Genome-wide survey of single-nucleotide polymorphisms reveals fine-scale population structure and signs of selection in the threatened Caribbean elkhorn coral, *Acropora palmata*

Meghann D Durante[1], Iliana B Baums[1]

[1]Department of Biology, Pennsylvania State University, University Park, PA, USA

Corresponding author:
Iliana B. Baums

Email address: baums@psu.edu

Abstract

The advent of next-generation sequencing tools has made it possible to conduct fine-scale surveys of population differentiation and genome-wide scans for signatures of selection in non-model organisms. Such surveys are of particular importance in sharply declining coral species, since knowledge of population boundaries and signs of local adaptation can inform restoration and conservation efforts. Here, we use genome-wide surveys of single-nucleotide polymorphisms in the threatened Caribbean elkhorn coral, *Acropora palmata*, to reveal fine-scale population structure and infer the major barrier to gene flow that separates the eastern and western Caribbean populations between the Bahamas and Puerto Rico. The exact location of this break had been subject to discussion because two previous studies based on microsatellite data had come to differing conclusions. We investigate this contradiction by analyzing an extended set of 11 microsatellite markers including the five previously employed and discovered that one of the original microsatellite loci is apparently under selection. Exclusion of this locus reconciles the results from the SNP and the microsatellite datasets. Scans for outlier loci in the SNP data detected 13 candidate loci under positive selection, however there was no correlation between available environmental parameters and genetic distance. Together, these results suggest that reef restoration efforts should use local sources and utilize existing functional variation among geographic regions in *ex situ* crossing experiments to improve stress resistance of this species.

Introduction

There is an ongoing debate about the importance of local recruitment and barriers to gene flow in marine species. Many marine species reproduce via planktonic larvae and strong ocean currents have the potential to carry propagules over long distances. However, a high degree of self-recruitment has been found in a range of species with planktonic larval duration being a poor predictor of genetic structure (Selkoe & Toonen 2011). The development of cheap genome-scale genotyping is poised to open a new chapter in this discussion (Peterson et al. 2012; Toonen et al. 2013; Wang et al. 2012). American eels for example show panmixia in their central breeding ground in the North Atlantic but single nucleotide polymorphism (SNP) genotyping of adults along the Eastern seaboard revealed local differentiation (Gagnaire et al. 2012). Thus, a well-mixed pool of larvae sorted into environmental niches and so resulted in a structured adult population.

SNPs are ubiquitous throughout the genome, located in coding and non-coding regions, and each locus has a maximum of four alleles (the four bases). This is in contrast to microsatellite loci that consist of tandem repeats, in which allelic variation is determined by the number of tandem repeats and thus can be large. The limited number of alleles at each SNP locus requires a larger number of loci to be assayed to achieve the same power of detecting population genetic structure as a panel of microsatellite loci (Morin et al. 2009; Ryman et al. 2006). The advent of reduced representation sequencing methods have made it possible to develop and assay a large number of SNP loci at a reasonable cost (Altshuler et al. 2000; Hoffberg et al. 2016). Recently, Genotyping by Sequencing (GBS) data including 4,764 SNPs in *A cervicornis* identified population structure within the Florida Reef tract (Willing et al. 2012) where microsatellite markers did not (Baums et al. 2010). Other flavors of reduced representation sequencing methods (Drury et al. 2016; Toonen et al. 2013; Wang et al. 2012) have yielded information on population structure, and genetic diversity in reef building corals (Drury et al. 2016; Howells et al. 2016a).

Genome-scale genotyping can provide insights into genetic diversity within functional regions of the genome that may be under selection (those genomic regions that code for proteins or regulate transcription of genes). These regions are not commonly surveyed even though they are of interest to conservation managers who want to understand how much capacity there is in a species to adapt to changing conditions (Becks et al. 2010). Statistical methods have been

developed that allow scanning of SNP loci for signatures of selection. Despite the risk of generating false positive results (Vilas et al. 2012), these methods yield candidate loci that should be substantiated by further testing (Renaut et al. 2011; Sork et al. 2016). The same methods can be used to scan microsatellite loci for signatures of selection (Nielsen et al. 2006; Vasemägi et al. 2005), however, power is often limited by the small number of assayed loci.

Acropora palmata is one of a few Caribbean coral species whose population genetic structure has been thoroughly investigated on local and range-wide scales (Baums et al. 2014b; Baums et al. 2005b; Baums et al. 2006a). A range-wide survey of *A. palmata* population genetic structure using five coral specific polymorphic microsatellite markers showed that *A. palmata* stands are structured into two long-separated populations (Baums et al. 2005a). While most reefs are self-recruiting, *A. palmata* stands are not inbred and harbor high genetic diversity at these microsatellite loci (Baums et al. 2005b). Bio-physical modeling identified a transient feature in the Mona Passage important in restricting present-day gene flow between the eastern and western population (Baums et al. 2006b). However, it is unclear whether the eastern and western populations differentiated initially due to selection. Subsequent denser sampling of *A. palmata* along the Antilles Island Arc raised the possibility of a hybrid zone across Puerto Rico rather than a clear-cut break between the eastern and western Caribbean at the Mona Passage (Fig 1, Mège et al. 2014).

We sought to refine the location of the east-west population divide and test for the presence of finer scale population differentiation in *A. palmata* by developing a large number of SNP markers. We assayed genome-wide SNPs in archived samples from two geographic regions in the western *A. palmata* population (Bahamas and Florida) and two geographic regions in the eastern population [Puerto Rico and the U.S. Virgin Islands (USVI)]. We then compared the results to population structure derived from eleven microsatellite loci. We further aimed to produce a more comprehensive estimate of genetic diversity across the genome using SNPs and screened loci for signatures of selection.

Materials & Methods

SAMPLE COLLECTION

Colonies of *A. palmata* were collected between 2002 and 2010 and previously genotyped (Baums et al. 2014b; Baums et al. 2005b). Unique genets were selected from our database for a

total of 24 samples from each of four geographic regions; the Bahamas, Florida, Puerto Rico and the US Virgin Islands (USVI). The goal was to have eight samples from three different reefs within each geographic region, however this was not always possible either due to small sample sizes from a particular reef or low clonal diversity of a reef. In those cases, we selected additional unique genets from nearby reefs. See Table 1A for detailed sample information.

We used an extended set of samples to compare the population genetic structure ascertained via microsatellite genotyping to the SNP results. This extended set of samples included 260 samples from six geographic regions; Belize, Florida, Puerto Rico, the USVI, and Curacao (Table 1B). Note that not all SNP-genotyped samples were included in the microsatellite dataset.

LIBRARY PREPARATION

Coral tissue samples were extracted from ethanol preserved samples using DNeasy Blood & Tissue Kit (QIAGEN, Hilden, Germany) with the following modifications. Time of incubation in the extraction buffer was increased to 16-20 hours and two 100 µl elutions were performed, the second of which was kept for library production as this fraction contained the high molecular weight DNA. Extracted DNA was then treated with 0.01 mg of RNase A (10 mg/ml, Amresco Solon, OH). Extraction concentrations ranging from 500 ng to 6 µg were double-digested with 10 units of each of the restriction enzymes MluCI (^AATT) and NlaIII (CATG^) (New England Biolabs, Ipswich, MA) following the protocol described by Peterson et al. (2012). Digestions were purified using 1.5X Ampure beads (Beckman Coulter Inc, Brea, CA) and quantified on a Qubit® fluorometer (Life Technologies, Carlsbad, CA). Digested DNA was standardized to 100 ng for each sample before adaptor ligation. Samples were identified with eight 6-bp indices on the NlaIII (rare-cutter) P1 adapter (Supplemental Table 1). Samples were pooled into 12 libraries and then size selected in the range of 200-800 bp on a Pippin-Prep (Sage Science, Beverly, MA). Next, Illumina flow-cell annealing sequences, unique multiplexing indices and sequencing primer annealing regions were added through PCR amplification to the MluCL cut end (See Peterson et al. 2012, Protocol S1, Fig 1). The libraries were enriched with 12 amplification cycles in four separate PCR reactions for each library containing 10 µl of Phusion High-Fidelity PCR Master Mix with HF Buffer (New England Biolabs, Ipswich, MA), 2 µl of each amplification primer, 1 µl of library DNA and 5 µl of water (total 20 µl). Samples were

pooled into four libraries each containing 24 samples (Table 2, Supplemental Table 1). Each library was sequenced on one lane of Illumina HiSeq 2000 sequencer (paired-end, 2x150 bp) at the Pennsylvania State Genomics Core Facility. There were two libraries sequenced on each chip. See Peterson et al. (2012) Supplemental File 1 for a detailed protocol. Radseq methods have been used successfully in scleractinian corals (Combosch & Vollmer 2015; Dimond et al. 2017; Forsman et al. 2017) and other marine invertebrates (Lal et al. 2016; Reitzel et al. 2013).

RAW SEQUENCE FILTERING

Raw sequence reads were filtered using the process_radtags in the pipeline STACKS 1.21 (Catchen et al. 2013; Catchen et al. 2011). Barcodes and the RAD-Tag cut sites were identified to de-multiplex the pooled data into individual samples (Suppl Table 2). Reads were discarded that had low quality (with an average raw phred score <10 within a 15-base pair sliding-window), adapter contamination, and uncalled bases. Since all indices differed by at least 2 bp, it was possible to correct and retain any index that differed by a single bp from an expected index.

ASSEMBLY

Processed sequences were then aligned to the *Acropora digitifera* genome (V1.0) (Shinzato et al. 2011) with BOWTIE2 (Langmead & Salzberg 2012) within the GALAXY (Bedoya-Reina et al. 2013; Blankenberg et al. 2014) framework using end-end read alignment settings in order to remove symbiont and other associated microorganisms. After alignment, paired-end sequencing BAM files were assembled in the ref_map.pl pipeline in STACKS 1.30 with the following parameters. Each paired-end sequencing set was run separately through STACKS to compare results (designated Read1 and Read2) in a one-way ANOVA. The ANOVA used each paired-end read as a technical replicate of the same genomic region. We did this to assess whether we would retrieve similar estimates of F_{IS} and heterozygosity from both reads, as expected.

The number of raw reads required to report a stack was m=5. The number of mismatches allowed between loci when building the catalog was n=4. SNPs with a log-likelihood of less than -10 were removed as reads with poor log-likelihoods tend to have sequencing error and/or low coverage. Two of the barcodes (TCGAT and CGATC) had few sequence reads across all four geographic regions with all Illumina lanes being affected, and samples with these barcodes were removed before assembly in STACKS.

Sequencing reads are available under NCBI BioProject ID PRJNA407327.

GENOME COVERAGE

BEDTOOLS (Quinlan & Hall 2010) was used to create a histogram of genome coverage for each sample from the BOWTIE2 BAM format alignment files. All positions with a depth of coverage greater to or equal to 20 were combined into a single bin in the histogram. Data from all geographic regions were averaged (excluding samples with barcodes TCGAT and CGATC) and a cumulative distribution of sequencing coverage was then plotted in SIGMAPLOT v12.

POPULATION GENETIC STATISTICS

We explored values for several parameters relevant to population genetic analyses. In the Populations module in STACKS 1.30 we required a locus to be present in all regions for all analyses (option –p = 4). For each locus, we then set the minimum percentage of individuals in a region required to have data for that locus to 40% or 60% (option –r). Further, we set the minimum minor allele frequency (MAF) required to process a nucleotide site at a locus (option –min_maf) to 0.025, 0.05 and 0.075. A *p*-value correction was applied to F_{ST} scores, so that if a F_{ST} score was not significantly different from 0 (according to Fisher's Exact Test) the value was set to 0. Additionally, only one random SNP from any RAD locus was written to the STRUCTURE export file in order to prevent linked loci from being processed. Read 1 and Read 2 STRUCTURE export files were combined and duplicate loci removed randomly between reads. F_{ST} (*p*-value<0.05) was calculated in STACKS. F_{IS} and F_{ST} distributions are included in the supplement (Supplemental Figure 1, 2).

