The Pennsylvania State University

The Graduate School

GENOMIC AND TRANSCRIPTOMIC INSIGHTS ON THE *ORBICELLA* SPECIES COMPLEX

A Dissertation in

Biology

by

Ana María González Angel

Submitted in Partial Fulfillment

of the Requirements

for the Degree of

Doctor of Philosophy

December 2019

The dissertation of Ana María González Angel was reviewed and approved* by the following:

James Marden

Professor of Biology

Chair of Committee

Mónica Medina

Professor of Biology

Dissertation Advisor

István Albert

Research Professor of Bioinformatics

Stephen W. Schaeffer

Professor of Biology

Associate Department Head of Graduate Education

*Signatures are on file in the Graduate School

ABSTRACT

In this dissertation I explore different aspects of the ecological speciation and genomics of marine organisms. First, I review the literature and explore different factors, particularly, depth, that isolate marine populations. These factors can be sufficiently strong to enhance divergence among populations, and so, lead to reproductive isolation. I discuss numerous examples but focus on the well documented case of *Orbicella,* a Caribbean coral genus that is environmentally segregated and has evolved at least two mechanisms of reproductive isolation, temporal isolation and gamete recognition. In Chapter 2, I explore the genomes of these corals to infer how different their protein coding ortholog groups are and if gene content is reflected in the differentiation in these taxa. Overall, I found the genomes of *Orbicella* sister species are extremely similar and other factors may be responsible for their differentiation such as gene expression, gene silencing, differential transcription factor activity or SNP presence. In Chapter 3, I study the temporal isolation by assessing the gene expression profiles of two sister *Orbicella* species during and after the moment of spawning, which occurs only one time per year. Although their genomes are very similar, gene expression profiling suggests these species use their genetic toolkit very differently. Minimal overlap was found in the differentially expressed genes (DEG) involved in the spawning behavior, and the ones that do not overlap have different identities and putative functions while many others are remain uncharacterized and unknown. Further studies including more timepoints are need to address the rhythmicity of the DEG putatively responsible for the allochronic assortative mating (or timing of gamete release) that occurs in *Orbicella*. Other aspects of the reproductive barriers such as gamete recognition and hybrid inviability also require attention.

Overall, the studies of the genomic and cellular elements responsible for the prezygotic isolation in *Orbicella* are still nascent and warrant more work to unravel their complexity and consequences in ecology and evolutionary history of this group.

Table of Contents

List of Figures……..viii

List of Tables x

Acknowledgements xi

Chapter 1, Ecological Speciation in Corals……..1

Abstract 1

Introduction 2

Biodiversity in the Ocean 3

Speciation in the Ocean 5

Ecological Speciation 5

Adaptation Across Gradients in the Sea 7

Depth as a Driver of Ecological Speciation in Coral Reefs ……. ...…….9

Mechanisms of reproductive isolation among populations living in different habitats ……...…….....15

Spawning timing……...15

Sperm-egg recognition systems……18

Evolutionary Genomics of the Coral Speciation Process…….21

Conclusion……23

References……24

Chapter 2, Gene Orthology assessment in *Orbicella*……...41

Abstract…….41

Introduction……...41

Methods……..……..43

Results……...44

Discussion…….52

Conclusion…….……54

References…….……55

Chapter 3, Transcriptional insights of temporal isolation in broadcast spawning corals

Abstract…….59

Introduction……...60

Methods…….65

Results 69

Discussion 73

Conclusions 77

References……. 77

Chapter 4, Conclusion 84

Appendix A: *Cyphastrea* genome assembly supplementary information…… 88

Appendix B: Lists of summarized (p_{adj} <0.05) gene ontology groupings in the genomes of

Orbicella ……99

Appendix C: Lists of transcripts present in gametes and parental samples…….109

Appendix D: Lists of Differentially Expressed Genes in *Orbicella annularis* and *O. franksi*

……..118

LIST OF FIGURES

Figure 1-1: *O. annularis, O. franksi,* and *O. faveolata*. Photos are courtesy of Mónica Medina. ……10

Figure 1-2: We retrieved ortholog protein sequences from protein models from the genomes of four symbiotic cnidarians (*O. faveolata, Acropora digitifera, Stylophora pistillata,* and *Exaiptasia pallida*) and one asymbiotic cnidarian (*Nematostella vectensis*) using blast bidirectional best hit (BBHs) (Altschul et al.,1990). We did protein alignments and curation with ClustalW (Thompson et al.,1994) and Gblocks (Castresana, 2000). We built protein distance matrices using Hamming dissimilarities algorithm implemented in Ugene (Okonechnikov et al.,2012) (Fig 1). The heatmaps of the protein distances between different reproductive proteins in five Cnidarians are depicted. Green colors represent closer distances (fully conserved proteins equal to 1), while red colors represent more distant relationships (equals a value of 0.05). Grey indicates sequence absence. A) Comparison among proteins from sister *Orbicella* species B) Comparison of *O. faveolata, Acropora digitifera, Stylophora pistillata,* and *Exaiptasia pallida* genomes and *Nematostella vectensis*.....23

Figure 3.1: Sample collection time points sequenced in this study and experimental design. The orange boxes around the sampling time points for each species show the transcriptome comparisons we present here. The horizontal arrow at the bottom signals time progress and spawning occurred on September 14 2014.....69

Figure 3-2: Heatmaps showing the differentially expressed genes in *O. annularis* and *Orbicella franksi*. A) *Orbicella franksi* data B) *O. annularis* data. The samples in batches of four correspond to the biological replicates during either spawning or post-spawning. There are 117 in *O. franksi* and 48 DEG in *O. annularis*. Colors indicate expression such that yellow represents upregulation respect to blue hues and log2 fold change of Z scores. 74

Figure 3-3: Venn Diagrams showing the overlap in differentially gene expression profiles in three species during spawning and after spawning. Comparisons were based on UNIPROT/SWISSPROT identifier codes. Note than only the transcripts that had an annotation are plotted here (73/117 known to be DEG in *O. franksi*, 25/48 known to be DEG in *O. annularis*, 198/206 known to be DEG in *Acropora millepora*).....76

LIST OF TABLES

Table 1-1: Cases of ecological segregation in marine invertebrates. The segregation column refers to Depth (D), Latitudinal gradients (L), Habitat (H), Host Preference (HP), or Intertidal height (I). Most studies consider mesophotic environments as habitats beginning at 30 meters of depth, but sometimes this varies depending on author's criteria. For more details on mesophotic description, see (Laverick, et al.,2017))..... ..7

Table 2-1: Metrics of ortholog groups found in the genomes of *A. cervicornis, A. palmata, O. annularis, O. faveolata*, and *O franksi* …….48

Table 2-2: Quantitative overlap in orthologous groups.....49

Table 2-3: Species-specific ortholog groups in *O. franksi*. Annotations were obtained by aligning to NCBI, BLAST P program, non-redundant database. The bold sequences are the description sequence of the orthologous group and it is also the one that corresponds to the annotation boxes.....53

Table 2-4: Species-specific ortholog groups in *O. faveolata.* Annotations were obtained by aligning to NCBI, BLAST P program, non-redundant database. The bold sequences are the description sequence of the orthologous group and it is also the one that corresponds to the annotation boxes. 54

Table 3.1: Compiled data from genomes available at the time of the analyses ……70

ACKNOWLEDGEMENTS

I want to thank my doctoral supervisor, Monica Medina, for introducing me to the wonderful world of *Orbicella* and guiding me through this journey. I am also very grateful to my committee members, all of which helped me greatly at different times. Jim Marden told me to "think of yourself as a detective. This is when science becomes an art" inspiring me to find art in my work. István Albert and his team were instrumental in getting my bioinformatics skills and confidence started because "if you know bioinformatics you are already in the top 1% of the population". Finally, I also thank Professor Schaeffer who mentored me and even advocated for the allocation of financial assistance from the Department of Biology towards the end of my degree. Overall, I had a critical, constructive and caring committee for which I am very grateful.

I also want to thank the Medina Lab members that I have crossed paths with. We shared moments of hard work, good and bad times. I learned very much from my conversations with Bishoy, Aki and Viridiana. Carlos Prada in particular became a mentor and close friend.

I want to thank my parents for their consistent love and support. I also want to thank my friends in town and around the globe for our countless moments of friendship and support that I will cherish forever. And, finally, this was all possible given the financial support I received primarily from Colciencias, but also Mónica's start-up funds, her NSF funds, and the Department of Biology at Penn State.

