Genotypic similarity among algal symbionts corresponds to associations with closely related coral hosts

Hannah G. Reich[1,2*], Sheila A. Kitchen[1,3*], Kathryn H. Stankiewicz[1], Meghann Devlin-Durante[1], Nicole D. Fogarty[4], Iliana B. Baums[1*]

[1]Department of Biology, The Pennsylvania State University, University Park, PA 16801 USA

[2]Present address: Department of Biological Sciences, University of Rhode Island, Kingston, RI 02881, USA

[3]Present address: Division of Biology and Biological Engineering, California Institute of Technology, Pasadena, CA 91125

[4]Department of Biology and Marine Biology, Center for Marine Science, University of North Carolina Wilmington, Wilmington, NC 28409

Correspondence: hgreigh16@gmail.com; sak3097@caltech.edu; baums@psu.edu

Keywords: Acropora, hybrid, niche diversification, single-nucleotide polymorphisms, Symbiodiniaceae, symbiosis

Running title: Allelic variation of an endangered coral symbiont

Abstract

Mutualisms where hosts are coupled metabolically to their symbionts often exhibit high partner fidelity. Most reef-building corals form obligate symbioses with specific species of photosymbionts, dinoflagellates in the family Symbiodiniaceae, despite needing to acquire

symbionts early in their development from environmental sources. Three Caribbean acroporids (*Acropora palmata*, *A. cervicornis*, and their hybrid *A. prolifera*) are geographically sympatric across much of their range in the greater Caribbean, but often occupy different depth and light habitats. Both species and their hybrid associate with *Symbiodinium 'fitti'*, a genetically diverse species of symbiont that is specific to these hosts. Since the physiology of the dinoflagellate partner is strongly influenced by light (and therefore depth), we investigated whether *S. 'fitti'* populations from each host source were differentiated genetically. We generated shallow genome sequences of acroporid colonies sampled from across the Caribbean. Single Nucleotide Polymorphisms (SNPs) among *S. 'fitti'* strains were identified by aligning sequences to a ~600 Mb draft assembly of the *S. 'fitti'* genome, assembled from an *A. cervicornis* metagenome. Phylogenomic and multivariate analyses revealed that allelic variation among *S. 'fitti'* partitioned to each host species, as well as their hybrid, rather than by biogeographic origin. This is particularly noteworthy because the hybrid, *A. prolifera*, has a sparse fossil record and may be of relatively recent origin. Many of the SNPs putatively under selection were non-synonymous mutations predicted to alter protein efficiency. Differences in allele frequency among *S.* '*fitti*' populations from each host taxon may correspond to distinct phenotypes that thrive in the different cellular environments found in each acroporid. The non-random sorting among genetically diverse strains, or genotypes, to different hosts could be the basis for lineage diversification via disruptive selection, leading to ecological specialization and ultimately speciation.

Introduction

Ecosystem services provided by coral-dinoflagellate mutualisms rival the contributions of other widely studied symbioses (ex: tubeworms and bacteria, yeast and termites, plant and fungi). In the coral-dinoflagellate mutualism, each partner benefits as the coral receives photosynthetic sugars from their dinoflagellate symbionts and the algal symbiont receives nutrients and protection in return (Trench 1979). Though decades of investigations have probed the causes and consequences of coral-algal dysbiosis (ex: coral bleaching), we are still gathering information on their basic biology (Cziesielski, Schmidt-Roach, & Aranda 2019). Specifically, little is known about the intraspecies level of co-evolutionary dynamics between host and their symbionts. The interdependence of the partners adds complexity to the system as each partner is selected in the context of the other.

Understanding the mechanisms that promote the apparent fidelity of reef-building corals towards one endosymbiotic dinoflagellate species (Symbiodiniaceae), despite having the opportunity to horizontally acquire symbionts, is important in light of rapid climate change. This is because the ability of corals to cope with heat stress by "shuffling" their endosymbiont communities to a more heat tolerant lineage (Little, Oppen, & Willis 2004; Silverstein, Cunning, & Baker 2017) may be limited. Corals have co-evolved with the Symbiodiniaceae since the Jurassic period (LaJeunesse et al. 2018) and, over time, may have become uniquely adapted to their symbionts (Goulet 2006; LaJeunesse et al. 2018; LaJeunesse et al. 2004; Lewis, Chan, & LaJeunesse 2019; Parkinson, Coffroth, & LaJeunesse 2015b; Thornhill, Fitt, & Schmidt 2006a). While juvenile corals often host several symbiont species, this community wanes over time to the dominant symbiont species (Abrego, Van Oppen, & Willis 2009; Coffroth, Goulet, & Santos 2001; Poland & Coffroth 2017). This suggests that corals are most compatible with their dominant symbiont species and foreign pairings might be maladaptive, at least under current

conditions (Cunning, Gillette, Capo, Galvez, & Baker 2015; Pettay, Wham, Smith, Iglesias-Prieto, & LaJeunesse 2015).

A central question in coral reef science is whether coral-dinoflagellate symbioses can adapt to increasing sea surface temperatures over ecological time scales, especially if shuffling of symbiont partners is restricted (Goulet 2006). Genetic variation present within algal species has remained largely unstudied as a source that may fuel such adaptation (Buerger et al. 2020; Parkinson, Banaszak, Altman, LaJeunesse, & Baums 2015a). The potential for genetic variation within algal species to fuel adaptation to changing conditions can be assessed in the laboratory via experimental evolution experiments where algal strains are selected over several generations under heat stress conditions (Baker et al. 2018; Buerger et al. 2020; Chakravarti & van Oppen 2018). In one instance, Symbiodiniaceae strains adapted to heat stress selection *in vitro* but once introduced into the coral partner, gains were not always retained highlighting the complexity of adaptation in the context of mutualistic partners (Buerger et al. 2020).

Alternatively, field studies may assess the long-term influence of selective factors such as a strong light gradient on the genetic variation of Symbiodiniaceae by taking advantage of depth stratification found in their coral hosts (Bongaerts et al. 2015a; Serrano et al. 2016). While sharing a geographic range, Caribbean *Acropora* species often differentiate across a depth and light gradient; *A. cervicornis* occupies a lower light habitat (~10 m depth) relative to its high-light dwelling (~3 m depth) sibling species *A. palmata* and their hybrid *A. prolifera* (~1 m depth; (Fogarty 2012; Goreau 1959; LaJeunesse 2002). All three taxa harbor the dinoflagellate endosymbiont *Symbiodinium 'fitti'* (ITS2 type A3), which is distinct from other *Symbiodinium* A3 lineages found in giant clams and other cnidarians (Kemp et al. 2015; Lee et al. 2015; Pinzón et al. 2015; Shoguchi et al. 2018). The variation of morphology between the three taxa ranges

92 from broad, moose antler branches to thinner, stag antler branches results in differences in the
93 flow and light field within and around the colonies (Enríquez, Méndez, Hoegh-Guldberg, &
94 Iglesias-Prieto 2017; Gladfelter 1983; Gladfelter 2007). Therefore, the persistence of *S. 'fitti'* in
95 three host taxa at a range of depths across a large geographic region provides a unique
96 opportunity to study how evolutionary history, geography, natural ecology, and biophysical
97 parameters may contribute to adaptation and co-evolution in the coral holobiont.
98 The ecological and evolutionary dynamics between host and symbiont species are
99 influenced by differences in their reproduction and dispersal strategies (Reviewed in Thornhill,
100 Howells, Wham, Steury, & Santos 2017). Caribbean acroporid corals reproduce via production
101 of meiotic, planktonic larvae and also disperse locally via fragmentation. Gene flow is restricted
102 between eastern and western Caribbean region for *A. palmata* and *A. cervicornis* (Baums, Miller,
103 & Hellberg 2005, 2006). Within each region, further population structure is observed but the
104 specifics differ between species with *A. cervicornis* showing generally more fine-scale
105 differentiation then *A. palmata* (Baums et al. 2005, 2006; Devlin-Durante & Baums 2017;
106 Hemond & Vollmer 2010; Kitchen et al. 2019; Vollmer & Palumbi 2002, 2006). *A. palmata* and
107 *A. cervicornis* have been present in the fossil record since the late Pliocene (~2.6-3.6 Mya)
108 whereas the hybrid *A. prolifera* is mostly absent from the fossil record (Budd & Johnson 1999;
109 McNeill, Budd, & Borne 1997; Precht, Vollmer, Modys, & Kaufman 2019). Although *A.*
110 *prolifera* produces viable eggs and sperm, molecular analyses indicates that F2 adults are very
111 rare or absent while backcrosses with either parent species occur occassionally (Kitchen et al.
112 2020; Van Oppen, Willis, Vugt, & Miller 2000; Vollmer & Palumbi 2002).
113 The population structure and genotypic diversity of *S. 'fitti'* has received less attention
114 and at a coarser level of genomic resolution (Baums et al. 2019; Baums, Devlin-Durante, &

LaJeunesse 2014; Thornhill et al. 2017). Despite that, higher levels of population genetic structure are documented in *S. 'fitti'* when compared to one of its hosts, *A. palmata* (Baums et al. 2014). *S. 'fitti'* cells divide mitotically within the host and water column dispersal appears limited (Fitt & Trench 1983; Thornhill et al. 2017). Accordingly, the majority of Caribbean acroporids colonies host a single strain of *S. 'fitti'* and maintain long-term fidelity to that strain (Baums et al. 2014; O'Donnell, Lohr, Bartels, Baums, & Patterson 2018). However, sexual reproduction in Symbiodiniaceae has not been completely ruled out as a reproductive strategy because meiotic machinery has been detected in genomic data (Bellantuono, Dougan, Granados-Cifuentes, & Rodriguez-Lanetty 2019; Chi, Parrow, & Dunthorn 2014; Levin et al. 2016; Shah, Chen, Bhattacharya, & Chan 2020) and recombination is evident in population genetic data (Baums et al. 2014). These contrasting life-history strategies may contribute to the higher levels of population structure of *S. 'fitti'* compared to their acroporid hosts throughout the Caribbean (Baums et al. 2014; Thornhill et al. 2017).

Here, we describe fine-scale genetic differences in *S. 'fitti'* strains across its three host taxa spanning the geographic distribution of the mutualism. A draft *S. 'fitti'* genome assembly was constructed from *A. cervicornis* metagenomic sequences and compared to other Symbiodiniaceae genomic resources. Variation in genome-wide single nucleotide polymorphisms (SNPs) were investigated in *S. 'fitti'* and scanned for mutations that may change protein structure and function. Lastly, the potential biological and evolutionary ramifications of the allelic composition of *S. 'fitti'* are discussed.

Methods

Sample Collection, Sequencing, and Assembly

Tissue was collected for genome sequencing from 76 acroporids spanning the geographic distribution of *S. 'fitti'* (Fig. 1, Table S1; Kitchen et al. 2019). High molecular weight DNA was isolated from each coral tissue sample using the Qiagen DNeasy kit (Qiagen, Valencia, CA) without prior enrichment for *S. 'fitti'*. Of these samples, one specimen for each species from Florida (*A. cervicornis* CFL14120 and *A. palmata* PFL1012) was 'deeply' sequenced (~150x coverage; Kitchen et al. 2019). Paired-end short insert (550 nt) sequencing libraries of the two deeply sequenced samples were constructed with 1.8-2 µg sample DNA and the TruSeq DNA PCR-Free kit (Illumina, San Diego, CA). The remaining 74 paired-end short insert (350 nt) sequencing libraries were constructed using 100 ng sample DNA and the TruSeq DNA Nano kit (Illumina, San Diego, CA) with coverage between 8-40x (Kitchen et al. 2019). Deep- and shallow-sequence libraries were pooled separately and sequenced on either Illumina HiSeq 2500 or HiSeq 4000 (Table S1, Illumina, San Diego, CA).

Sequencing adaptors and low-quality base calls (Phred score < 25) from the 3' end of the deeply sequenced *A. cervicornis* metagenome reads were trimmed using cutadapt v 1.6 (Martin 2011). After initial filtering, processed reads shorter than 50 bp were discarded and PCR duplicates removed using FastUniq v. 1.1 (Xu et al. 2012). A series of filtering steps were completed to identify the fraction of reads originating from *A. cervicornis* and *Symbiodinium* spp. First, a modified approach similar to Blobology, which compares sequence homology, read coverage and GC content, was performed (Kumar, Jones, Koutsovoulos, Clarke, & Blaxter 2013). Contigs from a preliminary genome assembly created with SOAPdenovo2 v0.4 (parameters -K 95 –R) were compared for homology against the coral *Acropora digitifera* (NCBI: GCF_000222465.1) and symbiont *Breviolum minutum* (OIST: symbB.v1.0.genome.fa) genomes, and NCBI nucleotide database nt using megablast (evalue 1e-5; Altschul et al. 1997;

Shinzato et al. 2011; Shoguchi et al. 2013). The sequence matches to the nt database for contigs that had no match to either coral or symbiont genome, were used to create a local contamination database to further screen the reads (Luo et al. 2015).