CLUSTERING ANALYSES

Clustering analyses for the SNP and microsatellite analysis were performed in the program STRUCTURE 2.3.4 (Falush et al. 2003; Hubisz et al. 2009) using the admixture model with correlated allele frequencies. The analysis included the following parameters: 100,000 burn-in iterations and 1,000,000 Markov chain Monte Carlo repetitions, with and without a population prior, for a total of three replicates for each value of K. K values ranged from 2 to 5. The most likely value for K was determined by CLUMPAK (Kopelman et al. 2015) BEST K which uses either the Evanno method (Evanno et al. 2005) or LN(PR(X|K) values to identify the K for which PR(K=K) is the highest as described in STRUCTURE's manual section 5.1. Results of the three structure runs were merged with CLUMPAK (Kopelman et al. 2015). Based on our exploration of

minor allele frequency (MAF) cut off values and the percent of individuals per geographic region allowed to miss a locus (%M), we report results for MAF = 0.05 and for a %ML = 60 % in the main text (Fig 2). STRUCTURE clustering analyses for minor allele frequencies cutoffs of 0.025 and 0.075 are included in the supplement (Supplemental Figure 3). STRUCTURE clustering analysis when the minor allele frequency cutoff was 0.05 and when outlier loci were removed, is also included in the supplement (Supplemental Figure 4). STRUCTURE clustering analysis when the minor allele frequency cutoff was 0.05 and when a locus must be present in at least 40% of individuals in a geographic region, is included in the supplement (Supplemental Figure 5). PCA clustering analysis, for SNPs and microsatellites, using adegenet (Jombart 2008) is included in the supplement (Supplemental Figure 6).

Previously genotyped samples (n=260) at 10 and 11 microsatellite markers (181, 182, 192, 207, 0585, 0513, 2637, 007, 9253, 5047, with and without locus 166) (Baums et al. 2009; Baums et al. 2005a) were also analyzed with STRUCTURE 2.3.4 (Falush et al. 2003; Hubisz et al. 2009) using the admixture model with correlated allele frequencies (See Table 1 for sample information). The analysis included the following parameters; 100,000 burn-in iterations and 1,000,000 MCMC repetitions, with and without a population prior, for a total of 3 replicates for each value of K. K values tested ranged from 2 to 7.

MANTEL TESTS

Data on temperature, salinity, dissolved oxygen (ml/l), and phosphates was downloaded from the World Ocean Atlas 2013 (WOA13 V2, Suppl Table 3). Silicates and nitrates were not used as there was not sufficient data for all locations. For the Bahamas, Puerto Rico, and the USVI the geographic center point among several sampling sites was used because reefs were further apart than in Florida. For all data, the statistical mean of the annual average of years 1955-2012 and depths of 0-10 m was used. Grid sizes were 1/4° for temperature and salinity, and 1° for dissolved oxygen (ml/l), and phosphates (µmol/l) (Supplemental Table 2). SPSS V22 was used to calculate a dissimilarity matrix expressed as the Euclidean distances between geographic regions based on the above environmental data. To obtain a single GPS location for each geographic region we had to average the latitude and longitude coordinates for all samples in each region. Then GenAlEx v6.501 (Peakall & Smouse 2006) was used to calculate a pairwise geographic distance matrix between the four geographic regions. GenAlEx v6.501 (Peakall &

Smouse 2006) was used to calculate Mantel multi-comparison tests between the geographic distance matrix, F_{ST} pairwise matrix between geographic regions from STACKS, and the environmental dissimilarity matrix.

OUTLIER ANALYSIS

Two independent methods were applied to identify putative loci under selection. The first program used was LOSITAN (Antao et al. 2008) which utilizes the method of Beaumont and Nichols (1996a) to identify loci under selection based on the joint distributions of expected heterozygosity and F_{ST} under an island model of migration. The following settings were used for the SNP and the microsatellite datasets. The neutral mean setting was selected in which during an initial run (100,000 simulations), a candidate subset of selected loci (outside the 95 % confidence interval) were identified and removed. Then the distribution of neutral F_{ST} was computed using 100,000 simulations and a bisection approximation algorithm (Antao et al. 2008), with the following options, force mean F_{ST}, infinite alleles mutation model, and a confidence interval 0.99. A FDR < 0.1 correction for multiple testing was applied. Loci outside the upper and lower confidence areas were identified as candidates affected by positive and balancing selection, respectively (Supplemental Table 4). All geographic regions were analyzed together. Outliers identified as being under balancing selection were not considered as these are more likely to be false positives (Lotterhos & Whitlock 2014). The positive outlier loci ($p<0.01$) were blasted against the NCBI nr, UniProt, and Trembl databases with parameters of expected value = 0.00001, gap opening penalty = 11, gap extension penalty = 1, length of initial exact match (word size) = 6 and scoring matrix = BLOSUM62 using BLASTX 2.2.32+ (Altschul et al. 1997).

The STACKS exported GENEPOP dataset was also reformatted with PGDSPIDER version 2.0.5.2 (Lischer & Excoffier 2012) to a GESTE file. The method of Foll and Gaggiotti (2008) was performed using BAYESCAN 2.0 (http://www-leca.ujf-grenoble.fr/logiciels.html). For each locus, the probability of it being under selection was inferred using the Bayes factor (BF). Based on Jeffreys' (1961) scale of evidence, a log10 BF of 1.5–2.0 is interpreted as "strong evidence" for departure from neutrality at that locus and corresponds to a posterior probability between 0.97-0.99. For our analysis, the estimation of model parameters was set as 20 pilot runs of 5,000 iterations each, followed by 50,000 iterations.

Results

Summary statistics

Illumina sequencing of the RAD libraries generated 49.3 million reads per pool of eight samples, averaging 6.2 million 150 bp reads per sample prior to quality filtering. After quality filtering, 4.99 million reads per sample (81%) were retained on average (Table 2). Pools had similar numbers of reads after processing (mean = 39.9 million per pool, SD = 4.95 million, one-way ANOVA, $F = 2.638$, $p > 0.1$). The average % GC content for Read 1 and 2 was 41.7 and 39.6, respectively. The percentage of polymorphic sites per genomic region varied little among geographic regions, from 0.150 to 0.173 % (Table 3). The average observed heterozygosity in variant sites was 22%. Overall F_{IS} values, when considering all sites with a minor allele frequency cutoff a ≥ 0.05, were close to 0 and hence provided no evidence of inbreeding (Table 3). However, when only considering variant positions within the region of Florida, F_{IS} values were negative ($F_{is} = -0.0086$), indicating an excess of heterozygosity. Using the two paired-end read sets as replicates, a one-way ANOVA was performed for each variable (Table 4). Summary statistics for all geographic regions were found to be similar. Alignment of *A. palmata* SNPs to the published *A. digitifera* genome indicated that on average, 2.5% percent of the *A. digitifera* genome had sequence coverage at a stack depth of 5 (Supplemental Figure 7). All four geographic regions produced similar sequence coverage.

Population genetics

A total of 390 SNPS were identified after filtering and including a minor allele frequency cutoff a ≥ 0.05 (Table 3). This included 219 for Read 1 and 176 for Read 2 from the paired-end sequencing (5 SNPs were identical between reads and only considered once). Analysis of Molecular Variance (AMOVA) revealed patterns of genetic differentiation among geographic regions (Table 4). This was also evident when the 307 SNPs (analysis included only one SNP per 150 bp locus) were subjected to a multi-locus clustering analysis in STRUCTURE. Samples from Florida clustered first, followed by the Bahamas at K=3. Puerto Rico and the USVI were not distinguishable until K=4, (Fig 2). CLUMPAK BEST K (Kopelman et al. 2015) indicated that K=3 was the most likely K-value, after both the Evanno method and LN(PR(X|K) values, regardless of whether the geographic region was used as a prior.

To compare to the SNP analysis, microsatellite data from samples collected in six regions were analyzed in STRUCTURE using the geographic region as a prior (Supplemental Table 5). At K=2, a western (including Belize, Florida, Bahamas and Puerto Rico) and an eastern cluster (including the USVI and Curacao) was evident (Fig 3A). At K=3, an isolation-by-distance like pattern was apparent in the western cluster (Fig 3B). K=4 was the most likely K-value, after both the Evanno method and LN(PR(X|K) values, based on 11 microsatellite markers (Kopelman et al. 2015). Florida and Belize grouped as one cluster, and Puerto Rico and the Bahamas as the second, with the USVI as the third and Curacao as an admixed fourth cluster (Fig 3C).

According to the outlier analysis in LOSITAN, microsatellite locus 166 was identified as a potential outlier and thus possibly under selection. It was therefore excluded from the analysis in STRUCTURE. This resulted in more comparable results to the SNP analysis with the most likely K-value being 3, after both the Evanno method and LN(PR(X|K) values (Kopelman et al. 2015). Again, the first separation was between a western and an eastern cluster, however this time Puerto Rico assigned to the eastern cluster with an isolation-by-distance like pattern appearing between the west and east (Fig 3D). At the most likely K of 3, Curacao now formed a separate cluster. At K=4, the Bahamas started to separate from the remainder of the western region similar to what was observed in the SNP clustering analysis (Fig 3E).

Environmental drivers of population structure

A Mantel test showed a significant positive relationship in the SNP dataset between pairwise F_{ST} values and geographic distance (R^2= 0.65, p=0.05) consistent with the microsatellite results (10 loci) from the Florida, Bahamas, Puerto Rico, and Curacao samples only (Fig4C, Fig4D). Correlations between environmental factors including average temperature, salinity, dissolved oxygen, and pairwise F_{ST} values or geographic distance were not significant (Fig4A, Fig4B).

Loci under selection

BAYESCAN and LOSITAN identified 2 and 12 SNPs (Supplemental Table 4) that showed signs of positive selection when including all four geographic regions, one of which was identified by both programs (a total of 13 unique loci identified between both programs). Outliers accounted for 3.3% of the total SNPs, consistent with other studies in which F_{ST} outlier loci have represented a substantial fraction of the total loci investigated (2-10%) (Nosil et al.

2009). Annotation of the candidate loci proved difficult as only 23% produced significant hits when queried against the NCBI NR database, Uniprot, and Trembl; with two of the hits being annotated as unconventional myosin-IXb isoform X7 and tyrosine-protein kinase transmembrane receptor ROR1-like. Screening of the microsatellite loci identified locus 166 as an outlier under positive selection, yet no annotation information of this locus is currently available.

Discussion

Comparison with previous *Acropora* gene flow studies

The previous range-wide survey of *A. palmata* population genetic structure using five, presumed neutrally evolving microsatellite markers showed that while most reefs are self-recruiting, *A. palmata* stands are not inbred and harbor high microsatellite genetic diversity (Baums et al. 2005b). Furthermore, *A. palmata* stands were structured into two long-separated populations, one in the eastern and one in the western Caribbean (Baums et al. 2005b). Here, we report that genome-wide SNPs (MAF $\geq$ 0.05) resolved further population structure in the endangered reef-building coral, *A. palmata* from Florida to the USVI compared to previous microsatellite-based analyses.