Chapter 1

Ecological Speciation in Corals

Ana M. González, Carlos A. Prada, Viridiana Ávila, and Mónica Medina

A. M. González, V. Ávila, M. Medina (*) Department of Biology, Pennsylvania State University, University Park, PA, USA e-mail: mum55@psu.edu

C. A. Prada Department of Biological Sciences, University of Rhode Island, Kingston, RI, USA

Marjorie F. Oleksiak and Om P. Rajora (eds.), Population Genomics: Marine Organisms, Population Genomics [Om P. Rajora (Editor-in-Chief)], https://doi.org/10.1007/13836_2018_35,

Abstract

The ocean is generally a homogenous environment with few geographic barriers that allow populations to connect over hundreds of kilometers, increasing gene flow and slowing down diversification and the formation of species. However, biodiversity in the ocean is vast across thousands of kilometers and even within single individuals (e.g., coral colonies). Species diversity peaks at coral reef ecosystems, which house at least one quarter of the marine biodiversity. Why are these systems so diverse? How do species differentiate despite rampant genetic connectivity? One possibility to explain biodiversity hotspots in the ocean, along with physical barriers, is through ecological factors.

Populations can diverge if they specialize ecologically, reducing interbreeding, which can lead to reproductive isolation. We reviewed cases of speciation in coral reefs with emphasis on those driven by ecological factors. We find few studies in coral research using genomic approaches to understand the genetics of reproductive isolation. We propose the cases of the coral *Orbicella* spp. and the octocoral *Eunicea* spp. as ideal examples to study ecological speciation in corals.

Introduction

The study of species formation is not only critical for enhancing marine conservation, but it is also one of the major topics of interest in evolutionary biology (Darwin 1909; Mayr 1942; Coyne and Orr 2004; Nosil and Feder 2012). Species form when reproductive isolation (RI) develops, preventing the breeding of groups of organisms and leading to genetically differentiated populations (Mayr 1942). Reproductive isolation is regarded as a fundamental process for species generation. One way to achieve RI is as a by-product of geographical isolation. For example, the rise of the Isthmus of Panama roughly 3 million years ago resulted in profound oceanographic (redirection of currents in the Gulf of Mexico and interoceanic closure) and biological impacts including isolation of populations on either side, preventing gene flow and eventually generating thousands of new sister species on either side of the Isthmus (Lessios 1979; O'Dea et al. 2016). Alternatively, though not mutually exclusive, adaptation to different habitats could result in RI and speciation via ecological factors (Rundle and Nosil 2005; Prada et al. 2008; Prada and Hellberg 2013).

Speciation via geographical barriers has been traditionally emphasized in terrestrial taxa (Coyne and Orr 2004). Physical barriers act as hard boundaries that block gene flow

among populations, allowing divergence and the formation of new species. As most populations in terrestrial systems are fragmented across landscapes with little genetic connectivity and restricted gene flow, geographical isolation is often found to be the causative agent for species divergence. In contrast, marine species often disperse across hundreds of kilometers as planktonic larvae, enhancing gene flow among populations and hindering population differentiation and speciation. For marine species in which gene flow persists over large geographical scales, the formation of species may be largely the result of ecologically based divergent selection.

Here we review studies of ecological speciation in marine environments. We highlight those studies that benefited from incorporating modern genomic tools and multidisciplinary work involving ecology, morphology, behavior, experimental, and evolutionary biology. We also favored coral systems given corals' ecological relevance and our own expertise.

Biodiversity in the Ocean

Speciation in the sea has been prolific and has resulted in over 243,000 species (WoRMS Editorial Board 2018) with an abundant presence of undescribed and unrecognized cryptic species that could boost biodiversity estimates to at least tenfold (Sala and Knowlton 2006). For example, according to May (1994), 32 of 33 animal phyla occur in the sea, 21 of which are exclusively marine, whereas only 12 phyla occur on land, and only 1 is exclusive to land.

Biodiversity in the sea is not only astonishing based on the number of species but also the uniqueness of body plans, which partly reflects the action of natural selection in these systems (Knowlton, 2010). Marine biodiversity is particularly rich in coral reefs,

which contain one quarter of all species in the ocean (Reaka-Kudla 2005; Sala and Knowlton 2006). Coral reef ecosystems occur, whereby hermatypic corals grow large colonies that form complex 3D networks of living tissue and a myriad of niches for other species, creating a marine biodiversity hotspot (Birkeland 2015). In the Caribbean alone, researchers have recorded 12,000 marine species, though this is likely an underestimation considering only a few islands in the Caribbean have been explored and the lack of taxonomic expertise for certain groups (Miloslavich et al. 2010). Coral reefs are ecologically important as corals store carbon in their skeleton and thus act as CO_2 sinks, alleviating carbon dioxide in the atmosphere and regulating other biogeochemical cycles such as sulfur (Raina et al. 2013). In addition, healthy reefs provide both ecosystem services by mitigating beach erosion from storms and hurricanes and economic services including industries in tourism, fisheries, jewelry, aquarium hobbies, and aquaculture (Spalding et al. 2004).

One of the properties of the biodiversity on coral reefs is that it is highly stratified with different kinds of organisms occupying different habitats (Montaggioni and Braithwaite 2009). For example, plate-like corals are often found in deep areas below 25 m at the reef drop off zone. Branching corals need more light and are more resistant to wave action being found in the reef crest and fore reef areas such as *Acropora cervicornis* and *A. palmata* in the Caribbean. Massive corals are often found at intermediate habitats such as some *Orbicella* (formerly known as *Montastraea*) (Goreau, 1959). Such segregation of coral species along reef habitats is also reflected at finer scales with a plethora of sister species often occupying different habitats (Knowlton 1993). The co-occurrence of these sister marine species with high dispersal capabilities poses a challenge for evolutionary biologists trying to understand how new species emerge without obvious geographic isolation.

Speciation in the Ocean

Speciation has been largely studied on land, where reproductive isolation is often achieved due to physical barriers such as rivers and mountains that isolate populations and generate new species (Coyne and Orr 2004; Mayr 1954; Morris-Pocock et al. 2016; Hayes and Sewlal 2004; Ceccarelli et al. 2016). While speciation via geographical isolation occurs in the ocean (Lessios et al. 2001), the scarcity of physical barriers suggests this mode of isolation does not operate as widely as on land (Palumbi 1994). Contrary to terrestrial systems, many marine organisms engage in external fertilization and have planktonic larvae that can disperse hundreds of kilometers, connecting populations across vast distances (Lessios and Robertson 2006, 2013; Roberts 1997).

The dynamics among populations in the sea differs from that on land in at least two ways: (1) there is higher gene flow among populations and (2) populations sustain larger number of individuals (i.e., larger effective population sizes). Gene flow and population size influence the rate of speciation. Increased gene flow delays genomic differentiation and speciation. Similarly, larger populations take longer for drift to fix, further slowing diversification and speciation. Apart from geographical isolation, environmental differentiation often generated by physical variation can influence the formation of species in the sea. Adaptation of populations across these environmental changes such as gradients of light, temperature, and depth can cause ecologically based divergent selection.

Ecological Speciation

During ecological speciation, RI is achieved by divergent natural selection acting on ecologically segregated populations even when dispersal is not an impediment to

random mating (Rundle and Nosil 2005). In these instances, speciation appears to have occurred due to natural selection acting on genes responsible for ecological traits. Even when gene flow is absent during divergence, ecological speciation can accelerate the process because different alleles may be fixed under different environments under natural selection (Schluter 2009). Similarly, because local adaptation generates alternative states in different environments, when nascent species come into contact, they are less likely to reproduce because both extrinsic and intrinsic factors reduce gene flow (Doebeli 2005; van Doorn et al. 2009). Ecological speciation research has provided evidence that RI can happen rapidly in both plants and animals (Savolainen et al. 2006; Barluenga et al. 2006) producing parallel patterns across taxa and geographical regions (Østbye et al. 2005; Derome and Bernatchez 2006; Quesada et al. 2007; Schluter 2009). Given that the ocean is one of the most stratified systems on earth, ecological divergence may play a fundamental role in promoting speciation in marine taxa with high dispersal potential (Table 1). In fact, ecological segregation is widespread among closely related marine species with genetic differences often detected between habitat-segregated populations with overlapping ranges (Brazeau and Harvell 1994; Carlon and Budd 2002; Levitan et al. 2004; Prada et al. 2008) and adaptation of alternative ecotypes occurring even within meters in species with dispersal potential of hundreds of kilometers (Prada and Hellberg 2014). Segregated marine broadcast spawners often differ in the timing of spawning, which can lead to temporal RI (Knowlton et al. 1997). Thus, habitat segregation has the potential to link ecological and reproductive traits, increasing the likelihood of isolation (van Doorn et al. 2009). This generates assortative mating, which, coupled with habitat specificity, provides conditions where ecological differentiation can drive speciation.

