Title

Deep Amplicon Sequencing Quantitatively Detected Mixed Community Assemblages of Symbiodinium in Orbicella faveolata and Orbicella franksi

Permalink

https://escholarship.org/uc/item/31n4975j

Author

Green, Elizabeth

Publication Date

2014

Peer reviewed|Thesis/dissertation

UNIVERSITY OF CALIFORNIA,
MERCED

Deep Amplicon Sequencing Quantitatively Detected Mixed Community Assemblages of
Symbiodinium in *Orbicella faveolata* and *Orbicella franksi*

THESIS

submitted in partial satisfaction of the requirements
for the degree of

MASTER OF SCIENCE

in

Quantitative and Systems Biology

by

Elizabeth A. Green

Committee in charge:
David Ardell, chair
Miriam Barlow
Mónica Medina
Michele Weber

2014

The thesis of Elizabeth A. Green is approved, and it is acceptable in quality and form for

publication on microfilm and electronically:

Miriam Barlow

Mónica Medina

Michele Weber

David Ardell   Chair

University of California, Merced
2014

Dedication

This thesis is dedicated to my loving and supportive husband, Colten Green.

Table of Contents

LIST OF FIGURES

Page

LIST OF TABLES

ACKNOWLEDGEMENTS

I also would like to acknowledge all my lab members for their continued patience and willingness to provide wisdom. Bishoy Kamel provided insight to troubleshoot laboratory complications and bioinformatics support all of which was always delivered with an encouraging smile and accompanying parable. Collin Closek for sharing his organization tricks and lending a listening ear throughout my entire graduate experience. Erika Diaz for her permeating low stress personality to keep me calm and keep my priorities aligned. Anke Klueter for her encouraging work ethic and willingness to share her expertise. Michele Weber never ceased to lend a practical approach to overcome obstacles and impart her *Symbiodinium* expertise. I also acknowledge previous lab members Emmanuel Buschiazzo without which I never would have met Dr. Medina and Kristen Marhaver for editing documents and encouraging bold approaches when optimizing laboratory procedures.

I would also like to thank my committee members Dr. David Ardell and Dr. Miriam Barlow for devoting an exhaustive amount of time and energy to this thesis and me as a graduate student.

Lastly, I acknowledge my advisor, Dr. Mónica Medina who had faith in me even when I did not. Her advising skills have nurtured me as an aspiring scientist and imparted a renewed level of confidence in myself as an individual. Her patience and guidance throughout this entire process will never be forgotten.

ABSTRACT

Deep Amplicon Sequencing Quantitatively Detected Mixed Community Assemblages of
Symbiodinium in *Orbicella faveolata* and *Orbicella franksi*

by

Elizabeth A. Green[1]

Master of Science in Quantitative and System Biology

University of California, Merced

Professor Mónica Medina[2]

1. Quantitative and Systems Biology, University of California, Merced, 5200 North
    Lake Road, Merced, CA 95343, USA

2. Biology Department, Pennsylvania State University, 326 Mueller, University
    Park, PA 16801, USA

The taxonomic diversity of endosymbiont *Symbiodinium* species and the
physiological contributions they confer to coral host species are challenging to assess. It
seems likely the genetic diversity of obligate *Symbiodinium* species modulates the
physiological performance of reef building corals. Here we used deep amplicon
sequencing to reveal high genetic diversity of the endosymbiotic
dinoflagellate, *Symbiodinium*, in coral hosts by amplifying ribosomal DNA internal
transcribed spacer region 2 (ITS-2). We revealed quantitative proportions of
mixed *Symbiodinium* populations within the *Orbicella* species complex collected from
three Caribbean locations. *Symbiodinium* diversity at the Flower Garden Banks
for *Orbicella franksi* and *Orbicella faveolata* demonstrated coral host species contained
Symbiodinium clade B type B1 and detected five genetically variable B1 *Symbiodinium*
types, three of which may be endemic to the Flower Garden Banks locations. Our second

dataset assessed *Symbiodinium* diversity in *O. faveolata* collected from Puerto Morelos, Mexico, previously genotyped using restriction fragment length polymorphism (RFLP) and direct sequencing. We detected an additional background population of clade A, type A13. Additionally we showed one colony hosted a mixed infection of *Symbiodinium* genotypes B1 and A13 where RFLP and direct sequencing assigned genotypes C7 and B1. Our third dataset assessed *Symbiodinium* diversity across a depth gradient in *O. faveolata* collected from Curaçao using the Illumina MiSeq platform. We showed shallow water coral hosted a more complex mixed infection relative to hosts from deeper waters and detected *Symbiodinium* genotype G3, a genotype not commonly associated with scleractinian hosts. These results highlight the need for consistent molecular genotyping techniques to assess community assemblages of *Symbiodinium*-host relationships. This deep sequencing approach used to characterize cryptic genetic diversity of *Symbiodinium* will potentially contribute to the understanding of physiological variations among coral populations.

Introduction

Many tropical scleractinian corals rely on an obligate symbiosis with the dinoflagellate algae in the genus *Symbiodinium* to provide it with photosynthetic products critical for metabolic processes and to a lesser extent for calcification (Muscatine & Cernichiari 1969; Muscatine et al. 1984; reviewed by Trench 1979; Trench 1987). However, we are still learning about the establishment and plasticity of this complex symbiosis between host and *Symbiodinium*. Recent advancements in molecular genotyping techniques have revealed unique *Symbiodinium* genotypes within the nine accepted clades thereby increasing our knowledge of taxonomic diversity (Coffroth & Santos 2005; Pochon & Gates 2010). These novel *Symbiodinium* genotypes may impart different physiological benefits to the overall health and response mechanisms of coral hosts (Rowan et al. 1997; Sampayo et al., 2008; Voolstra et al. 2009).

In an established symbiosis the coral host provides shelter and inorganic nutrients for the obligate endosymbiont (Pearse & Muscatine 1971; reviewed by Trench 1979). *Symbiodinium* provide the host with photosynthetic products mainly in the form of glycerol (Muscatine 1967) but also organic carbon (Muscatine et al. 1984), amino acids (Swanson & Hoegh-Guldberg 1998) and energy for calcification (Hoegh-Guldberg 2004; Pearse & Muscatine 1971). Environmental factors such as increasing ocean acidification rates, reduced water quality from increased land based run off and higher degree heating days for longer durations contribute to coral stress (Smith & Buddemeier 1992). A malfunctioning symbiosis results in a stress response known as a bleaching event where brown photosynthetic algae are expelled resulting in a white coloration of the coral

(Brown 1997; Glynn 1993; Hoegh-Guldberg 1999). Much focus has been to delineate upper and lower thermal thresholds in order to predict bleaching events. However, defining one cause for a bleaching event remains inconclusive and bleaching is likely the result of a suite of variables contributing to the loss of photosynthetic algae (Fitt et al. 2001). Corals are extremely sensitive to thermal fluctuations showing one to two degrees Celsius above average seasonal temperatures for one to two weeks likely results in a bleaching response (Fitt et al. 2001; Glynn 1993; Hoegh-Guldberg 2004), but other biological contributors to overall reef health should also be considered when defining stress cues to predict bleach responses. Current climate projections predict severe increases in temperature for longer durations, increasing carbon emissions and ocean acidification rates, variable irradiance and salinity levels that exceed corals' ability to acclimate thereby contributing to more frequent bleaching events (Hoegh-Guldberg et al. 2007; Hoegh-Guldberg & Smith 1989). Whether or not corals are able to recover is a result of the duration and severity of the stress event (Lang et al. 1992). As a non-motile, long lived and slow growing colonial metazoan the ability of corals to acclimate quickly seems unlikely thereby decreasing biodiversity beyond repair (Hughes et al. 2003). A better understanding of the flexibility of symbiosis between corals capable of hosting a mixed infection (Douglas 1998; LaJeunesse et al. 2003) versus corals with strict specificity (Diekmann 2002; Sampayo et al. 2007) for one symbiont type will contribute to a better understanding of coral-*Symbiodinium* relationships to survive quick and extreme environmental perturbations.

Currently it is still challenging to understand global distributions of endosymbionts, however this knowledge is critical to assess coral reef resilience. It was initially thought there was only one *Symbiodinium* species, *S. microadriaticum* (Freudenthal 1962). The advent of DNA sequencing and molecular genotyping techniques revealed the complex taxonomic diversity of this genus. Currently there are nine accepted clades of *Symbiodinium* (A-I) (Pochon et al. 2010) and within each of these major clades are subclades. Clade A, the most ancient lineage likely diverged about 50 MYA (Pochon et al. 2006). Clades E, G, D and B were likely the next clades to diverge sometime throughout the Eocene and Oligocene epochs (Pochon et al. 2006). The youngest clades C and H diverged in the last 15 MYA (Pochon et al. 2006). It is essential to develop molecular genotyping methods capable of detecting diversity at the subclade level consistently and quantitatively across a diverse range of host species. The three coral host species within the *Orbicella annularis* species complex each typically host a mixed population of *Symbiodinium* species at different proportions (Rowan & Knowlton 1995; Rowan et al. 1997). The functional and genetic diversity of these *Symbiodinium* lineages and how they correlate with host physiology is unclear (Baker 2003; Knowlton & Rohwer 2003). *Symbiodinium* species have shown varying photosynthetic efficiency and saturation points suggesting coral host physiology is at least partially dependent on symbiotic interactions (Baums et al. 2010; DeSalvo et al. 2010; Fitt & Warner 1995; Warner et al. 1996). Some *Symbiodinium* physiological traits that benefit the coral host for some genotypes within clades have been proposed, but rarely are these consistent across all lineages within an entire clade and require more investigation. Clade A,

commonly found in shallow water corals, has been shown to endure higher light intensities and provide UV protection (Reynolds et al. 2008; Rowan et al. 1997; Toller et al. 2001). A few clade B genotypes have shown to withstand colder temperatures by returning to normal photosynthetic activity post extended cold stress events (Thornhill et al. 2008). Clade C, the most diverse lineage (Cantin et al. 2009; LaJeunesse 2005) has shown to confer more

^{14}C-labeled photosynthate to coral host resulting in increased calcification rates and carbon input (Cantin et al. 2009). Members of clade D have been researched for their opportunistic reputation to re-populate bleached corals and be more stress tolerant (Berkelmans & van Oppen 2006; Chen et al. 2003; Stat & Gates 2011). However, clade D also has shown negative long-term impacts such as reduced growth rates (Little et al. 2004; Stat & Gates 2011). Clades F, G, H and I typically associate more with foraminifera and will not be discussed further here (Pochon & Gates 2010; Pochon et al. 2004; Pochon et al. 2006; Pochon et al. 2001).

Detecting *Symbiodinium* taxonomic diversity across numerous diverse sites and host species will help to close the gap between the functional and genetic diversity of *Symbiodinium* species. This will allow us to expand our knowledge of coral physiology and the plasticity of symbiosis. Experiments prior to 1993 used coarse resolution genotyping techniques likely detecting only dominant genotypes in a mixed *Symbiodinium* population within corals (Baker 2003; LaJeunesse 2005; Mieog et al. 2007; Rowan & Powers 1992; Stat et al. 2006). Use of molecular genotyping techniques within the last two decades paired with growing databases of phylogenetic markers have

provided support for various hosts to house mixed *Symbiodinium* populations and more

consistently identified novel genetic haplotypes previously underestimated likely due to

coarse genotyping techniques (Baird et al. 2007; Baker & Romanski 2007; Fay & Weber

2012; LaJeunesse 2002; Rowan et al. 1997). The rapidly advancing field of next

generation sequencing (NGS) platforms provides an exciting future for increasing the

detection limits for low frequency *Symbiodinium* species in a mixed community (Kenkel

et al. 2013; Quigley et al. 2014). Detecting these novel haplotypes in mixed communities

can enhance our understanding of *Symbiodinium* physiology and their roles for hosts as

well as define their biogeographical distributions (Jones & Berkelmans 2010; Mieog et

al. 2009). This highlights the importance of using molecular genotyping techniques that

accurately quantify *Symbiodinium* diversity in order to evaluate the different coral-algal

symbioses and the ability of corals to survive increasingly stressful environmental

conditions.

This study applied targeted amplicon sequencing of the internal transcribed spacer

region 2 (ITS-2) nuclear ribosomal DNA to detect *Symbiodinium* diversity within the

endangered *Orbicella annularis* (IUCN 2011), formerly known as a member of the genus

Montastraea (Budd et al. 2012), species complex using the Roche 454 GS FLX and

Illumina MiSeq. We first assessed the *Symbiodinium* diversity at the Flower Garden

Banks between the Caribbean corals *O. faveolata* and *O. franksi* across two locations,

east and west, using the Roche 454 GS FLX platform to determine whether

Symbiodinium diversity would vary more relative to host genetic structure or geographic

location. We also evaluated *Symbiodinium* species diversity in *O. faveolata* from Puerto

Morelos, Mexico previously genotyped using restriction fragment length polymorphism (RFLP) and direct sequencing (DeSalvo et al. 2010) for a thermal stress-recovery experiment to compare genotyping approaches. Lastly, we genotyped *Symbiodinium* species in *O. faveolata* from Curaçao collected across a depth gradient using the Illumina MiSeq to investigate whether species diversity decreased with respect to depth.

Methods
Locations

The remote Flower Garden Banks (FGB) is a National Marine Sanctuary established in 1992 located 185 kilometers off the coast of Texas (27°54'N, 93°35'W East Flower Garden Banks and 27°53'N, 93°49'W West Flower Garden Banks) (Schmahl et al. 2008). The east and west banks are separated by 19 kilometers. Flower Garden Banks (FGB) are the most northern coral reefs in the Gulf of Mexico making it an important location to understand limits of latitudinal distributions of coral species (Schmahl et al. 2008). The FGB has lower diversity compared to other Caribbean reefs but higher coral cover ranging between 50% and 70% (Precht et al. 2005). FGB are also a deep reef with reef tops starting at 17 meters extending beyond 50 meters (Schmahl et al. 2008). Temperature variations expose reefs to their thermal minimums with ranges between 18°C and 30°C, contrary to most other Caribbean reefs which are studied for exceeding their thermal maximums (Schmahl et al. 2008). The Flower Garden Banks National Marine Sanctuary (FGBNMS) is still home to 24 shallow water (<50 meters) scleractinian coral species (Schmahl et al. 2008) and provides a unique opportunity to assess *Symbiodinium* distributions in some of the most pristine coral reefs throughout the Caribbean.

The National Autonomous University of Mexico (UNAM) located in Puerto

Morelos, Mexico provides easy access to the Mesoamerican reef tract. Reef crests and

back reef zones display an abundant coral composition with few recorded bleach

events (Rodríguez-Martinez, R. 1993).

The Curaçao Marine Park was established in 1983 in response to reports of

significantly declining coral reef health provided by the Caribbean Research and

Management of Biodiversity (CARMIBI) (CARMIBI researchstationcarmibi.org). Unlike

many other Caribbean reefs that are heavily influenced by land-based runoff, the island of

Curaçao has relatively few inhabited areas resulting in minimal pollution runoff

(CARMIBI researchstationcarmibi.org). The less populated north and west facing sides

of the island show reduced anthropogenic influences resulting in a diverse coral reef

system within swimming distance from the beach (CARMIBI

researchstationcarmibi.org). Additionally reefs exhibit coral cover greater than 70% in

some areas (CARMIBI researchstationcarmibi.org). Overall coral reef health near

CARMIBI has suffered from four recorded bleach events since the 1980s. Most recently

a devastating hurricane system left its mark on this diverse reef (Bries et al. 1999).

DNA extractions

One hundred ninety-seven 1cm *x* 1cm coral fragments were collected from the

outer edge of each colony at both east and west Flower Garden Banks (FGB) in August

2011 (*Orbicella faveolata*, n=96) and August 2012 (*O. franksi*, n=101) (Figure 1 and 2).