Reads were aligned with Bowtie2 v. 2.2.9 (parameters –q –fast; Langmead & Salzberg 2012) consecutively to the *A. digitifera* mitochondria (KF448535.1), followed by a concatenated set of three Symbiodiniaceae genomes (*Symbiodinium microadriaticum, Breviolum minutum, Fugacium;* Aranda et al. 2016; Lin et al. 2015; Shoguchi et al. 2013), and the contamination database. This filtering step, however, only aligned 0.28% of the reads from *A. cervicornis* to the Symbiodiniaceae genomes. Reads that mapped to Symbiodiniaceae genomes (n=1,004,992) were extracted and assembled using SPADES v3.9.1 with a multi-kmer approach (-k 21,33,55,77,99) (Bankevich et al. 2012). The reads that aligned to the contamination database were assembled separately with SPADES v3.9.1 as described above, resulting in three additional contigs that matched Symbiodiniaceae genomes through blast homology.

The filtered *A. cervicornis* reads were assembled with SoapDeNovo v2, followed by six rounds of gap filling using GapCloser v1.12 and scaffolding using both SSPACE v2.0 (rounds 1, 3, and 5) and LINKs v1.8.5 with *A. digitifera* scaffolds as the "long-reads" (-t 2 –d 3000 –k 25, round 2, 4 and 6) on alternate rounds (Boetzer & Pirovano 2014; Luo et al. 2015; Warren et al. 2015). After the first three rounds and then each subsequent round, the contigs/scaffolds were partitioned to either coral or symbiont based on the top scoring match (lowest e-value) against a local blast database containing three Symbiodiniaceae genomes and five cnidarian genomes (*Hydra*, *Hydractinia*, *Nematostella, Exaiptasia, Acropora digitifera*). If the scaffolds equally matched cnidarian and symbiont sequences or did not match either they were retained in the coral fraction. Scaffolds identified as Symbiodiniaceae after the six rounds as well as those

assembled with SPADES above were combined. Two additional rounds of scaffolding with LINKS using *S. microadriaticum* as "long-reads" followed by SSPACE and gap filling with GapCloser were performed. To remove any remaining sequences matching cnidarian sequences, a final round of scaffold partitioning was performed by comparing the scaffolds against the *Acropora* spp. genomes: *A. digitifera* (Shinzato et al. 2011), *A. palmata* (Kitchen et al. unpublished; http://baumslab.org/research/data), *A. cervicornis* (Kitchen et al. unpublished), *A. hyacinthus* (Barshis et al. 2013), *A. tenuis* (Liew, Aranda, & Voolstra 2016); and *Symbiodinium* spp. genomes: *S. microadriaticum* (Aranda et al. 2016), and *S. tridacnidorum* (Shoguchi et al. 2018).

Genome annotation and completeness

Genes were predicted using Augustus v 3.2.3 (Stanke et al. 2006; Stanke, Steinkamp, Waack, & Morgenstern 2004). Each predicted gene in *S. 'fitt*i' was queried against the NCBI nr, Uniprot SwissProt and trembl databases using blastx 2.6.0+ (max target seqs = 5, max hsps = 1, e-value = 1e-5; Altschul et al. 1997; Apweiler et al. 2004; Bairoch & Apweiler 1997; UniProt 2014). Gene models were also compared to the *S. microadriaticum* gene and protein predictions (NCBI: GCA_001939145.1) using blast (Altschul et al. 1997). To calculate assembly statistics and compare completeness to other Symbiodiniaceae genomes available at the time of analysis (*F. kawagutii, C. goreaui, B. minutum, S. microadriaticum, S. tridacnidorum*), we compared our *S. 'fitti'* assembly and the aforementioned assemblies using an online version of CEGMA with the eukaryote ortholog set executed by gVolante (https://gvolante.riken.jp; Simão, Waterhouse, Ioannidis, Kriventseva, & Zdobnov 2015). Orthofinder v2.2.1 with default settings (Emms &

206 Kelly 2015) was used to identify unique and shared orthogroups between *S. 'fitti'* and six other
207 Symbiodiniaceae species.
208
209 *S. 'fitti' infection status*
210 Presence of multiple *S. 'fitti'* strains within a coral host sample was determined using 12
211 *S. 'fitti'* specific microsatellite loci as described by Baums et al. (2014). *S. 'fitti'* is haploid, thus
212 samples with multiple alleles for any given *S. 'fitti'* microsatellite locus were deemed co-infected
213 and removed from downstream analyses (Table S1).
214
215 *Variant detection and filtering*
216 For the purpose of SNP analyses, the *S. 'fitti'* genome assembly based on the *A.*
217 *cervicornis* metagenome was used as a reference for variant calling of the deeply-sequenced *A.*
218 *palmata* and all shallow genome samples (Table S1). The 47 shallow *S. 'fitti'* and 1 deep
219 sequenced *A. palmata* "like" *S. 'fitti'* genome samples were aligned using BWA v0.7.15 (Li
220 2013). Samtools v1.4 was used to remove PCR duplicates from the BAM file and alignment
221 statistics were calculated using samtools *flagstat* (Table S1; Li et al. 2009). Variants were
222 gathered using samtools *mpileup* using the –ugAEf and –t AD,DP flags and called using bcftools
223 v1.4 using the haploid, –f GQ, and -vmO z flags (Li et al. 2009; Narasimhan et al. 2016). The
224 bcftools (Li et al. 2009; Narasimhan et al. 2016) –m2 –M2 –v snps flags were used to separate
225 SNPs from the output and the –v indels flag was used to remove indels from the output
226 (Narasimhan et al. 2016). High-quality SNPs and indels were characterized as variants with a
227 quality score over 200 and with no more than 20% of variant calls missing at a given site among

all samples (Danecek et al. 2011; Narasimhan et al. 2016). Only high-quality SNPs were used in subsequent analyses.

Population structure

The *psbA* minicircle was assembled from each sample to determine if the dominant algal partners amongst the three host taxa were all *S. 'fitti'* The psbA minicircle in the *S. 'fitti'* genome assembly was identified through blast searches against three psbA sequences from NCBI (JN557866.1 = *Symbiodinium* type A3, JX094319.1 = *Breviolum minutum*, and AJ884898.1= *B. faviinorum* (Barbrook, Visram, Douglas, & Howe 2006; Mungpakdee et al. 2014; Pochon, Putnam, Burki, & Gates 2012). The *psbA* minicircle for the remaining samples was assembled using two approaches. In the first approach, filtered and trimmed short-read sequences were mapped to *S. 'fitti' psbA* sequence (scaffold71443) using Bowtie 2 v2.3.4.1 (Langmead & Salzberg 2012) with the --sensitive mode parameter. Mapped reads were extracted using bedtools v2.26.0 (Quinlan & Hall 2010) and assembled using SPAdes v3.10.1 (Bankevich et al. 2012) with various kmer sizes (-k 21, 33, 55, 77 and 99). In the second approach, the *de novo* organelle genome assembler NOVOplasty was used (Dierckxsens, Mardulyn, & Smits 2016). The *S. 'fitti'* genome *psbA* sequence was used as the seed sequence to extract similar sequences from the original, unfiltered reads for each sample. A consensus sequence from the two approaches for each sample was created after manual alignment of the sequences using MEGA6 (Tamura, Stecher, Peterson, Filipski, & Kumar 2013).

Phylogenomic patterns of *S. 'fitti'* allelic composition were determined using a subset of the high-quality SNPs without missing data with the RAxML-NG v. 0.9.0 GTR+FO+G nucleotide model (Stamatakis 2014). The tree topology with the lowest likelihood score is

presented with nodal support from 100 bootstrap replicates (Stamatakis 2014). Population structure was evaluated using STRUCTURE v2.3.4 for the 58,813 high-quality SNPs (Pritchard, Stephens, & Donnelly 2000). Additionally, the R package poppr v2.1.0 was used to determine the multilocus genotype of each strain using high quality SNPs with different genetic distance thresholds ranging 10-20% (See table S1; Kamvar, Tabima, & Grünwald 2014; Kitchen et al. 2020) Clusters in multivariate space were detected using the *pca* function in PCAdapt (Luu, Bazin, & Blum 2017). An Analysis of Molecular Variance (AMOVA, poppr R package) was used for additional detection of population differentiation (Kamvar et al. 2014).

Determination of variants under selection

Two different methods were used to identify candidate loci under selection. BayeScan v2.1 is a Bayesian method that incorporates uncertainty of allele frequencies between populations with small sample sizes (Fischer, Foll, Excoffier, & Heckel 2011; Foll, Fischer, Heckel, & Excoffier 2010; Foll & Gaggiotti 2008). The default BayeScan settings were used for determining SNPs under selection when accounting for *S. 'fitti'* host, location, and host*location interactions. PCAdapt v4.0.3 was used to determine SNPs under selection without prior population information using the default settings (Knaus & Grünwald 2017). Outliers from BayeScan were determined as markers where FDR <0.05 and outliers from PCAdapt v4.0.3 were determined by q-values larger than the alpha value (0.05; Fischer et al. 2011; Foll et al. 2010; Foll & Gaggiotti 2008; Knaus & Grünwald 2017). Outlier loci with a Bayes probability of 1 and q-value of 0 which becomes infinite following logarithmic transformation and were removed from plotting. All statistics from SNPs under selection, their proximity to coding regions, sequence coverage,

and per SNP F_{ST} are in Tables S7 and S9. SnpEff v4.3 was used to predict downstream functional implications of all detected variants (De Baets et al. 2011).

Data and code availability

Raw data is publicly available on NCBI under SRA project PRJNA473816. Code for data analysis and figure generation is available on github (https://github.com/hgreich/Sfitti).

Results

Genome statistics and comparison to other Symbiodiniaceae

The *Symbiodinium 'fitti'* assembly has a total nucleotide length of over 600 Mb (601,782,011 bp) and contains 274,185 contigs/scaffolds (*A. cervicornis*-*S. 'fitti'* CFL14120; Fig. 1, Table S2). The *A. palmata*- *S. 'fitti'* (PFL1012) deeply-sequenced sample had 297,371,995 reads map to the *A. cervicornis*-*S. 'fitti'* (CFL14120) reference (19% mapping rate, 8.5% paired reads, 4.2% singleton reads; Table S1). The shallow-sequenced genome samples with one *S. 'fitti'* genotype had an average of 6,461,332 reads map to the reference (19.2% mapping rate, 8.8% paired reads, 4.3% singleton reads; Table S1). The GC content and number of ambiguous bases were comparable to the *S. microadriaticum* assembly at 50.24% and 4.82%, respectively (Fig. 1, Table S2; Aranda et al. 2016). The genome completeness was assessed by the identification of the 248 core genes queried using the CEGMA program. The *S. 'fitti'* assembly had 55 complete proteins, 79 complete + partial proteins, and 169 missing proteins, which is comparable to the other Symbiodiniaceae assemblies queried (Table S2). The average number of orthologs per core gene was ~1.4 for the Symbiodiniaceae assemblies (Table S2). Gene prediction of *S. 'fitti'* assembly revealed 24,286 gene models, however, many were incomplete

(i.e., missing start or stop codon; Table S2, S3). In the gene family analysis, 3,368 orthogroups were found to be shared by the Symbiodiniaceae assemblies excluding *S. 'fitti'* whereas 2,982 orthogroups were found to be shared by all Symbiodiniaceae assemblies (Fig. S1; including *S. 'fitti'*). Additionally, 1,898 orthogroups were uniquely shared by *S. 'fitti'* and its closest relative *S. tridacnidorum* whereas 1,357 orthogroups were shared by the three assemblies from the genus *Symbiodinium* (Fig. S1). *S. 'fitti'* had 11 orthogroups that were not shared with any other Symbiodiniaceae assemblies.

Variation of the allelic composition of Symbiodinium 'fitti'

Based on the analysis of *S. 'fitti'* specific microsatellite loci, the majority of shallow-sequenced samples harbored one strain of *S. 'fitti* (n= 47, 75.9% of *A. palmata*-*S. 'fitti'* and 82.6% of *A. cervicornis*- and *A. prolifera*-*S. 'fitti'*; Table S1) and were used for further analysis. A total of 2,505,230 SNPs and 569,337 indels were identified between all samples. Of these, 58,538 SNPs and 1,874 indels were considered high-quality (Fig. 1). The 58,538 high-quality SNPs represent a range of per SNP fixation levels from 0 to 1 (Fig. S2). The Transition: Transversion ratio of the high-quality SNPs was 1.79 and did not vary by host species (Table S1). Of the high-quality SNPs, 12,780 (21.8%) occurred in coding regions (Table S4). The majority of the SNPs (87.5%) in coding regions matched other Symbiodiniaceae genomic resources (primarily the closely related species, *S. microadriaticum*). Multi-locus genotype (MLG) filtering of the 58,538 "high quality" SNPs indicated each sample represented a unique MLG (strain), consistent with the microsatellite analysis, and was retained for downstream analyses (Table S1). Additional filtering to remove variants with missing data resulted in 6,813 high-quality SNPs, hereafter called conservative SNPs. After this procedure of quality filtering

SNPs and setting a stringent missing data threshold, the average read coverage increased from 1.53 to 11.3 per SNP (Table S1; average 655.3% increase).