Adaptation Across Gradients in the Sea

Variation in the distribution of physical and ecological factors creates environmental niches. Some of the most dissimilar niches occur at opposite ends of temperature gradients across latitudes and depth ranges of light availability and between salinity levels at fresh-to-seawater across estuaries (Table 1-1). Populations often cope with this environmental variation by adapting to different niches across these gradients, and this divergent selection across such environments creates the condition for ecological speciation.

One of the first described examples of marine speciation driven by ecological factors is that of the sponge *Chondrilla* cf. *nucula* inhabiting mangroves and coral reefs (Duran and Rützler 2006). This species displays a different morphology and coloration respective to the environment it inhabits, but, more importantly, populations from the same habitat, even if separated across vast distances, are more genetically alike than populations from different habitats found locally (Duran and Rützler 2006). Similarly, the habitat differentiation of the mobile fish *Halichoeres* spp. between coastal and more oceanic habitats has also been found to be reflected in significantly high genetic divergence (Rocha et al. 2005). Ecologically segregated populations will be genetically similar to populations in their same ecological niche even if separated by great distances while being strongly divergent from closer populations that are in different ecological niches.

Table **1-1**: Cases of ecological segregation in marine invertebrates. The segregation column refers to Depth (D), Latitudinal gradients (L), Habitat (H), Host Preference (HP), or Intertidal height (I). Most studies consider mesophotic environments as habitats

beginning at 30 meters of depth, but sometimes this varies depending on author's criteria. For more details on mesophotic description, see (Laverick, et al.,2017)).

Taxa ID	Common name	Segregation by	Environment	Marker	Reference
Favia fragum	coral	D	sea grass beds vs reef	morphometrics and allozymes	(Carlon & Budd, 2002)
Neilonella salicensis	bivalve	D	deep ocean	nuclear (28S and calmodulin intron) and mitochondrial (cytochrome c oxidase subunit 1)	(Glazier & Etter, 2014)
Callogorgia	octocoral	D	deep ocean	morphometrics and mitochondrial barcode (cox1+igr1+mtMutS) + microsatellites	(Quattrini et al., 2013; Quattrini, Baums, Shank, Morrison, & Cordes, 2015)
Orbicella	coral	D	shallow and mesophotic habitat	microsatellites	(Weil & Knowlton, 1994)
Agaricia and Symbiodinium	coral	D	shallow and mesophotic habitat	coral morphometrics and mitochondrial (*atp6*), and ITS2 for *Symbiodinium*	(Bongaerts et al., 2013)
Seriatopora hystrix	coral	D	shallow and mesophotic habitat	molecular (mtDNA and ITS2-DGGE) and photo-physiological	(Bongaerts et al., 2011)
Eunicella singularis	octocoral	D	shallow and mesophotic habitat	microsatellites and ITS1	(Costantini et al., 2016)
Eunicea flexuosa	octocoral	D	shallow habitat	morphometrics, nuclear (18S) and mitochondrial (msh1)	(Prada, Schizas, & Yoshioka, 2008)
Corallium rubrum	octocoral	D	mesophotic habitat	microsatellites	(Costantini et al., 2011)
Briareum asbestinum	octocoral	D	shallow and mesophotic habitat	allozymes	(Brazeau & Harvell, 1994)
Ophiotrix	bristle stars	D	intertidal and subtidal (>100m) temperate waters	Cytochrome oxidase I and 16S rRNA	(Taboada, S., Pérez-Portela, 2016)
Plexaura	octocoral	D	feeding strategy	Feeding behavior and morphometrics	(Lasker, Gottfried, & Coffroth, 1983)
Symbiodinium B1B184 associated with Gorgonia ventalina (coral)	algae	D	shallow habitat	microsatellites	(Kirk, Andras, Harvell, Santos, & Coffroth, 2009)
Symbiodinium and its associated with Seriatopora hystrix (coral)	algae	D	shallow habitat	rDNA ITS2	(Van Oppen, Bongaerts, Underwood, Peplow, & Cooper, 2011)
Madracis pharensis and Symbiodinium	coral	D	shallow and mesophotic habitat	mitochondrial (mtDNA: nad5) and two nuclear (nDNA: ATPSa and SRP54)	(Frade et al., 2010)
Cliona delitrix	sponge	D	shallow habitat	ecology (substratum and habitat preference)	(Halperin, Chaves-Fonnegra, & Gilliam, 2016)
Nucella	limpet	D and L	rocky shores	mtDNA COI	(González-Wevar, Nakano, Cañete, & Poulin, 2011)
Celleporella hyalina	bryozoo	H	intertidal	fitness experiments	(Hughes, 1992)
Actinia tenebrosa	anemone	H	intertidal: rock pools and boulder habitats	allozymes and microsatellite markers	(Sherman & Ayre, 2008)
Littorina subrotundrata	snail	H	salt marsh and rocky intertidal ecotypes	cytochrome B	(Kyle & Boulding, 1998)
Chondrilla nf. Nucula	sponge	H	mangal or coral reef	morphometrics and mitochondrial (COI)	(Duran & Rützler, 2006)
Phestilla	nudibranch	HP	coral reef, different coral host	mtDNA (COI) and rDNA 16S	(Faucci, Toonen, & Hadfield, 2007)
Synalpheus	snapping shrimp	HP	sponge host preference	ecology (substratum selection and demography), morphometrics and allozymes	(Duffy, 1992, 1996)
Amphitoe longimana	amphipod	HP	seeweeds	mtCOI and nuclear ITS1	(Sotka, Wares, & Hay, 2003)
Elysia viridis	sea slug	HP	host and feeding preference of temperate macroalgae	ecology tests	(Trowbridge & Todd, 2017)
Ampithoidae, Biancolinidae,Hyalidae and Hyalidae	amphipod	HP	coral, alga or sponge host	ecology tests	(Poore, Watson, de Nys, Lowry, & Steinberg, 2000)
Littorina saxatilis	snail	I	rocky intertidal	morphometrics and allozymes	(Johannesson, Johannesson, & Rolan-Alvarez, 1993)
Cellana	limpet	I	intertidal	mtDNA (12S, 16S, COI) nDNA (ATPSβ, H3)	(Bird, Holland, Bowen, & Toonen, 2011)
Acrocnida	bristle stars	I	intertidal and subtidal temperate waters	allozymes and mtCOI	(Muths, Davoult, Gentil, & Jollivet, 2006)
Petrolisthes	porcelain crabs	I	rocky intertidal	thermal tolerance	(Stillman, 2002)
Collisella	limpets	I	rocky intertidal	ecology (substratum preference and predation)	(Mercurio, Palmer, & Lowell, 1985)
Notoacmea	limpet	I	exposed shores or mudflat segregation	mtCOI and nuclear ITS1	(Nakano & Spencer, 2007)
Acanthina monodon	snail	L	intertidal	morphometrics (mtCOI was done in a different study and reported no genetic difference)	(Sepúlveda & Ibáñez, 2012)

Depth as a Driver of Ecological Speciation in Coral Reefs

Among sibling species in the sea, over 50% of the divergences involved depth as a driving factor, even though not all comparisons included depth (Knowlton, 1993). Sympatric sibling species in the sea commonly occupy depth niches differentially (Knowlton, 1993). Depth co-varies with light, water motion, sediment transport and many other physical and chemical factors. Variation in the interaction of these factors produces dissimilar distribution of resources, favoring combinations of traits that result in fitness differences among habitat segregated populations (Prada et al. 2008; Prada and Hellberg 2013). Along with physiological changes to match the environments at different depths, depth-segregated marine broadcast-spawners often differ in the timing of spawning, which can lead to temporal RI (Knowlton et al. 1997). Depth segregation has the potential to link ecological and reproductive traits, increasing the likelihood of speciation (Felsentein 1981; Tomaiuolo, et al. 2007;van Doorn, et al. 2009).

