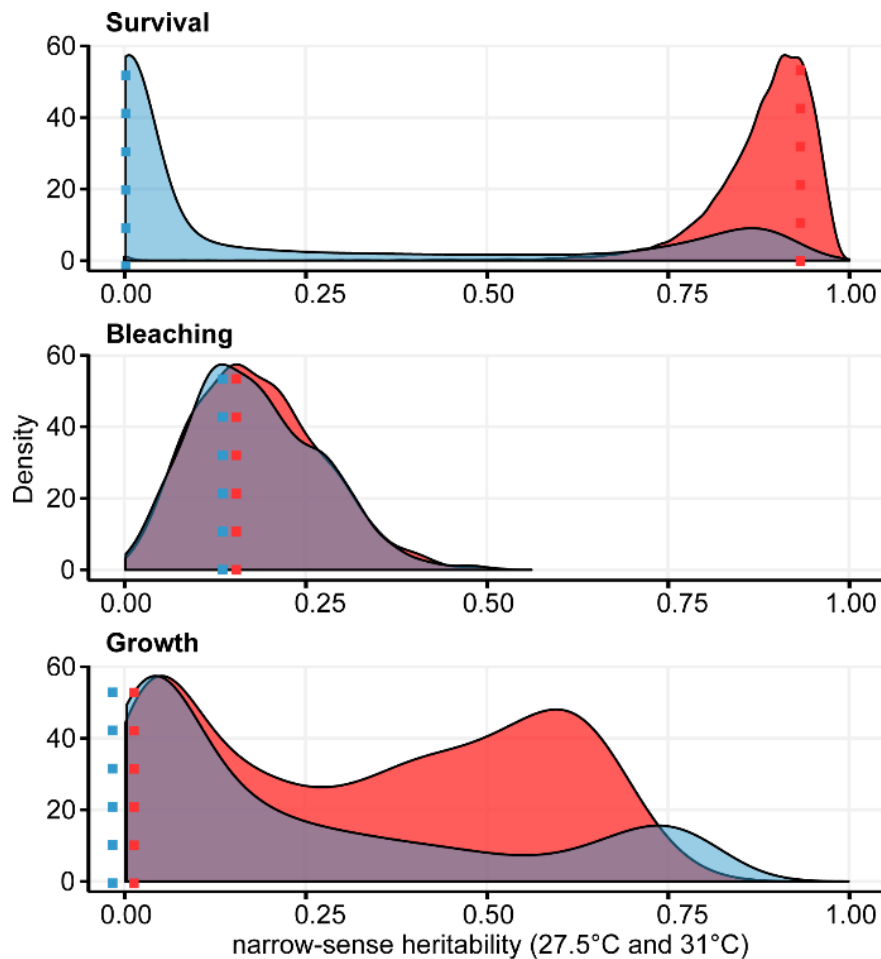
Figure 4. Narrow-sense heritability (h^2) of survival, bleaching and growth under ambient (blue: 27.5°C) and heat stress (red: 31°C) temperature conditions in juvenile corals infected with *Cladocopium goreaui* . Dashed lines correspond to mean h^2 of the distribution.

Hosted file

Quigleyetal._SNPs_AGF2017_MErefs.docx available at https://authorea.com/users/320037/articles/449661-genome-wide-snp-analysis-reveals-an-increase-in-adaptive-genetic-variation-through-selective-breeding-of-coral