Sample depth ranged from 21 to 23 meters. Tissue was preserved in 96% ethanol and

stored at room temperature. The coral holobiont is comprised of a coral host, obligate

dinoflagellate species in the genus *Symbiodinium* and an associated microbiome that includes bacteria, archaea, viruses and fungi (Rowher et al. 2002) and likely other undiscovered organisms. Instead of isolating *Symbiodinum*, holobiont DNA extraction began by covering cores with a tissue digest buffer (100 mM NaCl, 10 mM Tris-Cl pH 8.0, 25 mM EDTA pH 8.0, 0.5% SDS, 0.1 mg ml^{-1} Proteinase K, and 1 µgml^{-1} RNaseA) for one hour at 42°C followed by a standard phenol-chloroform extraction protocol (Chomczynski & Sacchi 2006; Davies et al. 2013). All data regarding the FGB study will furthermore be referred to as the FGB dataset.

Thirty *O. faveolata* individuals from Puerto Morelos, Quintana Roo near Mexico (20°52 '28.77"N 86°51'04.53"W) were collected on July 31st, 2007 and used for a thermal

stress-recovery experiment (DeSalvo et al. 2010) (Figure 1). DNA was re-extracted from ground coral powder using the Power Plant® (MoBio Laboratories, Carlsbad, CA) DNA isolation kit for eleven individuals (Control samples: 1-3 and 5, Bleached samples: 1,2,4,5 Recovery samples: 1-3). DNA from the original experiment was used for control sample four and bleached sample three, which also used a Power Plant® (MoBio Laboratories, Carlsbad, CA) DNA isolation kit for DNA extractions (Table 1). These will furthermore be referred to as the Mexico dataset.

Five *O. faveolata* colonies from three water depth gradients, shallow (7-7.3 meters), medium (11.5-13 meters) and deep (15.5-16.5 meters) were collected from West Point (12°37'72"N, 69°16'13"W) in Curaçao in April 2011 (n=15) (Figure 1, Table 2). Coral cores 20 centimeters in diameter coral were flash-frozen in a dry shipper and stored

at -80°C. Coral tissue was ground using a mortar and pestle. DNA was extracted from coral tissue using a modified protocol for the Power Plant® DNA isolation kit (MoBio Laboratories, Carlsbad, CA) in August 2013. Approximately 100 mg of coral powder was homogenized with Power Plant Bead Solution and 10 U/μL of Ready-Lyse Lysozyme. Samples were placed on a rotator block at room temperature for ten minutes. Twenty mg/mL of Proteinase K was added and samples were incubated for 90 minutes at 65°C. Two sizes of zirconia/silica beads (400 mg each of 0.1 mm and 0.5 mm) were added and homogenized for 30 s using a Mini- BeadBeater-16 (BioSpec Product Catalog #607). Remnant purification steps followed the Power Plant® (MoBio) protocol. Yields were quantified using a NanoDrop 2000 spectrophotometer (Thermo Scientific). These will furthermore be referred to as the Curaçao dataset.

Host Genotype Analysis

FGB Dataset

One hundred ninety-three FGB coral hosts successfully amplified at nine microsatellite loci following methods described in Davies (et al. 2013). Amplicons were analyzed on ABI 3130XL capillary sequencer at the University of Texas, Austin Genomic Sequencing and Analysis Facility using a ROX-labeled size standard. Files were downloaded from the facility and uploaded into GeneMarker 1.90 (SoftGenetics) to call alleles (Davies et al. 2013). Allele calls were binned using a customized Perl script (Supplementary Data 1) and formatted for downstream analysis. STRUCTURE (version 2.3.4) output (q-scores) (Falush et al. 2003; Falush et al. 2007; Hubisz et al. 2009; Pritchard et al. 2000) was used to avoid selecting potential hybrid colonies. Hybrids from

the *Obicella annularis* species complex have been reported in literature (Budd & Pandolfi 2004; Fukami et al. 2004; Szmant et al. 1997). Only individuals with greater than 80% posterior probability of belonging to one of the two major STRUCTURE-derived clusters were retained (73 individuals of *O. faveolata* and 101 individuals of *O. franksi*) (Foster et al. 2012). Sixty of these, fifteen colonies of *O. faveolata* and fifteen *O. franksi* from both east and west FGB, were chosen for *Symbiodinium* ITS-2 genotyping. Five replicate STRUCTURE (version 2.3.4) analyses were run with a burn-in of 300,000 steps, 10^6 Markov-Chain Monte Carlo (MCMC) iterations, starting with a uniform alpha for degree of admixture, uncorrelated allele frequencies and no location prior. STRUCTURE results were used as input to STRUCTURE HARVESTER to select the optimal number of clusters (K) (Earl & vonHoldt 2012; Evanno et al. 2005). Output files from STRUCTURE HARVESTER were then used as input for CLUMPP (Jakobsson & Rosenberg 2007) to combine results of replicate runs and compute a weighted average. CLUMPP outputs were used as input to DISTRUCT for plotting (Rosenberg 2004). Each species, *O. faveolata* and *O. franksi*, was then assessed with independent STRUCTURE analyses following the same parameters to look for within species differentiation (Foster et al. 2012). An analysis of molecular variance (AMOVA) was implemented in GenAlEx (version 6.5) to assess genetic differentiation by computing pairwise F_{ST} for species and sites (Peakall & Smouse 2012).

Mexico Dataset

Thirteen *O. faveolata* individuals were selected from a previous thermal stress experiment, five control individuals, five bleached individuals and three recovery individuals.

Curaçao Dataset

Coral hosts were genotyped using the same nine microsatellite loci and protocol as for the FGB dataset (Davies et al. 2013). No clones, 100% identical alleles across the nine loci, were detected. The same STRUCTURE (version 2.3.4) parameters were run as for the FGB dataset (Falush et al. 2003; Falush et al. 2007; Hubisz et al. 2009; Pritchard et al. 2000). STRUCTURE results were also analyzed following the same protocol as the FGB dataset, including STRUCTURE HARVESTER (Earl & vonHoldt 2012; Evanno et al. 2005), CLUMPP (Jakobsson & Rosenberg 2007) and DISTRUCT (Rosenberg 2004). An AMOVA was conducted in GenAlEx (version 6.5) to assess genetic differentiation by calculating F_{ST} within species between the three depth gradients (Peakall & Smouse 2012).

Amplification of *Symbiodinium* Internal Transcribed Spacer 2 region (ITS-2) for deep amplicon sequencing

FGB and Mexico Dataset

Sixty individual hosts from the FGB, 30 *Orbicella faveolata* and 30 *O. franksi*, and 13 *O. faveolata* individuals from Mexico were amplified for ITS-2 and submitted to the Genomics Sequencing and Analysis Facility at the University of Texas at Austin for targeted amplicon sequencing in January 2013. *Symbiodinium* specific ITS-2 primers,

ITS-Dino-forward (5'-GTGAATTGCAGAACTCCGTG-3') (Pochon et al. 2001) and

its2rev2-reverse (5'-CCTCCGCTTACTTATATGCTT-3') (Stat et al. 2009) were PCR

amplified yielding a target amplicon of 300 base pairs. Each 30 µL PCR reaction

contained 13.3 µL of water, 3.0 µL 10 *x ExTaq* HS buffer, 0.2 mM dNTP,

0.75 U (0.15 µL of 5 U/µL) *ExTaq* HS polymerase (Takara Biotechnology),

0.375 U (0.15 µL of 2.5 U/µL) *Pfu* polymerase (Agilent Technologies), 0.2 µM final

primer concentration and 50 ng DNA template (Kenkel et al. 2013). A DNA Engine

Tetrad 2 Thermal Cycler (Bio-Rad, Hercules, CA, USA) was used for all amplifications.

The following PCR protocol was used: 20 cycles of 94°C for five minutes, 95°C for 40

seconds, 59°C for two minutes, 72°C for one minute and final extensions of 72°C for five

minutes. Individuals were loaded into a two percent agarose gel for 15 minutes at 180

volts. All individuals were amplified to the same faint band intensity in order to prevent

over or under representation of PCR products, termed "cycle-checking". Cycle conditions

(95°C for 40 seconds, 59°C for two minutes, 72°C for one minute) were added to

individuals until similar faint band resolution was obtained for every individual. For

example, five cycles were added to individuals showing no band after the initial 20 cycles

of PCR. Cycle numbers were recorded once a faint band was obtained. Individuals that

would not amplify by 35 cycles were repeated with a lower starting template

concentration of 20 ng of DNA. Six individuals did not amplify by 35 cycles. For these

six individuals the starting template was reduced to 20 ng for 25 cycles until desired band

resolution was acquired. Every individual was run on one two percent agarose gel to

compare band intensities. Over-amplified individuals were repeated using fewer cycles

and faint individuals were run for additional cycles (Supplementary Table 1). All individuals amplified by 34 cycles during cycle checking except one west FGB *O. faveolata* that was removed from the analysis. One east FGB *O. faveolata* was removed from the analysis due to insufficient amounts of DNA template.

PCR products were cleaned up using GeneJET PCR purification kit (Fermentas Life Sciences). An equal amount of binding buffer and 100% isopropanol was added to remaining PCR products to maintain a 1:1 ratio. Samples were mixed and transferred to a GeneJET^{TM} purification column. All centrifugation steps were completed at max speed (13,200 rpm on Eppendorf microcentrifuge 5415R). Final products were eluted into 30 µL of elution buffer. The first 12 cleaned up samples were quantified using NanoDrop 2000 spectrophotometer (Thermo Scientific). Replicate quantifications of the same individuals gave conflicting results. Six individuals were run on a two percent agarose gel to check presence of remaining PCR product post clean-up steps. DNA quantification was not conducted due to time constraints and malfunctioning NanoDrop 2000 spectrophotometer (Thermo Scientific). Modifications for future protocol users would include quantifying DNA post PCR clean-up and diluting DNA to equal concentrations for each individual prior to assigning barcodes.

New 30.3 µL PCR reactions were performed to attach previously designed RAPID adaptors specific for 454 GS FLX. The adaptors were designed as follows: reverse barcoded primer sequence (A-RAPID primer+**unique barcode**+its2rev2primer) and forward B-RAPID primer (B-RAPID primer+ITS-Dino) (Figure 3). Each reaction contained 17.6 µL water, 0.2 mM dNTP, 3 µL 10 *x ExTaq* HS buffer, 0.75 U (0.15 µL of

5 U/µL) *ExTaq* HS polymerase (Takara Biotechnology), 0.375 U (0.15 µL of 2.5 U/µL)

Pfu Polymerase (Agilent Technologies), 50 ng of cleaned PCR product, 0.33 µM (1 µL of

10 µM concentration) of 454 B-RAPID ITS2-forward (5'-

CCTATCCCCTGTGTGCCTTGAGAGACGHC+GTGAATTGCAGAACTCCGTG-3')

and 0.33 µM (1 µL of 10 µM concentration) of 454 A-RAPID ITS2 adaptor with unique

barcode (5'-

CCATCTCATCCCTGCGTGTCTCCGACGACT+**TGTAGCGC**+CCTCCGCTTACTT

ATATGCTT-3', barcode sequence is in bold) (Kenkel et al. 2013). Reverse barcodes

were assigned to each individual (Supplementary Table 2). PCR was performed on a

DNA Engine Tetrad 2 Thermal Cycler (Bio-Rad, Hercules, CA, USA) under the

following conditions to attach barcodes: 95°C for five minutes, four cycles of 95°C for

30 seconds, 59°C for 30 seconds, 72°C for one minute and incubated for 72°C for five

minutes. All barcoded individuals were run on one two percent agarose gel for 15

minutes at 180 volts to confirm barcodes annealed. The band intensities from the

resulting image were used to pool equal representations of each individual. Three pools

were created based on band intensity (Supplementary Figure 1). For example, 10 µL of

individuals with average band intensity were added to a pool, 5 µL of individuals with

bright bands and 15 µL of individuals with faint bands were added to each pool. Two

pools of thirty unique barcodes contained a mix of *O. franksi* and *O. faveolata* from east

FGB and west FGB. The third pool contained the thirteen Mexico individuals. Pools were

ethanol precipitated following the protocol described in Kenkel (et al. 2013) with the

following modifications. Pools were incubated for 60 minutes at minus 20°C followed by

a 30 minute 13,200 rpm centrifugation (Eppendorf microcentrifuge 5415D) at 4°C.

Supernatant was removed and pellet washed with 500 uL of 80% ethanol followed by a

five minute 13,200 rpm centrifugation (Eppendorf microcentrifuge 5415D) at 4°C.

Supernatant was removed and pellets were air dried for 15 minutes. Pellets were re-

suspended in 30 μL of milli-Q water. Pools were loaded onto a one percent SYBR green

(Invitrogen) gel and run at 100 volts, 270 milliAmps for 45 minutes. Bands were

immediately excised post electrophoresis using a blue-light box, macerated in 25 μL of

milli-Q water and soaked overnight at 4°C. Samples were centrifuged for two minutes at

max speed. Supernatant was removed and stored at minus 20°C until submitted to the

University of Texas at Austin Genome Sequencing and Analysis Facility (GSAF) for

Roche 454 GS FLX sequencing. Two thousand minimum target reads per barcoded

individual were requested.

Curaçao Dataset

Amplification of ITS-2 followed the same methods as above for fifteen *O.*

faveolata hosts. All individuals amplified by 25 cycles during cycle checking. One pool

of fifteen RAPID barcoded individuals was created and ethanol precipitated

(Supplementary Table 3). Final clean up steps were completed at the University of Texas

at Austin GSAF and an additional conversion library preparation step to anneal one

MiSeq index to the pool. Ten thousand target minimum reads were requested per

individual to be sequenced on the Illumina MiSeq. However, in order to expedite

sequencing and fill the plate additional coverage was split across submissions yielding

46,000 reads per individual.

Bioinformatics

Trimming reads for FGB and Mexico Datasets

A fasta file of adaptors was created and converted to a database using legacy

BLAST package tool 'formatdb'. Raw *sff* files for 58 individuals from FGB and 13

individuals from the Mexico dataset were downloaded from Fourierseq at the University

of Texas at Austin. Uniquely barcoded individual reads were extracted and trimmed with

custom Perl scripts to remove adaptors, barcodes and low quality reads (Kenkel et al.

2013). Briefly, raw *sff* files were converted to *fna* files without any quality trimming

(sffinfo). Using the adaptors database and custom Perl script (Supplementary Data 1)

primer location was mapped and removal verified. The coordinates of the adaptors were

saved and returned to the original *sff* file (sfffile). The binary *sff* files were then converted

to fasta file format (sffinfo).

usearch pipeline for Flower Garden Banks

All reads less than 290 base pairs were removed using custom Perl scripts

(Supplementary Data 1). Replicate sequences, using full length matching, were removed

and assigned a size annotation for the number of times that sequence existed in the

database using *usearch -derep_fulllength*. Reads were sorted by size annotation using

usearch -sortbysize. The clustering algorithm *usearch* was then used to cluster reads into

operational taxonomic units (OTUs) using *usearch –cluster_otus* (Edgar 2010; Edgar

2013). Operational taxonomic units were defined with the following criteria: all OTUs

must have greater than three percent difference from each other and sequences assigned

to each OTU must have greater than 97% similarity. Quality trimmed reads were mapped

to OTUs using SHRiMP2 (David et al. 2011). Of 153 OTUs, five OTUs had a median count exceeding one (i.e., were detected in more than half of all samples) and were retained. These OTUs were used as query sequences for BLASTn (Altschul et al. 1990) to identify the *Symbiodinium* genotypes. OTUs were aligned using Clustal Omega online server (version 1.2.0) (Goujon et al. 2010; McWilliam et al. 2013; Sievers et al. 2011). Alignments were viewed and manually trimmed using SeaView (version 4.4.2) (Gouy et al. 2010) to see the variations in the five retained *Symbiodinium* genotypes compared to previously published *Symbiodinium* genotypes.

usearch pipeline for Mexico

The *usearch* pipeline for the Mexico dataset was the same as the FGB dataset with the following modifications. Reads longer than 350 bp were removed prior to clustering reads into OTUs using *fastx_trimmer* (Pearson et al. 1997). When removing replicate sequences all sequences less than 200 bases were removed. When reads were sorted by size all size annotations of two or less were also removed. The *usearch* provided python script 'fasta_number.py' assigned the OTU prefix to the cumulative fasta file of sequences. All seven OTUs were retained and used as query sequences for BLASTn (Altschul et al. 1990) and OTUs were locally blasted against the downloaded ITS-2 GeoSymbio database (409 sequences) (Franklin et al. 2012 (http://sites.google.com/site/geosymbio/)). Fasta files of OTUs and their top hit from the local database were created for each *Symbiodinium* clade detected. Fasta files were aligned using online server Clustal Omega (Goujon et al. 2010; McWilliam et al. 2013; Sievers et al. 2011(http://www.ebi.ac.uk/Tools/msa/clustalo/)). Sequences for the 13

individuals were mapped to the created OTU database using *usearch –usearch_global*

with 97% identity. The *usearch* provided python script 'uc2otutab.py' was slightly

modified to create a tab-delimited output showing the number of sequences each

individual mapped to each OTU database reference (Edgar 2010; Edgar 2013).

 usearch pipeline for Curaçao Dataset

Paired fastq files were joined using ea-utils *fastq-join* (Aronesty 2011; Aronesty

2013; Pearson et al. 1997). Adaptors and barcodes were removed using a custom Perl

script (Supplementary Data 1). Reads were trimmed to 330 base pairs using fastx toolkit

(version 0.0.13) *fastx_trimmer* (Pearson et al. 1997). The *usearch* clustering pipeline and

BLAST analysis was followed as previously described for the Mexico dataset with the

following modifications (Edgar 2010; Edgar 2013). Only OTUs with greater than 0.1% of

the total reads were retained yielding eight *Symbiodinium* genotypes. Then total quality

reads for each coral host individual (n=15) were mapped to the selected eight OTUs.