Patterns of host-specificity and biogeography within S. 'fitti'

The phylogeny of the *psbA* minicircle non-coding region revealed little differentiation between symbiont strains with respect to their host taxa, confirming that *S. 'fitti'* is one species (Fig. S2). The AMOVA corroborated that most of the variation among *S. 'fitti'* is at the within-species level (86.1+%; Table S5). Variation attributed to host species explained 11.6% of the components of covariance (sigma = 87.0) and then variation among the geographic location of each host explained 2.3% of the components of covariance (sigma = 17.1; Table S5). Variation among geographic locations explained a negative amount of the components of covariance (-5.4%, sigma = -39.3) whereas variation among the host taxa at the various geographic locations explained 16.3% of the components of covariance (sigma = 117.6; Table S5).

Consistent with the AMOVA results, samples clustered primarily by host taxon rather than geographic origins in a Principal Component Analysis (PCA) with the high-quality SNPs and in the Maximum Likelihood tree with the conservative SNPs (Figs. 2, 3). In the PCA, 16% of variance was explained by PC1 whereas 13.6% was explained by PC2 (Fig. 2). Within each host, there was some indication of biogeographic partitioning in the phylogeny but not in the PCA (Figs. 2, 3; Table S6). The *S. 'fitti'* associated with *A. prolifera* were found intermediate to the parental species in the Maximum Likelihood tree and clustered loosely together but were more similar to the *A. palmata S. 'fitti'* (Fig. 2). Analysis of STRUCTURE output using the *delta K* method (Evanno, Regnaut, & Goudet 2005) identified three clusters as the most likely K (Table S6). The three clusters largely corresponded to host taxa (Fig. 4).

342

343 *SNPs under selection*

344 Of the high-quality SNPs, 4,987 (8.5%) were determined as selection outliers by
345 PCAdapt (Fig. 5). When BayeScan accounted for host identity, location of host, and host by
346 location interaction, 217, 5 and 197 selection outliers were identified, respectively (n= 370 SNPs;
347 Fig. 5; Table S7). Additionally, 339 selection outlier SNPs were shared between the two
348 programs (Fig. 5; Table S7). 103 outlier loci identified by BayeScan had a Bayes probability of 1
349 and q-value of 0 which becomes infinite following logarithmic transformation and were therefore
350 removed from the Manhattan plot (Fig. 5) but were reported in table S7. For each set of SNPs
351 under selection, a subset was found in coding regions (899 from PCAdapt, 19 from BayeScan, 14
352 from both callers; Table S7). The 14 outliers in coding regions that were shared by both callers
353 were found in the coding regions of Putative cytosolic oligopeptidase, tankyrase-like proteins,
354 alpha-agarase, and uncharacterized proteins (Table S7).

355

356 *Predicted functional implications of SNPs within S. 'fitti'*

357 Of all high-quality SNPs, SnpEff identified 60,373 modifier/non-coding variants
358 (84.43%), 3,629 moderate/mostly harmless variants (5.08%), 7,451 low impact variants that
359 might change protein efficiency/effectiveness (10.42%), and 51 highly disruptive SNPs (0.07%).
360 Of the predicted mutations, SnpEff identified 3,644 non-synonymous mutations (32.87%), 32
361 premature stop codon/nonsense mutations (0.29%), and 7,410 synonymous mutations (66.84%;
362 Table S8). The aforementioned mutations are predicted to cause 11,086 codon changes and 3,676
363 amino acid changes (Tables S8, S9). SnpEff predicted the 14 aforementioned outlier mutations

364 of Putative cytosolic oligopeptidase, tankyrase-like proteins, and alpha-agarase as modifier
365 variants found in the introns (Tables S7-S10).
366
367 **Discussion**
368 The population dynamics and evolutionary history of reef-building corals are relatively
369 well studied compared to the dinoflagellate symbionts they harbor. However, selection acts on
370 both partners and differences in life-history characteristics between algae and corals suggest that
371 the spatial and temporal scale of adaptation may differ. This would have consequences for our
372 understanding of how they may adapt to rapidly changing climates. Genomic data of Caribbean
373 acroporids reveals fine-scale population structure within the host taxa (Devlin-Durante & Baums
374 2017; Kitchen et al. 2019; Kitchen et al. 2020). Previous analyses of microsatellite loci
375 demonstrated that *S. 'fitti'* gene-flow scales were smaller than its host *Acropora palmata* (Baums
376 et al. 2014), however, this study jointly analyzed SNP data from all three Caribbean acroporid
377 host taxa. We showed that the allelic composition of sympatric *S. 'fitti'* populations are
378 partitioned by host taxon and describe two potential scenarios that would lead to this result (Figs.
379 2-4). Differentiation of *S. 'fitti'* by host taxon implies partner selectivity, which may be the result
380 of coevolution (Scenario 1). Alternatively, coevolution and partner selectivity *per se* (i.e. via
381 specific recognition of genetic variants among partners) may not explain the patterns of allelic
382 variation. Differences in the micro-environment (light, depth, nutrient availability) associated
383 with the habitat preferences of the host taxa may drive symbiont differentiation without specific
384 recognition interactions or coevolution (Scenario 2). However, it is likely a combination of these
385 mechanisms that drives the coevolution of S. *'fitti'* strains within Caribbean acroporids. In either
386 case, *S. 'fitti'* genetic diversity is tied to that of its endangered hosts.

Host selectivity of symbiont strains in a horizontal symbiont transmission

Broadcast spawning coral species acquire symbiotic partners horizontally each generation during the aposymbiotic larval stage (Baird, Guest, & Willis 2009). Thus, it is perplexing that coral lineages remain specific to a single symbiont strain, despite the potential to select a novel symbiont that might expand its physiological capacity and range of suitable habitats (Barneah, Weis, Perez, & Benayahu 2004; Chan, Lewis, Neely, & Baums 2019). Though initial intra-family, intra-genus, and intra-specific symbiont diversity is observed during early months of development, the diversity wanes and reverts to the dominant (adult) symbiont species shortly thereafter (Abrego et al. 2009; Poland & Coffroth 2017). Specificity between adult broadcast spawning corals and their symbionts is commonly observed at the species level with multi-marker and microsatellite approaches (Chan et al. 2019; LaJeunesse 2001; Lewis et al. 2019). Here, we add population genomic analyses using SNP data to reveal that sub-species level partitioning occurs between *S. 'fitti'* and its three Caribbean acroporid hosts (Figs. 2-4). Though adult colonies of Caribbean acroporids primarily associate with *S. 'fitti',* they can occasionally harbor other genera of Symbiodiniaceae (*Breviolum* spp., *Durusdinium trenchii* and *Cladocopium* spp.), but these associations are often transitory and revert to *S. 'fitti'* over time (Baums, Johnson, Devlin-Durante, & Miller 2010; Thornhill, LaJeunesse, Kemp, Fitt, & Schmidt 2006b).

The degree of selectivity of the cnidarian host when accepting a symbiotic partner likely has a role in maintaining the long-term specificity in these mutualisms. Host selectivity is in part modulated by cell recognition pathways such as lectin-glycan interactions and might contribute to the coevolution of *S. 'fitti'* and its respective host acroporids (Davy, Allemand, & Weis 2012;

Logan, LaFlamme, Weis, & Davy 2010; Parkinson et al. 2018; Weis, Reynolds, deBoer, & Krupp 2001; Wood-Charlson, Hollingsworth, Krupp, & Weis 2006). Though the *exact* role of cell-signaling in host selectivity has yet to be fully described, the specificity between partners is maintained by specialization pressures (LaJeunesse et al. 2018; Parkinson et al. 2018; Wood-Charlson et al. 2006). The ~160 million years following the widespread adaptive radiation of stony corals and Symbiodiniaceae has resulted in their current inter-dependence on one another (LaJeunesse et al. 2018). Consequently, the various symbiotic pairings remain adapted to meet the unique biochemical and metabolic demands of each host microenvironment (Barott, Venn, Perez, Tambutté, & Tresguerres 2015; Sogin, Anderson, Williams, Chen, & Gates 2014). The different morphologies and corallite structures possessed by each acroporid results in different light scattering properties and ultimately, light availability to the resident endosymbiont (Enríquez et al. 2017). These differences likely lead to strong selection pressures unique to each acroporid taxa. Ultimately, the specificity between *S. 'fitti'* strains and their acroporid hosts could be maintained by its ability (or lack thereof) to meet the metabolic needs of its host and adapt to the microenvironment created by its host.

The role of symbiont selectivity as a driver of S. 'fitti' *intraspecific genomic variation*

Akin to role of host selectivity, that of the symbiont is also important for upholding partner specificity in symbioses. The diversity and abundance of Symbiodiniaceae in the water column is not necessarily proportional to their counterparts partaking in mutualisms with reef-building corals (Cunning, Yost, Guarinello, Putnam, & Gates 2016; Littman, van Oppen, & Willis 2008; Manning & Gates 2008). However, some lineages of Symbiodiniaceae possess the ability to establish symbiosis with non-coral invertebrates (Cunning et al. 2016; Decelle et al.

433 2018). The total diversity of *S. 'fitti'* likely spans beyond strains that inhabit acroporids and may
434 incorporate free-living conspecifics living in the water column, sediments, etc. Therefore, the
435 slight differences in allelic composition of each *S. 'fitti'* strain may be a result of the differential
436 host preference of the available symbionts (Figs. 2-4). However, symbiotic *S. 'fitti'* strains might
437 be attracted to the different microbial composition and abundance in the water column adjacent
438 to a coral colony that constitute the 'ecosphere' surrounding each acroporid (Weber, González-
439 Díaz, Armenteros, & Apprill 2019). Similarly, the food associated with each 'ecospheres' may
440 attract different Symbiodiniaceae (Pollock et al. 2018; Weber et al. 2019). Putative intraspecific
441 variation in the swimming availability and chemosensory responses of *S. 'fitti'* may also, in part,
442 dictate which Symbiodiniaceae persist in each ecosphere (Fitt 1984; Fitt 1985; Fitt, Chang, &
443 Trench 1981; Kamykowski, Reed, & Kirkpatrick 1992). Future experimental validation of
444 intraspecific variation in *S. 'fitti'* swimming and chemosensory ability and how it pertains to
445 selectivity of their acroporid hosts (and 'ecospheres') will shed light on how inter-partner
446 specificity is maintained.
447
448 *Coevolution as a driver of* S. 'fitti' *intraspecific genomic variation*
449 Coevolution is the process by which two interacting species reciprocally adapt to each
450 other (sensu Janzen 1980). The coevolution of cnidarian-dinoflagellate mutualisms, in part, has
451 resulted in the long-term fidelity between partners (Figs. 2-4; LaJeunesse et al. 2018; Stanley
452 2006). Though the fossil record supports the coevolution of these mutualisms (Muscatine et al.
453 2005; Stanley 2006), experimental follow up and verification has received far less attention.
454 Advances in phylogenomics reveal extensive differentiation of genomic features and gene family
455 enrichment when comparing symbiotic and free-living *Symbiodinium* spp., a potential biproduct

of coevolving with their hosts (González-Pech, Bhattacharya, Ragan, & Chan 2019a; González-Pech, Ragan, & Chan 2017; González-Pech et al. 2019b). Similarly, population genetic analyses of host and symbiont reveals widespread long-term partnerships between several cnidarian species and their endosymbiotic dinoflagellates, which might be the biproduct of coevolution (Baums et al. 2014; O'Donnell et al. 2018; Poland & Coffroth 2017; Thornhill et al. 2006a; Thornhill et al. 2017; Thornhill et al. 2006b). Therefore, the correspondence of the allelic variation of *S. 'fitti'* to its host acroporids is likely, in part, the result of coevolution. The STAGdb genotyping array (SNPchip) can be harnessed to experimentally verify the contributions of coevolution, host specificity, and symbiont selectivity to the evolutionary dynamics of *Acropora*-*S. 'fitti'* symbioses' (Kitchen et al. 2020).