Two of the best-studied Caribbean systems in which ecological factors seem to have driven speciation across depths are the common *Orbicella* species (formerly known as *Montastraea annularis* complex) and the octocoral *Eunicea flexuosa*. The *Orbicella* genus is one of the major reef building groups in the Caribbean and includes three species: *O. faveolata, O. annularis* and *O. franksi* (Knowlton, et al. 1992)(Figure 1-1).

Figure 1-1: *O. annularis, O. franksi,* and *O. faveolata.* Photos are courtesy of Mónica Medina.

Multiple sources of evidence including behavior, genetics and ecology, have shown that each species tends to occupy different habitats (Fukami et al.,2004; Knowlton et al.,1992; Lopez, et al. 1999; Weil & Knowlton, 1994). In addition, the *Orbicella* species also correspond to distinct ecotypes that segregate by depth (Budd, Fukami, Smith, & Knowlton, 2012). For example, *O. franksi* are found at greater depths (>20m), *O. faveolata* is located at intermediate depths, and *O. annularis* is more common in shallower depths (<10m). They overlap at intermediate depths (Pandolfi & Budd, 2008; Weil & Knowlton, 1994) and are ecotypically differentiated by coral colony morphology (columnar, massive or bumpy), which likely provides ecological advantages to each species in its own depth. In fact, genome sequencing provides evidence that the extinction of previous *Orbicella* spp. created a niche gap in which modern *Orbicella* species have thrived, enabling ecological segregation of modern taxa (Prada et al.,2016). Therefore, the columnar morphology of *O. annularis* allows colonies to growth faster and better compete in shallow habitats with high sediment transport. The more massive form of *O. franksi* allows this coral to increase its area perpendicular to the reception of light, which is scarce in deeper environments. Such morphological differences are

adaptive and allow the corals to perform better in their native habitats than in non-native habitats (in the case of *O. annularis*, performance is best in shallow habitats versus deep habitats) (Pollock *et al*. in prep).

Similar to the *Orbicella* species, *Eunicea flexuosa* shows two genetically distinct, depth-segregated ecotypes that also match morphological differentiation consistent with local adaptation. Although the two distinct morphotypes used to be attributed to phenotypic plasticity, a study that used reciprocal transplantation and molecular markers (nuclear and mitochondrial) found evidence that these morphotypes are not only ecologically but also genetically distinct despite living in sympatry which explains why the morphological characters are consistently fixed for shallow (<5m) and deep (>17m) populations (Prada, et al. 2008). Moreover, studies have shown that sympatric populations of *Eunicea* segregate by depth and that migration is limited between shallow and deep zones, suggesting that survival is higher for native genotypes from each niche than for foreign recruits (Prada & Hellberg, 2013, 2014). Species in this genus take approximately 15 years to reach sexual maturity. By then, immigrant inviability operates on incoming larvae weeding out unfit colonies and selecting for locally adapted ones. In a typical case of ecological speciation, populations of *Eunicea* at different depth zones are fully segregated genetically when living in sympatry yet populations of each depth specialist maintain high levels of gene flow across the Caribbean (Prada & Hellberg, 2013, 2014). As corals in general delay sexual maturity for years to decades, selection operates for a long time (i. e. long prereproductive selection), resulting in high immigrant filtering efficiency across habitats before reproduction promoting RI (Prada & Hellberg, 2013). Both depth-segregated specialists harbor distinct *Symbiodinium* symbiont species that they select from the water column and remain host-specific even after reciprocal

transplantation, suggesting algal specificity may be a factor in the ecological segregation of *Eunicea* (Prada et al.,2014).

There are a few cases of segregation by depth across scleractinian corals varying in their degree of speciation: from little divergence (population polymorphism) to fully resolved species. *Favia fragum* corals from Panama are thought to be a case of recent speciation. Although there is some overlap at shallower depths (≤ 1 m), segregation of two *F. fragum* morphotypes is clear, and each morphotype is found at a particular depth (≤ 1 m vs 3 m) (Carlon & Budd, 2002). Polyp morphometrics and allozyme analyses suggest that segregation can be explained by an incipient speciation process with incomplete lineage sorting (divergence-with-gene-flow model) likely due to ecological division and RI since these corals are mostly self-crossing (Carlon & Budd, 2002).

In the case of *Seriatopora hystrix* from the Great Barrier Reef, depth segregation is present along the reef slope where ecotypes are exclusive to certain depth ranges. These ecotypes also establish stable symbioses with *Symbiodinium*, suggesting local adaptation to each particular depth niche in both algae and coral (Bongaerts et al.,2011). Similarly to other cases of depth segregation samples from the same local reef area are much more genetically similar to distant regional samples at the same depth, than to samples within the same area but at a different depths (Van Oppen, et al. 2011). Niche diversification based on depth has been reported in the Mediterranean octocoral *Corallium rubrum,* which is separated in two populations within the 20-70 m gradient it inhabits with a population boundary at 40-50 meters of depth (Costantini et al.,2011). Another Mediterranean coral, *Eunicella singularis,* has two morphotypes corresponding to a shallow and a deep niche that are in fact isolated genetically (Costantini et al.,2016). In both cases, it was hypothesized that the thermocline

may prevent deeper larvae from migrating to shallow water populations (Costantini et al. 2011; Costantini et al. 2016). Similarly, in Florida and United States Virgin islands (USVI), populations of *Porites astreoides* experience low vertical connectivity attributed in some cases (for example, Dry Tortugas) to mesoscale eddies that result in segregation of shallow and deep populations, whereas gene flow is high between Florida and USVI (which is almost 2000 km) despite that these corals release competent larvae that typically settle close to the parental colonies (Serrano et al.,2016). In addition to segregation in the coral host, associated *Symbiodinium* is also segregated by depth (clade A and C inhabit shallow and deep waters respectively) (Serrano et al.,2016).

Iglesias-Prieto and collaborators (2004) found vertical distribution in corals depended on the *Symbiodinium* each coral species hosts. Two depth segregated coral species, *Pocillopora verrucosa* (shallow) and *Pavona gigantea* (deep) harbor a unique algal composition based on ITS2 marker profiling: *Pavona* harbors *Symbiodinium* type C1 and *Pocillopora* harbors *Symbiodinium* type D1. Light-depth segregation in *Symbiodinium* is so strong for these two species that it can alone determine the coral host niche segregation regardless of environmental conditions, and therefore influence niche diversification. Genetic evidence supports two depth-associated lineages of the Caribbean coral *Madracis pharensis* that host different algal symbionts. Shallow corals host *Symbiodinium* type B7 whereas deep corals host *Symbiodinium* type B15 (Frade et al.,2010). A similar study of five *Agaricia* coral species found depth segregation in the coral host and host specificity with the algal populations (Bongaerts et al.,2013). And more recently, genome wide genotyping by RAD sequencing determined that reduced gene flow between depth segregated *Agaricia fragilis* resulted in genome wide indicating high selective pressure to depth adaptation despite

symbiont type (all *A. fragilis* studied hosted the same algal type) (Bongaerts et al.,2017). Interestingly, however, in the same study *Stephanocoenia intersepta* from the same reef showed no genetic structure between different depths suggesting that each species has unique natural histories and generalizations are hard to support (i.e. deep reef refugia hypothesis) (Bongaerts et al.,2017).

Octocorals are also known to occur at particular depth niches with specific *Symbiodinium* algal symbioses. *Gorgonia ventalina,* is an abundant Caribbean species that shows *Symbiodinium* genetic segregation based on depth (Kirk, Andras, Harvell, Santos, & Coffroth, 2009). In addition to the role of proteins underlying spawning behavior and fertilization in corals, other factors are influential in the process of speciation. For example, the presence and maintenance of dinoflagellate algal symbionts is key in determining the ecological niche of a given species (Bongaerts et al.,2011; Iglesias-Prieto et al.,2004; Prada & Hellberg, 2014). Genomic and trancriptomic tools have informed the ecology of coral-algal symbiosis. For example, in the case of bleaching stress *Orbicella faveolata* and *Acropora hyacinthus* transcriptomic data suggest coral physiology remains disturbed for months even after *Symbiodinium* recovery (Pinzón et al.,2015; Thomas & Palumbi, 2017). An intriguing possibility is that such physiological stress may be differently handled by corals occupying different niches and containing different symbionts (Parkinson et al.,2016). Another key finding is that transmembrane transport, response to oxidative stress and UV radiation protection genes are enriched in *Symbiodinium* genomes and transcriptomes, which are presumably necessary to maintain the symbiosis (González-Pech, et al. 2017). It remains to be seen if the evolution of these transmembrane proteins differs between species and populations occupying different habitats with varying light levels such as across depth gradients.