It was recently suggested that the East-West divide of *A. palmata* lies not in the Mona Passage (Baums et al. 2005b; Baums et al. 2006b) but rather to the east of Puerto Rico (Fig 1, Mège et al. 2014). The 307 SNPs analyzed here confirm earlier findings that Puerto Rico and the USVI regions are more similar to each other than Puerto Rico is to either the Bahamas or Florida without imposing any priors in a STRUCTURE analysis (MAF$\geq$0.05). However, it is not always possible to determine, with confidence, the correct clustering solution that accurately reflects genetic population structure when there is an underlying isolation by distance pattern (Frantz et al. 2009). We show here that there is significant isolation by geographic distance from Florida to the USVI when using presumably neutrally evolving SNP and microsatellite loci. Interestingly, inclusion of microsatellite locus 166, flagged as being an outlier locus, obscured this isolation by distance pattern (Supplemental Figure 8). Therefore, locus 166 is a strong candidate for a locus under selection (or it is linked to a locus under selection) and its functional significance might prove a fruitful subject for future studies (Nielsen et al. 2006).

An east-west Caribbean divide was also evident in the corals *Orbicella annularis* (Foster et al. 2012) and *Acropora cervicornis* (Vollmer & Palumbi 2007). An additional barrier to gene flow in *A. palmata* was reported by Porto-Hannes et al. (2014) between Venezuela and the Mesoamerican Barrier Reef System utilizing four of the microsatellites markers.

The total number of SNPs (n=307) retained for population genetic analysis was lower than expected. This was due to more than a 10-fold increase in the number of fragments retrieved from the genome digest using the enzymes MluCI (^AATT) and NlaIII (CATG^) compared to what was predicted from an in-silico restriction of an incomplete draft genome of *A. palmata* (Baums, unpublished). The in-silico restriction predicted 19,067 to the actual 322,425 (read 1) and 276,753 (read 2) fragments retrieved. This under-prediction was most likely due to an early, incomplete genome draft and unknown genome size at the time of this study (Herrera et al. 2015). A larger set of SNP loci may reveal additional finer scale structure in *A. palmata* across the Caribbean. However, this may not necessarily be the case. In a study that used three orders of magnitude more loci (905,561 SNPs) failed to reveal population structure in *A. digitifera* collected from the Ryukyu Archipelago of Japan using Bayesian clustering based methods (Shinzato et al. 2015). Low coverage, 5X in this study, is also a concern however this depth of coverage has been used in other non-model species (Babbucci et al. 2016; Blanco-Bercial & Bucklin 2016; Laporte et al. 2016). Yet, in the coral *Platygyra daedalea,* 5x coverage was sufficient to assign samples to two distinct clusters based on their geographic origin, the Persian Gulf or Sea of Oman and was consistent with their 20x coverage data set (Howells et al. 2016b).

A Mantel test showed a significant positive relationship between the SNP-derived pairwise F_{ST} values and geographic distance ($r^2 = 0.65$, $p = 0.05$) consistent with the microsatellite results (10 loci) from the Florida, Bahamas, Puerto Rico, and Curacao samples (Fig4C, Fig4D). This may be due to Wright's Isolation-By-Distance (IBD) process however Mantel tests are prone to false positives as the test assumes spatial independence of the data (Meirmans 2012). Nevertheless, genetic variability is structured in geographic space.

Correlations between environmental factors including average temperature, salinity, dissolved oxygen, and pairwise F_{ST} values or geographic distance were not significant (Fig4A, Fig4B). It should be noted that the environmental data had a resolution of ¼ to 1 degree latitude, an equivalent of about 28-111 km, whereas the genetic data was collected on much smaller

spatial scales. For example, in Florida, sampled reefs were often less than 10 km apart, and the distance between Sand Island Reef and French Reef is only 2.6 km. (Supplemental Table 2). Here, reefs often harbor just one or a few *A. palmata* genets (albeit represented by many colonies) making it challenging to obtain the needed >25 genets per population recommended for F_{ST} analyses on a scale of a few km. Thus, genets were pooled over geographic regions to match the scale of the environmental data and yield sample sizes of at least 25 per location. Yet, significant micro-environmental differences among colonies growing on the same reef have been documented (Drury et al. 2017; Gorospe & Karl 2010). Therefore, landscape genetic approaches that may reveal environmental drivers of population differentiation (Manel et al. 2003) must await higher resolution environmental data and, perhaps , a greater number of SNP loci.

Genetic diversity indices in *A. palmata*

Several factors could account for negative F_{IS} values including negative assortative mating, if a species is outcrossed and lacks selfed progeny or there is a selection pressure that favors the most heterozygous genets. Of our samples, 49 out of 96 were ramets of larger genets. *A. palmata* colonies fragment frequently; the branches regrow into new colonies resulting in stands of genetically identical colonies (Baums et al. 2006a). [Note that samples included here all represented distinct genets]. Asexual reproduction could explain the excess of heterozygosity in *A. palmata* within the Florida region (see Balloux et al. 2003; Carlon 1999; Delmotte et al. 2002). Excess hetereozygosity has been observed in other clonal organisms. For example, significant negative F_{IS} values in a partially clonal but self-incompatible wild cherry tree was explained in part by asexual reproduction (Stoeckel et al. 2006).

Nucleotide diversity is a measure of a species' genetic diversity and varies predictably with life history (Hamrick & Godt 1996; Romiguier et al. 2014). Because *A. palmata* populations experienced dramatic losses in the 1980s and therefore may now have reduced genetic diversity we compared *A. palmata*'s nucleotide diversity to the diversity found in other species. The nucleotide diversity π, describes the degree of nucleotide polymorphism in populations and can be calculated based on variant sites only or on variant and non-variant sites combined. In acroporids, estimates range from 0.007-0.022 (Macdonald et al. 2011) in *A. austere* to 0.09 in *A. cervicornis* (Drury et al. 2016). In other Cnidaria, estimates range from 0.00403 in *Aiptasia* (Bellis et al. 2016) to 0.0065 in *Nematostella* (Putnam et al. 2007). Synonymous nucleotide diversity ranged from 0.012 –0.020 in transcriptomes from three gorgonian species (Romiguier

et al. 2014). Average pairwise nucleotide diversity in other metazoans include *Drosophila pseudoobscura* (0.0024-0.0179, Kulathinal et al. 2009) and *Homo sapiens* (0.000751, Sachidanandam et al. 2001). Our estimates of nucleotide diversity (including variant and non-variant sites) was 0.0004 for all geographic regions, an order of magnitude lower than in other cnidarians. Further, based on a survey of 374 individual transcriptome-derived SNPs from 76 non-model animal species, the level of nucleotide diversity found in *A. palmata* is well below that predicted for a long-lived species, with small propagule size and large adult size (Romiguier et al. 2014). This low nucleotide diversity could be due to either a relatively small long-term effective population size, a severe bottleneck associated with a selective sweep (Ellegren & Galtier 2016), the small number of SNPs included in this study (Fischer et al. 2017) or the RAD-tag method (Arnold et al. 2013). In addition, we find that Florida is the least genetically diverse geographic region when comparing nucleotide diversity in variant sites only (0.203, Table 3), as would be expected in a marginal environment (Baums 2008; Baums et al. 2014a; Cahill & Levinton 2016; Eckert et al. 2008). This is in contrast to Drury et. al, which found samples of the congener *A. cervicornis* from Florida to be higher in SNP nucleotide diversity than those from the Dominican Republic (Drury et al. 2016). Increased sampling of the genome as well as analysis of historical samples may shed light on whether the low nucleotide diversity in *A. palmata* is due to technical issues, the recent population bottleneck or unrelated causes.

Allelic richness of microsatellite data correlates better with genome-wide estimates of genetic diversity based on SNPs than heterozygosity (Fischer et al. 2017) and allelic richness is more sensitive to recent population bottlenecks than heterozygosity (Allendorf 1986). Average microsatellite-based allelic richness in 14 Indo-Pacific *Acropora* corals was 4.96 overall and 6.21 in the five geographically widespread species (calculated based on Table 6 in Richards & Oppen 2012) which compares favorably with an average allelic richness of 8.49 in *A. palmata* found here. Thus, allelic richness of microsatellite loci remains high in Caribbean *A. palmata* despite recent population declines and the documented loss of alleles in Florida (Williams et al. 2014).

To resolve the contradictory findings with respect to genetic diversity based on micorsatellites and SNPs, future studies should include several thousand SNPs assayed in samples from across the species range. This approach may provide more conclusive data on the impact of recent population declines on overall genetic diversity in *A. palmata*.

Genes under positive selection

Thirteen loci out of 395 were identified as being under positive selection in *Acropora palmata*. Detecting regions of the genome under selection is difficult, and statistical detection methods are prone to different rates of type 1 and type 2 errors. Further, LOSITAN and BAYESCAN often identify different loci as being under selection (Narum & Hess 2011). LOSITAN identifies outliers based on the joint distributions of F_{ST} and expected heterozygosity under an island model of migration (Beaumont & Nichols 1996b). Whereas, BAYESCAN uses a hierarchical Bayesian method of Foll and Gaggiotti (2008), which has been modified based on the approach proposed by Beaumont and Balding (2004). Lotterhos et al claim that many of the published F_{ST} outliers based on FDIST2 and BAYESCAN are probably false positives however, their results show that these false positives are mostly in balancing selection and we did not include outliers identified as being under balancing selection for this reason (Lotterhos & Whitlock 2014). In a comparison of F_{ST} outlier tests, FDIST2 and BAYESCAN appeared to provide the most power, depending on the scenario, and BAYESCAN had fewest false positives (Narum & Hess 2011). Here, one locus 80994_17 (Digitifera scaffold gi|342271542|dbj|BACK01025553.1|, basepair = 5143) out of 13 was identified by both programs, therefore we consider this locus to be a strong candidate for being under selection and the other loci as possible candidates. However, STACKS locus 80994_17 was not annotated, a common occurrence even for transcribed loci in corals, where typically a third or less of genes have annotation (Meyer et al. 2011; Polato et al. 2010).

One of the SNP loci identified as being under positive selection was annotated as a *tyrosine-protein kinase transmembrane receptor ROR1-like*. ROR receptor protein is associated with the nervous system in the fruit fly *Drosophila* (Wilson et al. 1993), nematode *C. elegans* (Francis et al. 2005), and sea slug *Aplysia californica* (McKay et al. 2001). Functional analysis of *cam-1*, a gene that encodes for a ROR kinase in *C. elegans*, demonstrated roles in both the orientation of polarity in asymmetric cell division and axon outgrowth, and the ability to guide migrating cells (Forrester et al. 1999). The role of *ROR1* receptors in Cnidaria is unknown although studies in *Hydra* suggest a function in regulating cell specification and tissue morphogenesis (Bertrand et al. 2014; Krishnapati & Ghaskadbi 2014; Lange et al. 2014).

Another SNP identified as being under positive selection was located in the gene annotated as unconventional *myosin-IXb isoform X7*, a Rho GTPase-activating protein (RhoGAP) that is essential for coordinating the activity of Rho GTPases. Invertebrates are thought to contain a single myosin class IX gene (the exception is *Drosphilia* which has none) whereas most vertebrates have two with fishes having four (Liao et al. 2010). In general, Rho GTPases control the assembly and organization of the actin cytoskeleton which includes many functions such as cell adhesion, contraction and spreading, migration, morphogenesis, and phagocytosis. Little is known about the function of myosin-IX in invertebrates. However, a recent study in which *Orbicella faveolata* were exposed to immune challenges identified Unconventional myosin-IXb as a transcript that was significantly correlated with melanin protein activity (Fuess et al. 2016). In humans, Myosin-IXb is highly expressed in tissues of the immune system such as the lymph nodes, thymus, and spleen and also in imm cells like dendritic cells, macrophages and CD4 + T cells (Wirth et al. 1996). Myosin-IXb knockout mice showed impaired recruitment of monocytes and macrophages when exposed to a chemoattractant demonstrating that Myosin-IXb has an important function in innate immune responses *in vivo* (Hanley et al. 2010). Because statistical screens for loci under selection carry a high rate of false positive results, further experimental evidence is necessary before these loci can be considered targets of selection.