Statistical Analysis

R Studio (version 3.0.2) (R Developmental Core Team 2013) was used to analyze

sequences as a count data set using 'DESeq' package (Anders & Huber 2010).

Variance-stabilized data was created for a principle components analysis (PCA) (Anders

& Huber 2010). The total number of reads mapping to the five reference OTUs was used

as a sample size factor for each individual and variance-stabilizing transformation was

performed using empirical dispersion estimates (function estimateDispersions, options

sharingMode="gene-est-only"). The principle components analysis was performed using

the library 'vegan' (Oksanen et al. 2013). The differences in OTU representation between

species and sites were estimated jointly for all OTUs based on raw counts data using

Poisson-lognormal generalized linear mixed model, following the methodology

developed for quantitative PCR data (Matz et al. 2013). The model included fixed effects

of OTU, OTU:site, and OTU:species:site, plus the scalar random effect of sample. The

model was fitted using MCMCglmm function (Hadfield 2010). The results were

extracted and visualized using HPDplotBygeneBygroup function from the MCMC qpcr

package (Matz et al. 2013).

Figure 1. General location map showing collection sites for the three datasets, Flower

Garden Banks, Mexico and Curaçao. Map created using online ESRI web mapping

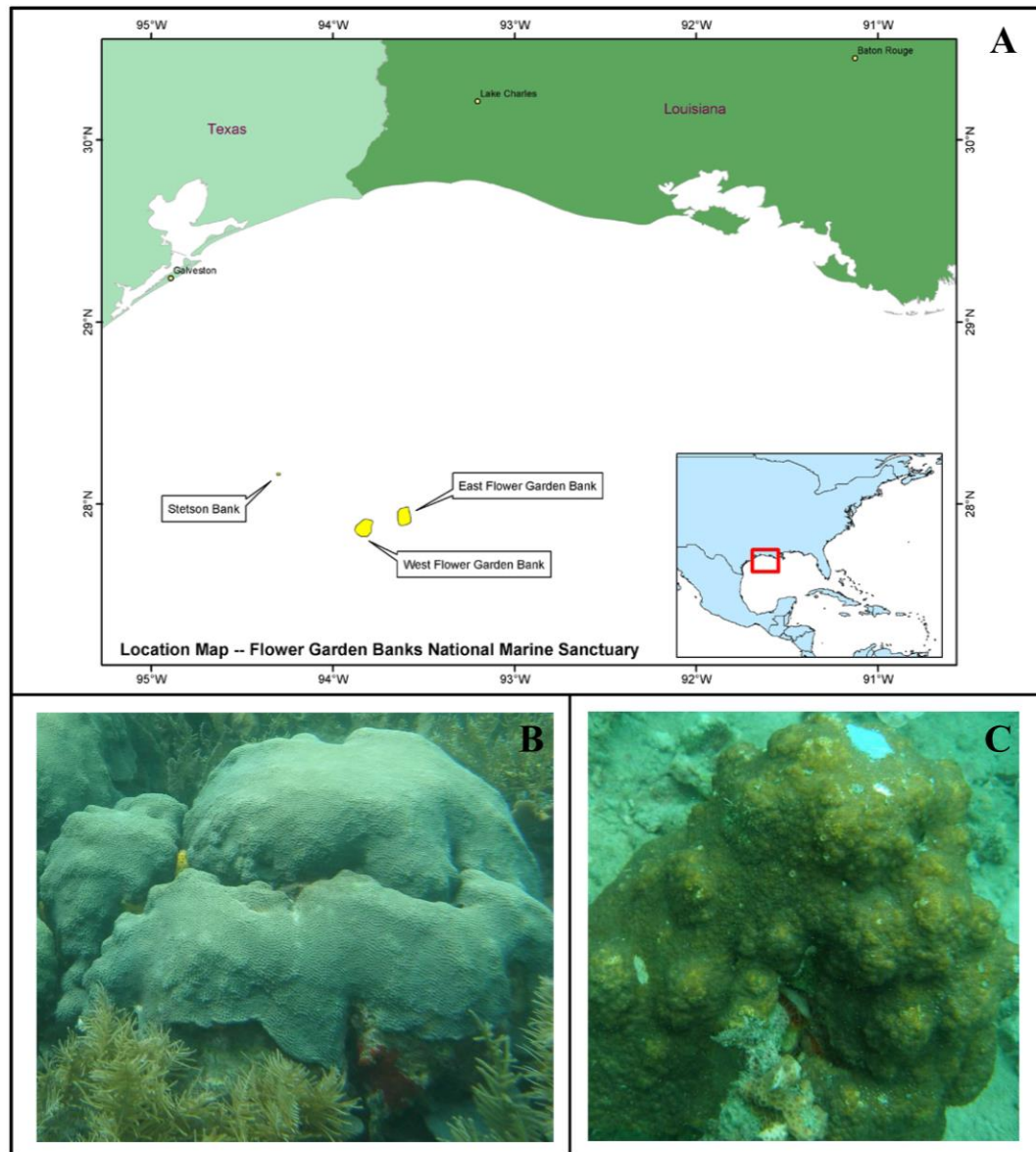service (http://www.arcgis.com/home/webmap/viewer.html)

Figure 2. A. Location of Flower Garden Banks National Marine Sanctuary, Credit:

USGS (http://pubs.usgs.gov/of/2003/of03-002/html/FGB_figs.htm) **B.** *Orbicella*

faveolata from Panama, Credit: Mónica Medina **C.** *Orbicella franksi* from Panama,

Credit: Mónica Medina

Table 1. Summary of Mexico samples that required DNA extractions from coral powder

ground for the original experiment.

	re-extracted DNA from original experiment (yes/no)
Control 1	yes
Control 2	yes
Control 3	yes
Control 4	no
Control 5	yes
Bleach 1	yes
Bleach 2	yes
Bleach 3	no
Bleach 4	yes
Bleach 5	yes
Recovery 1	yes
Recovery 2	yes
Recovery 3	yes

Table 2. List of Curaçao samples showing the depths each coral host individual was collected.

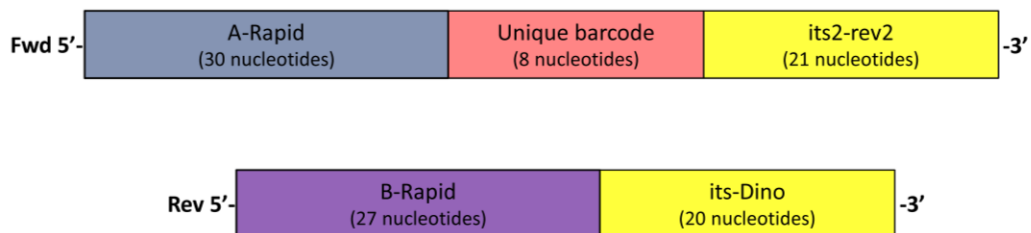|	Coral Host ID	Depth (meters)
Shallow	WP 13	7.01
	WP 14	7.32
	WP 15	7.32
	WP 17	7.01
	WP 21	7.01
Middle	WP 7	13.00
	WP 8	12.10
	WP 10	11.50
	WP 11	11.58
	WP 12	11.50
Deep	WP 1	16.46
	WP 2	15.54
	WP 3	15.54
	WP 5	15.24
	WP 6	16.46

Figure 3. Barcode design showing how coral host individuals were uniquely labeled and subsequently pooled.

Results

FGB Dataset

STRUCTURE analysis detected genetic differences between the two coral

species, and slightly stronger genetic structure in *Orbicella faveolata* between the east

and west FGBs than for *O. franksi* (Figure 4). The optimal number of clusters (K) for

each species was two, except for the independent analysis of *Orbicella faveolata* (n=73)

which showed a delta K of three (Supplementary Figure 2). This result was further

assessed by AMOVA analysis. AMOVA results comparing F_{ST} between species and sites

showed significant genetic differentiation between the two host species collected at each

site (Table 3). AMOVA analysis also showed stronger genetic structure in *O. faveolata*

between east and west FGB compared to *O. franksi* (Table 3).

Symbiodinium ITS-2 sequence coverage yielded 170,349 raw reads for 58

individuals, averaging 2,937 reads per individual (Table 4). Reads shorter than 290 base

pairs were removed leaving 122,867 reads. Of the 122,867 reads 20,260 unique

sequences remained. Clustering the unique sequences yielded 153 OTUs. Mapping the

original filtered reads to these 153 OTUs revealed that only five of the OTUs were

detected in more than half of all coral colonies sequenced. Only these five OTUs,

hereafter referred to as haplotypes, were analyzed further. Haplotype II was by far the

most dominant accounting for 94% of all reads (Figure 5). Quantitative analysis using a

generalized linear mixed model revealed that haplotypes IV and V were significantly

($P_{MCMC}<0.001$) less abundant at the west bank. Haplotype V was also significantly

(P_{MCMC}<0.001) less abundant in *O. faveolata* relative to *O. franksi* (Figure 6 and Table 5).

All five haplotypes top BLAST hit were *Symbiodinium* clade B type B1 (JN 558059.1) (Pochon et al. 2012), identified as *Symbiodinium minutum* (AF 333511.1) (LaJeunesse et al. 2012). After trimming, haplotypes I and II matched B1 (JN 558059.1, AF 333511.1) (LaJeunesse et al. 2012; Pochon et al. 2012) with 100% identity, whereas the remaining three haplotypes did not have an exact match in the database (Supplementary Figure 3 (Gouy et al. 2010)). Haplotype III differed from B1 (JN 558059.1, AF 333511.1) (LaJeunesse et al. 2012; Pochon et al. 2012) by a 13 base pair deletion. Haplotype IV differed from B1 (JN 558059.1, AF 333511.1) (LaJeunesse et al. 2012; Pochon et al. 2012) by a ten base pair insertion. Haplotype V differed from B1 (JN 558059.1, AF 333511.1) (LaJeunesse et al. 2012; Pochon et al. 2012) by a nine base pair deletion. These indels did not occur in homopolymer repeats and are not likely the result of sequencing error (Margulies et al. 2005).

The first component (PC1) from the PCA explained 40.83% of the variation and principle component two (PC2) explained 22.63% of the variation. Retaining the first two components met Kaiser's criterion (Kaiser 1960) and explained 63.46% of the variation. The samples were visibly partitioned with respect to the sampling locality along PC1 (Figure 7).

Mexico Dataset

All 13 individuals amplified by 25 cycles. Deep amplicon sequencing yielded 27,761 raw reads, averaging 1,028 reads per individual. Removing all reads longer than

350 bp and shorter than 200 bp left 6,909 sequences of which only 1,597 were unique. Clustering the unique sequences resulted in seven OTUs and all were retained. Mapping efficiency averaged 98.3% across the 27,739 reads mapped to the seven reference sequences.

Deep amplicon sequencing showed increased sensitivity to detect background populations of *Symbiodinium* species compared to RFLP and clone-direct sequencing. All previously detected *Symbiodnium* types were detected (A3, B1, C7) in addition to revealing a fourth *Symbiodinium* type A13 (JN 558096.1) (Pochon et al. 2012) and potentially a low frequency *Symbiodinium* type B37 (GU 907650) (Finney et al. 2010). Following the terminology of DeSalvo (et al. 2010) a control individual refers to a colony collected for the thermal stress experiment that was maintained at constant thermal conditions. A bleached individual refers to an individual subjected to thermal stress. Post the thermal stress, water temperature was reduced to average temperatures termed a recovery period. A recovery individual refers to the fragment collected during the recovery period. A representative *Symbiodinium* clade A haplotype, either A3 or A13, was the dominant type detected in all recovery coral host individuals (Figure 8). Deep amplicon sequencing revealed a different *Symbiodinium* community in contrast to previously published RFLP and direct sequencing results for one colony. For one of the colonies, where RFLP and direct sequencing identified a mixed *Symbiodinium* community comprised of 82% C7 and 18% B1 haplotypes, our targeted amplicon sequencing results detected 59.8% B1 and 38.7% A13 *Symbiodinium* haplotypes.

Curaçao Dataset

STRUCTURE ANALYSIS did not detect genetic differences in *O. faveolata* between the three collection depths (Figure 9). STRUCTURE HARVESTER showed a delta K of three for all analyses (Supplementary Figure 4). AMOVA results comparing F_{ST} between the three depth gradients showed no significant genetic differentiation between shallow and deep colonies (Table 6). However, significant genetic differentiation (F_{ST} =0.167) was detected between shallow and middle water depth colonies (Table 6). Significant (*p-value* <0.02) genetic structure was also detected between middle and deep-water colonies (Table 6). Deep amplicon sequencing of *Symbiodinium* ITS-2 yielded 698,089 raw reads for 15 individuals, averaging 46,000 reads per individual. After stitching paired reads together and removing reads longer than 350 bp, 548,121 reads remained, of which only 84,593 were longer than 200 bp and unique. Clustering the unique sequences yielded 30 OTUs. Only OTUs with at least 548 sequences mapped to them, i.e. greater than 0.1% of the total quality reads (n=548,121), were analyzed further, following the detection thresholds reported in Quigley et al. 2014. Mapping efficiency averaged 98.95% across the 548,121 reads mapped to the eight reference sequences.

The eight OTUs observed in corals from Curaçao represented *Symbiodinium* genotypes from clades A, B, C and G. *Symbiodinium* diversity was slightly reduced in deeper ranging hosts relative to shallow ranging collected hosts. Three coral hosts of the five collected from shallow water showed a more diverse *Symbiodinium* community

relative to deep collected coral hosts (Figure 10). *Symbiodinium* clade C genotypes C7 (AF499797) (DeSalvo et al. 2010) and C12 (AF 499801) (LaJeunesse 2002) were the most dominant *Symbiodinium* types across twelve of the fifteen coral hosts irrespective of water depth. All *Symbiodinium* clade A genotype A13=A1.1 (AF 333504) (LaJeunesse et al. 2009) were low frequency genotypes across all individuals from depth gradients. Clade B *Symbiodinium* genotypes were detected in three middle depth coral hosts and were a dominant genotype in three shallow collected coral hosts. *Symbiodinium* clade G genotype G3 (AM 748600) (Pochon et al. 2007) was detected as a dominant genotype in one coral host collected from the middle water depth range.