Ecological speciation is not exclusive to shallow environments as species also segregate along the deep ocean as well. Three deep sea sibling species in the octocoral genus *Callogorgia* also segregate by depth and by the specific environment associated to each depth (mostly explained by temperature, salinity and calcite saturation) with little overlap, indicating high depth specialization (Quattrini et al.,2013). In particular, genetic evidence from *Callogorgia delta* indicates that these octocorals segregate locally within species and are more responsive to depth than geographical distance supporting the depth-differentiation hypothesis at the species level (Quattrini et al. 2015).

Mechanisms of reproductive isolation among populations living in different habitats

Spawning timing

Adaptation to depth results in temporal reproductive isolation (Levitan et al.,2004). Coral spawning varies across depths with corals in shallow areas perceiving sunset earlier than deeper water colonies, thereby resulting in differential timing of spawning (Knowlton et al. 1997). The best case studied involves the *Orbicella* species (i.e. *O. annularis* mostly on shallow waters, *O. franksi* mostly on deep waters and *O. faveolata* in both shallow and deep waters). *O. franksi* spawns approximately two hours after sunset, whereas *O. annularis* and *O. faveolata* spawn 3:40 hours and 4:00 hours after sunset, respectively (Levitan et al. 2004; Levitan et al. 2011). This 2 hour window is ample to avoid cross-fertilization between *O. franksi* and *O. annularis* as gametes dilute and age quickly in the water column; and the overlap between *O. annularis* and *O. faveolata* does produce successful crosses at least in the

laboratory (Levitan et al. 2004; Levitan et al. 2011). In *Orbicella*, adaptation to different depths causes the development of RI due to timely species-specific gamete release events (Weil and Knowlton 1994; Levitan et al. 2011). Spawning times are sufficiently different to prevent hybridization even when corals are found in sympatry, yet conspecifics will spawn at their corresponding time. Since *O. franksi* is the earliest spawner, any unfertilized leftover gametes will drift and age by the time *O. annularis* spawns suggesting chances of successful hybridization are slim (Levitan et al. 2004; Levitan et al. 2011). Interspecific crosses induced in the laboratory yield much lower fertilization rates than intraspecific crosses (Levitan et al. 2004; Levitan et al. 2011). There is a correlation between genotype and timing of spawning in *Orbicella* corals (Levitan et al. 2011). Furthermore, depth isolated groups from the same species will spawn at comparable times indicating a strong species-specific spawning behavior as seen in *O. franksi, O. faveolata,* as well as other corals such as *M. cavernosa* and *Diploria strigosa* (Villinski, 2003; Vize, 2006).

The underlying genomic architecture of spawning behavior is partially understood. Heritable genomic components responsible for spawning behavior are thought to be associated with circadian clock networks that are triggered differently during spawning time. Some factors such as light exposure, onset of sunset, pheromones, tidal and osmotic pressure have been attributed to influence the timing of coral spawning (Baird et al. 2009; Knowlton, 1993). The majority of the evidence supports that spawning in corals is photoregulated and possibly under the influence of circadian rhythm genes (Kaniewska et al. 2015). Circadian rhythm gene networks are composed of highly conserved proteins in metazoans (Reitzel, et al.,2010), yet are known to play a role in RI between species. Because most proteins involved in biorhythms detected in corals are transcription factors (Levy et al.,2007; Shoguchi et al.,2013), it is likely

that timing of spawning and divergence in spawning time among populations and species is controlled at the transcriptional level. The *O. faveolata* genome has revealed the presence of approximately 18 circadian rhythm protein families that are likely involved in controlling spawning time in corals.

Some of the genes implicated in differential timing of spawning are responsive to blue light from lunar irradiance (Gorbunov & Falkowski, 2002), and evidence from *Acropora millepora* corals supports that at least two blue-light-sensing photoreceptor genes (cryptochromes *cry1* and *cry2*) are responsive to the moon light phases in this species (Levy et al.,2007). Studies show that gene expression measured using ESTs of *cry2* was increased in full moon nights as opposed to new moon nights indicating this gene may be operating the circadian clock thereby, participating in the regulation of spawning timing, although the involvement of other genes (like opsins) involved cannot be ruled out (Levy et al.,2007). There is not a clear understanding of what triggers spawn timing behavior in corals. It may be linked to a direct response to a light cues such as darkness (i.e. if the cue is shifted, the behavior shifts), or it could be operating under an entrained biological clock (i.e. if cue is shifted or removed the behavior continues in a rhythmic manner for some time). Most likely, at least in *Orbicella* spp. Sunset is the trigger that "starts the countdown" to spawning timing. Current studies of the transcriptome network that operates the temporal isolation behavior in *Orbicella franksi* and *Orbicella annularis* indicate a strong species-specific difference in the genes differentially expressed though these genes underlie similar functions (González et al.,*in prep*).

Sperm-egg recognition systems

In addition to differential timing of spawning, corals reproductively isolate via chemical variations in the proteins involved with sperm-egg interactions, which mediate whether fertilization is possible. After spawning and before fertilization, gametes must find and recognize each other as compatible. Gamete recognition and compatibility is crucial for successful reproduction. The sperm and egg of compatible individuals chemically recognize each other via the interaction of proteins on their surfaces (Vacquier, 1998). These reproductive proteins ultimately permit fertilization thus ensure RI in most marine broadcast spawners. Proteins responsible for gamete interactions are best known in sea urchins, abalone and turban snail species, although many eukaryote taxa are known to have reproductive proteins (Pujolar and Pogson 2011; Palmer et al. 2013; Hellberg et al. 2012; Lima and McCartney 2013; Clark, et al. 2006). Reproductive proteins are known to be among the fastest evolving proteins (Metz, et al. 1998; Swanson and Vacquier 2002). In the case of rapid evolution of reproductive proteins, and especially those involved in gamete recognition, adaptive evolution has been attributed to a series of inter- and intra-specific fertilization conflicts that seem to constantly favor rapid protein change, especially in external fertilizers (Vacquier & Swanson, 2011).

One hypothesis for the evolution of sperm-egg proteins in marine organisms is reinforcement which prevents prezygotic contact in sympatry by controlling gamete recognition such that eggs select for conspecific sperm (known as conspecific sperm precedence) or assortative mating (Marshall, et al. 2002; Fogarty et al. 2012; Palumbi 1999). This is the case of *Echinometra oblonga* and *Echinometra* sp. C, which may interbreed in no

choice crosses but that do not hybridize naturally. The eggs of these species also select for conspecific sperm (Geyer & Palumbi, 2005). These proteins tend to evolve rapidly and are attributed the ability to explain rapid speciation in marine systems even in sympatry (Geyer & Palumbi, 2003; Palumbi, 2009). In cases that gamete recognition fails to prevent all hybridization, ecological factors such as habitat or depth segregation, temporal and/or gametic isolation may aid maintaining prezygotic isolation (Lessios 2007). Morphological features in gametes (sperm shape, egg structure and size), motility limitations, and even chemical cues (pheromones) may also operate as prezygotic barriers in broadcast marine spawners (Wolstenholme 2004; Levitan 2006; Manier and Palumbi 2008; Marks et al. 2008).

An additional hypothesis for the evolution of sperm-egg proteins is sexual conflict. Intraspecific crossings are limited to the fertilization of an egg with a single sperm, since polyspermy (the fertilization of one egg by more than one sperm) leads to embryo death. As a result, sexual conflict arises between eggs and sperm, such that eggs have mechanisms to avoid polyspermy while sperm competition results in mechanisms to overcome the egg barriers. This is a sperm-density dependent scenario as rare alleles have higher fertilization rates when sperm density is high whereas more common alleles have higher fertilization rates when sperm density is low (Levitan and Ferrell 2006). In some organisms like mammals, birds, and echinoderms, eggs are able to block polyspermy after one sperm comes in successful contact with the egg (reviewed in Karr, et al. 2008). Other species modify the egg receptors to reduce the chances of insemination by multiple sperm, while sperm receptors are constantly being modified in order to fertilize eggs at all costs, a way of sexual conflict (reviewed in (Levitan 2010)). Sexual selection can also operate through cryptic female choice, which occurs when eggs prefer certain sperm surface alleles resulting in higher fertilization

rates for those allele carriers (Eberhard, 1996). Fertilization is highly dependent on density and genotype frequency of both sperm and eggs, therefore to understand the evolution of reproductive isolation based on gamete recognition proteins studies are more fruitful when observations are taken in the context of the organism's ecology (Levitan and Ferrell 2006; Palumbi 2009).