Restoration implications

Restoration efforts should proceed under the assumption that *A. palmata* harbors a significant amount of population structure requiring close matches of collection and outplant sites. Hybridization of *A. palmata* from different geographic regions may or may not result in heterosis depending on sexual compatibility, but would be worth pursuing in an *ex situ* setting to enable close monitoring of offspring performance under elevated temperatures (van Oppen et al. 2015). With respect to the sharply declining Florida colonies, these findings underline the need to manage and restore Florida's *A. palmata* as an isolated, genotypically depleted geographic region (Williams et al. 2014).

Acknowledgements

Thanks to PSU genome sequencing facility for expert library preparation and sequencing. N Polato contributed to the study design. We gratefully acknowledge the efforts by our

489 collaborators around the Caribbean that have contributed samples over the years. We thank the
490 editor and three reviewers for insightful comments.

491

492

Figure legends

Figure 1 *Acropora palmata* samples were obtained from throughout the Caribbean and north-west Atlantic range. Previous studies found a genetic break between the western and the eastern Caribbean but disagreed on the assignment of Puerto Rico to the western (long dashed line) or the eastern (short dashed line) population.

Figure 2 Bayesian cluster analysis of 307 SNP loci from *Acropora palmata* (n = 96). Reefs within geographic regions 1-4 sorted by latitude: Florida, Bahamas, Puerto Rico, US Virgin Islands. Analysis included only one SNP per locus after combining Read 1 and Read 2. Shown is the probability of membership (y-axis) in a given cluster for each sample (x-axis) assuming values of K=2 (a), K = 3 (b), and K = 4 (c). The most probable K was 3 (b) for the minor allele frequency corrected SNPs based on the mean estimated log probability of the data at a given K (3 replicate runs per K, +/- 1 standard deviation).

Figure 3 Bayesian cluster analysis of microsatellite data from *Acropora palmata* (n = 260). Panels (a-c). Analysis of 11 microsatellite loci with the most probable K being 4. Panels (d-e). Exclusion of the outlier locus 166 resulted in the analysis of 10 microsatellites with the most probable K being 3. Shown is the probability of membership (y-axis) in a given cluster for each sample (x-axis) assuming values of K=2 (a, d), K = 3 (b, e), and K = 4 (c, f).

Figure 4 MANTEL matrix correlation test between genetic (F_{st}), environmental (Euclidean) and geographic distances (km). . *Acropora palmata* samples from four geographic regions (Florida, Bahamas, Puerto Rico and USVI) were genotyped with 307 SNP (a-c) or 10 neutral microsatellite markers (d). Panel (a) $y = 0.0107x + 0.0104$, $R^2 = 0.610$, $p = 0.09$. Panel (b) $y=0.002x + 0.4175$, $R^2= 0.101$, $p = 0.21$. Panel (c) $y=0.000007x + 0.0098$. $R^2= 0.648$, p-value=0.05. Panel (d) $y=0.000007x + 0.0027$. $R^2= 0.69$, $p = 0.04$.

Supplemental Figure 1 SNP-derived F_{IS} distribution of *Acropora palmata* in four geographic regions.

Supplemental Figure 2 SNP-derived F_{ST} distributions of *Acropora palmata* in pairwise geographic region comparisons. Panel A: Bahamas vs Florida. Panel B: Bahamas vs Puerto Rico.

Panel C: Bahamas vs USVI. Panel D: Florida vs Puerto Rico. Panel E: Florida vs USVI. Panel F: Puerto Rico vs USVI.

Supplemental Figure 3 Minor allele frequency (MAF) cut of values of 0.025 (a-c) and 0.075 (d-f) yield similar Bayesian population clustering solutions to MAF = 0.05 (Fig 2 main text). Shown is the probability of membership (y-axis) in a given cluster for each sample (x-axis) assuming values of K=2 (a, d), K = 3 (b, e), and K = 4 (c, f). For each K, three replicate runs were performed. The most probable K was always 3 based on the mean estimated log probability of the data at a given K and the Evanno method. Setting the minor allele frequency to 0.025 (a-c) yields 632 SNPs (analysis included only one SNP per locus) after combining Read 1 and Read 2. Setting the minor allele frequency to 0.075 (d-f) yields 213 SNPs (analysis included only one SNP per locus) after combining Read 1 and Read 2. Reefs within geographic regions sorted by latitude.

Supplemental Figure 4 Excluding F_{ST} outliers from *Acropora palmata* SNP data does not change the choice of best K based on Bayesian cluster analysis with STRUCTURE. Shown is the probability of membership (y-axis) in a given cluster for each sample (x-axis) assuming values of K=2 (a), K = 3 (b), and K = 4 (c). Analysis of 299 SNPs (analysis included only one SNP per locus) after combining Read 1 and Read 2. The most probable K was 3 (B). Minor allele frequency cut of value was 0.05 and percentage of missing data per geographic region was 40%. Reefs within geographic regions are sorted by latitude.

Supplemental Figure 5 Decreasing the minimum percentage of individuals in a region required to have data for that locus from 60% to 40% does not change the choice of best K based on Bayesian cluster analysis with STRUCTURE. Figure shows results of a STRUCTURE analysis of 4793 SNPs (analysis included only one SNP per locus) after combining Read 1 and Read 2 and assuming a K of 2 (A), 3 (B) or 4 (C). The minor allele frequency was set to ≥0.05. The most probable K was 3 (B) based on the mean estimated log probability of the data at a given K and the Evanno method (3 replicate runs per K, +/- 1 standard deviation). Reefs within geographic regions are sorted by latitude.

Supplemental Figure 6 Principal coordinate analysis (PCA) of single nucleotide polymorphism (SNP, a) and microsatellite data (b) from *Acropora palmata*. Red= Florida, blue =Bahamas, yellow=Puerto Rico, green=USVI, purple=Belize, and orange = Curacao.

Supplemental Figure 7 Cumulative distribution of sequencing coverage of Bowtie2 aligned *Acropora palmata* reads to the *A. digitifera* genome using BEDTOOLS. On average 2.5% percent of the *A. digitifera* genome had sequence coverage at a stack depth of 5.

Supplemental Figure 8 Mantel matrix correlation test between genetic and geographic distances. *Acropora palmata* samples from four geographic regions (Florida, Bahamas, Puerto Rico and USVI) were genotyped with 11 microsatellite markers, including the outlier locus, 166. $y= 5*10^{-6}x + 0.0137$. $R^2 = 0.1147$, p-value=0.1.

Tables

Table 1 *Acropora palmata* colonies included in the SNP (A) and microsatellite (B) analyses. Samples were obtained from 3 – 6 (A) or more (B) reefs in four (A) and six (B) geographic regions in the Caribbean/north-west Atlantic. Given are latitude and longitude in decimal degrees (WGS84). NA = not available.

Table 2 RAD-tag sequencing summary table of *Acropora palmata* samples.

Table 3 Summary statistics for Read 1 and Read 2 combined. % PL = percent polymorphic loci, Obs Hom = observed homozygosity, Obs Het = observed heterozygosity, StdErr = standard error, Exp = expected. F_{IS} calculations with and without minor allele frequency restrictions. Calculated by STACKS 1.30.

Table 4 Pairwise F_{ST} calculated from STACKS 1.3. Read 1 and 2 combined (duplicated stacks between reads removed, MAF≥0.05). Considered were loci (n=390) present in all geographic regions.

Supplemental Table 1 DD-Rad sequencing**.** There were 12 pools with 8 unique barcodes in each. The Database ID is a unique identifier for each coral specimen. Given is also the total number of ramets for each genet in the Baumslab database. The indices are short DNA sequences that uniquely identify products in the final libraries.

Supplemental Table 2 Example code for STACKS.

Supplemental Table 3 GPS coordinates in decimal degrees (WGS84) for the World Ocean Atlas 2013 (WOA13 V2) environmental data, averaged for a geographic region.

Supplemental Table 4 Outlier SNPs identified by programs LOSITAN and BAYESCAN. Stacks locus_bp is the STACKS program locus ID with the SNP location basepair after the underscore. Read category indicates whether the outlier SNP was found in read 1 or read 2 or the paired-end sequencing run. *A. dig* scaffold identifies the *A. digitifera* scaffold where the *A. palmata* STACKS locus aligned to, followed by the basepair location in the next column. S start= sequence start. S end = sequence end. Lositan P = Lositan P- values (Simulated F_{ST} < sample F_{ST}). Bayescan q = Bayescan q-values. Ns = not significant.

.

Supplemental Table 5 Microsatellite genotypes at five loci for the samples included in the SNP dataset. The Database ID is a unique identifier for each coral specimen. Given is also is the geographic region and reef the samples were collected from, microsatellite allele call 1 and 2 for loci 166, 181, 182, 192, and 207 (in basebairs), and the coral genet ID. The last four columns show how the radseq samples were pooled and gives the inline barcode, the Illumina index, the Illumina sequence lane and chip.

607 References

608

609 Allendorf FW. 1986. Genetic drift and the loss of alleles versus heterozygosity. *Zoo Biology* 5:181-190.
610 10.1002/zoo.1430050212
611 Altschul SF, Madden TL, Schäffer AA, Zhang J, Zhang Z, Miller W, and Lipman DJ. 1997. Gapped BLAST
612 and PSI-BLAST: a new generation of protein database search programs. *Nucleic acids research*
613 25:3389-3402.
614 Altshuler D, Pollara VJ, Cowles CR, Van Etten WJ, Baldwin J, Linton L, and Lander ES. 2000. An SNP map
615 of the human genome generated by reduced representation shotgun sequencing. *Nature*
616 407:513-516.
617 Antao T, Lopes A, Lopes RJ, Beja-Pereira A, and Luikart G. 2008. LOSITAN: A workbench to detect
618 molecular adaptation based on a Fst-outlier method. *Bmc Bioinformatics* 9:323.
619 Arnold B, Corbett-Detig RB, Hartl D, and Bomblies K. 2013. RADseq underestimates diversity and
620 introduces genealogical biases due to nonrandom haplotype sampling. *Molecular Ecology*
621 22:3179-3190.
622 Babbucci M, Ferraresso S, Pauletto M, Franch R, Papetti C, Patarnello T, Carnier P, and Bargelloni L.
623 2016. An integrated genomic approach for the study of mandibular prognathism in the
624 European seabass (Dicentrarchus labrax). *Scientific reports* 6.
625 Balloux F, Lehmann L, and de Meeus T. 2003. The population genetics of clonal and partially clonal
626 diploids. *Genetics* 164:1635-1644.
627 Baums IB. 2008. A restoration genetics guide for coral reef conservation. *Molecular Ecology* 17:2796-
628 2811. doi:10.1111/j.1365-294X.2008.03787.x
629 Baums IB, Devlin-Durante M, Laing BA, Feingold J, Smith T, Bruckner A, and Monteiro J. 2014a. Marginal
630 coral populations: the densest known aggregation of Pocillopora in the Galápagos Archipelago is
631 of asexual origin. *Frontiers in Marine Science* 1:59.
632 Baums IB, Devlin-Durante MK, Brown L, and Pinzón JH. 2009. Nine novel, polymorphic microsatellite
633 markers for the study of threatened Caribbean acroporid corals. *Molecular Ecology Resources*
634 9:1155-1158.
635 Baums IB, Devlin-Durante MK, and LaJeunesse TC. 2014b. New insights into the dynamics between reef
636 corals and their associated dinoflagellate endosymbionts from population genetic studies.
637 *Molecular Ecology* 23:4203-4215. 10.1111/mec.12788
638 Baums IB, Hughes CR, and Hellberg MH. 2005a. Mendelian microsatellite loci for the Caribbean coral
639 *Acropora palmata*. *Marine Ecology - Progress Series* 288:115-127.
640 Baums IB, Johnson ME, Devlin-Durante MK, and Miller MW. 2010. Host population genetic structure and
641 zooxanthellae diversity of two reef-building coral species along the Florida Reef Tract and wider
642 Caribbean. *Coral Reefs* 29:835-842. 10.1007/s00338-010-0645-y
643 Baums IB, Miller MW, and Hellberg ME. 2005b. Regionally isolated populations of an imperiled
644 Caribbean coral, *Acropora palmata*. *Molecular Ecology* 14:1377-1390.
645 Baums IB, Miller MW, and Hellberg ME. 2006a. Geographic variation in clonal structure in a reef building
646 Caribbean coral, *Acropora palmata*. *Ecological Monographs* 76:503-519.
647 Baums IB, Paris CB, and Cherubin LM. 2006b. A bio-oceanographic filter to larval dispersal in a reef-
648 building coral. *Limnology and Oceanography* 51:1969-1981.
649 Beaumont MA, and Balding DJ. 2004. Identifying adaptive genetic divergence among populations from
650 genome scans. *Molecular Ecology* 13:969-980.
651 Beaumont MA, and Nichols RA. 1996a. Evaluating loci for use in the genetic analysis of population
652 structure. *Proceedings of the Royal Society of London Series B-Biological Sciences* 263:1619-
653 1626.