Table 3. Analysis of Molecular Variance (AMOVA) Fixation Index (F_{ST}) values showing slight genetic differentiation between *Orbicella faveolata* and *Orbicella franksi*, among *Orbicella faveolata* within the two geographic locations and no significant genetic differentiation among *Orbicella franksi* within the two geographic locations. (n=174)

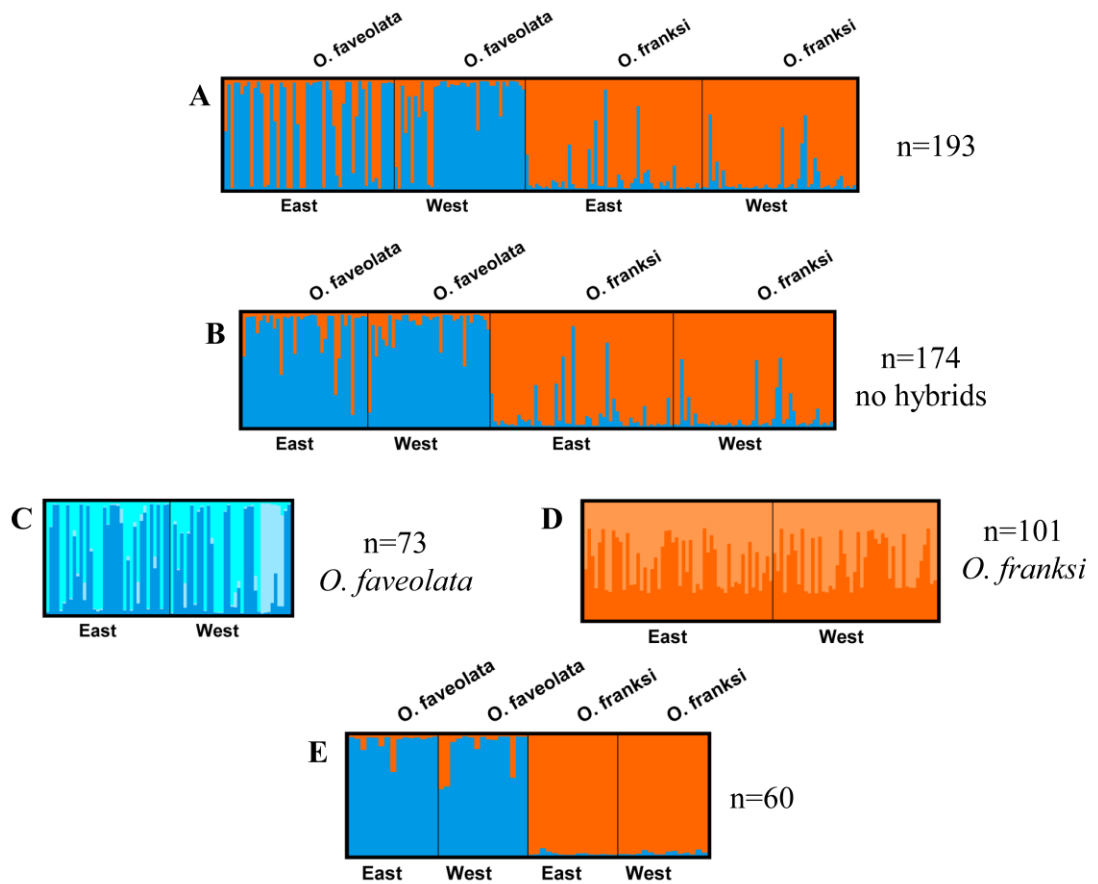|	F_{ST}	*p-value*
O. faveolata vs *O. franksi*	0.069	0.001
O. franksi East vs West	0	0.529
O. faveolata East vs West	0.009	0.016

Figure 4. DISTRUCT plots of all STRUCTURE analyses. **A.** All samples from *Orbicella faveolata* and *Orbicella franksi* collected from east and west Flower Garden Banks National Marine Sanctuary (K=2, n=193) **B.** Same as A but any potential hybrids were removed (K=2, n=174) **C.** *Orbicella faveolata* only and potential hybrids removed (K=3, n=73) **D.** *Orbicella franksi* only and potential hybrids removed (K=2, n=101) **E.** The selected 60 (K=2), *Orbicella faveolata* (n=30) and *Orbicella franksi* (n=30)

Table 4. Summary of sequence coverage and mapping efficiency for Flower Garden

Banks dataset sorted by geographic location and species from using the 454 GS FLX

platform.

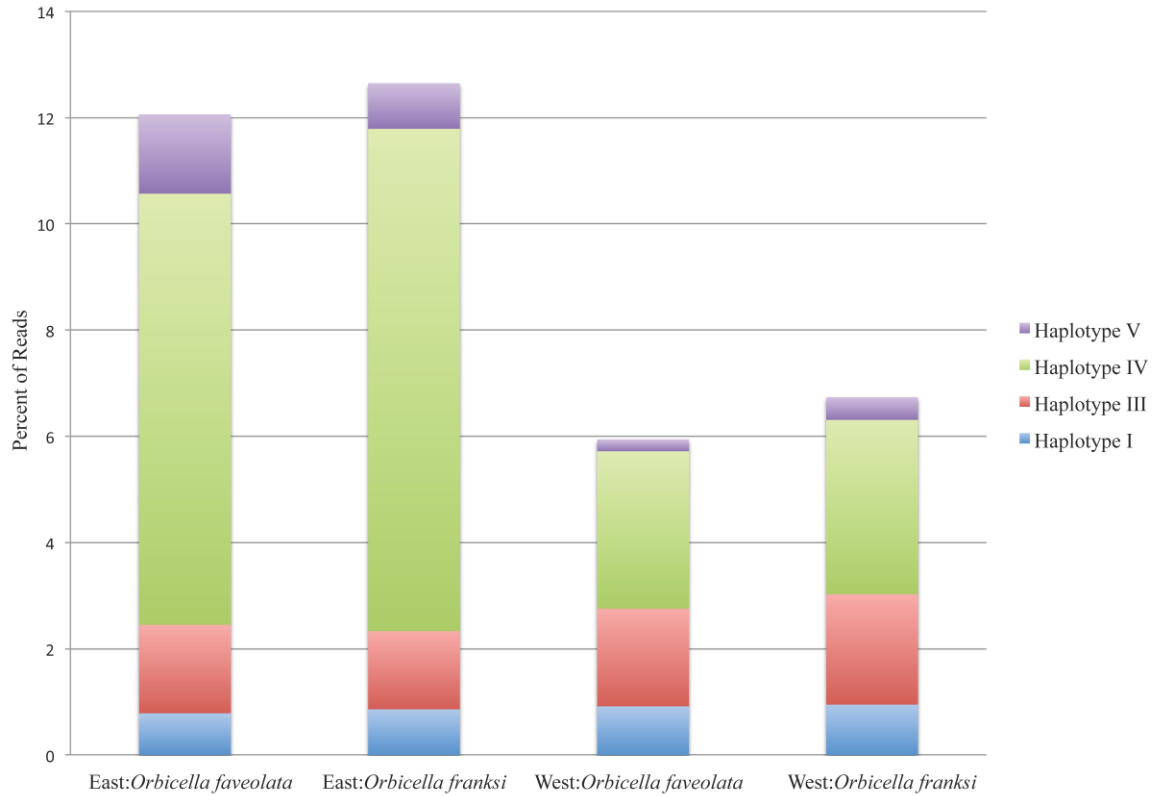|	**Raw Read Number**	**Trimmed Reads**	**Mapped Reads**	**Mapping Efficiency**
East	95,478	68,670	68,637	100%
West	74,871	54,197	54,175	100%
O. faveolata	74,840	53,938	53,913	100%
O. franksi	95,509	68,929	68,899	100%
TOTAL	170,349	122,867	122,812	100%

Figure 5. Only showing the percent of reads mapped to the four minor *Symbiodinium* B1 haplotypes by geographic location and coral host species. The dominant B1 haplotype II, not shown here, was used 93.26% across all coral host individuals. (East: *Orbicella faveolata* = 27,121 sequences, East: *Orbicella franksi* = 40,078, West: *Orbicella faveolata* = 26,143 sequences, West: *Orbicella franksi* = 27,376 sequences)
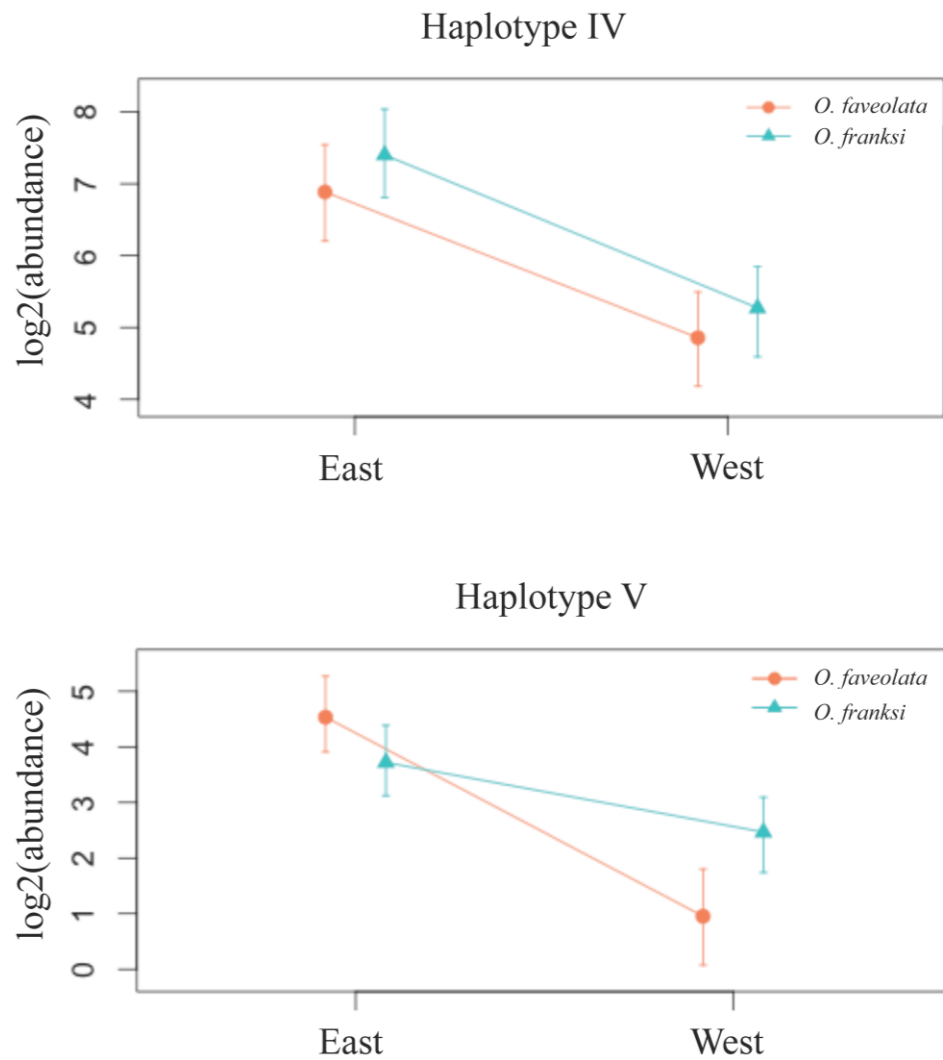
Figure 6. Distribution of abundance (log2 transformed) of *Symbiodinium* haplotypes IV and V detected in east and west Flower Garden Banks from Poisson-lognormal modeling. Haplotypes I, II and III did not have significant effects and are not shown here.

Table 5. Table summarizing the significant results from Poisson-lognormal Generalized
Linear Models (GLMs). Haplotypes IV and V are significantly less abundant at west
Flower Garden Bank compared to the east Flower Garden Bank (P_{MCMC} <0.001).
Haplotype V is also significantly less abundant in *Orbicella faveolata* relative to
Orbicella franksi (P_{MCMC} <0.001).

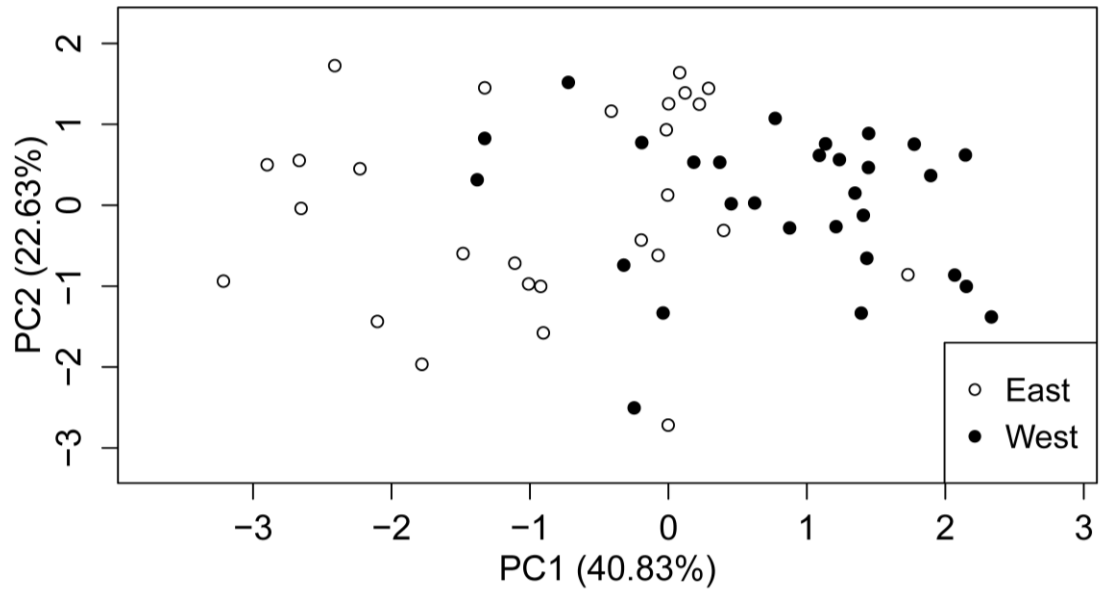|	posterior mean	lower 95% CI	upper 95% CI	effective sample size	*p-value* MCMC
Haplotype IV:West	-1.407299	-2.0212	-0.698313	1059.1	<0.001
Haplotype V:West	-2.486064	-3.26447	-1.698435	719.1	<0.001
Haplotype V:West:*O.franksi*	1.611213	0.650416	2.66628	811.4	<0.001

Figure 7. Principle Components Analysis (PCA) plot showing samples partition by geographic location. The count data set (n=56) was variance stabilized transformed (VST) prior to PCA analysis. Principle Component 1 (PC1) explains 40.83% of the variation and Principle Component 2 (PC2) explains 22.63% of the variation.
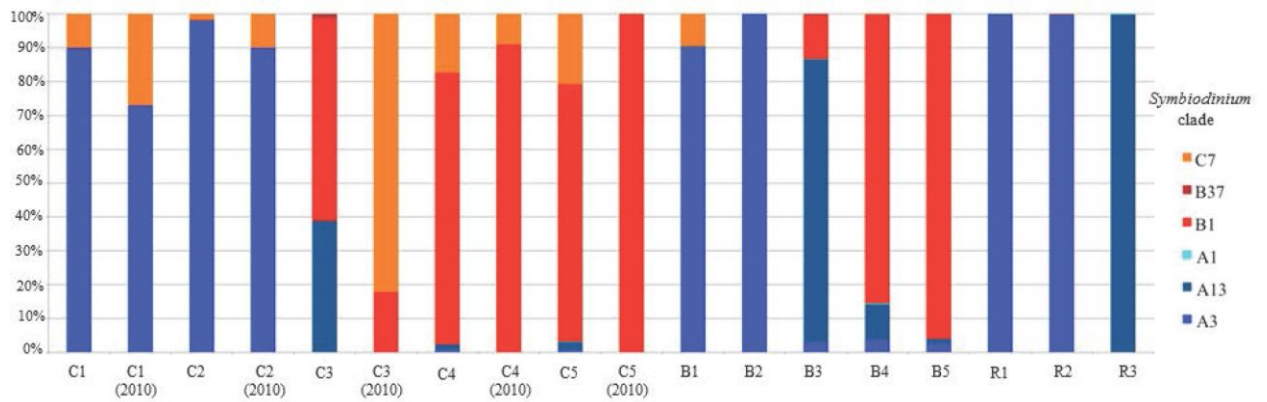
Figure 8. Percentage of sequences mapped to each *Symbiodinium* type detected throughout a thermal stress experiment. Read counts for the two *Symbiodinium* type B1 haplotypes from the reference operational taxonomic unit (OTU) database were summed to show only the dominant types detected. (B= Bleach individuals, C= Control individuals, R= Recovery individuals, # denotes colony number, C#(2010)= direct sequencing results from DeSalvo et al. 2010)
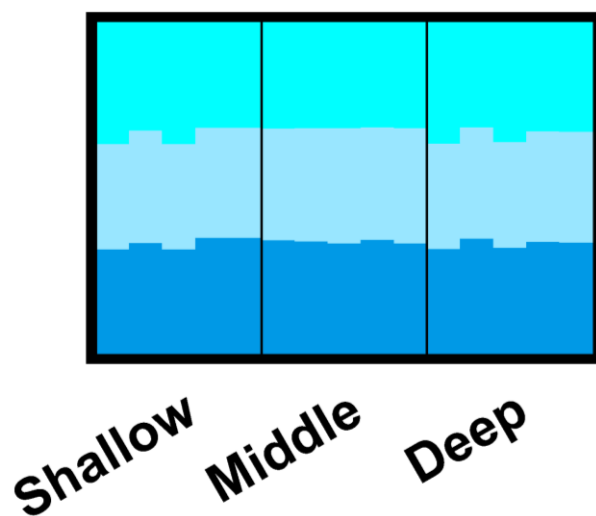
Figure 9. DISTRUCT plots of STRUCTURE analysis for *Orbicella faveolata* from the

Curaçao dataset (K=3, n=15)

Table 6. Analysis of Molecular Variance (AMOVA) Fixation Index (F_{ST}) values showing

genetic differentiation between water depth gradients for the Curaçao dataset.