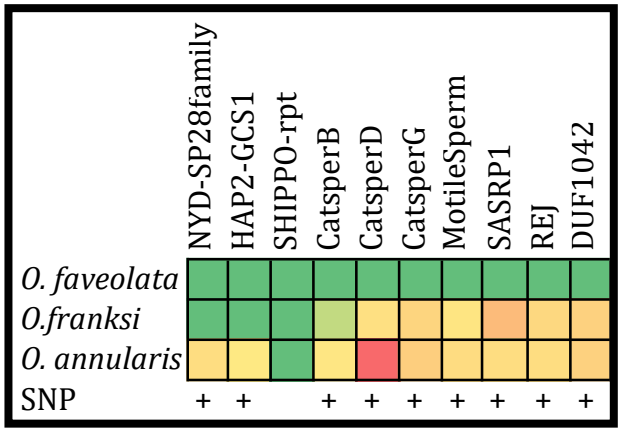
In summary, in external fertilizers like sea urchins, snails and other invertebrates, gamete recognition proteins play a key role regulating egg-sperm interactions, reproductively isolating taxa and given their fast evolution, facilitating speciation. It is unknown if gamete recognition proteins are present in corals but low fertilization rates in self-fertilization trials (Szmant, et al, 1997) and interspecific crosses suggests they may occur and mediate fertilization (Knowlton, et al. 1997). However, it is known, that asymmetric conspecific sperm precedence exists in *Orbicella* such that the early spawner *O. franksi* shows strong preference towards sperm of its own species whereas the late spawner *O. annularis* does not show strong preference in choice experiments for either species (Fogarty et al.,2012). The ecology of these species should be taken into account considering that by the time *O. annularis* spawns, leftover *O. franksi* sperm may just be too diluted and old to naturally fertilize fresh *O. annularis* eggs. When the spawning times overlap in the case of *O. annularis* and *O. faveolata*, gametes are incompatible as shown by unsuccessful laboratory cross experiments, hence preventing hybridization even when sibling congenerics are found in sympatry (Levitan et al. 2004; Szmant et al. 1997).

Evolutionary Genomics of Coral and the Speciation Process

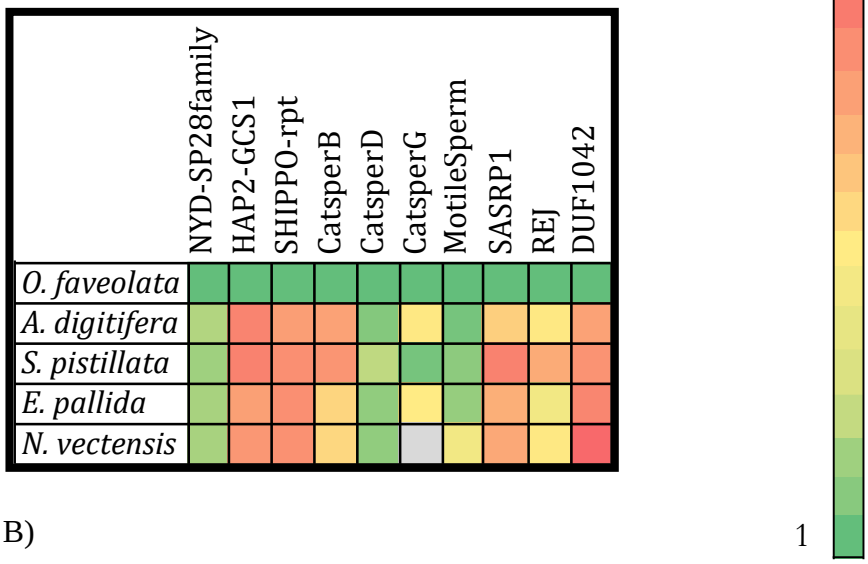
Comparative genomic research is now feasible due to the evolution of "next" generation sequencing platforms and the growing myriad of respective sophisticated analysis tools. Areas of interest within the scope of model systems have devoted attention to Genome-Wide Association Studies (GWAS). In the case of cnidarians, and particularly corals, some studies now incorporate these new technologies. Genome wide genotyping has been used to assess fine population genetics and diversity in a physical range. Genome wide data suggest *Acropora palmata* populations seem to segregate by geography (Devlin-Durante & Baums, 2017), yet *Orbicella* species segregate by depth (Carlos Prada, unpublished). This technique has also shown the lack of genetic difference in *Acropora digitifera* from Japanese reefs (Shinzato, et al. 2015).

The life histories of *Eunicea* and *Orbicella* species present a great natural experiment to study how prezygotic barriers operate in long lived broadcast spawning corals. The highly continuous genome of *Orbicella faveolata* allows the study of evolution of sperm-egg recognition proteins in corals (Prada et al. 2016). Our preliminary analysis in *Orbicella* corals indicates that substantial sequence divergence exists across candidate reproductive proteins. Figure 1A illustrates that *CatsperD,* a sperm mobility protein (Chung et al.,2011), is highly dissimilar between *O. faveolata* and *O. annularis.* We hypothesize *CatsperD* may contribute to prezygotic barriers since sperm need to swim to reach the egg and different motilities elicit different mechanical responses in the egg layers (Levitan 2000). The second molecule with substantial differences between *Orbicella* species is the Receptor for Egg Jelly protein (REJ), which is a known sperm-egg binding protein of the acrosomal reaction in sea urchins (Moy et

al. 1996; Karr, et al.,2009). These candidate proteins may be partially responsible for RI in these species.



A)



B)

Figure **1-2**: We retrieved ortholog protein sequences from protein models from the genomes of four symbiotic cnidarians (*O. faveolata, Acropora digitifera, Stylophora pistillata,* and *Exaiptasia pallida*) and one asymbiotic cnidarian (*Nematostella vectensis*) using blast

bidirectional best hit (BBHs) (Altschul et al.,1990). We did protein alignments and curation with ClustalW (Thompson et al.,1994) and Gblocks (Castresana, 2000). We built protein distance matrices using Hamming dissimilarities algorithm implemented in Ugene (Okonechnikov et al.,2012) (Fig 1). The heatmaps of the protein distances between different reproductive proteins in five Cnidarians are depicted. Green colors represent closer distances (fully conserved proteins equal to 1), while red colors represent more distant relationships (equals a value of 0.05). Grey indicates sequence absence. A) Comparison among proteins from sister *Orbicella* species B) Comparison of *O. faveolata, Acropora digitifera, Stylophora pistillata,* and *Exaiptasia pallida genomes* and *Nematostella vectensis.*

CONCLUSION

Environmental gradients often drive genetic segregation in marine populations and ecological speciation is common in the sea. One of the main examples of ecological speciation in the ocean is depth segregation on coral reefs. Organisms that harbor photosynthetic symbionts such as scleractinian corals and octocorals, are bound to physiological requirements of both host and algal symbionts. These requirements are often quite distinct due to restrictions of light penetration into the benthos, ultimately leading to reproductive isolation among populations along this depth gradient. In species with delayed reproduction such as corals, selection acts for years to decades and effectively removes unfit individuals. Adaptation to depth in these systems is tied to reproductive isolation as light cues drive gamete release timing providing temporal isolation. The rapid evolution of sperm-egg recognition proteins provides an additional prezygotic isolating barrier to maintain and

generate biodiversity in the sea. Genomic tools are enhancing our understanding of genetic variants associated with local adaptation as well as elucidating the molecular mechanisms driving reproductive isolation and speciation in the sea. The use of multidisciplinary research that combines genomic approaches with field biology promises to close gaps in our understanding of ecological genomics and marine speciation.

REFERENCES

Aronson, R., Bruckner, A., Moore, J., Precht, B. & E. Weil. 2008. *Montastraea annularis*. The IUCN Red List of Threatened Species 2008: e.T133134A3592972. http://dx.doi.org/10.2305/IUCN.UK.2008.RLTS.T133134A3592972.enAccessed on **31 May 2018**.

Aronson, R., Bruckner, A., Moore, J., Precht, B. & E. Weil. 2008. *Montastraea faveolata*. The IUCN Red List of Threatened Species 2008: e.T133373A3712432. http://dx.doi.org/10.2305/IUCN.UK.2008.RLTS.T133373A3712432.en Accessed on **31 May 2018**.