Beaumont MA, and Nichols RA. 1996b. Evaluating loci for use in the genetic analysis of population structure. *Proceedings of the Royal Society of London B: Biological Sciences* 263:1619-1626.

Becks L, Ellner SP, Jones LE, and Hairston NG, Jr. 2010. Reduction of adaptive genetic diversity radically alters eco-evolutionary community dynamics. *Ecology Letters* 13:989-997. 10.1111/j.1461-0248.2010.01490.x

Bedoya-Reina O, Ratan A, Burhans R, Kim H, Giardine B, Riemer C, Li Q, Olson T, Loughran T, vonHoldt B, Perry G, Schuster S, and Miller W. 2013. Galaxy tools to study genome diversity. *GigaScience* 2:17.

Bellis ES, Howe DK, and Denver DR. 2016. Genome-wide polymorphism and signatures of selection in the symbiotic sea anemone Aiptasia. *BMC Genomics* 17:160.

Bertrand S, Iwema T, and Escriva H. 2014. FGF signaling emerged concomitantly with the origin of Eumetazoans. *Molecular Biology and Evolution* 31:310-318.

Blanco-Bercial L, and Bucklin A. 2016. New view of population genetics of zooplankton: RAD-seq analysis reveals population structure of the North Atlantic planktonic copepod Centropages typicus. *Molecular Ecology* 25:1566-1580.

Blankenberg D, Von Kuster G, Bouvier E, Baker D, Afgan E, Stoler N, Team G, Taylor J, and Nekrutenko A. 2014. Dissemination of scientific software with Galaxy ToolShed. *Genome Biology* 15:3. 10.1186/gb4161

Cahill AE, and Levinton JS. 2016. Genetic differentiation and reduced genetic diversity at the northern range edge of two species with different dispersal modes. *Molecular Ecology* 25:515-526.

Carlon DB. 1999. The evolution of mating systems in tropical reef corals. *Trends in Ecology & Evolution* 14:491-495.

Catchen J, Hohenlohe PA, Bassham S, Amores A, and Cresko WA. 2013. Stacks: an analysis tool set for population genomics. *Molecular Ecology* 22:3124-3140. 10.1111/mec.12354

Catchen JM, Amores A, Hohenlohe P, Cresko W, and Postlethwait JH. 2011. Stacks: Building and Genotyping Loci De Novo From Short-Read Sequences. *G3: Genes, Genomes, Genetics* 1:171-182. 10.1534/g3.111.000240

Combosch DJ, and Vollmer SV. 2015. Trans-Pacific RAD-Seq population genomics confirms introgressive hybridization in Eastern Pacific Pocillopora corals. *Molecular phylogenetics and evolution* 88:154-162.

Delmotte F, Leterme N, Gauthier JP, Rispe C, and Simon JC. 2002. Genetic architecture of sexual and asexual populations of the aphid *Rhopalosiphum padi* based on allozyme and microsatellite markers. *Molecular Ecology* 11:711-723. 10.1046/j.1365-294X.2002.01478.x

Dimond JL, Gamblewood SK, and Roberts SB. 2017. Genetic and epigenetic insight into morphospecies in a reef coral. *Molecular Ecology*.

Drury C, Dale KE, Panlilio JM, Miller SV, Lirman D, Larson EA, Bartels E, Crawford DL, and Oleksiak MF. 2016. Genomic variation among populations of threatened coral: *Acropora cervicornis*. *Bmc Genomics* 17:286. 10.1186/s12864-016-2583-8

Drury C, Manzello D, and Lirman D. 2017. Genotype and local environment dynamically influence growth, disturbance response and survivorship in the threatened coral, Acropora cervicornis. *PLoS One* 12:e0174000.

Eckert C, Samis K, and Lougheed S. 2008. Genetic variation across species' geographical ranges: the central–marginal hypothesis and beyond. *Molecular Ecology* 17:1170-1188.

Ellegren H, and Galtier N. 2016. Determinants of genetic diversity. *Nature Reviews Genetics* 17:422-433.

Epstein N, Bak RPM, and Rinkevich B. 2001. Strategies for Gardening Denuded Coral Reef Areas: The Applicability of Using Different Types of Coral Material for Reef Restoration. *Restoration Ecology* 9:432-442. 10.1046/j.1526-100X.2001.94012.x

Evanno G, Regnaut S, and Goudet J. 2005. Detecting the number of clusters of individuals using the software STRUCTURE: a simulation study. *Molecular Ecology* 14:2611-2620.

Falush D, Stephens M, and Pritchard JK. 2003. Inference of population structure using multilocus genotype data: Linked loci and correlated allele frequencies. *Genetics* 164:1567-1587.

Fischer MC, Rellstab C, Leuzinger M, Roumet M, Gugerli F, Shimizu KK, Holderegger R, and Widmer A. 2017. Estimating genomic diversity and population differentiation – an empirical comparison of microsatellite and SNP variation in *Arabidopsis halleri*. *Bmc Genomics* 18:69. 10.1186/s12864-016-3459-7

Fogarty ND. 2012. Caribbean acroporid coral hybrids are viable across life history stages. *Marine Ecology Progress Series* 446:145-159.

Foll M, and Gaggiotti O. 2008. A genome-scan method to identify selected loci appropriate for both dominant and codominant markers: a Bayesian perspective. *Genetics* 180:977-993.

Forrester WC, Dell M, Perens E, and Garriga G. 1999. A *C. elegans* Ror receptor tyrosine kinase regulates cell motility and asymmetric cell division. *Nature* 400:881-885.

Forsman Z, Knapp I, Tisthammer K, Eaton D, Belcaid M, and Toonen R. 2017. Coral hybridization or phenotypic variation? Genomic data reveal gene flow between Porites lobata and P. Compressa. *Molecular phylogenetics and evolution* 111:132-148.

Foster NL, Paris CB, Kool JT, Baums IB, Stevens JR, Sanchez JA, Bastidas C, Agudelo C, Bush P, Day O, Ferrari R, Gonzalez P, Gore S, Guppy R, McCartney MA, McCoy C, Mendes J, Srinivasan A, Steiner S, Vermeij MJA, Weil E, and Mumby PJ. 2012. Connectivity of Caribbean coral populations: complementary insights from empirical and modelled gene flow. *Molecular Ecology* 21:1143-1157. 10.1111/j.1365-294X.2012.05455.x

Francis MM, Evans SP, Jensen M, Madsen DM, Mancuso J, Norman KR, and Maricq AV. 2005. The Ror receptor tyrosine kinase CAM-1 is required for ACR-16-mediated synaptic transmission at the C. elegans neuromuscular junction. *Neuron* 46:581-594.

Frantz A, Cellina S, Krier A, Schley L, and Burke T. 2009. Using spatial Bayesian methods to determine the genetic structure of a continuously distributed population: clusters or isolation by distance? *Journal of Applied Ecology* 46:493-505.

Fuess LE, Weil E, and Mydlarz LD. 2016. Associations between transcriptional changes and protein phenotypes provide insights into immune regulation in corals. *Developmental & Comparative Immunology* 62:17-28.

Gagnaire PA, Normandeau E, Cote C, Hansen MM, and Bernatchez L. 2012. The Genetic Consequences of Spatially Varying Selection in the Panmictic American Eel (*Anguilla rostrata*). *Genetics* 190:725-U703. 10.1534/genetics.111.134825

Gorospe KD, and Karl SA. 2010. Small-scale spatial analysis of in situ sea temperature throughout a single coral patch reef. *Journal of Marine Biology* 2011.

Griffin JN, Schrack EC, Lewis K-A, Baums IB, Soomdat N, and Silliman BR. 2015. Density-dependent effects on initial growth of a branching coral under restoration *Restoration Ecology* 23:197-200. 10.1111/rec.12173

Griffin S, Spathias H, Moore DM, Baums IB, and Griffin BA. 2012. Scaling up *Acropora* nurseries in the Caribbean and improving techniques. In: Yellowlees D, and Hughes TP, eds. *Proceedings of the 12th International Coral Reef Symposium*. Townsville, Australia: James Cook University.

Hamrick JL, and Godt M. 1996. Effects of life history traits on genetic diversity in plant species. *Philosophical transactions: biological sciences*:1291-1298.

Hanley PJ, Xu Y, Kronlage M, Grobe K, Schön P, Song J, Sorokin L, Schwab A, and Bähler M. 2010. Motorized RhoGAP myosin IXb (Myo9b) controls cell shape and motility. *Proceedings of the National Academy of Sciences* 107:12145-12150. 10.1073/pnas.0911986107