Depth Range	F_{ST}	*p-value*
Shallow:Middle	0.167	0.002
Shallow:Deep	0.061	0.063
Middle:Deep	0.079	0.012
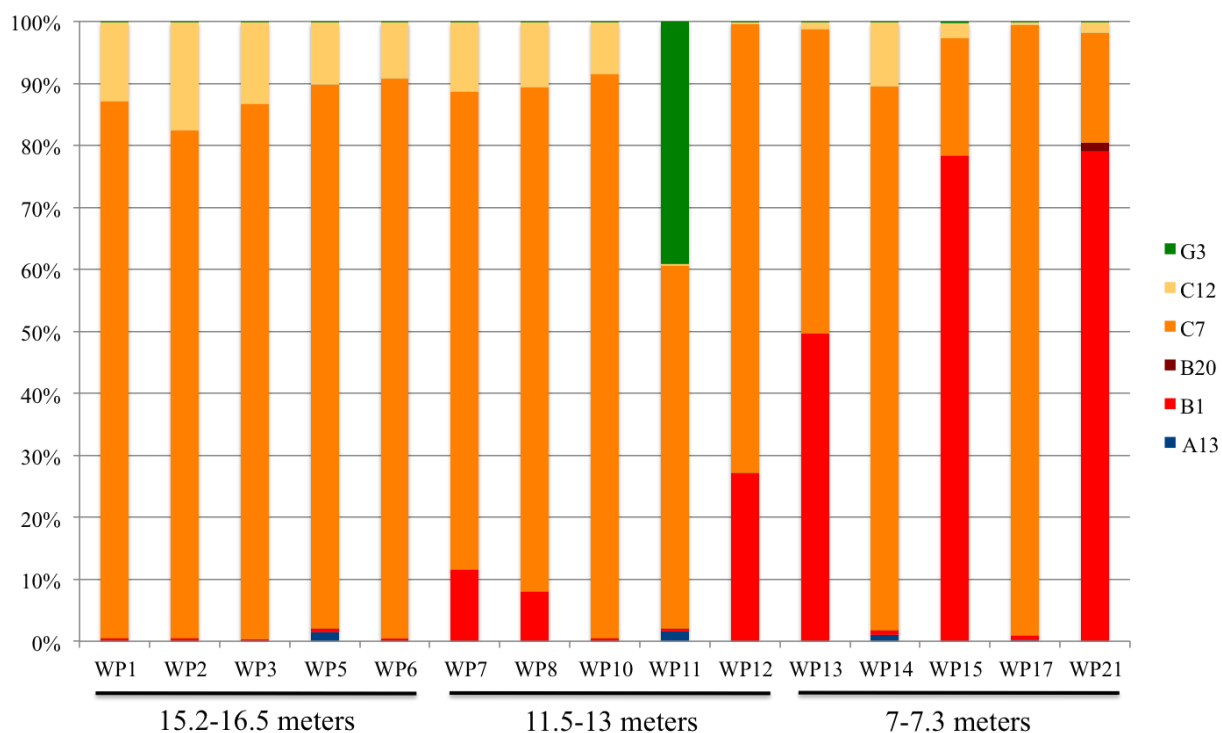
Figure 10. Percentage of sequences mapped to each *Symbiodinium* genotype detected in each *Orbicella faveolata* host colony for the Curaçao dataset summarized by water depth category (n=15). Read counts for the two *Symbiodinium* type B1 haplotypes from the reference operational taxonomic unit (OTU) database were summed to show only the dominant types detected. (WP=West Point, # denotes the sample identification number, legend shows all detected *Symbiodinium* clade types)

Discussion

FGB host genotyping significance

Nine recently developed microsatellite markers (Davies et al. 2013) were used for host genotyping to distinguish the two host species, *Orbicella faveolata* and *O. franksi*, since this species complex has been shown to hybridize (Budd & Pandolfi 2004; Fukami et al. 2004; Szmant et al. 1997). Multiple analyses were conducted to confirm that the two host species in this study did not include individuals showing evidence of recent introgression (Figure 4, Table 1). Species in the *Orbicella annularis* complex vary considerably morphologically however genetically these differences are not as pronounced making species identification in the *Orbicella annularis* species complex challenging (Fukami et al. 2004). Continued use of these recently developed nine loci (Davies et al. 2013), the eight previously developed microsatellite loci (Lopez et al. 1999; Severance et al. 2004) and continued efforts in high resolution marker development will advance detection limits to confidently assign species in the *O. annularis* species complex and ultimately lead to a better understanding of host connectivity patterns.

Monotypic symbiont population at FGB

In this study, deep amplicon sequencing was used to detect *Symbiodinium* species diversity within *O. faveolata* and *O. franksi* at east and west FGB using ITS-2. Both *Orbicella* species hosted clade B all broadly related to genotype B1, the most prevalent *Symbiodinium* genotype within Caribbean coral hosts (Baker 2003; LaJeunesse 2002; LaJeunesse et al. 2003). Interestingly other assessments of *Symbiodinium* diversity in *Orbicella* species throughout the Caribbean have shown mixed populations of species

ranging from clade A to clade D (Rowan & Knowlton 1995; Rowan et al. 1997; Thornhill et al. 2006; Toller et al. 2001). A variety of environmental factors have been proposed to explain *Symbiodinium* distributions, including but not limited to depth, irradiance levels, latitudinal location and temperature. Our results for the FGB *O. faveolata* and *O. franksi* show specificity for *Symbiodinium* clade B, which paralleled findings of fewer mixed infections in corals from deeper environments (LaJeunesse 2002). Corals from the FGB likely experience lower thermal minimums relative to the rest of the Caribbean (Schmahl et al. 2008; Thornhill et al. 2008) and these corals represent the northernmost latitudinal reef in the Gulf of Mexico (LaJeunesse & Trench 2000). However, we acknowledge faster evolving loci, such as microsatellites, may reveal more fine scale genetic diversity within ITS-2 clade B between *O. faveolata* and *O. franksi* at the FGB and other geographic locations (Finney et al. 2010; Pettay & LaJeunesse 2007; Santos et al. 2004).

Our results showed little genetic divergence between the two coral host species *O. faveolata* and *O. franksi* and a monotypic *Symbiodinium* population of only clade B types broadly related to B1. Previous studies have shown structure in *Symbiodinium* communities and host species across different habitat types (Bongaerts et al. 2010). As a consequence, we hypothesize *O. faveolata* and *O. franksi* at the FGB do not host a more diverse community of *Symbiodinium* species because of the lack of genetic divergence at the host level and the broadly similar environmental conditions at both banks.

Furthermore our results did show that within *Symbiodinium* type B1, haplotypes IV and V were significantly diminished at the west FGB. Additionally, haplotype V was significantly more diminished in *O. faveolata* compared to *O. franksi*. This result is

interesting since the east and west FGB are only separated by 19 kilometers and experience similar environmental conditions (Schmahl et al. 2008). Previous studies have shown habitat type correlates to coral communities and symbionts (Bongaerts et al. 2010) suggesting that *Symbiodinium* genotypes affect host physiology (DeSalvo et al. 2010). Though physiological contributions of host and *Symbiodinium* populations were outside the scope of this study, we do show the importance of accurately detecting low frequency *Symbiodinium* genotypes to contribute to understanding the distributions of local community assemblages and how *Symbiodinium* genotypes affect host physiology.

Potential roles of mesophotic reefs

The roles of mesophotic reefs, reefs between 30 and 150 meters (Lesser et al. 2010), remain understudied. Previous studies suggest mesophotic reefs may supply host larvae for shallow water reef systems (Lesser et al. 2009). There is increasing interest to investigate possible connectivity patterns between shallow and deep reefs to understand the roles and ecology of deep ranging hosts and *Symbiodinium* genotypes from mesophotic coral ecosystems (Kahng et al. 2014; Lesser et al. 2009; Lesser et al. 2010). The FGB are one example of an understudied deeper reef with reduced anthropogenic influences, fewer recorded bleaching events and minimal total cover loss relative to other Caribbean reefs since monitoring began in the 1970s (Hickerson & Schmahl 2005). This presents a unique location for future studies to assess species diversity, correlate environmental factors with *Symbiodinium* distributions and investigate roles of mesophotic reefs. The less disturbed conditions at the FGB may play a role in the unique host-algal genotype combinations seen at the FGB between *Symbiodinium minutum* and

coral hosts *O. faveolata* and *O. franksi*. Perhaps these coral host-*Symbiodinium* combinations have been maintained over many generations. The potential connectivity patterns between these understudied reef systems and shallow water reefs or to other Caribbean reefs are an area of future research.

Plasticity of symbiosis

Two mechanisms have been postulated to explain the plasticity of symbiosis between host and symbiont termed "shuffling" and "switching". "Shuffling" is a change in the existing proportions of a mixed *Symbiodinium* infection whereby a dominant symbiont type may become reduced while a background, or cryptic, symbiont type becomes increasingly prevalent (Berkelmans & van Oppen 2006; Fay & Weber 2012; LaJeunesse et al. 2009; Rowan et al. 1997; Silverstein et al. 2012; Stat et al. 2006). "Switching" is when new exogenous *Symbiodinium* are acquired as the dominant type, also known as an "open" symbiotic system (Baker 2001; Buddemeier & Fautin 1993). In order to assess whether corals "switch" or "shuffle", we must consistently detect cryptic *Symbiodinium* diversity. Use of a quantitative molecular genotyping approach with high sensitivity will allow us to assess distribution patterns of *Symbiodinium*-host relationships ranging from global scales over regional to individual reef scales. By doing so, it will also become more feasible to examine changes in *Symbiodinium* composition over time and detect species shuffling as well as potential horizontal uptake with more fine spatio-temporal resolution. This presents an exciting future for contributing more *Symbiodinium* haplotypes to existing cumulative databases.

Using deep amplicon sequencing to detect species diversity

Multiple efforts were made to make this protocol high-throughput and avoid including PCR and sequencing errors (Kenkel et al. 2013). The two-step barcode approach reduces PCR bias by using as few cycles as possible (Berry et al. 2011). By annealing unique barcodes to each individual we pooled up to thirty individuals making this protocol high-throughput with reduced cost. We pooled equal representations of each individual after assigning barcodes to increase the likelihood of equal coverage across individuals. Quigley, KM (et al. 2014) verified sensitivity down to 0.1% with an increased target minimum coverage of 10,000 reads per individual. This protocol utilized one set of barcoded primers that allowed the detection of fine scale proportions of *Symbiodinium* diversity within all clades. An additional advantage of this technique is no a priori knowledge of *Symbiodinium* species diversity is required. There is an initial upfront cost associated with barcoded primers, however this method will become increasingly more high-throughput and cost effective. Future users can transition to Illumina platforms as more tags are released and read lengths increase. We can now investigate *Symbiodinium* diversity by multiplexing multiple loci into a single Illumina lane. This method appears to be high-throughput, cost effective, reproducible and capable of detecting low frequency species in a sample with a mixed *Symbiodinium* population (Kenkel et al. 2013). Future studies can apply this method to investigate other members of the coral holobiont (Rohwer et al. 2002), such as other algae, fungi, protists, bacteria, archaea and viruses.

Limitations of deep amplicon sequencing

While the sensitivity of using deep amplicon sequencing to detect species diversity offers many advantages caution should be applied. Not unlike any other utilized molecular genotyping technique, deep amplicon sequencing does not detect functional versus non-functional haplotypes. Our study identified three unique B1 ITS-2 types. Given the abundance of these haplotypes across both species and geographic locations we believe that these are natural sequences. Perhaps these haplotypes are specific for the FGB but this hypothesis requires broader applications of this molecular genotyping method throughout the Caribbean. We cannot conclude with certainty whether these haplotypes might be prospective pseudogenes maintained in the populations (Thornhill et al. 2007). It is unlikely that these indels result from sequencing errors since they are not in homopolymer repeats (Margulies et al. 2005). Given unknown whole and partial genome duplication events in *Symbiodinium* some of these reference haplotypes could potentially come from the same genome (Hou & Lin 2009). Empirical analyses may predict copy numbers but do not provide conclusive results for inter versus intra-genomic haplotypes. Future users should use caution when assigning haplotypes within *Symbiodinium* clades to reference sequences to avoid over estimating species diversity.

Mexico Discussion

Mexico dataset shows sensitivity of deep amplicon sequencing

These initial targeted amplicon sequencing results allowed the comparison of three molecular genotyping techniques, two of which targeted the *Symbiodinium* ITS-2 locus. The original experiment reported *Symbiodinium* diversity within *Orbicella*

faveolata colonies collected for a thermal stress experiment (DeSalvo et al. 2010). Coral

cores were thermally stressed and allowed to recover. The *Symbiodinium* communities

were assessed throughout the bleach and recovery periods using RFLP of 18s rRNA

(DeSalvo et al. 2010). The *Symbiodinium* diversity for five control cores was genotyped

at ITS-2 through direct sequencing (DeSalvo et al. 2010). The targeted amplicon

sequencing results provided here show the utility of NGS to more sensitively detect low

frequency species in mixed communities. We detected additional *Symbiodinium*

genotypes A13 and A1 using deep amplicon sequencing that were not seen using direct

sequencing. Additionally, we detected *Symbiodinium* genotype B37 (GU 907650.1)

(Finney et al. 2010) as a low frequency *Symbiodinium* genotype. However, the top

alignment score from NCBI BLAST results and local database were not the same. The

top scoring hit from NCBI BLAST for this OTU (315 bp) had 95% identity with a B1

genotype (JN 558059.1) (Pochon et al. 2012) with an E value of 1e-134 and score of 488

bits. Local BLAST results using GeoSymbio ITS-2 database (409 ITS-2 sequences)

matched this OTU with 98% identity to B37 (GU 907650.1) (Finney et al. 2010) with an

E value of 7e-136 and score of 473 bits. Therefore we believe the *Symbiodinium*

genotype explored here is likely to be a *Symbiodinium* B37 variant but could also be a

variant of a *Symbiodinium* B1 genotype and requires further investigation.

Deep amplicon sequencing and direct sequencing results identified different

Symbiodinium communities in one coral host used as a control in the previous thermal

stress experiment (DeSalvo et al. 2010). Our sequencing results showed similar

proportions of genotypes in twelve out of the thirteen samples studied in

DeSalvo (et al. 2010). However, in one sample DeSalvo (et al. 2010) showed

Symbiodinium genotype C7 was dominant and B1 was less abundant, in contrast to our

results that suggested *Symbiodinium* B1 was dominant and A13 was less abundant.