The special case of the F1 coral hybrid as habitat

The shared history (~2.6-3.6 million years of coexistence) between parents *A. palmata* and *A. cervicornis* with their symbiont, *S. 'fitti'* may have allowed sufficient time for co-evolutionary processes to play out (and so may help explain the strain differentiation of *S. 'fitti'* by host; Figs 2-4; Budd & Johnson 1999; McNeill et al. 1997). However, the situation differs for their first-generation hybrid, *A. prolifera*, which cannot directly respond to selection pressure from *S. 'fitti'* via differential successful sexual reproduction of its colonies because most are sterile (Vollmer & Palumbi 2002). Thus, any changes in the allele frequencies of *A. prolifera* are restricted to somatic mutations occurring within their lifetime which can be on the order of hundreds of years (Irwin et al. 2017). Though the fossil record of *A. prolifera* is rather sparse (McNeill et al. 1997), *S. 'fitti'* may have encountered these colonies over many thousands of

478 years as they are generated anew with each hybridization event. Thus, while a host co-adaptive
479 response is unlikely, *S. 'fitti'* may have evolved strains that preferentially colonize *A. prolifera*.
480
481 *Environmental differentiation as a driver of intraspecific genomic variation*
482 Environmental differences create variation in partner selectivity (Thrall, Hochberg,
483 Burdon, & Bever 2007). *S. 'fitti'* population dynamics are confounded with differences in host
484 habitat preferences including light, temperature, nutrient concentration, and food availability in
485 the water column (Crossland & Barnes 1983; Miller 1995; Terraneo et al. 2019; Williams et al.
486 2018). The inverse relationship between depth and light availability is a common driver of coral
487 and Symbiodiniaceae zonation (Bongaerts et al. 2015a; Bongaerts et al. 2015b; Bongaerts et al.
488 2017; Fogarty 2012; Goulet, Lucas, & Schizas 2019; LaJeunesse 2002; Serrano et al. 2014;
489 Serrano et al. 2016). Throughout much of their distribution, Caribbean acroporid species reside
490 in different habitats (Fogarty 2012; Goreau 1959). Specifically, *A. cervicornis* occupies a lower
491 light habitat (~10m depth) relative to its high-light dwelling (~3m depth) sibling species *A.*
492 *palmata* (Fogarty 2012; Goreau 1959; LaJeunesse 2002). Although the hybrid's depth
493 distribution often overlaps with *A. palmata,* it can also be found in less than 1m of water
494 (Fogarty 2012). Therefore, adaptation to different light availabilities may correspond to genomic
495 differentiation between the shallow *A. palmata- S. 'fitti'* and *A. prolifera- S. 'fitti'* versus deep *A.*
496 *cervicornis- S. 'fitti'* (Figs. 2-4; Finney et al. 2010; Kirk, Andras, Harvell, Santos, & Coffroth
497 2009). Microenvironments created by light attenuation at depth may lead to range-limited
498 dispersal of *S. 'fitti'* and modulate the available pool of symbionts (Finney et al. 2010; Serrano et
499 al. 2016). Furthermore, the differences in the skeletal morphology of acroporids result in
500 different light scattering properties that likely have profound effects on the light availability and

microenvironments for their resident *S. 'fitti'* (Enríquez et al. 2017; Gladfelter 1983; Gladfelter 2007). The different light regimes, skeletal features and flow fields associated with each acroporid may exert some selection pressure on symbiont strains.

Genomic basis for extended phenotypes in S. 'fitti' - *acroporid symbiosis*

The physiological capacity of the holobiont (coral and symbiont) hinges upon the specific partner pairings as well as external (environmental) drivers (Parkinson & Baums 2014). The large number of non-synonymous SNPs differentiating the *S. 'fitti'* strains among their hosts may demarcate the onset of eventual speciation (Fig. 5; Tables S7-10) although it is difficult to know what barriers to gene flow may exist that allow for such a process. Changes in amino acid sequences and protein efficiency resulting from these mutations may serve as the genomic basis causing intraspecific variation in physiological aptitude (Parkinson et al. 2015a; Parkinson & Baums 2014). Similarly, the candidate genes under selection identified by BayeScan and PCAdapt may underlie the strain differentiation by host species (Tables S7-10). Non-synonymous mutations in putative cytosolic oligopeptidase and alpha-agarase regions may result in the subtle alteration of zinc and calcium ion binding, respectively, which in turn likely contribute to variation in the physiological capacity of *S. 'fitti'* (Table S7-S10; Kmiec, Teixeira, Murcha, & Glaser 2016; Zhang et al. 2018). Further, these genotypic and phenotypic differences may facilitate the adaptation of each *S. 'fitti'* strain to the internal and external microenvironments associated with each host niche and meeting their metabolic and nutritional demands (Figs. 2-5; Tables S7-10; Hemond, Kaluziak, & Vollmer 2014; Muscatine, Porter, & Kaplan 1989; Reich, Rodriguez, LaJeunesse, & Ho 2020; Sogin et al. 2014).

Conclusion

We show here that the population genetic structure of *Symbiodinium 'fitti'* is, in part, explained by its host association. Because the host species occupy different habitats, we cannot yet disentangle the role of host versus depth as a potential driver of population genetic structure. However, the genomic resources for the *S. 'fitti'* - acroporid system described here can be used in future studies to determine whether, and to what degree, the observed variation of allelic composition is a result of host selectivity, symbiont selectivity, coevolution, environmental differentiation, or a combination of these mechanisms. The appreciation of population genetic structure and evolutionary dynamics of both coral holobiont partners will better inform the genomic underpinnings of their phenotypes and physiological capacity.

Acknowledgements

We thank the PSU genomics facility for assistance with library prep and sequencing. We thank Prof. Todd LaJeunesse for assistance with the *psbA* phylogeny. Funding for this project was supported by NSF-OCE-1537959 (to IBB) and NSF-OCE-1538469 (to NDF). HGR was supported through NSF-OCE-1636022 (to T. LaJeunesse). Permits for samples include Florida: CRF permit numbers CRF-2017-009, CRF-2017-012, NOAA FKNMS permit numbers FKNMS-2011-159-A4, FKNMS-2001-009, FKNMS-2014-148-A2, and FKNMS-2010-130-A, Belize: CITES Permit 0385, 7487 and 7488; Curacao: CITES Permit 16US784243/9 and 12US784243/9; and USVI Department of planning and natural resources, Division of fish and wildlife DFW14017T.

References

Abrego D., Van Oppen M., & Willis B. L. (2009) Onset of algal endosymbiont specificity varies among closely related species of Acropora corals during early ontogeny. *Molecular ecology* **18**, 3532-3543.

Altschul S. F., Madden T. L., Schäffer A. A., Zhang J., Zhang Z., Miller W., & Lipman D. J. (1997) Gapped BLAST and PSI-BLAST: a new generation of protein database search programs. *Nucleic acids research* **25**, 3389-3402.

Apweiler R., Bairoch A., Wu C. H., Barker W. C., Boeckmann B., Ferro S., . . . Magrane M. (2004) UniProt: the universal protein knowledgebase. *Nucleic acids research* **32**, D115-D119.

Aranda M., Li Y., Liew Y. J., Baumgarten S., Simakov O., Wilson M. C., . . . Voolstra C. R. (2016) Genomes of coral dinoflagellate symbionts highlight evolutionary adaptations conducive to a symbiotic lifestyle. *Scientific reports* **6**, 39734.

Baird A. H., Guest J. R., & Willis B. L. (2009) Systematic and biogeographical patterns in the reproductive biology of scleractinian corals. *Annual Review of Ecology, Evolution, and Systematics* **40**, 551-571.

Bairoch A., & Apweiler R. (1997) The SWISS-PROT protein sequence data bank and its supplement TrEMBL. *Nucleic acids research* **25**, 31-36.

Baker K. G., Radford D. T., Evenhuis C., Kuzhiumparam U., Ralph P. J., & Doblin M. A. (2018) Thermal niche evolution of functional traits in a tropical marine phototroph. *Journal of Phycology* **54**, 799-810.

Bankevich A., Nurk S., Antipov D., Gurevich A. A., Dvorkin M., Kulikov A. S., . . . Prjibelski A. D. (2012) SPAdes: a new genome assembly algorithm and its applications to single-cell sequencing. *Journal of Computational Biology* **19**, 455-477.

571 Barbrook A. C., Visram S., Douglas A. E., & Howe C. J. (2006) Molecular Diversity of
572 Dinoflagellate Symbionts of Cnidaria: The psbA Minicircle of Symbiodinium. *Protist*
573 **157**, 159-171.
574 Barneah O., Weis V. M., Perez S., & Benayahu Y. (2004) Diversity of dinoflagellate symbionts
575 in Red Sea soft corals: mode of symbiont acquisition matters. *Marine Ecology Progress*
576 *Series* **275**, 89-95.
577 Barott K. L., Venn A. A., Perez S. O., Tambutté S., & Tresguerres M. (2015) Coral host cells
578 acidify symbiotic algal microenvironment to promote photosynthesis. *Proceedings of the*
579 *National Academy of Sciences* **112**, 607-612.
580 Barshis D. J., Ladner J. T., Oliver T. A., Seneca F. O., Traylor-Knowles N., & Palumbi S. R.
581 (2013) Genomic basis for coral resilience to climate change. *Proceedings of the National*
582 *Academy of Sciences* **110**, 1387-1392.
583 Baums I. B., Baker A. C., Davies S. W., Grottoli A. G., Kenkel C. D., Kitchen S. A., . . . Shantz
584 A. A. (2019) Considerations for maximizing the adaptive potential of restored coral
585 populations in the western Atlantic. *Ecological Applications* **0**, e01978.
586 Baums I. B., Devlin-Durante M. K., & LaJeunesse T. C. (2014) New insights into the dynamics
587 between reef corals and their associated dinoflagellate endosymbionts from population
588 genetic studies. *Molecular ecology* **23**, 4203-4215.
589 Baums I. B., Johnson M. E., Devlin-Durante M. K., & Miller M. W. (2010) Host population
590 genetic structure and zooxanthellae diversity of two reef-building coral species along the
591 Florida Reef Tract and wider Caribbean. *Coral Reefs* **29**, 835-842.
592 Baums I. B., Miller M. W., & Hellberg M. E. (2005) Regionally isolated populations of an
593 imperiled Caribbean coral, Acropora palmata. *Molecular ecology* **14**, 1377-1390.

Baums I. B., Miller M. W., & Hellberg M. E. (2006) Geographic variation in clonal structure in a reef-building Caribbean coral, Acropora palmata. *Ecological Monographs* **76**, 503-519.

Bellantuono A. J., Dougan K. E., Granados-Cifuentes C., & Rodriguez-Lanetty M. (2019) Free-living and symbiotic lifestyles of a thermotolerant coral endosymbiont display profoundly distinct transcriptomes under both stable and heat stress conditions. *Molecular ecology* **28**, 5265-5281.

Boetzer M., & Pirovano W. (2014) SSPACE-LongRead: scaffolding bacterial draft genomes using long read sequence information. *BMC Bioinformatics* **15**, 211.

Bongaerts P., Carmichael M., Hay K. B., Tonk L., Frade P. R., & Hoegh-Guldberg O. (2015a) Prevalent endosymbiont zonation shapes the depth distributions of scleractinian coral species. *Royal Society open science* **2**, 140297.

Bongaerts P., Frade P. R., Hay K. B., Englebert N., Latijnhouwers K. R., Bak R. P., . . . Hoegh-Guldberg O. (2015b) Deep down on a Caribbean reef: lower mesophotic depths harbor a specialized coral-endosymbiont community. *Scientific reports* **5**, 7652.

Bongaerts P., Riginos C., Brunner R., Englebert N., Smith S. R., & Hoegh-Guldberg O. (2017) Deep reefs are not universal refuges: Reseeding potential varies among coral species. *Science Advances* **3**.

Budd A. F., & Johnson K. G. (1999) Origination preceding extinction during late Cenozoic turnover of Caribbean reefs. *Paleobiology* **25**, 188-200.

Buerger P., Alvarez-Roa C., Coppin C. W., Pearce S. L., Chakravarti L. J., Oakeshott J. G., . . . van Oppen M. J. H. (2020) Heat-evolved microalgal symbionts increase coral bleaching tolerance. *Science Advances* **6**, eaba2498.

616 Chakravarti L. J., & van Oppen M. J. H. (2018) Experimental Evolution in Coral
617 Photosymbionts as a Tool to Increase Thermal Tolerance. *Frontiers in Marine Science* **5**,
618 227.
619 Chan A. N., Lewis C. L., Neely K. L., & Baums I. B. (2019) Fallen Pillars: The Past, Present,
620 and Future Population Dynamics of a Rare, Specialist Coral-Algal Symbiosis. *Frontiers*
621 *in Marine Science* **6**, 218.
622 Chi J., Parrow M. W., & Dunthorn M. (2014) Cryptic Sex in Symbiodinium (Alveolata,
623 Dinoflagellata) is Supported by an Inventory of Meiotic Genes. *Journal of Eukaryotic*
624 *Microbiology* **61**, 322-327.
625 Coffroth M. A., Goulet T. L., & Santos S. R. (2001) Early ontogenic expression of specificity in
626 a cnidarian-algal symbiosis. *Mar Ecol Prog Ser* **222**.
627 Crossland C. J., & Barnes D. J. (1983) Dissolved nutrients and organic particulates in water
628 flowing over coral reefs at Lizard Island. *Marine and Freshwater Research* **34**, 835-844.
629 Cunning R., Gillette P., Capo T., Galvez K., & Baker A. (2015) Growth tradeoffs associated with
630 thermotolerant symbionts in the coral *Pocillopora damicornis* are lost in warmer oceans.
631 *Coral Reefs* **34**, 155-160.
632 Cunning R., Yost D. M., Guarinello M. L., Putnam H. M., & Gates R. D. (2016) Variability of
633 Symbiodinium Communities in Waters, Sediments, and Corals of Thermally Distinct
634 Reef Pools in American Samoa. *PloS one* **10**, e0145099.
635 Cziesielski M. J., Schmidt-Roach S., & Aranda M. (2019) The past, present, and future of coral
636 heat stress studies. *Ecology and Evolution* **9**, 10055-10066.
637 Danecek P., Auton A., Abecasis G., Albers C. A., Banks E., DePristo M. A., . . . Sherry S. T.
638 (2011) The variant call format and VCFtools. *Bioinformatics* **27**, 2156-2158.