748 Herrera S, Reyes-Herrera PH, and Shank TM. 2015. Predicting RAD-seq marker numbers across the
749 eukaryotic tree of life. *Genome biology and evolution* 7:3207-3225.
750 Hoffberg SL, Kieran TJ, Catchen JM, Devault A, Faircloth BC, Mauricio R, and Glenn TC. 2016. RADcap:
751 sequence capture of dual-digest RADseq libraries with identifiable duplicates and reduced
752 missing data. *Molecular Ecology Resources* 16:1264-1278. 10.1111/1755-0998.12566
753 Howells EJ, Abrego D, Meyer E, Kirk NL, and Burt JA. 2016a. Host adaptation and unexpected symbiont
754 partners enable reef-building corals to tolerate extreme temperatures. *Global Change Biology*
755 22:2702-2714. 10.1111/gcb.13250
756 Howells EJ, Abrego D, Meyer E, Kirk NL, and Burt JA. 2016b. Host adaptation and unexpected symbiont
757 partners enable reef-building corals to tolerate extreme temperatures. *Global change biology*.
758 Hubisz M, Falush D, Stephens M, and Pritchard J. 2009. Inferring weak population structure with the
759 assistance of sample group information. *Molecular Ecology Resources* 9:1322-1332.
760 Jeffreys H. 1961. *The theory of probability*. Oxford: Oxford University Press.
761 Jombart T. 2008. adegenet: a R package for the multivariate analysis of genetic markers. *Bioinformatics*
762 24:1403-1405.
763 Kopelman NM, Mayzel J, Jakobsson M, Rosenberg NA, and Mayrose I. 2015. Clumpak: a program for
764 identifying clustering modes and packaging population structure inferences across K. *Molecular*
765 *ecology resources* 15:1179-1191.
766 Krishnapati L-S, and Ghaskadbi S. 2014. Identification and characterization of VEGF and FGF from Hydra.
767 *International Journal of Developmental Biology* 57:897-906.
768 Kulathinal RJ, Stevison LS, and Noor MA. 2009. The genomics of speciation in Drosophila: diversity,
769 divergence, and introgression estimated using low-coverage genome sequencing. *PLoS Genet*
770 5:e1000550.
771 Lal MM, Southgate PC, Jerry DR, and Zenger KR. 2016. Fishing for divergence in a sea of connectivity:
772 The utility of ddRADseq genotyping in a marine invertebrate, the black-lip pearl oyster Pinctada
773 margaritifera. *Marine genomics* 25:57-68.
774 Lange E, Bertrand S, Holz O, Rebscher N, and Hassel M. 2014. Dynamic expression of a Hydra FGF at
775 boundaries and termini. *Development Genes and Evolution* 224:235-244.
776 Langmead B, and Salzberg SL. 2012. Fast gapped-read alignment with Bowtie 2. *Nature Methods* 9:357-
777 U354. Doi 10.1038/Nmeth.1923
778 Laporte M, Pavey SA, Rougeux C, Pierron F, Lauzent M, Budzinski H, Labadie P, Geneste E, Couture P,
779 and Baudrimont M. 2016. RAD sequencing reveals within-generation polygenic selection in
780 response to anthropogenic organic and metal contamination in North Atlantic Eels. *Molecular*
781 *Ecology* 25:219-237.
782 Liao W, Elfrink K, and Bähler M. 2010. Head of Myosin IX Binds Calmodulin and Moves Processively
783 toward the Plus-end of Actin Filaments. *Journal of Biological Chemistry* 285:24933-24942.
784 10.1074/jbc.M110.101105
785 Lischer HEL, and Excoffier L. 2012. PGDSpider: an automated data conversion tool for connecting
786 population genetics and genomics programs. *Bioinformatics* 28:298-299.
787 10.1093/bioinformatics/btr642
788 Lotterhos KE, and Whitlock MC. 2014. Evaluation of demographic history and neutral parameterization
789 on the performance of FST outlier tests. *Molecular Ecology* 23:2178-2192.
790 Macdonald AH, Schleyer M, and Lamb J. 2011. Acropora austera connectivity in the south-western
791 Indian Ocean assessed using nuclear intron sequence data. *Marine Biology* 158:613-621.
792 Manel S, Schwartz MK, Luikart G, and Taberlet P. 2003. Landscape genetics: combining landscape
793 ecology and population genetics. *Trends in Ecology & Evolution* 18:189-197.

McKay SE, Hislop J, Scott D, Bulloch AG, Kaczmarek LK, Carew TJ, and Sossin WS. 2001. *Aplysia ror* forms clusters on the surface of identified neuroendocrine cells. *Molecular and Cellular Neuroscience* 17:821-841.

Mège P, Schizas NV, Garcia Reyes J, and Hrbek T. 2014. Genetic seascape of the threatened Caribbean elkhorn coral, *Acropora palmata*, on the Puerto Rico Shelf. *Marine Ecology*:online early. 10.1111/maec.12135

Meirmans PG. 2012. The trouble with isolation by distance. *Molecular Ecology* 21:2839-2846.

Meyer E, Aglyamova G, and Matz M. 2011. Profiling gene expression responses of coral larvae (Acropora millepora) to elevated temperature and settlement inducers using a novel RNA-Seq procedure. *Molecular Ecology* 20:3599-3616.

Morin PA, Martien KK, and Taylor BL. 2009. Assessing statistical power of SNPs for population structure and conservation studies. *Molecular ecology resources* 9:66-73. 10.1111/j.1755-0998.2008.02392.x

Narum SR, and Hess JE. 2011. Comparison of FST outlier tests for SNP loci under selection. *Molecular ecology resources* 11:184-194.

Nielsen EE, Hansen MM, and Meldrup D. 2006. Evidence of microsatellite hitch-hiking selection in Atlantic cod (*Gadus morhua* L.): implications for inferring population structure in nonmodel organisms. *Molecular Ecology* 15:3219-3229. 10.1111/j.1365-294X.2006.03025.x

Nosil P, Funk DJ, and Ortiz-Barrientos D. 2009. Divergent selection and heterogeneous genomic divergence. *Molecular Ecology* 18:375-402. DOI 10.1111/j.1365-294X.2008.03946.x

Peakall R, and Smouse PE. 2006. GENALEX 6: genetic analysis in Excel. Population genetic software for teaching and research. *Molecular Ecology Notes* 6:288-295.

Peterson BK, Weber JN, Kay EH, Fisher HS, and Hoekstra HE. 2012. Double Digest RADseq: An Inexpensive Method for De Novo SNP Discovery and Genotyping in Model and Non-Model Species. *PLoS ONE* 7. 10.1371/journal.pone.0037135

Polato NR, Voolstra CR, Schnetzer J, DeSalvo MK, Randall CJ, Szmant AM, Medina M, and Baums IB. 2010. Location-specific responses to thermal stress in larvae of the reef-building coral Montastraea faveolata. *PLoS One* 5:e11221. 10.1371/journal.pone.0011221

Porto-Hannes I, Zubillaga AL, Shearer TL, Bastidas C, Salazar C, Coffroth MA, and Szmant AM. 2014. Population structure of the corals Orbicella faveolata and Acropora palmata in the Mesoamerican Barrier Reef System with comparisons over Caribbean basin-wide spatial scale. *Marine Biology*:1-18. 10.1007/s00227-014-2560-1

Putnam NH, Srivastava M, Hellsten U, Dirks B, Chapman J, Salamov A, Terry A, Shapiro H, Lindquist E, and Kapitonov VV. 2007. Sea anemone genome reveals ancestral eumetazoan gene repertoire and genomic organization. *Science* 317:86-94.

Quinlan AR, and Hall IM. 2010. BEDTools: a flexible suite of utilities for comparing genomic features. *Bioinformatics* 26:841-842. 10.1093/bioinformatics/btq033

Reitzel AM, Herrera S, Layden MJ, Martindale MQ, and Shank TM. 2013. Going where traditional markers have not gone before: utility of and promise for RAD sequencing in marine invertebrate phylogeography and population genomics. *Molecular Ecology* 22:2953-2970. 10.1111/mec.12228

Renaut S, Nolte AW, Rogers SM, Derome N, and Bernatchez L. 2011. SNP signatures of selection on standing genetic variation and their association with adaptive phenotypes along gradients of ecological speciation in lake whitefish species pairs (*Coregonus* spp.). *Molecular Ecology* 20:545-559.

Richards ZT, and Oppen MJH. 2012. Rarity and genetic diversity in Indo–Pacific *Acropora* corals. *Ecology and Evolution* 2:1867-1888. 10.1002/ece3.304

Rinkevich B. 2006. The coral gardening concept and the use of underwater nurseries: lessons learned from silvics and silviculture. *Coral reef restoration handbook*:291-302.

Romiguier J, Gayral P, Ballenghien M, Bernard A, Cahais V, Chenuil A, Chiari Y, Dernat R, Duret L, and Faivre N. 2014. Comparative population genomics in animals uncovers the determinants of genetic diversity. *Nature* 515:261.

Ryman N, Palm S, Andre C, Carvalho GR, Dahlgren TG, Jorde PE, Laikre L, Larsson LC, Palme A, and Ruzzante DE. 2006. Power for detecting genetic divergence: differences between statistical methods and marker loci. *Molecular Ecology* 15:2031-2045.

Sachidanandam R, Weissman D, Schmidt SC, Kakol JM, Stein LD, Marth G, Sherry S, Mullikin JC, Mortimore BJ, and Willey DL. 2001. A map of human genome sequence variation containing 1.42 million single nucleotide polymorphisms. *Nature* 409:928-933.

Schopmeyer SA, Lirman D, Bartels E, Byrne J, Gilliam DS, Hunt J, Johnson ME, Larson EA, Maxwell K, Nedimyer K, and Walter C. 2012. In Situ Coral Nurseries Serve as Genetic Repositories for Coral Reef Restoration after an Extreme Cold-Water Event. *Restoration Ecology* 20:696-703. 10.1111/j.1526-100X.2011.00836.x

Selkoe KA, and Toonen RJ. 2011. Marine connectivity: a new look at pelagic larval duration and genetic metrics of dispersal. *Marine Ecology Progress Series* 436:291-305. 10.3354/meps09238

Shinzato C, Mungpakdee S, Arakaki N, and Satoh N. 2015. Genome-wide SNP analysis explains coral diversity and recovery in the Ryukyu Archipelago. *Scientific reports* 5:18211.

Shinzato C, Shoguchi E, Kawashima T, Hamada M, Hisata K, Tanaka M, Fujie M, Fujiwara M, Koyanagi R, Ikuta T, Fujiyama A, Miller DJ, and Satoh N. 2011. Using the *Acropora digitifera* genome to understand coral responses to environmental change. *Nature* 476:320-323. 10.1038/nature10249

Sork VL, Squire K, Gugger PF, Steele SE, Levy ED, and Eckert AJ. 2016. Landscape genomic analysis of candidate genes for climate adaptation in a California endemic oak, *Quercus lobata*. *American Journal of Botany* 103:33-46. 10.3732/ajb.1500162

Stoeckel S, Grange J, Fernandez-Manjarres JF, Bilger I, Frascaria-Lacoste N, and Mariette S. 2006. Heterozygote excess in a self-incompatible and partially clonal forest tree species - *Prunus avium* L. *Molecular Ecology* 15:2109-2118. 10.1111/j.1365-294X.2006.02926.x

Toonen RJ, Puritz JB, Forsman ZH, Whitney JL, Fernandez-Silva I, Andrews KR, and Bird CE. 2013. ezRAD: a simplified method for genomic genotyping in non-model organisms. *PeerJ* 1:e203. 10.7717/peerj.203

van Oppen MJH, Oliver JK, Putnam HM, and Gates RD. 2015. Building coral reef resilience through assisted evolution. *Proceedings of the National Academy of Sciences* 112:2307-2313. 10.1073/pnas.1422301112

Vasemägi A, Nilsson J, and Primmer CR. 2005. Expressed Sequence Tag-Linked Microsatellites as a Source of Gene-Associated Polymorphisms for Detecting Signatures of Divergent Selection in Atlantic Salmon (*Salmo salar* L.). *Molecular Biology and Evolution* 22:1067-1076. 10.1093/molbev/msi093

Vilas A, PÉRez-Figueroa A, and Caballero A. 2012. A simulation study on the performance of differentiation-based methods to detect selected loci using linked neutral markers. *Journal of evolutionary biology* 25:1364-1376. 10.1111/j.1420-9101.2012.02526.x

Vollmer SV, and Palumbi SR. 2007. Restricted gene flow in the Caribbean staghorn coral *Acropora cervicomis*: Implications for the recovery of endangered reefs. *Journal of Heredity* 98:40-50.

Wang S, Meyer E, McKay JK, and Matz MV. 2012. 2b-RAD: a simple and flexible method for genome-wide genotyping. *Nature Methods* 9:808-+. 10.1038/nmeth.2023

887 Williams DE, Miller MW, and Baums IB. 2014. Cryptic changes in the genetic structure of a highly clonal
888 coral population and the relationship with ecological performance. *Coral Reefs* 33:595-606.
889 10.1007/s00338-014-1157-y
890 Willing E-M, Dreyer C, and van Oosterhout C. 2012. Estimates of Genetic Differentiation Measured by
891 F(ST) Do Not Necessarily Require Large Sample Sizes When Using Many SNP Markers. *PLoS One*
892 7:e42649. 10.1371/journal.pone.0042649
893 Wilson C, Goberdhan D, and Steller H. 1993. Dror, a potential neurotrophic receptor gene, encodes a
894 *Drosophila* homolog of the vertebrate Ror family of Trk-related receptor tyrosine kinases.
895 *Proceedings of the National Academy of Sciences* 90:7109-7113.
896 Wirth J, Jensen K, Post P, Bement W, and Mooseker M. 1996. Human myosin-IXb, an unconventional
897 myosin with a chimerin-like rho/rac GTPase-activating protein domain in its tail. *Journal of Cell*
898 *Science* 109:653-661.