Because the results were congruent for most samples, we are confident that the

discrepancy we revealed for this sample was likely caused by an error when selecting

cloned colonies for direct sequencing in the original experiment. Our results provide

support for the continued use of high-throughput molecular genotyping techniques as an

alternative to cloning techniques that may be more error prone.

Curaçao Discussion

These preliminary results are included here to show the sensitivity of targeted

amplicon sequencing using NGS to detect low frequency *Symbiodinium* types.

Additionally this dataset provides support for paired end sequencing results from an

alternative Illumina MiSeq platform, instead of the Roche 454 used for the FGB and

Mexico datasets and application of a unique pooling technique.

Curaçao coral host genotype significance

Following the same host genotype analysis as above for the FGB dataset multiple

efforts were conducted here to confirm individuals were not potential hybrids. Our results

showed the strongest genetic variation between shallow and middle water depth gradients

for *O. faveolata* (Figure 9, Table 6). We also showed significant (*p-value*<0.02) but small

genetic differentiation between middle and deep water collected corals (Table 6). One

recent coral host analysis has provided support for depth segregated coral host lineages

(Prada et al. 2013). Our host genotype results are only based on microsatellite loci. We

acknowledge the use of more conserved loci, such as other nuclear and mitochondrial regions, are required to understand whether *O. faveolata* also show depth segregated lineages. The slight genetic structure shown here may show stronger genetic differentiation between depth segregated colonies with the use of more conserved loci.

Depth correlated Symbiodinium diversity in Curaçao

       Symbiodinium communities evaluated from three water depth gradients using targeted amplicon sequencing of *Symbiodinium* rRNA ITS-2 showed three of the five *O. faveolata* coral host-associated *Symbiodinium* communities were more diverse in shallow water. *Symbiodinium* provide coral hosts with photosynthetic products (Muscatine & Cernichiari 1969; Muscatine et al. 1984; Trench 1987). Therefore it is not surprising irradiance levels have been shown to correlate with *Symbiodinium* communities (Toller et al. 2001; LaJeunesse 2002; Rowan and Knowlton 1995; Sampayo et al. 2007; LaJeunesse et al. 2003). Though lower *Symbiodinium* diversity in reduced irradiance environments has been shown (Toller et al. 2001), the *Symbiodinium* lineages detected in these low photic environments varied. *Symbiodinium* genotypes from clades B (Savage et al. 2002), C and E (Toller et al. 2001), have all been reported from *Orbicella* hosts in lower irradiance zones (Finney et al. 2010). Additionally no *Symbiodinium* genotype has shown consistent irradiance partitioning. For example, *Symbiodinium* B1j was seen in all *Orbicella* hosts from shallow water but only B7 was detected in *Madracis* hosts across all depths (Frade et al. 2008). Our results paralleled other findings of reduced *Symbiodinium* diversity in deeper, i.e. lower irradiance, habitats (Toller et al. 2001; LaJeunesse 2002; Rowan and Knowlton 1995). Despite a small sample size within each depth category

(n=5) these results provide support for the coral host-associated *Symbiodinium* genotypes to be partially correlated to the water depth of the sampled coral hosts. Though irradiance levels were not measured in this study, we hypothesize the *Symbiodinium* community compositions reported here may be correlated to photic zones. However, future investigations measuring the irradiance levels for these coral hosts are required to make this conclusion.

Symbiodinium C abundance throughout the Caribbean

Our results showed *Symbiodinium* type C7 was dominant in collected *O. faveolata* from deeper environments. These results parallel other findings of clade C genotypes detected in deeper environments (>3 meters) (Rowan & Knowlton 1995; Rowan et al. 1997; Toller et al. 2001; LaJeunesse 2002). Though a more diverse *Symbiodinium* community was observed in shallow water environments, these genotypes were less abundant compared to *Symbiodinium* C7. Despite *Symbiodinium* clade B types being the most prevalent throughout the Caribbean (Baker 2003; LaJeunesse 2002; LaJeunesse et al. 2003), we find a dominant *Symbiodinium* clade C genotype in *O. faveolata* at Curaçao. Other studies have reported a dominant *Symbiodinium* clade C type in coral hosts known to host a diverse community of *Symbiodinium* (Garren et al. 2006; LaJeunesse 2002). Therefore our results support previously published *Symbiodinium* mixed communities seen in Caribbean collected *O. faveolata* hosts (Rowan and Knowlton 1995; Rowan et al. 1997; Thornhill et al. 2006; Toller et al. 2001).

Symbiodinium clade G

Our preliminary assessment of *O. faveolata*-associated *Symbiodinium* communities detected a *Symbiodinium* clade G type G3. This is interesting since most clade G *Symbiodinium* types have been detected in foraminifera (Pawlowski et al. 2001; Pochon et al. 2007). Coral host-associated *Symbiodinium* G1 types have been identified around Cape Panwa in Thailand (LaJeunesse et al. 2010). Given the abundance of this *Symbiodinium* genotype we believe it is a true representation of the *Symbiodinium* community composition for this one coral host. Perhaps this *Symbiodinium* genotype has colonized a unique habitat in this host at this geographic location and may be contributing a unique fitness contribution for the coral holobiont. Further analyses will be necessary to reveal why only one of the fifteen individuals appeared to associate with *Symbiodinium* genotype G3 and its potential roles within the coral holobiont.

Summary

This study applied deep amplicon sequencing to assess *Symbiodinium* diversity at the Flower Garden Banks. Results showed coral hosts *Orbicella faveolata* and *O. franksi* only associated with *Symbiodinium* type B1, however three possibly endemic haplotypes were detected. Two of these haplotypes were significantly less abundant at the west FGB, one of which was also significantly less abundant in *O. faveolata* compared to *O. franksi*. Faster evolving loci, such as microsatellites for *Symbiodinium*, may show fine scale clade B community variations between host species or geographic locations. Continued use of

deep amplicon sequencing, not only with ITS-2 but with other loci, to assess *Symbiodinium* species diversity within multiple host species will generate a better understanding of these complex community assemblages.

These results also showed the utility of deep amplicon sequencing as an alternative molecular genotyping technique to detect low abundance species, as seen from the Mexico dataset. The rapidly advancing field of molecular genotyping techniques will continue to improve detection thresholds and contribute to our understanding of species diversity.

Results from the Curaçao dataset applied deep amplicon sequencing but with the Illumina MiSeq platform to assess *Symbiodinium* diversity across a depth gradient. We detect *Symbiodinium* clade G3 not typically associated with scleractinian hosts. Additionally we show increased *Symbiodinium* diversity in shallow water coral hosts compared to deep collected. As new sequencing kits are released offering increased read lengths the application of these sensitive molecular genotyping techniques will likely only increase. Sustained use will contribute to a better understanding of species diversity within mixed communities.

References

Altschul SF, Gish W, Miller W, Myers EW, and Lipman DJ. 1990. Basic local alignment search tool. *Journal of Molecular Biology* 215:403-410.

Anders S, and Huber W. 2010. Differential expression analysis for sequence count data. *Genome Biology* 11:R106.

Aronesty E. 2011. Command-line tools for processing biological sequencing data. *ea-utils*.

Aronesty E. 2013. Comparison of Sequencing Utility Programs. *The Open Bioinformatics Journal*.

Baird A, Cumbo V, Leggat W, and Rodriguez-Lanetty M. 2007. Fidelity and flexibility in coral symbioses. *Marine Ecology Progress Series* 347:307-309.

Baker AC. 2001. Ecosystems: Reef Corals bleach to survive change. *Nature* 411:765-766.

Baker AC. 2003. Flexibility and specificity in coral-algal symbiosis: Diversity, ecology, and biogeography of *Symbiodinium*. *Annual Review of Ecology, Evolution, and Systematics* 34:661-689.

Baker AC, and Romanski AM. 2007. Multiple symbiotic partnerships are common in scleractinian corals, but not in octocorals: Comment on Goulet (2006). *Marine Ecology Progress Series* 335:237-242.

Berkelmans R, and van Oppen MJH. 2006. The role of zooxanthellae in the thermal tolerance of corals: a 'nugget of hope' for coral reefs in an era of climate change. *Proceedings of the Royal Society B: Biological Sciences* 273:2305-2312.

Berry D, Ben Mahfoudh K, Wagner M, and Loy A. 2011. Barcoded Primers Used in Multiplex Amplicon Pyrosequencing Bias Amplification. *Applied and Environmental Microbiology* 77:7846-7849.

Bongaerts P, Riginos C, Ridgway T, Sampayo EM, van Oppen MJH, Englebert N, Vermeulen F, and Hoegh-Guldberg O. 2010. Genetic Divergence across Habitats in the Widespread Coral *Seriatopora hystrix* and Its Associated *Symbiodinium*. *PLoS ONE* 5:e10871.

Brown BE. 1997. Coral bleaching: causes and consequences. *Coral Reefs* 16:S129-S138.

Budd AF, Fukami H, Smith ND, and Knowlton N. 2012. Taxonomic classification of the reef coral family Mussidae (Cnidaria: Anthozoa: Scleractinia). *Zoological Journal of the Linnean Society* 166:465-529.

Budd AF, and Pandolfi JM. 2004. Overlapping species boundaries and hybridization within the *Montastraea "annularis"* reef coral complex in the Pleistocene of the Bahama Islands. *Paleobiology* 30:396-425.

Buddemeier RW, and Fautin DG. 1993. Coral Bleaching as an Adaptive Mechanism. *BioScience* 43:320-326.

Chomczynski P, and Sacchi N. 2006. The single-step method of RNA isolation by acid guanidinium thiocyanate-phenol-chloroform extraction: twenty-something years on. *Nat Protocols* 1.

Coffroth MA, and Santos SR. 2005. Genetic Diversity of Symbiotic Dinoflagellates in the Genus *Symbiodinium*. *Protist* 156:19-34.

David M, Dzamba M, Lister D, Ilie L, and Brudno M. 2011. SHRiMP2: Sensitive yet Practical Short Read Mapping. *Bioinformatics* 27:1011-1012.

Davies SW, Rahman M, Meyer E, Green EA, Buschiazzo E, Medina M, and Matz MV. 2013. Novel polymorphic microsatellite markers for population genetics of the endangered Caribbean star coral, *Montastraea faveolata*. *Marine Biodiversity* 43:167-172.

DeSalvo MK, Sunagawa S, Fisher PL, Voolstra CR, Iglesias-Prieto R, and Medina M. 2010. Coral host transcriptomic states are correlated with *Symbiodinium* genotypes. *Molecular Ecology* 19:1174-1186.

Earl D, and vonHoldt B. 2012. STRUCTURE HARVESTER: a website and program for visualizing STRUCTURE output and implementing the Evanno method. *Conservation Genetics Resources* 4:359-361.

Edgar RC. 2010. Search and clustering orders of magnitude faster than BLAST. *Bioinformatics* 26:2460-2461.

Edgar RC. 2013. UPARSE: highly accurate OTU sequences from microbial amplicon reads. *Nat Meth* 10:996-998.

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.

Falush D, Stephens M, and Pritchard JK. 2007. Inference of population structure using multilocus genotype data: dominant markers and null alleles. *Molecular Ecology Notes* 7:574-578.

Fay SA, and Weber MX. 2012. The Occurrence of Mixed Infections of *Symbiodinium* (Dinoflagellata) within Individual Hosts. *Journal of Phycology* 48:1306-1316.

Finney JC, Pettay D, Sampayo E, Warner M, Oxenford H, and LaJeunesse T. 2010. The Relative Significance of Host–Habitat, Depth, and Geography on the Ecology, Endemism, and Speciation of Coral Endosymbionts in the Genus *Symbiodinium*. *Microbial Ecology* 60:250-263.

Foster NL, Paris CB, Kool JT, Baums IB, Stevens JR, Sanchez JA, Bastidas C, Agudelo C, Bush P, Day O et al. . 2012. Connectivity of Caribbean coral populations: complementary insights from empirical and modelled gene flow. *Molecular Ecology* 21:1143-1157.

Franklin EC, Stat M, Pochon X, Putnam HM, and Gates RD. 2012. GeoSymbio: a hybrid, cloud-based web application of global geospatial bioinformatics and ecoinformatics for *Symbiodinium*–host symbioses. *Molecular Ecology Resources* 12:369-373.

Fukami H, Budd AF, Don RL, Jara J, Kersanach R, and Knowlton N. 2004. Geographic Differences in Species Boundaries among Members of the *Montastraea annularis* Complex Based on Molecular and Morphological Markers. *Evolution* 58:324-337.

Glynn PW. 1993. Coral reef bleaching: ecological perspectives. *Coral Reefs* 12:1-17.

Goujon M, McWilliam H, Li W, Valentin F, Squizzato S, Paern J, and Lopez R. 2010. A new bioinformatics analysis tools framework at EMBL-EBI. *Nucleic Acids Research* 38:W695-W699.

Gouy M, Guindon S, and Gascuel O. 2010. SeaView Version 4: A Multiplatform Graphical User Interface for Sequence Alignment and Phylogenetic Tree Building. *Molecular Biology and Evolution* 27:221-224.

Hadfield JD. 2010. MCMC Methods for Multi-Response Generalized Linear Mixed Models: The MCMCglmm R Package. *Journal of Statistical Software* 33:1-22.

Hickerson E, and Schmahl G. 2005. The state of coral reef ecosystems of the Flower Garden Banks, Stetson Bank and other banks in the northwestern Gulf of Mexico. pp. 201-221. In: Waddell J, ed. *The State of Coral Reef Ecosystems of the United States and Pacific Freely Associated States: 2005*: NOAA Technical Memorandum NOS NCCOS 11.

NOAA/NCCOS Center for Coastal Monitoring and Assessment's Biogeography Team, Silver Spring, MD. 522 pp.

Hoegh-Guldberg O. 1999. Climate change, coral bleaching and the future of the world's coral reefs. *Marine and Freshwater Research* 50:839-866.

Hoegh-Guldberg O. 2004. Coral reefs in a century of rapid environmental change. *Symbiosis* 37:1-31.

Hou Y, and Lin S. 2009. Distinct Gene Number-Genome Size Relationships for Eukaryotes and Non-Eukaryotes: Gene Content Estimation for Dinoflagellate Genomes. *PLoS ONE* 4:e6978.

Hubisz MJ, Falush D, Stephens M, and Pritchard JK. 2009. Inferring weak population structure with the assistance of sample group information. *Molecular Ecology Resources* 9:1322-1332.

Hughes TP, Baird AH, Bellwood DR, Card M, Connolly SR, Folke C, Grosberg R, Hoegh-Guldberg O, Jackson JBC, Kleypas J et al. . 2003. Climate Change, Human Impacts, and the Resilience of Coral Reefs. *Science* 301:929-933.

IUCN. 2011. IUCN Red List of Threatened Species. Version 2011.1. IUCN, Switzerland.

Jakobsson M, and Rosenberg NA. 2007. CLUMPP: a cluster matching and permutation program for dealing with label switching and multimodality in analysis of population structure. *Bioinformatics* 23:1801-1806.

Jones A, and Berkelmans R. 2010. Potential Costs of Acclimatization to a Warmer Climate: Growth of a Reef Coral with Heat Tolerant vs. Sensitive Symbiont Types. *PLoS ONE* 5:e10437.

Kahng S, Copus J, and Wagner D. 2014. Recent advances in the ecology of mesophotic coral ecosystems (MCEs). *Current Opinion in Environmental Sustainability* 7:72-81.

Kaiser HF. 1960. The Application of Electronic Computers to Factor Analysis. *Educational and Psychological Measurement* 20:141-151.

Kenkel CD, Goodbody-Gringley G, Caillaud D, Davies SW, Bartels E, and Matz MV. 2013. Evidence for a host role in thermotolerance divergence between populations of the mustard hill coral (*Porites astreoides*) from different reef environments. *Molecular Ecology* 22:4335-4348.

Knowlton N, and Rohwer F. 2003. Multispecies Microbial Mutualisms on Coral Reefs: The Host as a Habitat. *The American Naturalist* 162:S51-S62.

LaJeunesse TC. 2002. Diversity and community structure of symbiotic dinoflagellates from Caribbean coral reefs. *Marine Biology* 141:387-400.