639 Davy S. K., Allemand D., & Weis V. M. (2012) Cell Biology of Cnidarian-Dinoflagellate
640 Symbiosis. *Microbiology and Molecular Biology Reviews* **76**, 229-261.
641 De Baets G., Van Durme J., Reumers J., Maurer-Stroh S., Vanhee P., Dopazo J., . . . Rousseau F.
642 (2011) SNPeffect 4.0: on-line prediction of molecular and structural effects of protein-
643 coding variants. *Nucleic acids research* **40**, D935-D939.
644 Decelle J., Carradec Q., Pochon X., Henry N., Romac S., Mahé F., . . . de Vargas C. (2018)
645 Worldwide Occurrence and Activity of the Reef-Building Coral Symbiont Symbiodinium
646 in the Open Ocean. *Current Biology* **28**, 3625-3633.e3623.
647 Devlin-Durante M. K., & Baums I. B. (2017) Genome-wide survey of single-nucleotide
648 polymorphisms reveals fine-scale population structure and signs of selection in the
649 threatened Caribbean elkhorn coral, *Acropora palmata*. *PeerJ* **5**, e4077.
650 Dierckxsens N., Mardulyn P., & Smits G. (2016) NOVOPlasty: de novo assembly of organelle
651 genomes from whole genome data. *Nucleic Acids Research* **45**, e18-e18.
652 Emms D. M., & Kelly S. (2015) OrthoFinder: solving fundamental biases in whole genome
653 comparisons dramatically improves orthogroup inference accuracy. *Genome Biology* **16**,
654 157.
655 Enríquez S., Méndez E. R., Hoegh-Guldberg O., & Iglesias-Prieto R. (2017) Key functional role
656 of the optical properties of coral skeletons in coral ecology and evolution. *Proceedings of*
657 *the Royal Society B: Biological Sciences* **284**, 20161667.
658 Evanno G., Regnaut S., & Goudet J. (2005) Detecting the number of clusters of individuals using
659 the software STRUCTURE: a simulation study. *Molecular ecology* **14**, 2611-2620.
660 Finney J. C., Pettay D. T., Sampayo E. M., Warner M. E., Oxenford H. A., & LaJeunesse T. C.
661 (2010) The Relative Significance of Host–Habitat, Depth, and Geography on the

662 Ecology, Endemism, and Speciation of Coral Endosymbionts in the Genus
663 Symbiodinium. *Microbial Ecology* **60**, 250-263.
664 Fischer M. C., Foll M., Excoffier L., & Heckel G. (2011) Enhanced AFLP genome scans detect
665 local adaptation in high-altitude populations of a small rodent (Microtus arvalis).
666 *Molecular ecology* **20**, 1450-1462.
667 Fitt W. K. (1984) The role of chemosensory behavior of Symbiodinium microadriaticum,
668 intermediate hosts, and host behavior in the infection of coelenterates and molluscs with
669 zooxanthellae. *Marine Biology* **81**, 9-17.
670 Fitt W. K. (1985) CHEMOSENSORY RESPONSES OF THE SYMBIOTIC
671 DINOFLAGELLATE SYMBIODINIUM MICROADRIATICUM (DINOPHYCEAE)1.
672 *Journal of Phycology* **21**, 62-67.
673 Fitt W. K., Chang S. S., & Trench R. K. (1981) Motility patterns of different strains of the
674 symbiotic dinoflagellate Symbiodinium (= Gymnodinium) microadriaticum
675 (Freudenthal) in culture. *Bulletin of Marine Science* **31**, 436-443.
676 Fitt W. K., & Trench R. K. (1983) The relation of diel patterns of cell division to diel patterns of
677 motility in the symbiotic dinoflagellate Symbiodinium microadria ticum Freudenthal in
678 culture. *New Phytologist* **94**, 421-432.
679 Fogarty N. D. (2012) Caribbean acroporid coral hybrids are viable across life history stages.
680 *Marine Ecology Progress Series* **446**, 145-159.
681 Foll M., Fischer M. C., Heckel G., & Excoffier L. (2010) Estimating population structure from
682 AFLP amplification intensity. *Molecular ecology* **19**, 4638-4647.

683 Foll M., & Gaggiotti O. (2008) A Genome-Scan Method to Identify Selected Loci Appropriate
684 for Both Dominant and Codominant Markers: A Bayesian Perspective. *Genetics* **180**,
685 977-993.
686 Gladfelter E. H. (1983) Skeletal development inAcropora cervicornis. *Coral Reefs* **2**, 91-100.
687 Gladfelter E. H. (2007) Skeletal development in Acropora palmata (Lamarck 1816): a scanning
688 electron microscope (SEM) comparison demonstrating similar mechanisms of skeletal
689 extension in axial versus encrusting growth. *Coral Reefs* **26**, 883-892.
690 González-Pech R. A., Bhattacharya D., Ragan M. A., & Chan C. X. (2019a) Genome Evolution
691 of Coral Reef Symbionts as Intracellular Residents. *Trends in Ecology & Evolution*.
692 González-Pech R. A., Ragan M. A., & Chan C. X. (2017) Signatures of adaptation and symbiosis
693 in genomes and transcriptomes of Symbiodinium. *Scientific reports* **7**, 15021.
694 González-Pech R. A., Stephens T. G., Chen Y., Mohamed A. R., Cheng Y., Burt D. W., . . .
695 Chan C. X. (2019b) Structural rearrangements drive extensive genome divergence
696 between symbiotic and free-living <em>Symbiodinium</em&gt. *bioRxiv*,
697 783902.
698 Goreau T. F. (1959) The ecology of Jamaican coral reefs I. Species composition and zonation.
699 *Ecology* **40**, 67-90.
700 Goulet T. L. (2006) Most corals may not change their symbionts. *Marine Ecology Progress*
701 *Series* **321**, 1-7.
702 Goulet T. L., Lucas M. Q., & Schizas N. V. (2019) Symbiodiniaceae Genetic Diversity and
703 Symbioses with Hosts from Shallow to Mesophotic Coral Ecosystems. In: *Mesophotic*
704 *Coral Ecosystems* (eds. Loya Y, Puglise KA, Bridge TCL), pp. 537-551. Springer
705 International Publishing, Cham.

706 Hemond E. M., Kaluziak S. T., & Vollmer S. V. (2014) The genetics of colony form and
707 function in Caribbean *Acropora* corals. *BMC Genomics* **15**, 1133.
708 Hemond E. M., & Vollmer S. V. (2010) Genetic diversity and connectivity in the threatened
709 staghorn coral (Acropora cervicornis) in Florida. *PloS one* **5**, e8652-e8652.
710 Irwin A., Greer L., Humston R., Devlin-Durante M., Cabe P., Lescinsky H., . . . Baums I. B.
711 (2017) Age and intraspecific diversity of resilient *Acropora* communities in Belize. *Coral*
712 *Reefs* **36**, 1111-1120.
713 Janzen D. H. (1980) When is it coevolution?
714 Kamvar Z. N., Tabima J. F., & Grünwald N. J. (2014) Poppr: an R package for genetic analysis
715 of populations with clonal, partially clonal, and/or sexual reproduction. *PeerJ* **2**, e281.
716 Kamykowski D., Reed R. E., & Kirkpatrick G. J. (1992) Comparison of sinking velocity,
717 swimming velocity, rotation and path characteristics among six marine dinoflagellate
718 species. *Marine Biology* **113**, 319-328.
719 Kemp D. W., Thornhill D. J., Rotjan R. D., Iglesias-Prieto R., Fitt W. K., & Schmidt G. W.
720 (2015) Spatially distinct and regionally endemic Symbiodinium assemblages in the
721 threatened Caribbean reef-building coral Orbicella faveolata. *Coral Reefs* **34**, 535-547.
722 Kirk N. L., Andras J. P., Harvell C. D., Santos S. R., & Coffroth M. A. (2009) Population
723 structure of Symbiodinium sp. associated with the common sea fan, Gorgonia ventalina,
724 in the Florida Keys across distance, depth, and time. *Marine Biology* **156**, 1609-1623.
725 Kitchen S. A., Ratan A., Bedoya-Reina O. C., Burhans R., Fogarty N. D., Miller W., & Baums I.
726 B. (2019) Genomic Variants Among Threatened Acropora Corals. *G3:*
727 *Genes|Genomes|Genetics* **9**, 1633.

728 Kitchen S. A., Von Kuster G., Kuntz K. L. V., Reich H. G., Miller W., Griffin S., . . . Baums I.
729 B. (2020) STAGdb: a 30K SNP genotyping array and Science Gateway for Acropora
730 corals and their dinoflagellate symbionts. *Scientific reports* **10**, 12488.
731 Kmiec B., Teixeira P. F., Murcha M. W., & Glaser E. (2016) Divergent evolution of the M3A
732 family of metallopeptidases in plants. *Physiologia Plantarum* **157**, 380-388.
733 Knaus B. J., & Grünwald N. J. (2017) vcfr: a package to manipulate and visualize variant call
734 format data in R. *Molecular Ecology Resources* **17**, 44-53.
735 Kumar S., Jones M., Koutsovoulos G., Clarke M., & Blaxter M. (2013) Blobology: exploring
736 raw genome data for contaminants, symbionts and parasites using taxon-annotated GC-
737 coverage plots. *Frontiers in Genetics* **4**, 237.
738 LaJeunesse T. C. (2001) Investigating the biodiversity, ecology, and phylogeny of
739 endosymbiotic dinoflagellates in the genus Symbiodinium using the ITS region: in search
740 of a "species" level marker. *Journal of Phycology* **37**, 866-880.
741 LaJeunesse T. C. (2002) Diversity and community structure of symbiotic dinoflagellates from
742 Caribbean coral reefs. *Marine Biology* **141**, 387-400.
743 LaJeunesse T. C., Parkinson J. E., Gabrielson P. W., Jeong H. J., Reimer J. D., Voolstra C. R., &
744 Santos S. R. (2018) Systematic Revision of Symbiodiniaceae Highlights the Antiquity
745 and Diversity of Coral Endosymbionts. *Current Biology*.
746 LaJeunesse T. C., Thornhill D. J., Cox E. F., Stanton F. G., Fitt W. K., & Schmidt G. W. (2004)
747 High diversity and host specificity observed among symbiotic dinoflagellates in reef
748 coral communities from Hawaii. *Coral Reefs* **23**, 596-603.
749 Langmead B., & Salzberg S. L. (2012) Fast gapped-read alignment with Bowtie 2. *Nature*
750 *methods* **9**, 357.

751 Lee S. Y., Jeong H. J., Kang N. S., Jang T. Y., Jang S. H., & Lajeunesse T. C. (2015)
752 Symbiodinium tridacnidorum sp. nov., a dinoflagellate common to Indo-Pacific giant
753 clams, and a revised morphological description of Symbiodinium microadriaticum
754 Freudenthal, emended Trench & Blank. *European Journal of Phycology* **50**, 155-172.
755 Levin R. A., Beltran V. H., Hill R., Kjelleberg S., McDougald D., Steinberg P. D., & van Oppen
756 M. J. H. (2016) Sex, scavengers, and chaperones: transcriptome secrets of divergent
757 symbiodinium thermal tolerances. *Molecular biology and evolution*, msw119.
758 Lewis A. M., Chan A. N., & LaJeunesse T. C. (2019) New Species of Closely Related
759 Endosymbiotic Dinoflagellates in the Greater Caribbean have Niches Corresponding to
760 Host Coral Phylogeny. *Journal of Eukaryotic Microbiology*.
761 Li H. (2013) Aligning sequence reads, clone sequences and assembly contigs with BWA-MEM.
762 *arXiv preprint arXiv:1303.3997*.
763 Li H., Handsaker B., Wysoker A., Fennell T., Ruan J., Homer N., . . . Genome Project Data
764 Processing S. (2009) The Sequence Alignment/Map format and SAMtools.
765 *Bioinformatics* **25**, 2078-2079.
766 Liew Y. J., Aranda M., & Voolstra C. R. (2016) Reefgenomics.Org - a repository for marine
767 genomics data. *Database* **2016**.
768 Lin S., Cheng S., Song B., Zhong X., Lin X., Li W., . . . Morse D. (2015) The *Symbiodinium*
769 *kawagutii* genome illuminates dinoflagellate gene expression and coral symbiosis.
770 *Science* **350**, 691.
771 Little A. F., Oppen M. J., & Willis B. L. (2004) Flexibility in algal symbioses shapes growth in
772 reef corals. *Science* **304**, 1492-1494.

773 Littman R. A., van Oppen M. J. H., & Willis B. L. (2008) Methods for sampling free-living
774 Symbiodinium (zooxanthellae) and their distribution and abundance at Lizard Island
775 (Great Barrier Reef). *Journal of Experimental Marine Biology and Ecology* **364**, 48-53.
776 Logan D. D. K., LaFlamme A. C., Weis V. M., & Davy S. K. (2010) FLOW-CYTOMETRIC
777 CHARACTERIZATION OF THE CELL-SURFACE GLYCANS OF SYMBIOTIC
778 DINOFLAGELLATES (SYMBIODINIUM SPP.)1. *Journal of Phycology* **46**, 525-533.
779 Luo R., Liu B., Xie Y., Li Z., Huang W., Yuan J., . . . Liu Y. (2015) SOAPdenovo2: an
780 empirically improved memory-efficient short-read de novo assembler. *GigaScience* **4**,
781 s13742-13015.
782 Luu K., Bazin E., & Blum M. G. B. (2017) pcadapt: an R package to perform genome scans for
783 selection based on principal component analysis. *Molecular Ecology Resources* **17**, 67-
784 77.
785 Manning M. M., & Gates R. D. (2008) Diversity in populations of free-living Symbiodinium
786 from a Caribbean and Pacific reef. *Limnology and Oceanography* **53**, 1853.
787 Martin M. (2011) Cutadapt removes adapter sequences from high-throughput sequencing reads.
788 *EMBnet. journal* **17**, 10-12.
789 McNeill D. F., Budd A. F., & Borne P. F. (1997) Earlier (late Pliocene) first appearance of the
790 Caribbean reef-building coral Acropora palmata: Stratigraphic and evolutionary
791 implications. *Geology* **25**, 891-894.
792 Miller M. W. (1995) Growth of a temperate coral: effects of temperature, light, depth, and
793 heterotrophy. *Marine Ecology Progress Series* **122**, 217-225.