Aronson, R., Bruckner, A., Moore, J., Precht, B. & E. Weil. 2008. *Montastraea franksi*. The IUCN Red List of Threatened Species 2008: e.T133012A3542659. http://dx.doi.org/10.2305/IUCN.UK.2008.RLTS.T133012A3542659.en Accessed on **31 May 2018**.

Barluenga, Marta, Kai N Stölting, Walter Salzburger, Moritz Muschick, and Axel Meyer. 2006. “Sympatric Speciation in Nicaraguan Crater Lake Cichlid Fish.” *Nature* 439 (7077): 719–23. doi:10.1038/nature04325.

Birkeland, Charles, ed. 2015. *Coral Reefs in the Anthropocene*. *Life and Death Of Coral Reefs*. Springer.

Bongaerts, Pim, Pedro R Frade, Julie J Ogier, Kyra B Hay, Judith van Bleijswijk, Norbert Englebert, Mark J a Vermeij, Rolf P M Bak, Petra M Visser, and Ove Hoegh-Guldberg. 2013. "Sharing the Slope: Depth Partitioning of Agariciid Corals and Associated Symbiodinium across Shallow and Mesophotic Habitats (2-60 m) on a Caribbean Reef." *BMC Evolutionary Biology* 13 (1): 205. doi:10.1186/1471-2148-13-205.

Bongaerts, Pim, Cynthia Riginos, Ramona Brunner, Norbert Englebert, Struan R Smith, and Ove Hoegh-guldberg. 2017. "Deep Reefs Are Not Universal Refuges : Reseeding Potential Varies among Coral Species."

Bongaerts, Pim, Cynthia Riginos, Kyra B Hay, Madeleine J.H. van Oppen, Ove Hoegh-Guldberg, and Sophie Dove. 2011. "Adaptive Divergence in a Scleractinian Coral: Physiological Adaptation of Seriatopora Hystrix to Shallow and Deep Reef Habitats." *BMC Evolutionary Biology* 303 (October).

Brazeau, D. A., and C. D. Harvell. 1994. "Genetic Structure of Local Populations and Divergence between Growth Forms in a Clonal Invertebrate, the Caribbean Octocoral Briareum Asbestinum." *Marine Biology* 119 (1): 53–60. doi:10.1007/BF00350106.

Budd, Ann F., Hironobu Fukami, Nathan D. Smith, and Nancy Knowlton. 2012. "Taxonomic Classification of the Reef Coral Family Mussidae (Cnidaria: Anthozoa: Scleractinia)." *Zoological Journal of the Linnean Society* 166 (3): 465–529. doi:10.1111/j.1096-3642.2012.00855.x.

Carlon, David B., and Ann F. Budd. 2002. "Incipient Speciation Across a Depth Gradient in a

Scleractinian Coral ?" *Evolution* 56 (11): 2227–42.

Ceccarelli, F. Sara, Andres A Ojanguren-Affilastro, Martin J Ramírez, Jose A Ochoa, Camilo I Mattoni, and Lorenzo Prendini. 2016. "Andean Uplift Drives Diversification of the Bothriurid Scorpion Genus Brachistosternus." *Journal of Biogeography* 43: 1942–54. doi:10.1111/jbi.12760.

Chung, Jean-Ju, Betsy Navarro, Grigory Krapivinsky, Luba Krapivinsky, and David E Clapham. 2011. "A Novel Gene Required for Male Fertility and Functional CATSPER Channel Formation in Spermatozoa." *Nature Communications* 2 (January): 153. doi:10.1038/ncomms1153.

Clark, Nathaniel L, Jan E Aagaard, and Willie J Swanson. 2006. "Evolution of Reproductive Proteins from Animals and Plants." *Reproduction (Cambridge, England)* 131 (1): 11–22. doi:10.1530/rep.1.00357.

Costantini, F, Andrea Gori, Pablo Lopez-González, Lorenzo Bramanti, Sergio Rossi, Josep Maria Gili, and Marco Abbiati. 2016. "Limited Genetic Connectivity between Gorgonian Morphotypes along a Depth Gradient." *PLoS ONE* 11 (8): 1–20. doi:10.1371/journal.pone.0160678.

Costantini, F, S Rossi, E Pintus, C Cerrano, J M Gili, and M Abbiati. 2011. "Low Connectivity and Declining Genetic Variability along a Depth Gradient in Corallium Rubrum Populations." *Coral Reefs* 30 (4): 991–1003. doi:10.1007/s00338-011-0771-1.

Coyne, J. A., & Orr, H. A. (2004). *Speciation*. Sunderland, Mass: Sinauer Associates.

Darwin, Charles. 1909. *On the Origin of Species by Means of Natural Selection, or the*

Preservation of Favoured Races in the Struggle for Life. Cassell and Company.

Derome, N., and L. Bernatchez. 2006. "The Transcriptomics of Ecological Convergence between 2 Limnetic Coregonine Fishes (Salmonidae)." *Molecular Biology and Evolution* 23 (12): 2370–78. doi:10.1093/molbev/msl110.

Devlin-Durante, Meghann K., and Iliana B. Baums. 2017. "Genome-Wide Survey of Single-Nucleotide Polymorphisms Reveals Fine-Scale Population Structure and Signs of Selection in the Threatened Caribbean Elkhorn Coral, *Acropora Palmata*." *PeerJ* 5: e4077. doi:10.7717/peerj.4077.

Doebeli, M. 2005. "Adaptive Speciation When Assortative Mating Is Based on Female Preference for Male Marker Traits." *Journal of Evolutionary Biology* 18 (6): 1587–1600. doi:10.1111/j.1420-9101.2005.00897.x.

Duran, Sandra, and Klaus Rützler. 2006. "Ecological Speciation in a Caribbean Marine Sponge." *Molecular Phylogenetics and Evolution* 40 (1): 292–97. doi:10.1016/j.ympev.2006.02.018.

Eberhard, W. (1996) Female Control: Sexual Selection by Cryptic Female Choice (Princeton Univ. Press, Princeton), p. 501

Felsentein, Joseph. 1981. "Skepticism Towards Santa Rosalia, or Why Are There so Few Kinds of Animals?" *Evolution* 35 (1): 124–38.

Fogarty, ND, M Lowenberg, MN Ojima, N Knowlton, and DR Levitan. 2012. "Asymmetric Conspecific Sperm Precedence in Relation to Spawning Times in the Montastraea Annularis Species Complex (Cnidaria: Scleractinia)." *Journal of Evolutionary Biology*, 1–8. doi:10.1111/j.1420-9101.2012.02625.x.

Frade, P. R., M. C. Reyes-Nivia, J. Faria, J. A. Kaandorp, P. C. Luttikhuizen, and R. P M Bak. 2010. “Semi-Permeable Species Boundaries in the Coral Genus Madracis: Introgression in a Brooding Coral System.” *Molecular Phylogenetics and Evolution* 57 (3). Elsevier Inc.: 1072–90. doi:10.1016/j.ympev.2010.09.010.

Fukami, Hironobu, Ann F Budd, Don R Levitan, Javier Jara, Ralf Kersanach, and Nancy Knowlton. 2004. “Geographic Differences in Species Boundaries among Members of the Montastraea Annularis Complex Based on Molecular and Morphological Markers.” *Evolution; International Journal of Organic Evolution* 58 (2): 324–37.

Geyer, Laura B, and Stephen R Palumbi. 2005. “Conspecific Sperm Precedence in Two Species of Tropical Sea Urchins.” *Evolution; International Journal of Organic Evolution* 59 (1): 97–105. doi:10.1554/04-202.

González-Pech, Raúl A., Mark A. Ragan, and Cheong Xin Chan. 2017. “Signatures of Adaptation and Symbiosis in Genomes and Transcriptomes of Symbiodinium.” *Scientific Reports* 7 (1): 1–10. doi:10.1038/s41598-017-15029-w.

Gorbunov, Maxim Y., and Paul G. Falkowski. 2002. “Photoreceptors in the Cnidarian Hosts Allow Symbiotic Corals to Sense Blue Moonlight.” *Limnology and Oceanography* 47 (1): 309–15. doi:10.4319/lo.2002.47.1.0309.

Hayes, Floyd E, and Jo-anne N Sewlal. 2004. “The Amazon River as a Dispersal Barrier to Passerine Birds : Effects of River Width , Habitat and Taxonomy.” *Journal of Biogeography* 31 (11): 1809–18.