LaJeunesse TC, Loh WKW, Woesik Rv, Hoegh-Guldberg O, Schmidt GW, and Fitt WK. 2003. Low Symbiont Diversity in Southern Great Barrier Reef Corals, Relative to Those of the Caribbean. *Limnology and Oceanography* 48:2046-2054.

LaJeunesse TC, Parkinson JE, and Reimer JD. 2012. A genetics-based description of *Symbiodinium minutum* sp. nov. and *S. psygmophilum* sp. nov. (Dinophyceae), two dinoflagellates symbiotic with cnidaria. *Journal of Phycology* 48:1380-1391.

LaJeunesse TC, Smith RT, Finney J, and Oxenford H. 2009. Outbreak and persistence of opportunistic symbiotic dinoflagellates during the 2005 Caribbean mass coral 'bleaching' event. *Proceedings of the Royal Society B: Biological Sciences* 276:4139-4148.

LaJeunesse TC, and Trench RK. 2000. Biogeography of two species of *Symbiodinium* (Freudenthal) inhabiting the intertidal sea anemone *Anthopleura elegantissima* (Brandt). *The Biological Bulletin* 199:126-134.

Lesser MP, Slattery M, and Leichter JJ. 2009. Ecology of mesophotic coral reefs. *Journal of Experimental Marine Biology and Ecology* 375:1-8.

Lesser MP, Slattery M, Stat M, Ojimi M, Gates RD, and Grottoli A. 2010. Photoacclimatization by the coral *Montastraea cavernosa* in the mesophotic zone: light, food, and genetics. *Ecology* 91:990-1003.

Lopez JV, Kersanach R, Rehner SA, and Knowlton N. 1999. Molecular Determination of Species Boundaries in Corals: Genetic Analysis of the *Montastraea annularis* Complex Using Amplified Fragment Length Polymorphisms and a Microsatellite Marker. *The Biological Bulletin* 196:80-93.

Margulies M, Egholm M, Altman WE, Attiya S, Bader JS, Bemben LA, Berka J, Braverman MS, Chen Y-J, Chen Z et al. . 2005. Genome sequencing in microfabricated high-density picolitre reactors. *Nature* 437:376-380.

Matz MV, Wright RM, and Scott JG. 2013. No Control Genes Required: Bayesian Analysis of qRT-PCR Data. *PLoS ONE* 8:e71448.

McWilliam H, Li W, Uludag M, Squizzato S, Park YM, Buso N, Cowley AP, and Lopez R. 2013. Analysis Tool Web Services from the EMBL-EBI. *Nucleic Acids Research* 41:W597-600.

Mieog JC, Van Oppen MJH, Berkelmans R, Stam WT, and Olsen JL. 2009. Quantification of algal endosymbionts (Symbiodinium) in coral tissue using real-time PCR. *Molecular Ecology Resources* 9:74-82.

Muscatine L. 1967. Glycerol Excretion by Symbiotic Algae from Corals and Tridacna and Its Control by the Host. *Science* 156:516-519.

Oksanen J, Blanchet FG, Kindt R, Legendre P, Minchin PR, O'Hara RB, Simpson GL, Solymos P, Stevens MHH, and Wagner H. 2013. vegan: Community Ecology Package. R package v2.0-10 ed.

Peakall R, and Smouse PE. 2012. GenAlEx 6.5: genetic analysis in Excel. Population genetic software for teaching and research—an update. *Bioinformatics* 28:2537-2539.

Pearse VB, and Muscatine L. 1971. ROLE OF SYMBIOTIC ALGAE (ZOOXANTHELLAE) IN CORAL CALCIFICATION. *The Biological Bulletin* 141:350-363.

Pearson WR, Wood T, Zhang Z, and Miller W. 1997. Comparison of DNA Sequences with Protein Sequences. *Genomics* 46:24-36.

Pettay DT, and LaJeunesse TC. 2007. Microsatellites from clade B *Symbiodinium* spp. specialized for Caribbean corals in the genus *Madracis*. *Molecular Ecology Notes* 7:1271-1274.

Pochon X, and Gates RD. 2010. A new *Symbiodinium* clade (Dinophyceae) from soritid foraminifera in Hawai'i. *Molecular Phylogenetics and Evolution* 56:492-497.

Pochon X, Pawlowski J, Zaninetti L, and Rowan R. 2001. High genetic diversity and relative specificity among *Symbiodinium*-like endosymbiotic dinoflagellates in soritid foraminiferans. *Marine Biology* 139:1069-1078.

Pochon X, Putnam HM, Burki F, and Gates RD. 2012. Identifying and Characterizing Alternative Molecular Markers for the Symbiotic and Free-Living Dinoflagellate Genus *Symbiodinium*. *PLoS ONE* 7:e29816.

Precht W, ML R, GS B, and GP S. 2005. Establishment and initial analysis of deep reef stations (32–40m) at the East Flower Garden Bank. Dedicated Issue, Flower Garden Banks National Marine Sanctuary. *Gulf of Mexico Science* 21:124-127.

Pritchard JK, Stephens M, and Donnelly P. 2000. Inference of Population Structure Using Multilocus Genotype Data. *Genetics* 155:945-959.

Quigley KM, Davies SW, Kenkel CD, Willis BL, Matz MV, and Bay LK. 2014. Deep-Sequencing Method for Quantifying Background Abundances of *Symbiodinium* Types: Exploring the Rare *Symbiodinium* Biosphere in Reef-Building Corals. *PLoS ONE* 9:e94297.

R Developmental Core Team. 2013. R: A language and environment for statistical computing. *R Foundation for Statistical Computing*.

Rohwer F, Seguritan V, Azam F, and Knowlton N. 2002. Diversity and distribution of coral-associated bacteria. *Marine Ecology Progress Series* 243:1-10.

Rosenberg NA. 2004. DISTRUCT: a program for the graphical display of population structure. *Molecular Ecology Notes* 4:137-138.

Rowan R, and Knowlton N. 1995. Intraspecific diversity and ecological zonation in coral-algal symbiosis. *Proceedings of the National Academy of Sciences* 92:2850-2853.

Rowan R, Knowlton N, Baker A, and Jara J. 1997. Landscape ecology of algal symbionts creates variation in episodes of coral bleaching. *Nature* 388:265-269.

Santos SR, Shearer TL, Hannes AR, and Coffroth MA. 2004. Fine-scale diversity and specificity in the most prevalent lineage of symbiotic dinoflagellates (*Symbiodinium*, Dinophyceae) of the Caribbean. *Molecular Ecology* 13:459-469.

Schmahl GP, Hickerson EL, and Precht WF. 2008. Biology and Ecology of Coral Reefs and Coral Communities in the Flower Garden Banks Region, Northwestern Gulf of Mexico. In: Riegl BM, and Dodge RE, editors. Coral Reefs of the USA. Dordrecht: Springer. p 221-262.

Severance EG, Szmant AM, and A. Karl S. 2004. Microsatellite loci isolated from the Caribbean coral, *Montastraea annularis*. *Molecular Ecology Notes* 4:74-76.

Sievers F, Wilm A, Dineen D, Gibson TJ, Karplus K, Li W, Lopez R, McWilliam H, Remmert M, Söding J et al. . 2011. Fast, scalable generation of high-quality protein multiple sequence alignments using Clustal Omega. *Molecular Systems Biology* 7.

Silverstein RN, Correa AMS, and Baker AC. 2012. Specificity is rarely absolute in coral–algal symbiosis: implications for coral response to climate change. *Proceedings of the Royal Society B: Biological Sciences* 279:2609-2618.

Smith SV, and Buddemeier RW. 1992. Global Change and Coral Reef Ecosystems. *Annual Review of Ecology and Systematics* 23:89-118.

Stat M, Carter D, and Hoegh-Guldberg O. 2006. The evolutionary history of *Symbiodinium* and scleractinian hosts—Symbiosis, diversity, and the effect of climate change. *Perspectives in Plant Ecology, Evolution and Systematics* 8:23-43.

Stat M, Pochon X, Cowie ROM, and Gates RD. 2009. Specificity in communities of *Symbiodinium* in corals from Johnston Atoll. *Mar Ecol Prog Ser* 386:83-96.

Swanson R, and Hoegh-Guldberg O. 1998. Amino acid synthesis in the symbiotic sea anemone Aiptasia pulchella. *Marine Biology* 131:83-93.

Szmant AM, Weil E, Miller MW, and Colón DE. 1997. Hybridization within the species complex of the scleractinan coral *Montastraea annularis*. *Marine Biology* 129:561-572.

Thornhill D, LaJeunesse T, Kemp D, Fitt W, and Schmidt G. 2006. Multi-year, seasonal genotypic surveys of coral-algal symbioses reveal prevalent stability or post-bleaching reversion. *Marine Biology* 148:711-722.

Thornhill DJ, Kemp DW, Bruns BU, Fitt WK, and Schmidt GW. 2008. Correspondence between cold tolerance and temperate biogeography in a western Atlantic *Symbiodinium* (Dinophyta) lineage. *Journal of Phycology* 44:1126-1135.

Thornhill DJ, Lajeunesse TC, and Santos SR. 2007. Measuring rDNA diversity in eukaryotic microbial systems: how intragenomic variation, pseudogenes, and PCR artifacts confound biodiversity estimates. *Molecular Ecology* 16:5326-5340.

Toller WW, Rowan R, and Knowlton N. 2001. Zooxanthellae of the *Montastraea annularis* Species Complex: Patterns of Distribution of Four Taxa of *Symbiodinium* on Different Reefs and Across Depths. *The Biological Bulletin* 201:348-359.

Supplementary Information

Supplementary Table 1. Number of polymerase chain reaction (PCR) cycles for each *Oribcella faveolata* and *Orbicella franksi* hosts to obtain similar faint band intensity in both the Flower Garden Banks and Mexico dataset. (*O.franksi=Orbicella franksi*, *O. faveolata=Orbicella faveolata*, FGB=Flower Garden Banks)

PCR tube ID	Individual ID	Location	Species	Final cycle number
1	fran102	east FGB	*O. franksi*	23
2	fran105	east FGB	*O. franksi*	26
3	fran101	east FGB	*O. franksi*	28
4	fran216	east FGB	*O. franksi*	28
5	fran218	east FGB	*O. franksi*	30
6	fran221	east FGB	*O. franksi*	30
7	fran222	east FGB	*O. franksi*	28
8	fran224	east FGB	*O. franksi*	28
9	fran231	east FGB	*O. franksi*	32
10	fran232	east FGB	*O. franksi*	26
11	fran233	east FGB	*O. franksi*	26
12	fran236	east FGB	*O. franksi*	28
13	fran82	east FGB	*O. franksi*	27
14	fran90	east FGB	*O. franksi*	30
15	fran91	east FGB	*O. franksi*	28
16	fran106	west FGB	*O. franksi*	28
17	fran107	west FGB	*O. franksi*	28
18	fran112	west FGB	*O. franksi*	23
19	fran114	west FGB	*O. franksi*	28
20	fran115	west FGB	*O. franksi*	27
21	fran116	west FGB	*O. franksi*	28
22	fran117	west FGB	*O. franksi*	27
23	fran199	west FGB	*O. franksi*	23
24	fran200	west FGB	*O. franksi*	28
25	fran201	west FGB	*O. franksi*	23
26	fran202	west FGB	*O. franksi*	24
27	fran203	west FGB	*O. franksi*	24
28	fran210	west FGB	*O. franksi*	23
29	fran205	west FGB	*O. franksi*	29
30	fran198	west FGB	*O. franksi*	25

31	----	----	H$_2$O control	----
32	154	east FGB	*O. faveolata*	29
33	140	east FGB	*O. faveolata*	33
34	117	east FGB	*O. faveolata*	25
35	142	east FGB	*O. faveolata*	23
36	143	east FGB	*O. faveolata*	33
37	145	east FGB	*O. faveolata*	24
38	137	east FGB	*O. faveolata*	25
39	150	east FGB	*O. faveolata*	25
40	153	east FGB	*O. faveolata*	30
41	157	east FGB	*O. faveolata*	32
42	161	east FGB	*O. faveolata*	28
43	162	east FGB	*O. faveolata*	25
44	123	east FGB	*O. faveolata*	33
45	128	east FGB	*O. faveolata*	29
46	94	west FGB	*O. faveolata*	33
47	95	west FGB	*O. faveolata*	25
48	89	west FGB	*O. faveolata*	33
49	188	west FGB	*O. faveolata*	28
50	98	west FGB	*O. faveolata*	24
51	99	west FGB	*O. faveolata*	25
52	97	west FGB	*O. faveolata*	27
53	107	west FGB	*O. faveolata*	25
54	102	west FGB	*O. faveolata*	23
55	103	west FGB	*O. faveolata*	27
56	112	west FGB	*O. faveolata*	23
57	91	west FGB	*O. faveolata*	25
58	191	west FGB	*O. faveolata*	25
59	109	west FGB	*O. faveolata*	no amplification
60	111	west FGB	*O. faveolata*	28
61	TC1	Mexico	*O. faveolata*	21
62	TC2	Mexico	*O. faveolata*	21
63	TC3	Mexico	*O. faveolata*	21
64	TC4	Mexico	*O. faveolata*	22
65	TC5	Mexico	*O. faveolata*	21
66	TB1	Mexico	*O. faveolata*	21
67	TB2	Mexico	*O. faveolata*	21
68	TB3	Mexico	*O. faveolata*	22

69	TB4	Mexico	*O. faveolata*	25
70	TB5	Mexico	*O. faveolata*	25
71	TR1	Mexico	*O. faveolata*	21
72	TR2	Mexico	*O. faveolata*	20
73	TR3	Mexico	*O. faveolata*	21

Supplementary Table 2. A-Rapid barcode + ITS-2 sequence used to uniquely tag coral hosts. (its-DINO primer from Pochon et al. 2001)

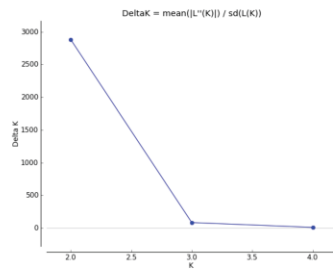| Barcode Primer Name	Name	Barcode	A-rapid + Barcode + Pochon ITS-2 Reverse Primer
16	Ar-ITS2R-16	TGTAGCGC	CCATCTCATCCCTGCGTGTCTCCGACGACT + TGTAGCGC + CCTCCGCTTACTTATATGCTT
17	Ar-ITS2R-17	ACACTCAC	CCATCTCATCCCTGCGTGTCTCCGACGACT + ACACTCAC + CCTCCGCTTACTTATATGCTT
18	Ar-ITS2R-18	ACATATAG	CCATCTCATCCCTGCGTGTCTCCGACGACT + ACATATAG + CCTCCGCTTACTTATATGCTT
19	Ar-ITS2R-19	ACGTGATC	CCATCTCATCCCTGCGTGTCTCCGACGACT + ACGTGATC + CCTCCGCTTACTTATATGCTT
20	Ar-ITS2R-20	AGTACAGA	CCATCTCATCCCTGCGTGTCTCCGACGACT + AGTACAGA + CCTCCGCTTACTTATATGCTT
21	Ar-ITS2R-21	AGTATCTC	CCATCTCATCCCTGCGTGTCTCCGACGACT + AGTATCTC + CCTCCGCTTACTTATATGCTT
22	Ar-ITS2R-22	ATACGCTG	CCATCTCATCCCTGCGTGTCTCCGACGACT + ATACGCTG + CCTCCGCTTACTTATATGCTT
23	Ar-ITS2R-23	ATCTAGTC	CCATCTCATCCCTGCGTGTCTCCGACGACT + ATCTAGTC + CCTCCGCTTACTTATATGCTT
24	Ar-ITS2R-24	CAGCGTAG	CCATCTCATCCCTGCGTGTCTCCGACGACT + CAGCGTAG + CCTCCGCTTACTTATATGCTT
25	Ar-ITS2R-25	CGCACGAG	CCATCTCATCCCTGCGTGTCTCCGACGACT + CGCACGAG + CCTCCGCTTACTTATATGCTT
26	Ar-ITS2R-26	CGTGTGCG	CCATCTCATCCCTGCGTGTCTCCGACGACT + CGTGTGCG + CCTCCGCTTACTTATATGCTT
27	Ar-ITS2R-27	CTAGATAC	CCATCTCATCCCTGCGTGTCTCCGACGACT + CTAGATAC + CCTCCGCTTACTTATATGCTT
28	Ar-ITS2R-28	CTGTCGAC	CCATCTCATCCCTGCGTGTCTCCGACGACT + CTGTCGAC + CCTCCGCTTACTTATATGCTT
29	Ar-ITS2R-29	TATGCACG	CCATCTCATCCCTGCGTGTCTCCGACGACT + TATGCACG + CCTCCGCTTACTTATATGCTT
30	Ar-ITS2R-30	TCGCGCTA	CCATCTCATCCCTGCGTGTCTCCGACGACT + TCGCGCTA + CCTCCGCTTACTTATATGCTT
88	Ar-ITS2R-88	GCAGTACG	CCATCTCATCCCTGCGTGTCTCCGACGACT + GCAGTACG + CCTCCGCTTACTTATATGCTT
89	Ar-ITS2R-89	GCGATCGT	CCATCTCATCCCTGCGTGTCTCCGACGACT + GCGATCGT + CCTCCGCTTACTTATATGCTT
90	Ar-ITS2R-90	GCGCTATA	CCATCTCATCCCTGCGTGTCTCCGACGACT + GCGCTATA + CCTCCGCTTACTTATATGCTT
91	Ar-ITS2R-91	GTACAGAT	CCATCTCATCCCTGCGTGTCTCCGACGACT + GTACAGAT + CCTCCGCTTACTTATATGCTT
92	Ar-ITS2R-92	GTAGCTCT	CCATCTCATCCCTGCGTGTCTCCGACGACT + GTAGCTCT + CCTCCGCTTACTTATATGCTT
93	Ar-ITS2R-93	GTATAGTG	CCATCTCATCCCTGCGTGTCTCCGACGACT + GTATAGTG + CCTCCGCTTACTTATATGCTT