794 Mungpakdee S., Shinzato C., Takeuchi T., Kawashima T., Koyanagi R., Hisata K., . . . Shoguchi
795 E. (2014) Massive Gene Transfer and Extensive RNA Editing of a Symbiotic
796 Dinoflagellate Plastid Genome. *Genome biology and evolution* **6**, 1408-1422.
797 Muscatine L., Goiran C., Land L., Jaubert J., Cuif J.-P., & Allemand D. (2005) Stable isotopes
798 (δ¹³C and δ¹⁵N) of organic matrix
799 from coral skeleton. *Proceedings of the National Academy of Sciences of the United*
800 *States of America* **102**, 1525.
801 Muscatine L., Porter J. W., & Kaplan I. R. (1989) Resource partitioning by reef corals as
802 determined from stable isotope composition. *Marine Biology* **100**, 185-193.
803 Narasimhan V., Danecek P., Scally A., Xue Y., Tyler-Smith C., & Durbin R. (2016)
804 BCFtools/RoH: a hidden Markov model approach for detecting autozygosity from next-
805 generation sequencing data. *Bioinformatics* **32**, 1749-1751.
806 O'Donnell K. E., Lohr K. E., Bartels E., Baums I. B., & Patterson J. T. (2018) Acropora
807 cervicornis genet performance and symbiont identity throughout the restoration process.
808 *Coral Reefs* **37**, 1109-1118.
809 Parkinson J. E., Banaszak A. T., Altman N. S., LaJeunesse T. C., & Baums I. B. (2015a)
810 Intraspecific diversity among partners drives functional variation in coral symbioses.
811 *Scientific reports* **5**, 15667.
812 Parkinson J. E., & Baums I. B. (2014) The extended phenotypes of marine symbioses: ecological
813 and evolutionary consequences of intraspecific genetic diversity in coralâ€“algal
814 associations. *Frontiers in microbiology* **5**, 445.
815 Parkinson J. E., Coffroth M. A., & LaJeunesse T. C. (2015b) New species of Clade B
816 *Symbiodinium* (Dinophyceae) from the greater Caribbean belong to different functional

guilds: *S. aenigmaticum* sp. nov., *S. antillogorgium* sp. nov., *S. endomadracis* sp. nov., and *S. pseudominutum* sp. nov. *Journal of Phycology* **51**, 850-858.

Parkinson J. E., Tivey T. R., Mandelare P. E., Adpressa D. A., Loesgen S., & Weis V. M. (2018) Subtle Differences in Symbiont Cell Surface Glycan Profiles Do Not Explain Species-Specific Colonization Rates in a Model Cnidarian-Algal Symbiosis. *Frontiers in microbiology* **9**, 842.

Pettay D. T., Wham D. C., Smith R. T., Iglesias-Prieto R., & LaJeunesse T. C. (2015) Microbial invasion of the Caribbean by an Indo-Pacific coral zooxanthella. *Proceedings of the National Academy of Sciences of the United States of America* **112**, 7513-7518.

Pinzón J. H., Kamel B., Burge C. A., Harvell C. D., Medina M., Weil E., & Mydlarz L. D. (2015) Whole transcriptome analysis reveals changes in expression of immune-related genes during and after bleaching in a reef-building coral. *Royal Society Open Science* **2**.

Pochon X., Putnam H. M., Burki F., & Gates R. D. (2012) Identifying and characterizing alternative molecular markers for the symbiotic and free-living dinoflagellate genus Symbiodinium. *PloS one* **7**, e29816-e29816.

Poland D. M., & Coffroth M. A. (2017) Trans-generational specificity within a cnidarian–algal symbiosis. *Coral Reefs* **36**, 119-129.

Pollock F. J., McMinds R., Smith S., Bourne D. G., Willis B. L., Medina M., . . . Zaneveld J. R. (2018) Coral-associated bacteria demonstrate phylosymbiosis and cophylogeny. *Nature Communications* **9**, 4921.

Precht W. F., Vollmer S. V., Modys A. B., & Kaufman L. (2019) Fossil Acropora prolifera (Lamarck, 1816) reveals coral hybridization is not only a recent phenomenon. *Proceedings of the Biological Society of Washington* **132**, 40-55.

840 Pritchard J. K., Stephens M., & Donnelly P. (2000) Inference of population structure using
841 multilocus genotype data. *Genetics* **155**, 945-959.
842 Quinlan A. R., & Hall I. M. (2010) BEDTools: a flexible suite of utilities for comparing genomic
843 features. *Bioinformatics* **26**, 841-842.
844 Reich H. G., Rodriguez I. B., LaJeunesse T. C., & Ho T.-Y. (2020) Endosymbiotic
845 dinoflagellates pump iron: differences in iron and other trace metal needs among the
846 Symbiodiniaceae. *Coral Reefs*.
847 Serrano X., Baums I., O'Reilly K., Smith T., Jones R., Shearer T., . . . Baker A. (2014)
848 Geographic differences in vertical connectivity in the Caribbean coral *Montastraea*
849 *cavernosa* despite high levels of horizontal connectivity at shallow depths. *Molecular*
850 *ecology* **23**, 4226-4240.
851 Serrano X., Baums I. B., Smith T. B., Jones R. J., Shearer T. L., & Baker A. C. (2016) Long
852 distance dispersal and vertical gene flow in the Caribbean brooding coral *Porites*
853 *astreoides*. *Scientific reports* **6**.
854 Shah S., Chen Y., Bhattacharya D., & Chan C. X. (2020) Sex in Symbiodiniaceae
855 dinoflagellates: genomic evidence for independent loss of the canonical synaptonemal
856 complex. *Scientific reports* **10**, 9792.
857 Shinzato C., Shoguchi E., Kawashima T., Hamada M., Hisata K., Tanaka M., . . . Satoh N.
858 (2011) Using the *Acropora digitifera* genome to understand coral responses to
859 environmental change. *Nature* **476**, 320-323.
860 Shoguchi E., Beedessee G., Tada I., Hisata K., Kawashima T., Takeuchi T., . . . Shinzato C.
861 (2018) Two divergent Symbiodinium genomes reveal conservation of a gene cluster for
862 sunscreen biosynthesis and recently lost genes. *BMC genomics* **19**, 458.

Shoguchi E., Shinzato C., Kawashima T., Gyoja F., Mungpakdee S., Koyanagi R., . . . Fujiwara M. (2013) Draft assembly of the Symbiodinium minutum nuclear genome reveals dinoflagellate gene structure. *Current Biology* **23**, 1399-1408.

Silverstein R. N., Cunning R., & Baker A. C. (2017) Tenacious D: Symbiodinium in clade D remain in reef corals at both high and low temperature extremes despite impairment. *The Journal of Experimental Biology*.

Simão F. A., Waterhouse R. M., Ioannidis P., Kriventseva E. V., & Zdobnov E. M. (2015) BUSCO: assessing genome assembly and annotation completeness with single-copy orthologs. *Bioinformatics* **31**, 3210-3212.

Sogin E. M., Anderson P., Williams P., Chen C.-S., & Gates R. D. (2014) Application of 1H-NMR Metabolomic Profiling for Reef-Building Corals. *PloS one* **9**, e111274.

Stamatakis A. (2014) RAxML version 8: a tool for phylogenetic analysis and post-analysis of large phylogenies. *Bioinformatics* **30**, 1312-1313.

Stanke M., Keller O., Gunduz I., Hayes A., Waack S., & Morgenstern B. (2006) AUGUSTUS: ab initio prediction of alternative transcripts. *Nucleic acids research* **34**, W435-W439.

Stanke M., Steinkamp R., Waack S., & Morgenstern B. (2004) AUGUSTUS: a web server for gene finding in eukaryotes. *Nucleic acids research* **32**, W309-W312.

Stanley G. D. (2006) Photosymbiosis and the Evolution of Modern Coral Reefs. *Science* **312**, 857.

Tamura K., Stecher G., Peterson D., Filipski A., & Kumar S. (2013) MEGA6: molecular evolutionary genetics analysis version 6.0. *Molecular Biology and Evolution* **30**, 2725-2729.

885 Terraneo T. I., Fusi M., Hume B. C. C., Arrigoni R., Voolstra C. R., Benzoni F., . . . Berumen M.
886 L. (2019) Environmental latitudinal gradients and host-specificity shape Symbiodiniaceae
887 distribution in Red Sea Porites corals. *Journal of Biogeography* **46**, 2323-2335.
888 Thornhill D. J., Fitt W. K., & Schmidt G. W. (2006a) Highly stable symbioses among western
889 Atlantic brooding corals. *Coral Reefs* **25**, 515-519.
890 Thornhill D. J., Howells E. J., Wham D. C., Steury T. D., & Santos S. R. (2017) Population
891 genetics of reef coral endosymbionts (Symbiodinium, Dinophyceae). *Molecular ecology*
892 **26**, 2640-2659.
893 Thornhill D. J., LaJeunesse T. C., Kemp D. W., Fitt W. K., & Schmidt G. W. (2006b) Multi-
894 year, seasonal genotypic surveys of coral-algal symbioses reveal prevalent stability or
895 post-bleaching reversion. *Marine Biology* **148**, 711-722.
896 Thrall P. H., Hochberg M. E., Burdon J. J., & Bever J. D. (2007) Coevolution of symbiotic
897 mutualists and parasites in a community context. *Trends in Ecology & Evolution* **22**, 120-
898 126.
899 Trench R. K. (1979) The cell biology of plant-animal symbiosis. *Annual Review of Plant*
900 *Physiology* **30**, 485-531.
901 UniProt C. (2014) UniProt: a hub for protein information. *Nucleic acids research* **43**, D204-
902 D212.
903 Van Oppen M. J. H., Willis B. L., Vugt H. V., & Miller D. J. (2000) Examination of species
904 boundaries in the Acropora cervicornis group (Scleractinia, Cnidaria) using nuclear DNA
905 sequence analyses. *Molecular ecology* **9**, 1363-1373.
906 Vollmer S. V., & Palumbi S. R. (2002) Hybridization and the Evolution of Reef Coral Diversity.
907 *Science* **296**, 2023.

908 Vollmer S. V., & Palumbi S. R. (2006) Restricted Gene Flow in the Caribbean Staghorn Coral
909 Acropora cervicornis: Implications for the Recovery of Endangered Reefs. *Journal of*
910 *Heredity* **98**, 40-50.
911 Warren R. L., Yang C., Vandervalk B. P., Behsaz B., Lagman A., Jones S. J. M., & Birol I.
912 (2015) LINKS: Scalable, alignment-free scaffolding of draft genomes with long reads.
913 *GigaScience* **4**, 35.
914 Weber L., González-Díaz P., Armenteros M., & Apprill A. (2019) The coral ecosphere: A
915 unique coral reef habitat that fosters coral-microbial interactions. *Limnology and*
916 *Oceanography* **9999**, 1-16.
917 Weis V. M., Reynolds W. S., deBoer M. D., & Krupp D. A. (2001) Host-symbiont specificity
918 during onset of symbiosis between the dinoflagellates Symbiodinium spp. and planula
919 larvae of the scleractinian coral Fungia scutaria. *Coral Reefs* **20**, 301-308.
920 Williams G. J., Sandin S. A., Zgliczynski B. J., Fox M. D., Gove J. M., Rogers J. S., . . . Smith J.
921 E. (2018) Biophysical drivers of coral trophic depth zonation. *Marine Biology* **165**, 60.
922 Wood-Charlson E. M., Hollingsworth L. L., Krupp D. A., & Weis V. M. (2006) Lectin/glycan
923 interactions play a role in recognition in a coral/dinoflagellate symbiosis. *Cellular*
924 *Microbiology* **8**, 1985-1993.
925 Xu H., Luo X., Qian J., Pang X., Song J., Qian G., . . . Chen S. (2012) FastUniq: a fast de novo
926 duplicates removal tool for paired short reads. *PloS one* **7**, e52249.
927 Zhang W., Xu J., Liu D., Liu H., Lu X., & Yu W. (2018) Characterization of an α-agarase from
928 Thalassomonas sp. LD5 and its hydrolysate. *Applied Microbiology and Biotechnology*
929 **102**, 2203-2212.
930

Data availability

Sequences are under NCBI SRA PRJNA473816. Code for data analysis and figure generation is available on github (https://github.com/hgreich/Sfitti).

Author contributions

Conceived the project: SAK, IBB, HGR. Obtained funding: IBB, NDF. Mentorship: IBB, SAK. Field collections of corals: SAK, IBB, NDF. Molecular work: SAK, MDD. Bioinformatic analyses: HGR, SAK, KHS. Wrote the paper: HGR, SAK, IBB, NDF, KHS.