899

900

Table 1 *Acropora palmata* colonies included in the SNP (A) and microsatellite (B) analyses. Samples were obtained from 3 – 6 (A) or more (B) reefs in four (A) and six (B) geographic regions in the Caribbean/north-west Atlantic. Given are latitude and longitude in decimal degrees (WGS84). NA = not available.

A)

Region	***Reef***	***Count of Samples***	***Latitude***	***Longitude***
Florida	Sand Island	6	25.018093	-80.368472
	French	8	25.03393	-80.34941
	Little Grecian	1	25.118433	-80.31715
	Horseshoe	1	25.139467	-80.29435
	Elbow	8	25.143628	-80.257927
Bahamas	Little Ragged Island	1	22.15375	-75.687208
	Adelaine Cay	8	22.173372	-75.703016
	Elkhorn Cay	2	22.328253	-75.783228
	Johnson Cay	3	22.33312	-75.77892
	Nairn Cay	8	22.35199	-75.79612
	Middle Beach	2	23.781199	-76.10391
Puerto Rico	San Cristobal	8	17.56493	-67.04515
	Rincon	6	18.21007	-67.15849
	Tres Palmas	2	18.350133	-67.266333
	La Cordillera	8	18.368522	-65.571678
US Virgin Islands	Tague Bay	8	17.763867	-64.613397
	Hawksnest Bay	8	18.347183	-64.780775
	Johnsons Reef	8	18.361733	-64.7743
Grand Total		96		

B)

Region	***Reef***	***Count of Samples***	***Latitude***	***Longitude***
Florida	Horseshoe	1	25.1395	-80.294
	Little Grecian	1	25.1184	-80.317
	Sand Island	6	25.0179	-80.369
	Western Sambo	6	24.4799	-81.719
	Rock Key	4	24.456	-81.86
	Dry Tortugas	1	24.6209	-82.868
	Marker 3	1	25.3733	-80.16

Region	***Reef***	***Count of Samples***	***Latitude***	***Longitude***
	Boomerang Reef	1	25.3525	-80.179
	Carysfort	4	25.2219	-80.211
Bahamas	Great Iguana	19	26.7075	-77.154
	Middle Beach	2	23.7812	-76.104
	Charlies Beach	1	23.7808	-76.104
	Black Bouy	1	23.8022	-76.146
	Bock Cay	1	23.8075	-76.16
	Little Darby	2	23.8474	-76.209
	Rocky Dundas	1	24.2788	-76.539
	Halls Pond	2	24.3539	-76.57
	LSI	3	23.7691	-76.096
	Little Ragged Island	1	22.1538	-75.687
	Adelaine Cay	1	22.1734	-75.703
	Johnson Cay	1	22.3331	-75.779
	Nairn Cay	4	22.352	-75.796
Puerto Rico	San Cristobal	14	17.5649	-67.045
	Rincon	24	18.2101	-67.159
	Aurora	3	17.9425	-66.871
	Paraguera	1	17.997	-67.052
USVI	Hawksnest Bay	6	18.3472	-64.781
	Johnsons Reef	12	18.3617	-64.774
	Haulover Bay	13	18.3489	-64.677
	Buck Island	14	18.2774	-64.894
	Flat Key	4	18.317	-64.989
	Hans Lollik	4	18.4019	-64.906
	Sapphire	6	18.3333	-64.85
	Botany	3	18.3572	-65.036
Belize	unknown	3	NA	NA
	Bugle Caye	1	NA	NA
	Curlew	5	16.7909	-88.083
	Gladden	1	16.4401	-88.192
	Glovers Atoll	3	NA	NA
	GSTF1	5	16.5499	-88.05
	GSTF12	7	16.5499	-88.05
	Larks Caye	1	NA	NA
	Laughing Bird Caye	4	16.4367	-88.199
	Loggerhead	2	NA	NA

Region	***Reef***	***Count of Samples***	***Latitude***	***Longitude***
	Sandbores	3	16.7791	-88.118
	Carrie Bow	13	16.8021	-88.082
Curacao	Blue Bay	7	12.1352	-68.99
	Boka Patrick	8	12.2873	-69.043
	Directors Bay	2	12.0664	-68.8603
	East Point	4	12.0407	-68.783
	PuntuPicu	9	12.0831	-68.896
	Red Bay	2	12.1355	-68.99
	Sea Aquarium	9	12.0838	-68.896
	Water Factory	3	12.1085	-68.9528
Sum		260		

911

912

913

914 **Table 2** RAD-tag sequencing summary table of *Acropora palmata* samples.

915

	Region	*Pool*	*Coral colonies*	*Lane*	*Total Reads*	*Retained Reads after processing*	*Average number of retained sequence reads per sample*	*Standard Deviation*
West	Bahamas	B1	8	2	50,900,230	41,199,646	5,149,956	1,915,875
		B2	8	2	56,097,984	45,237,633	5,654,704	1,853,265
		B3	8	2	58,379,852	47,706,860	5,963,358	2,734,261
	Florida	F1	8	1	50,925,548	39,750,070	4,968,759	1,681,820
		F2	8	1	48,752,776	42,036,153	5,254,519	4,422,737
		F3	8	1	49,942,322	38,611,895	4,826,487	2,518,097
East	Puerto Rico	P1	8	1	43,979,338	36,237,997	4,529,750	4,166,551
		P2	8	1	55,267,402	47,235,081	5,904,385	4,096,287
		P3	8	1	47,324,190	34,835,445	4,354,431	3,117,707
	USVI	U1	8	2	40,616,766	33,170,324	4,146,291	2,187,597
		U2	8	2	43,215,386	34,291,498	4,286,437	1,187,166
		U3	8	2	45,849,098	38,439,719	4,804,965	1,555,938
		Sum	96		591,250,892	478,752,321		

916

Table 3 Summary statistics for Read 1 and Read 2 combined. % PL = percent polymorphic loci, Obs Hom = observed homozygosity, Obs Het = observed heterozygosity, StdErr = standard error, Exp = expected. F_{IS} calculations with and without minor allele frequency restrictions. Calculated by STACKS 1.30.

		Bahamas	**Florida**	**Puerto Rico**	**USVI**
All positions: variant and fixed	Total Sites	200425	200425	200425	200425
	Variant Sites	390	390	390	390
	Private Alleles	2	1	0	2
	% PL	0.1732	0.1497	0.1694	0.1668
	F_{IS}	0.00005	0	0	0.00005
	Nucleotide diversity (π)	0.0004	0.0004	0.0004	0.0004
Variant positions only	Obs Hom	0.7728	0.7874	0.7791	0.7815
	Std Err	0.0164	0.0164	0.0154	0.0154
	Obs Het	0.2273	0.2126	0.2210	0.2186
	Std Err	0.0164	0.0164	0.0154	0.0154
	Exp Hom	0.7832	0.8050	0.7919	0.7916
	Exp Het	0.2169	0.1951	0.2081	0.2085
	F_{IS}	0.02235	**-0.0086**	0.0035	0.02065
	Nucleotide diversity (π)	0.2254	0.2034	0.2174	0.21705

Table 4 Pairwise F_{ST} comparisons of geographic regions based on SNP (A) and microsatellite (B) data.

A)

	Bahamas	Florida	Puerto Rico	USVI
Bahamas				
Florida	0.018			
Puerto Rico	0.013	0.022		
USVI	0.018	0.022	0.009	

B)

	Belize	Florida	Bahamas	Puerto Rico	USVI	Curacao
Belize						
Florida	0.0040					
Bahamas	0.0115	0.0097				
Puerto Rico	0.0206	0.0153	0.0063			
USVI	0.0206	0.0174	0.0098	0.0037		
Curacao	0.0240	0.0138	0.0181	0.0173	0.0208	

932

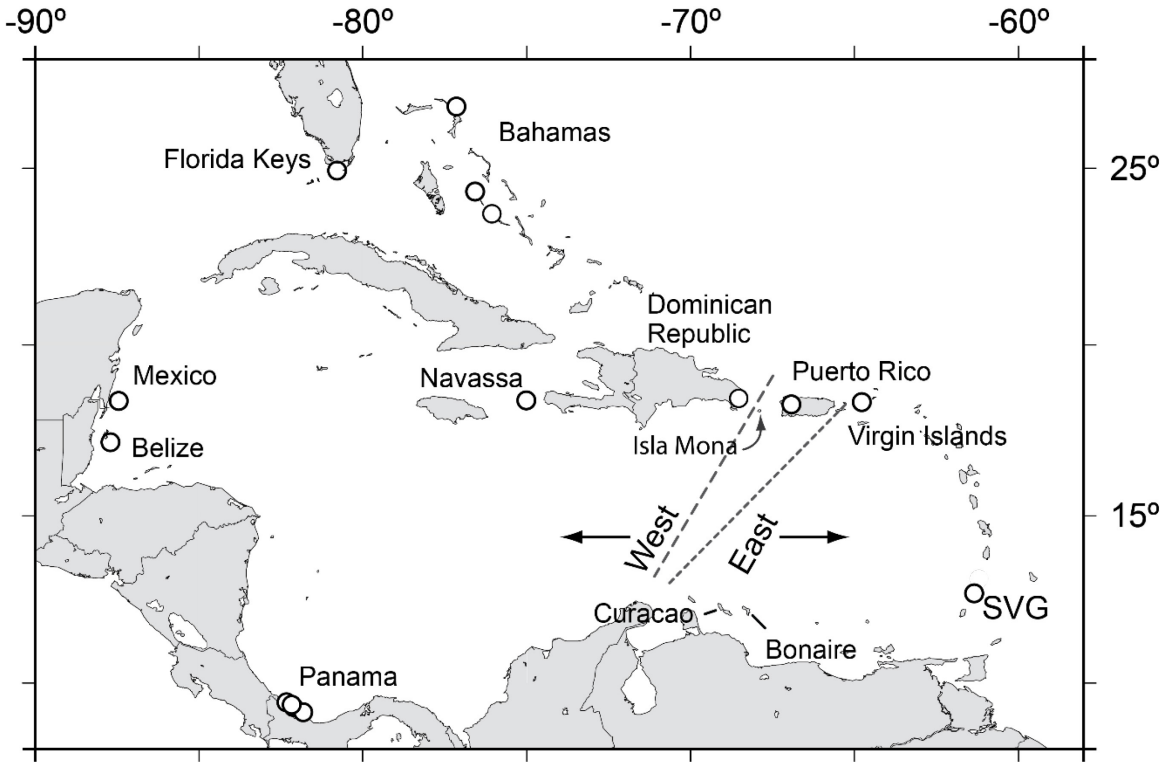

Figure 1 *Acropora palmata* samples were obtained from throughout the Caribbean and north-west Atlantic range. Previous studies found a genetic break between the western and the eastern Caribbean but disagreed on the assignment of Puerto Rico to the western (long dashed line) or the eastern (short dashed line) population.