94	Ar-ITS2R-94	GTCAGCGA	CCATCTCATCCCTGCGTGTCTCCGACGACT + GTCAGCGA + CCTCCGCTTACTTATATGCTT
95	Ar-ITS2R-95	GTCGCAGT	CCATCTCATCCCTGCGTGTCTCCGACGACT + GTCGCAGT + CCTCCGCTTACTTATATGCTT
96	Ar-ITS2R-96	GTCTCACG	CCATCTCATCCCTGCGTGTCTCCGACGACT + GTCTCACG + CCTCCGCTTACTTATATGCTT
97	Ar-ITS2R-97	GTGACTCA	CCATCTCATCCCTGCGTGTCTCCGACGACT + GTGACTCA + CCTCCGCTTACTTATATGCTT
98	Ar-ITS2R-98	TACACGCT	CCATCTCATCCCTGCGTGTCTCCGACGACT + TACACGCT + CCTCCGCTTACTTATATGCTT
99	Ar-ITS2R-99	TACGATAT	CCATCTCATCCCTGCGTGTCTCCGACGACT + TACGATAT + CCTCCGCTTACTTATATGCTT
100	Ar-ITS2R-100	TAGACAGA	CCATCTCATCCCTGCGTGTCTCCGACGACT + TAGACAGA + CCTCCGCTTACTTATATGCTT
101	Ar-ITS2R-101	TAGTACTC	CCATCTCATCCCTGCGTGTCTCCGACGACT + TAGTACTC + CCTCCGCTTACTTATATGCTT
102	Ar-ITS2R-102	TATATGTC	CCATCTCATCCCTGCGTGTCTCCGACGACT + TATATGTC + CCTCCGCTTACTTATATGCTT

Supplementary Figure 1. Gel image displaying barcoded individuals for Flower Garden

Banks and Mexico datasets. Three pools were created to submit for deep amplicon

sequencing. Band intensities were used to try to obtain equal sequencing coverage across
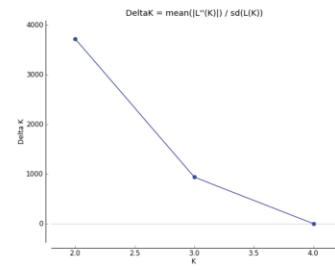
all individuals in each pool.
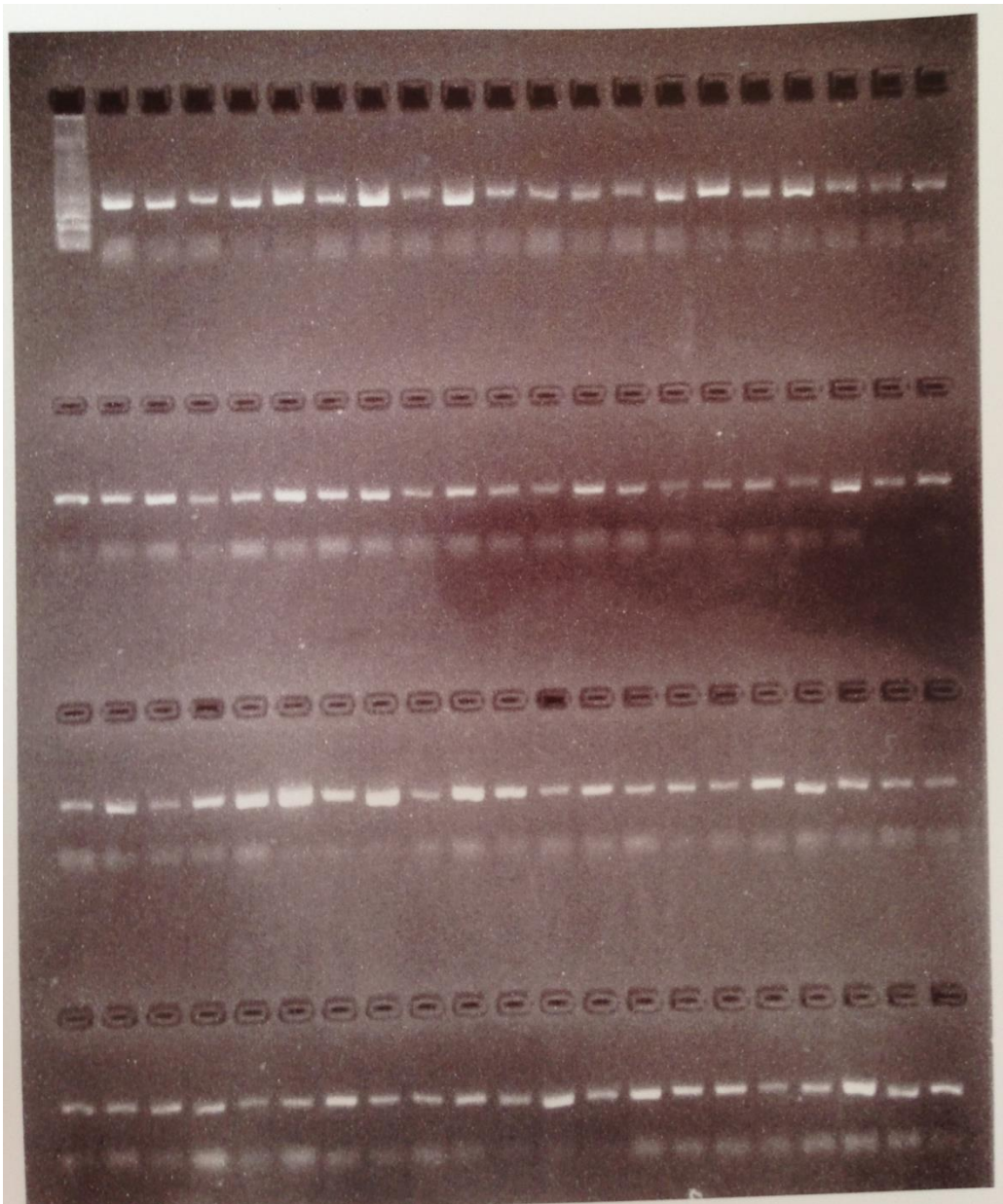
Supplementary Table 3. Number of polymerase chain reaction (PCR) cycles for

Oribcella faveolata hosts to obtain similar faint band intensity in the Curaçao dataset and

the subsequent barcode numbers annealed to each coral host sample.

	Coral Host ID	Final Cycle Number	Barcode Number
	WP 13	25	19
	WP 14	25	18
Shallow	WP 15	25	20
	WP 17	25	16
	WP 21	25	17
	WP 7	25	24
	WP 8	25	21
Middle	WP 10	25	23
	WP 11	25	22
	WP 12	25	25
	WP 1	25	28
	WP 2	25	29
Deep	WP 3	25	30
	WP 5	25	27
	WP 6	25	26
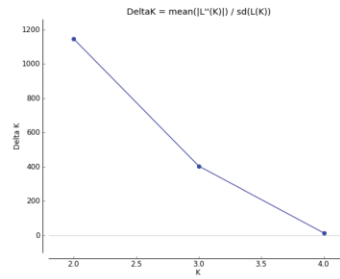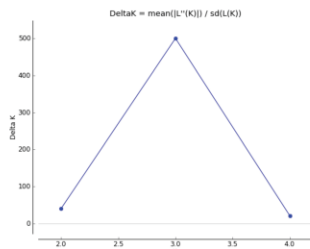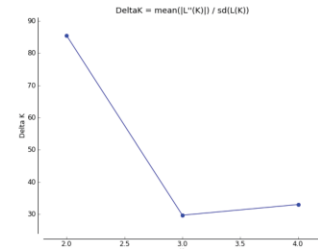
Supplementary Figure 2. Delta K figures from STRUCTURE HARVESTER for all collected *Orbicella faveolata* collected coral hosts (n=193), with potential hybrids removed (n=174), the selected 60 coral hosts, *Orbicella faveolata* (n=73) and *Orbicella franksi* (n=101).
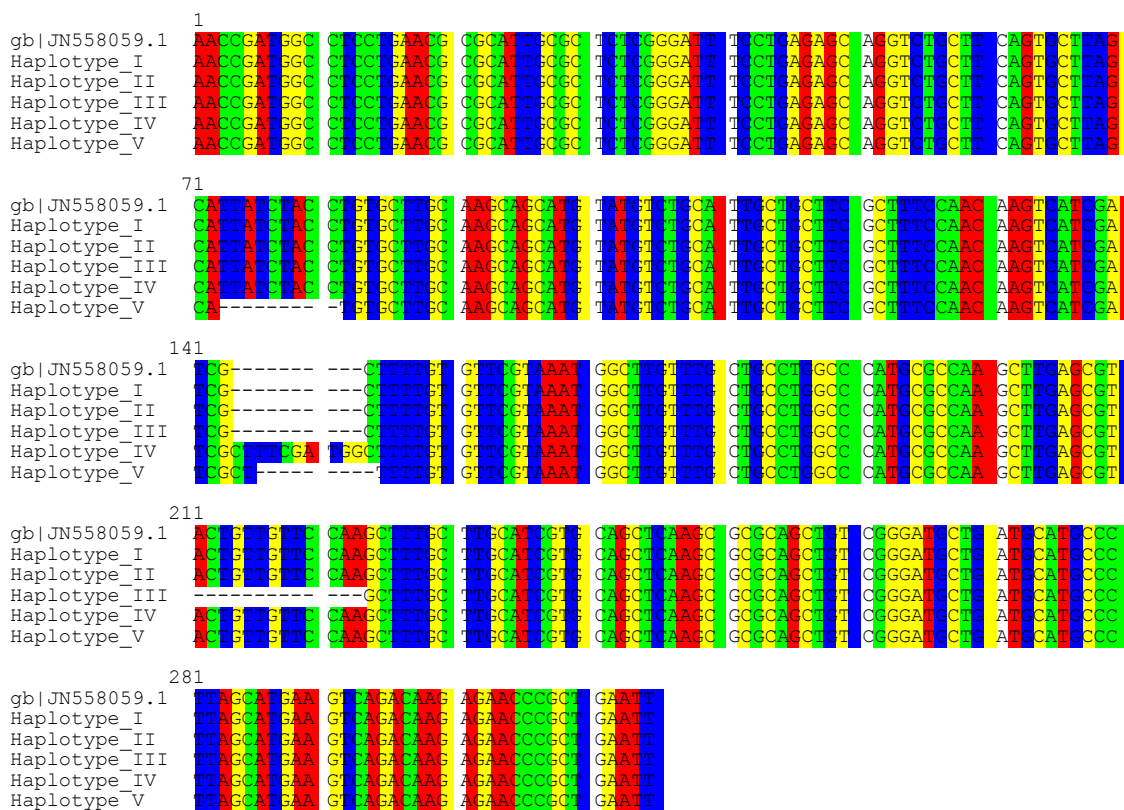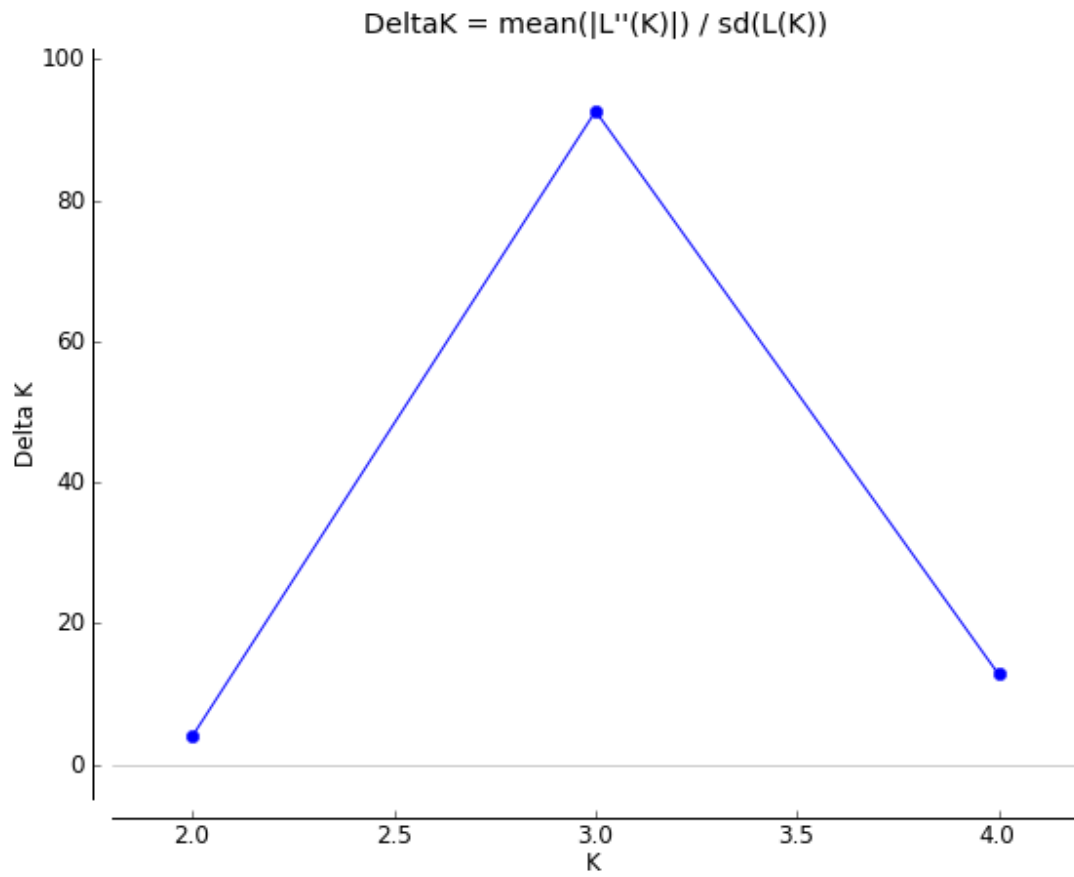
Supplementary Figure 3. Clustal Omega alignment of Flower Garden Banks

Symbiodinium B1 five reference haplotypes and the previously published *Symbiodinium*

B1 (JN 558059.1) displayed in SeaView.

Supplementary Figure 4. Delta K figure from STRUCTURE HARVESTER for the

Curaçao dataset. (K=3, n=15)