Tables

Table S1: Sample information including host taxa, location, sampling depth, sequencing platform, and accession number. Per sample information on read counts, mapping rates, and SNP summary statistics are also included. Multi-locus genotypes from 12 microsatellite loci, 58,538 "high quality" SNPs, and 6,813 "conservative" SNPs with no missing data are included.

Table S2: Genome assembly summary statistics for *Symbiodinium 'fitti'* sample PFL14120.

Table S3: Annotation information for the 24,000+ genes in *Symbiodinium 'fitti.'*

Table S4: Annotation information and selection outlier statistics for *Symbiodinium 'fitti'* SNPS (12,700) that are predicted to fall in coding regions.

954 Table S5: AMOVA of indicates that ~12% of *Symbiodinium 'fitti'* variation is due to host taxon
955 whereas negative % variation is due to location.
956
957 Table S6: Summary statistics from Structure Harvester including the Evanno's *delta* K method
958 which predicted three main clusters.
959
960 Table S7: Selection outlier statistics and per SNP summary statistics loci identified as selection
961 outliers by BayeScan (370) and PCAdapt (4,987). 307 selection outlier SNPs were shared
962 between the two programs.
963
964 Table S8: Summary statistics for the predicted downstream effects of SNPs (generated by the
965 SNPeff program). 60,373 (84.43%) modifier/non-coding, 3,629 (5.08%) moderate/mostly
966 harmless, 7,451 (10.42%) low impact variants that might change protein efficiency/effectiveness,
967 and 51 (0.07%) highly disruptive SNPs were identified.
968
969 Table S9: Predicted downstream effects (SnpEff) and annotation information for variants in
970 coding regions.
971
972 Table S10: Predicted downstream effects (SnpEff) and annotation information for 12 variants in
973 coding regions that were identified as selection outliers by both PCAdapt and BayeScan.
974
975 **Figures**
976

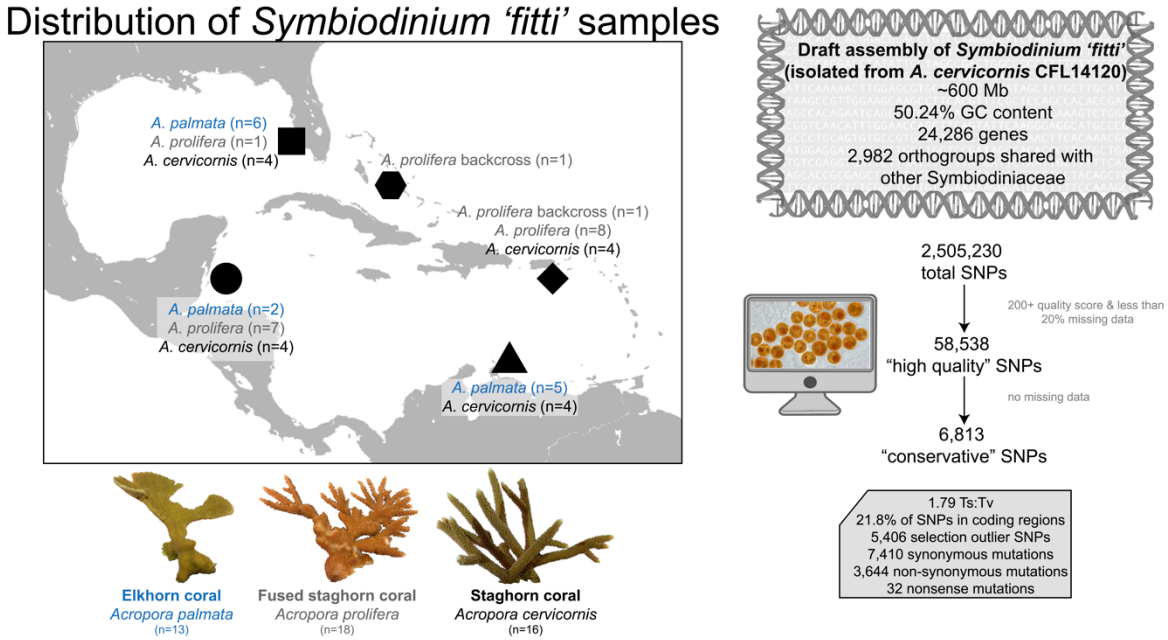
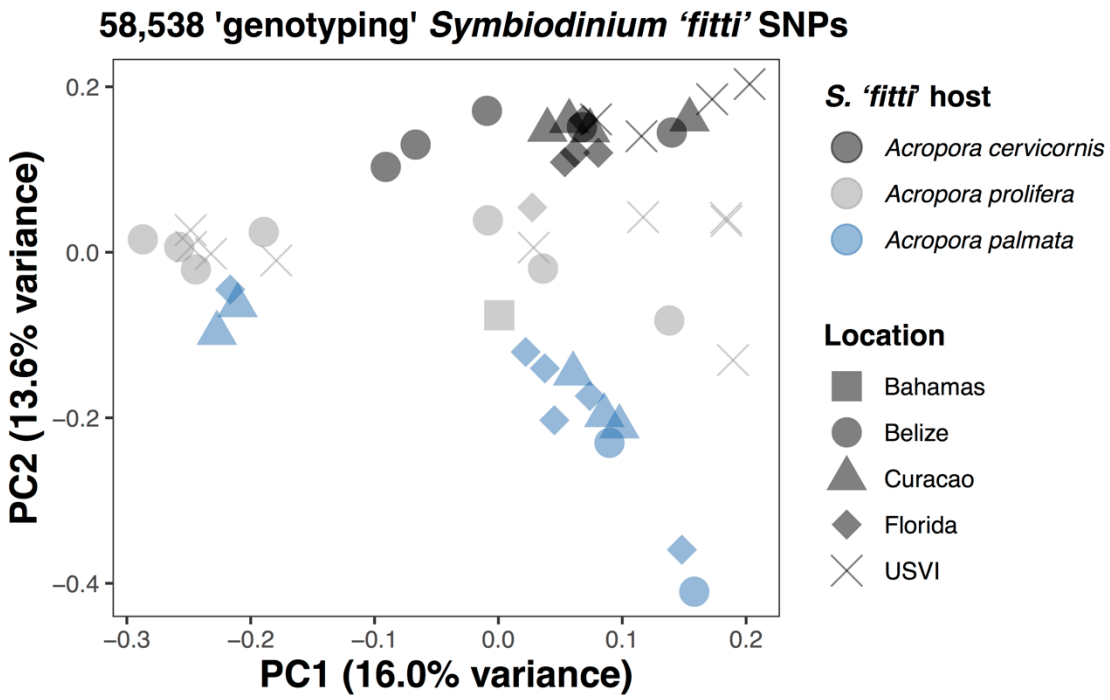


Fig. 1: Sampling design and summary statistics of population genomic approach used to characterize *Symbiodinium 'fitti'* across three coral hosts. *Acropora palmata* (n = 13), *Acropora cervicornis* (n = 16), and *Acropora prolifera* (n = 18) samples used for shallow genome sequencing spanned their geographic distribution. Summary statistics for the 'deeply-sequenced' CFL 14120 *A. cervicornis – S. 'fitti'* draft genome assembly including overall length, %GC content, # of genes, and shared gene families with other Symbiodiniaceae. Summary statistics and visual depiction of quality filtering work flow that was employed to identify high quality variants and those that are under selection. Coral images from N. Fogarty and I. Baums.



991

992 Fig. 2: Principal component analysis (PCA) of 58,583 genotyping *Symbiodinium 'fitti'* SNPs

993 illustrates genomic differentiation by host taxon. Coral images from N. Fogarty and I. Baums.

994

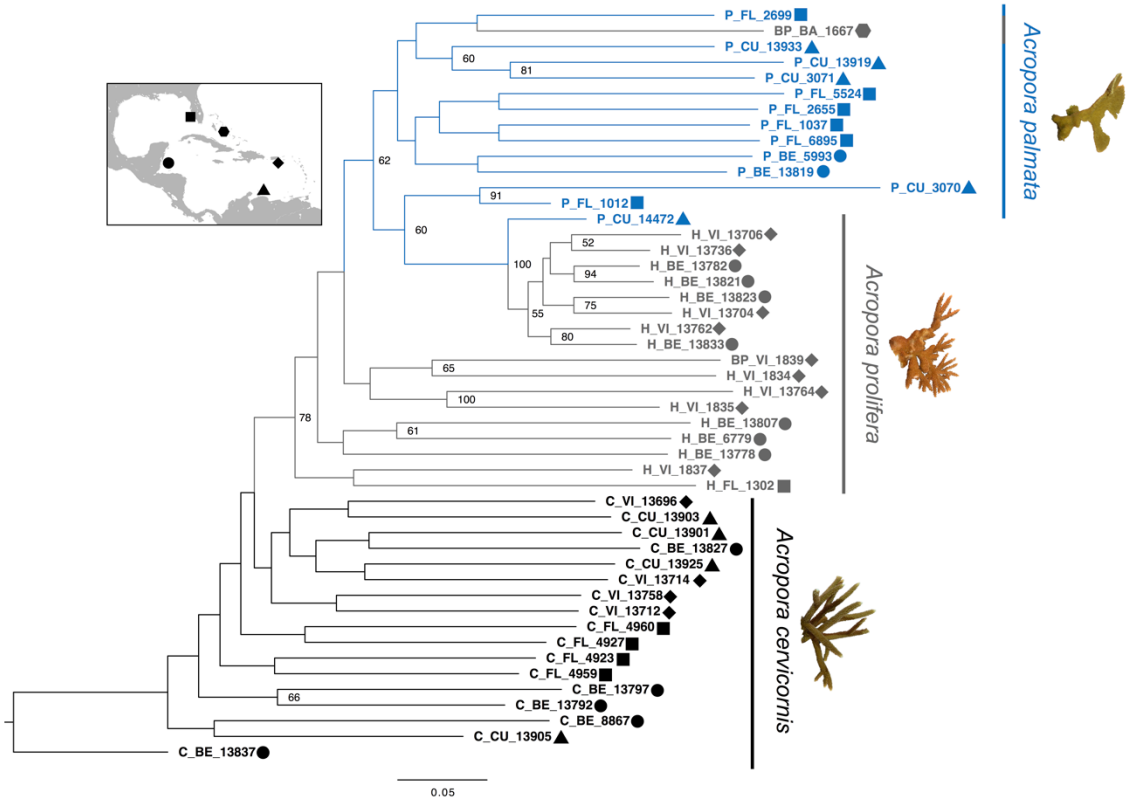
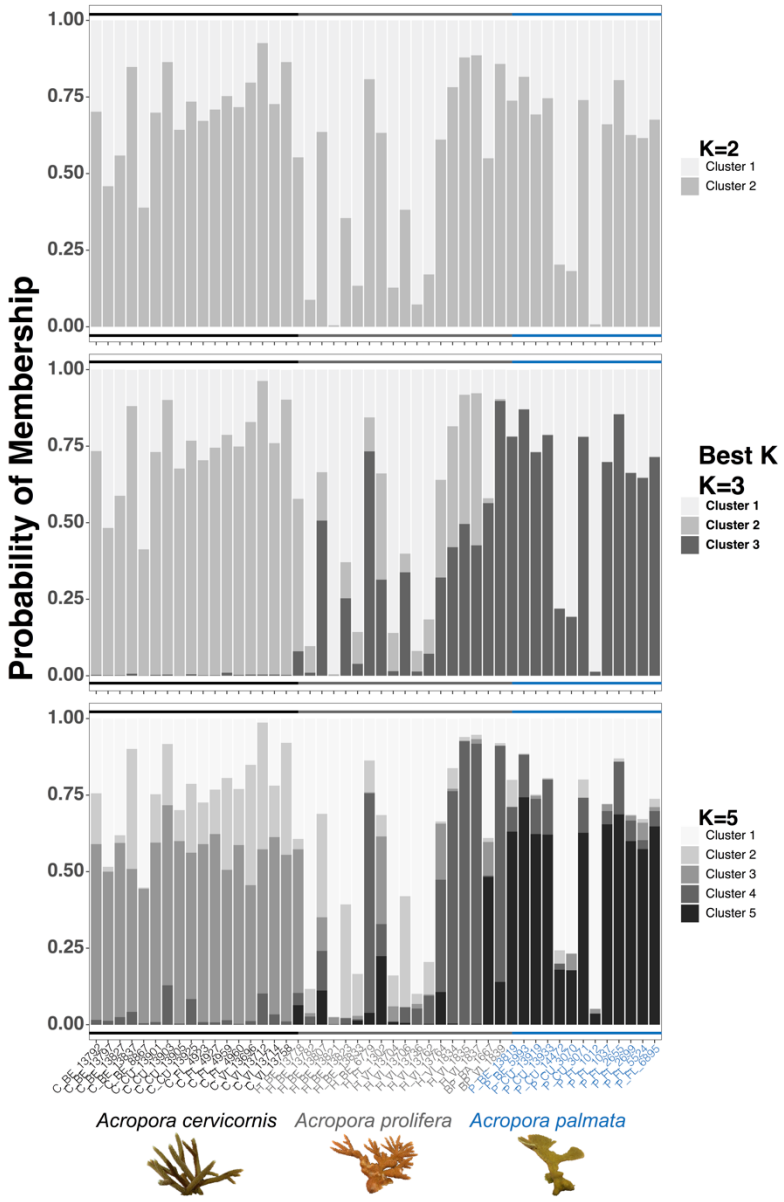


Fig. 3: Allelic composition of *Symbiodinium 'fitti'* is at the sub-species level. RAxML (Maximum Likelihood) phylogeny of 6,813 "conservative" genotyping *S. 'fitti'* SNPs without missing data and 100 bootstrap replicates illustrate *S. 'fitti'* genomic differentiation by host taxon.