933

934

935

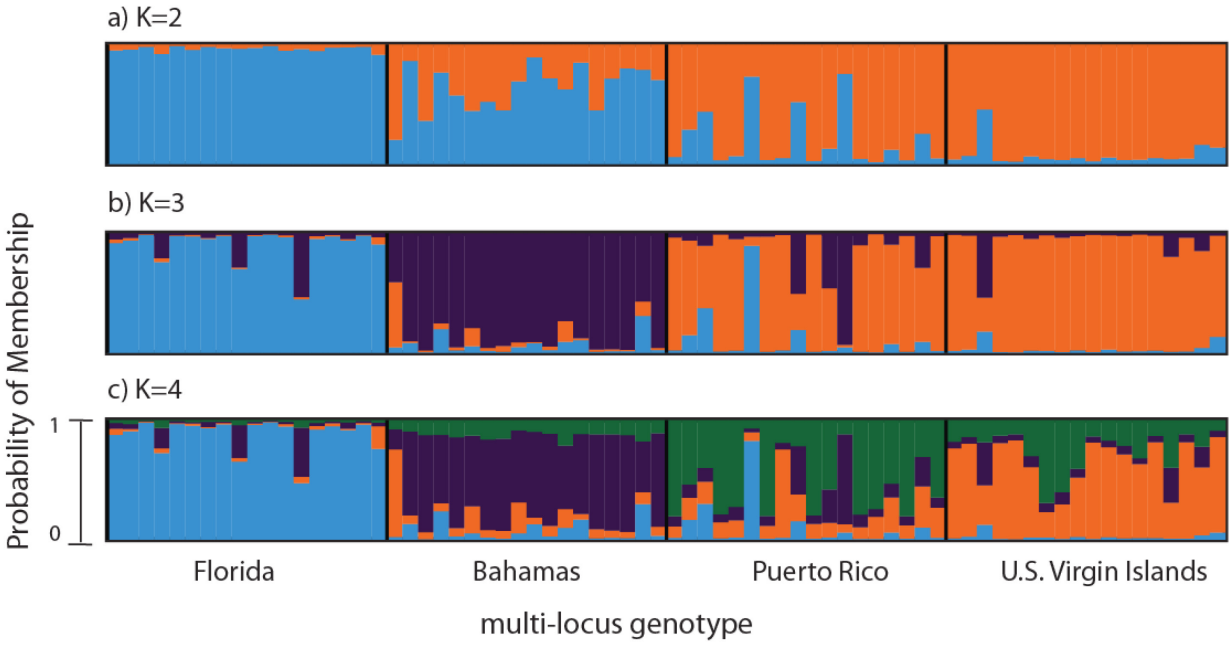

Figure 2 Bayesian cluster analysis of 307 SNP loci from *Acropora palmata* (n = 96). Reefs within geographic regions 1-4 sorted by latitude: Florida, Bahamas, Puerto Rico, US Virgin Islands. Analysis included only one SNP per locus after combining Read 1 and Read 2. Shown is the probability of membership (y-axis) in a given cluster for each sample (x-axis) assuming values of K=2 (a), K = 3 (b), and K = 4 (c). The most probable K was 3 (b) for the minor allele frequency corrected SNPs based on the mean estimated log probability of the data at a given K (3 replicate runs per K, +/- 1 standard deviation).

936
937
938
939

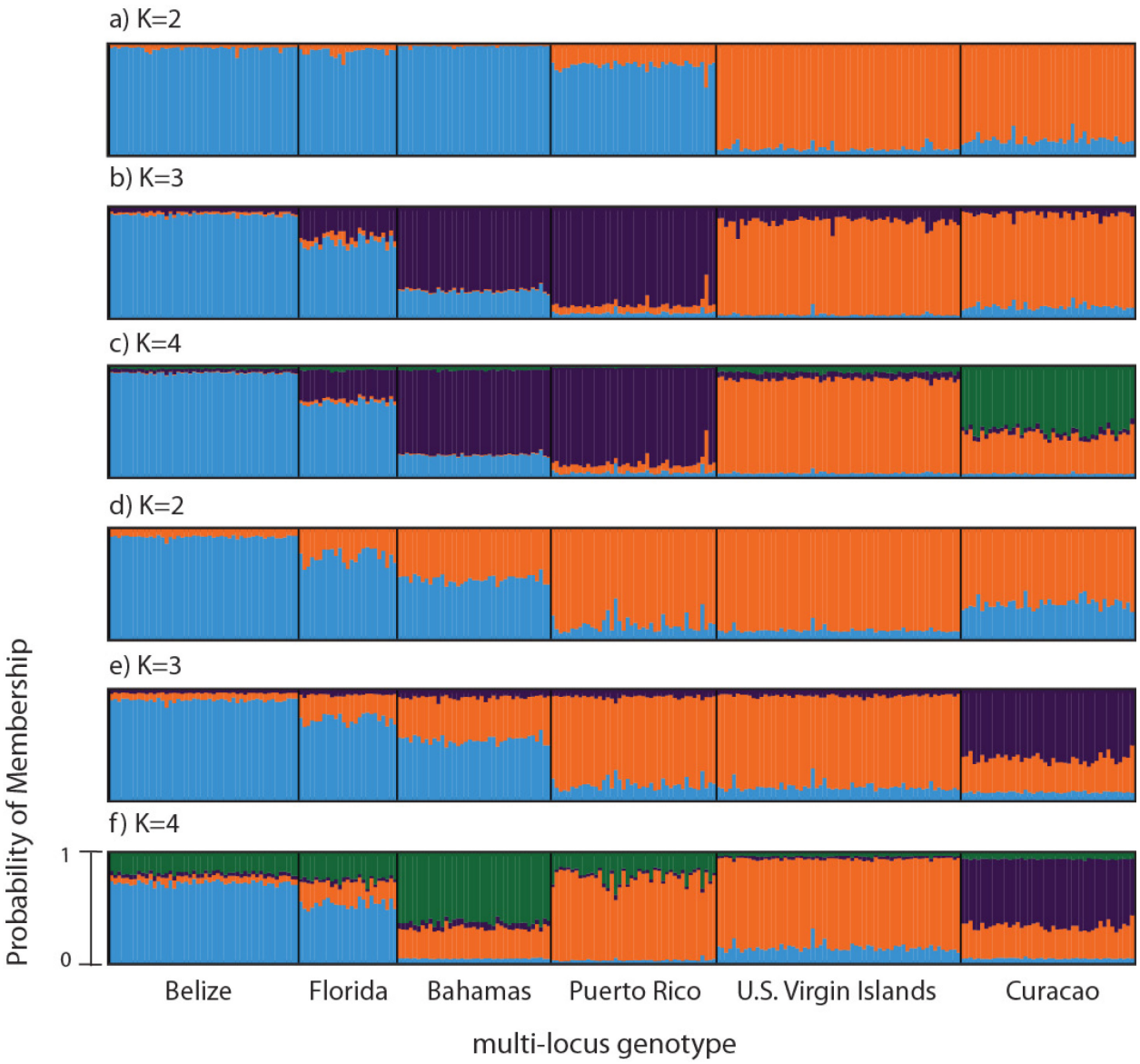

Figure 3 Bayesian cluster analysis of microsatellite data from *Acropora palmata* (n = 260). Panels (a-c): Analysis of 11 microsatellite loci with the most probable K being 4. Panels (d-e): Exclusion of the outlier locus 166 resulted in the analysis of 10 microsatellites with the most probable K being 3. Shown is the probability of membership (y-axis) in a given cluster for each sample (x-axis) assuming values of K=2 (a, d), K = 3 (b, e), and K = 4 (c, f).

940

941

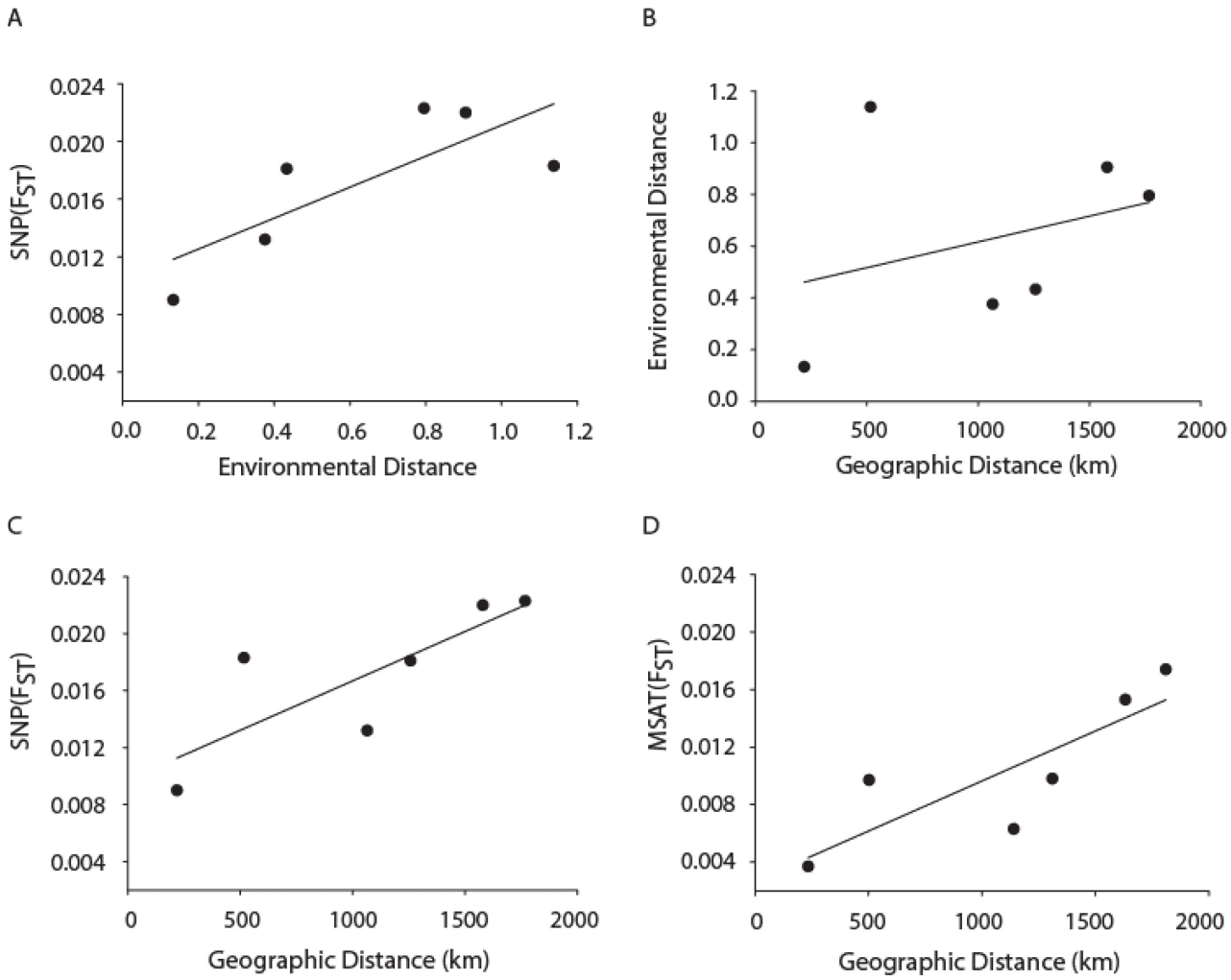

Figure 4 MANTEL matrix correlation test between genetic and geographic distances, and environmental parameters as calculated by a dissimilarity matrix expressed as the Euclidean distances between geographic regions based on environmental data. *Acropora palmata* samples from four geographic regions (Florida, Bahamas, Puerto Rico and USVI) were genotyped with 307 SNP (a-c) or 10 neutral microsatellite (d) markers. Panel (a) y = 0.0107x + 0.0104, R^2 = 0.610, p = 0.09. Panel (b) y=0.002x + 0.4175, R^2= 0.101, p = 0.21. Panel (c) y=0.000007x + 0.0098. R^2= 0.648, p-value=0.05. Panel (d) y=0.000007x + 0.0027. R^2= 0.69, p = 0.04.

942
943