1008

1009 Fig. 4: *Symbiodinium 'fitti'* strain population assignment aligns with host taxon. Probability of

1010 membership predicted by STRUCTURE for 58,583 "high quality" *Symbiodinium 'fitti'* SNPS.

1011 K=3 was determined as best K. K=2 and K=5 are also presented for comparison. These results

1012 illustrate that *S. 'fitti'* membership clusters largely correspond to their host acroporid taxon.

1013 Coral images from N. Fogarty and I. Baums.

1014

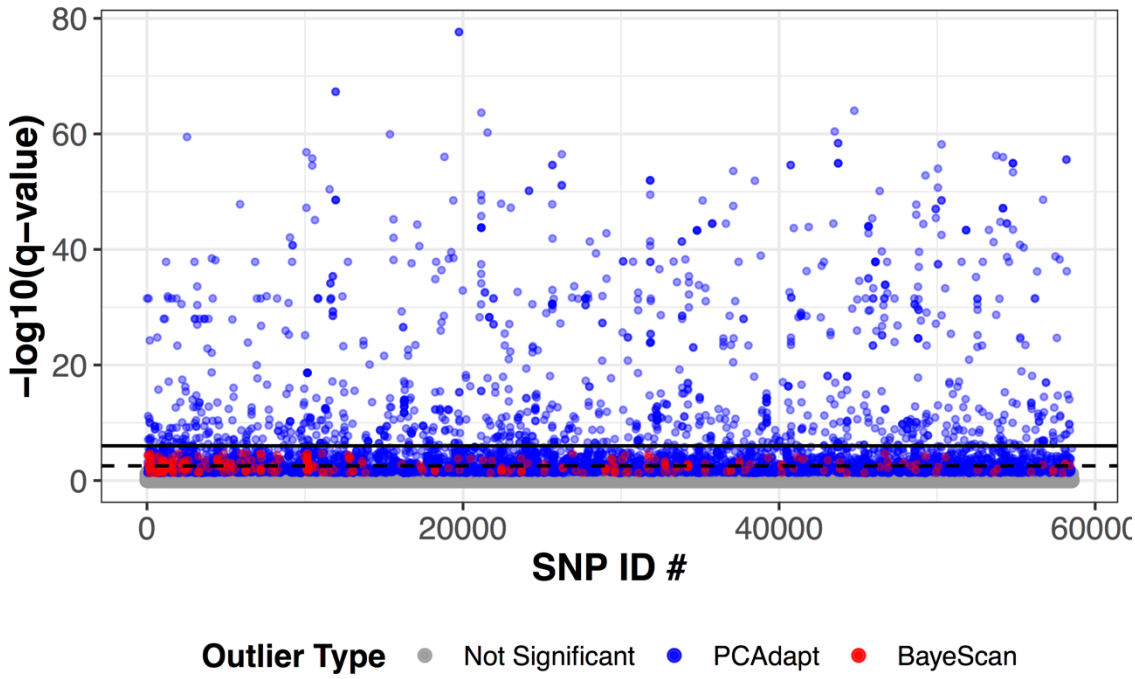


Fig. 5: Genetic variants of *S. 'fitti'* showing signatures of selection. Manhattan plot of –log10 transformed q-values for 58,583 genotyping SNPs. SNPs are highlighted by outlier detection program (PCAdapt loci highlighted in blue; BayeScan loci highlighted in red). 339 selection outlier SNPs were shared between the two programs. 103 outlier loci identified by BayeScan had a Bayes probability of 1 and q-value of 0 which becomes infinite following logarithmic transformation and were therefore removed from the plot. These loci were also had high BayeScan fixation levels between host and location. The dashed line represents the 5% FDR adjustment threshold (2.54) and whereas the solid line represents the 0.05 Bonferroni correction threshold (6.01).

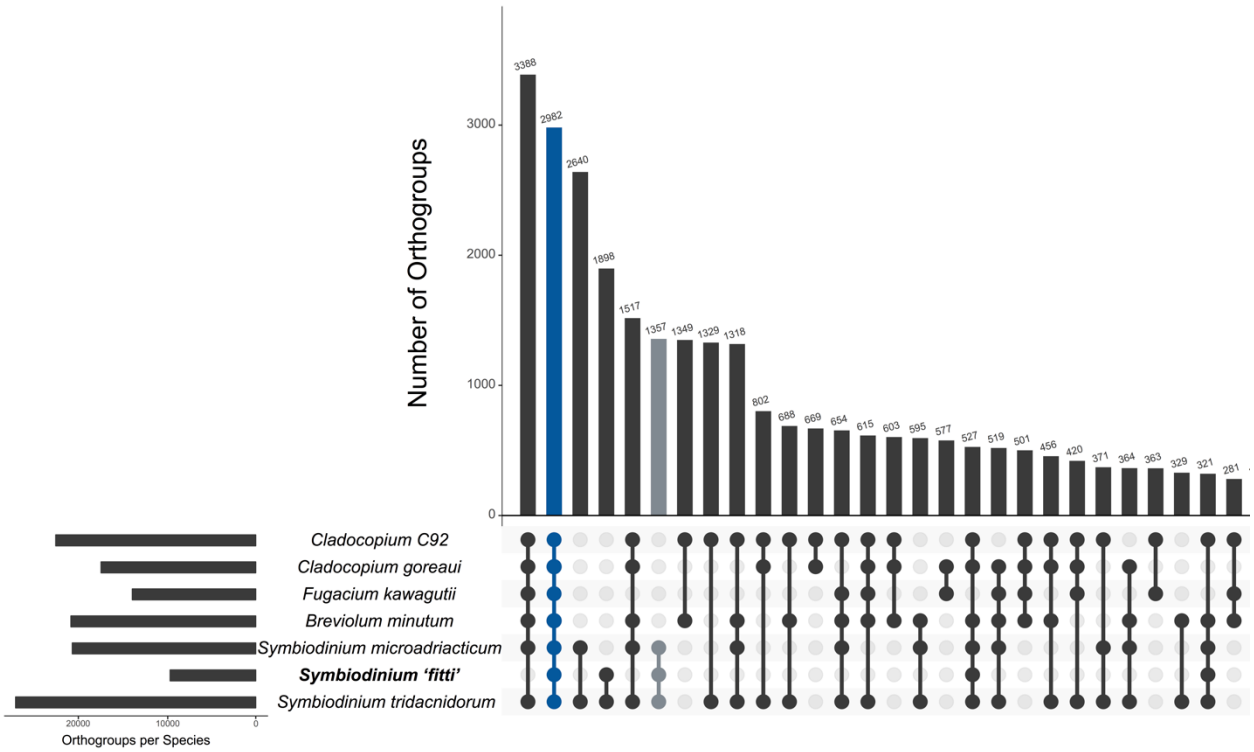
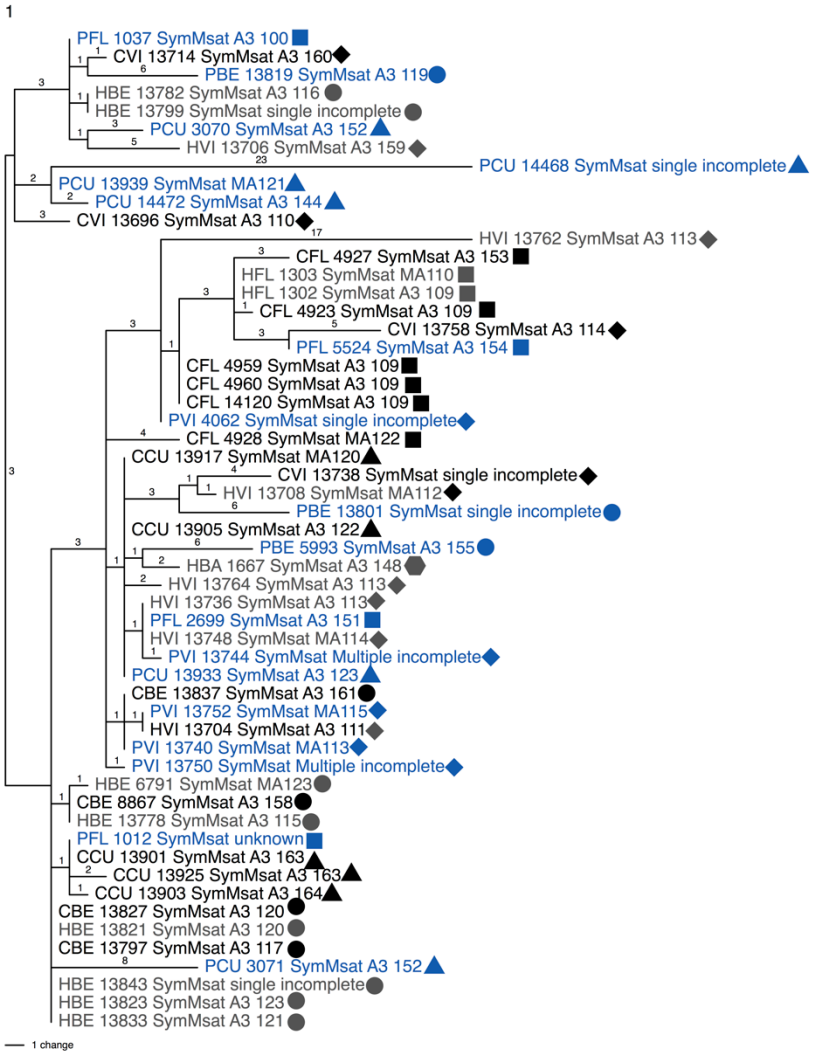


Fig. S1: Gene families shared between different Symbiodiniaceae lineages (including *Symbiodinium 'fitti'*).



1043

1044 Fig. S2: Intragenomic variation of *Symbiodinium 'fitti'* is at the sub-species level. Phylogeny of

1045 the *psbA* non-coding region, a commonly used marker used to help delimit *Symbiodinium* species

1046 indicates that *S. 'fitti'* in its three different hosts are all one species.

1047

1048

1049

1050

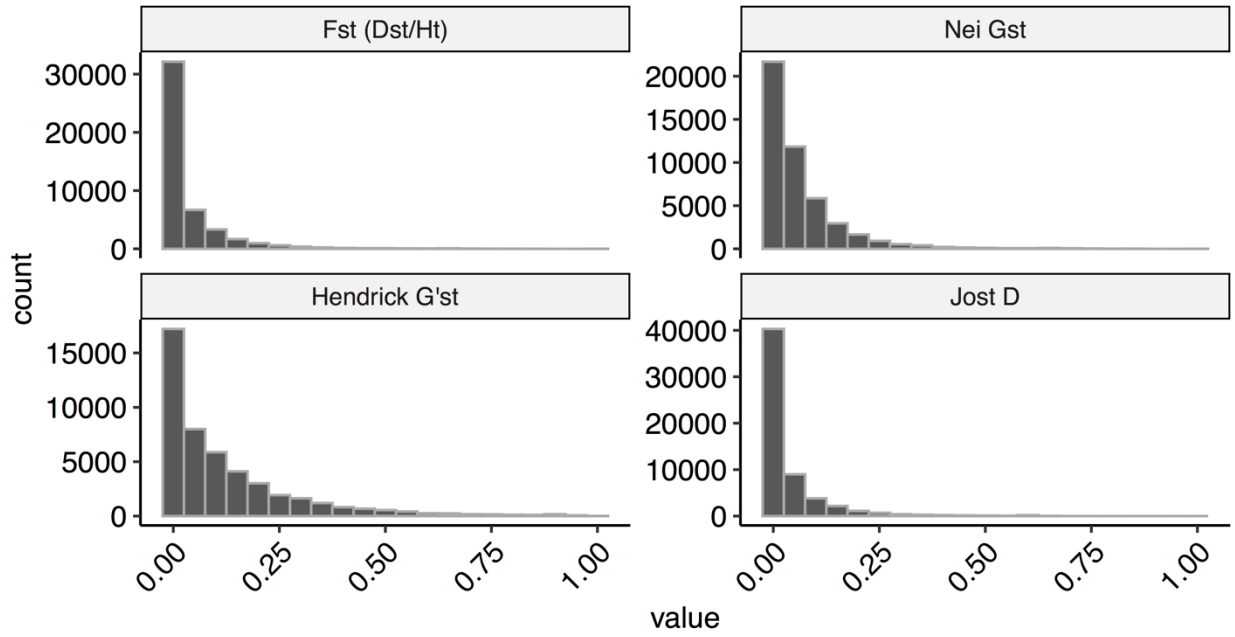


Fig. S3: Histogram of per SNP fixation levels for 58,538 genotyping SNPs.