Hellberg, Michael E, Alice B Dennis, Patricia Arbour-Reily, Jan E Aagaard, and Willie J Swanson. 2012. “The Tegula Tango: A Coevolutionary Dance of Interacting, Positively

Selected Sperm and Egg Proteins." *Evolution* 66 (6): 1681–94. doi:10.1111/j.1558-5646.2011.01530.x.

Iglesias-Prieto, R, V H Beltrán, T C LaJeunesse, H Reyes-Bonilla, and P E Thomé. 2004. "Different Algal Symbionts Explain the Vertical Distribution of Dominant Reef Corals in the Eastern Pacific." *Proceedings. Biological Sciences / The Royal Society* 271 (1549): 1757–63. doi:10.1098/rspb.2004.2757.

Kaniewska, Paulina, Shahar Alon, Sarit Karako-Lampert, Ove Hoegh-Guldberg, and Oren Levy. 2015. "Signaling Cascades and the Importance of Moonlight in Coral Broadcast Mass Spawning." *ELife* 4: 1–14. doi:10.7554/eLife.09991.

Karr, Timothy L., William J. Swanson, and Rhonda R. Snook. 2009. "The Evolutionary Significance of Variation in Sperm-Egg Interactions." In *Sperm Biology*, edited by Tim Birkhead, Dave Hosken, and Scott Pitnick, First Edit, 305–65. Elsevier Ltd. doi:10.1016/B978-0-12-372568-4.00008-2.

Kirk, Nathan L., Jason P. Andras, C. Drew Harvell, Scott R. Santos, and Mary Alice Coffroth. 2009. "Population Structure of Symbiodinium Sp. Associated with the Common Sea Fan, Gorgonia Ventalina, in the Florida Keys across Distance, Depth, and Time." *Marine Biology* 156 (8): 1609–23. doi:10.1007/s00227-009-1196-z.

Knowlton, Nancy. 1993. "SIBLING SPECIES IN THE SEA." *Annual Review of Ecology and Systematics* 24 (1993): 189–216.

Knowlton, Nancy, JL Mate, HM Guzman, R Rowan, and J Jara. 1997. "Direct Evidence for Reproductive Isolation among the Three Species of the Montastraea Annularis Complex in Central America (Panama and Honduras)." *Marine Biology* 127: 705–11.

Knowlton, Nancy, Ernesto Weil, Lee A. Weigt, and Hector M.. Guzmán. 1992. "Sibling Species in Montastraea Annularis, Coral Bleaching, and the Coral Climate Record." *Science (New York, N.Y.)* 255 (5042): 330–33. doi:10.1126/science.255.5042.330.

Laverick, Jack H, Dominic A Andradi-brown, and Alex D Rogers. 2017. "Using Light-Dependent Scleractinia to Define the Upper Boundary of Mesophotic Coral Ecosystems on the Reefs of Utila , Honduras," 1–18.

Lessios, H. A., and D. R. Robertson. 2013. "Speciation on a Round Planet: Phylogeography of the Goatfish Genus Mulloidichthys." *Journal of Biogeography* 40 (12): 2373–84. doi:10.1111/jbi.12176.

Lessios, H.A. 1979. "Use of Panamanian Sea Urchins to Test the Molecular Clock." *Nature*. doi:10.1038/280599a0.

Lessios, H. 2007. "Reproductive Isolation between Species of Sea Urchins." *Bulletin of Marine Science* 81 (2): 191–208.

Lessios, H A, and D R Robertson. 2006. "Crossing the Impassable: Genetic Connections in 20 Reef Fishes across the Eastern Pacific Barrier." *Proceedings. Biological Sciences / The Royal Society* 273 (1598): 2201–8. doi:10.1098/rspb.2006.3543.

Lessios, Harilaos, B D Kessing, and J S Pearse. 2001. "Population Structure and Speciation in Tropical Seas: Global Phylogeography of the Sea Urchin Diadema." *Evolution; International Journal of Organic Evolution* 55 (5): 955–75. doi:10.1111/j.0014-3820.2001.tb00613.x.

Levitan, D. R. 2000. "Sperm Velocity and Longevity Trade off Each Other and Influence Fertilization in the Sea Urchin Lytechinus Variegatus." *Proceedings of the Royal Society*

B: Biological Sciences 267 (1443): 531–34. doi:10.1098/rspb.2000.1032.

Levitan, Don R. 2006. “The Relationship between Egg Size and Fertilization Success in Broadcast-Spawning Marine Invertebrates.” *Integrative and Comparative Biology* 46 (3): 298–311. doi:10.1093/icb/icj025.

———. 2010. “Sexual Selection in External Fertilizers.” In *Evolutionary Behavioral Ecology*, 365–79.

Levitan, Don R, and David L Ferrell. 2006. “Selection on Gamete Recognition Proteins Depends on Sex, Density, and Genotype Frequency.” *Science* 312 (5771): 267–69.

Levitan, Don R, Nicole Fogarty, Javier Jara, KE Lotterhos, and Nancy Knowlton. 2011. “Genetic, Spatial, and Temporal Components of Precise Spawning Synchrony in Reef Building Corals of the Montastraea Annularis Species Complex.” *Evolution* 65 (5). Wiley Online Library: 1254–70.

Levitan, Don R, Hironobu Fukami, Javier Jara, David Kline, Tamara M McGovern, Katie E McGhee, Cheryl a Swanson, and Nancy Knowlton. 2004. “Mechanisms of Reproductive Isolation among Sympatric Broadcast-Spawning Corals of the Montastraea Annularis Species Complex.” *Evolution; International Journal of Organic Evolution* 58 (2): 308–23.

Levy, O, L Appelbaum, W Leggat, Y Gothlif, D.C Hayward, D.J Miller, and O Hoegh-Guldberg. 2007. “Light-Responsive Cryptochromes from a Simple Multicellular Animal, the Coral Acropora Millepora.” *Science* 318 (5849): 467–70.

Lima, Thiago G., and Michael a. McCartney. 2013. “Adaptive Evolution of M3 Lysin-A Candidate Gamete Recognition Protein in the Mytilus Edulis Species Complex.”

Molecular Biology and Evolution 30 (12): 2688–98. doi:10.1093/molbev/mst165.

Lopez, J V, R Kersanach, S A Rehner, and N Knowlton. 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 (1): 80–93.

Manier, Mollie K, and Stephen R Palumbi. 2008. "Intraspecific Divergence in Sperm Morphology of the Green Sea Urchin, Strongylocentrotus Droebachiensis: Implications for Selection in Broadcast Spawners." *BMC Evolutionary Biology* 8 (1): 283. doi:10.1186/1471-2148-8-283.

Marks, Jessica A, Christiane H Biermann, Walter F Eanes, Harald Kryvi, Jessica A Marks, Christiane H Biermann, Walter F Eanes, and Harald Kryvi. 2008. "Divergence between Pacific and Atlantic Oceans Linked References Are Available on JSTOR for This Article : Sperm Polymorphism Within the Sea Urchin Strongylocentrotus Droebachiensis : Divergence Between Pacific and Atlantic Oceans" 215 (2): 115–25.

Marshall, Jeremy L., Michael L. Arnold, and Daniel J Howard. 2002. "Reinforcement: The Road Not Taken." *Trends in Ecology & Evolution* 17 (12): 558–63.

May, Robert M. 1994. "Biological Diversity: Differences between Land and Sea." *Phil. Trans. R. Soc. Lon. B* 343 (1303): 105–11.

Mayr, E. (1942). *Systematics and the origin of species from the viewpoint of a zoologist*. New York: Columbia university press.

Mayr, Ernst. 1954. "Geographic Speciation in Tropical Echinoids." *Evolution* 8 (1): 1–18.

Metz, E C, R Robles-Sikisaka, and V D Vacquier. 1998. “Nonsynonymous Substitution in Abalone Sperm Fertilization Genes Exceeds Substitution in Introns and Mitochondrial DNA.” *Proceedings of the National Academy of Sciences USA* 95 (18): 10676–81. doi:10.1073/pnas.95.18.10676.

Miloslavich, Patricia, Juan Manuel Díaz, Eduardo Klein, Juan José Alvarado, Cristina Díaz, Judith Gobin, Elva Escobar-Briones, et al. 2010. “Marine Biodiversity in the Caribbean: Regional Estimates and Distribution Patterns.” *PloS One* 5 (8): e11916. doi:10.1371/journal.pone.0011916.

Montaggioni, Lucien, and C.J.R. Braithwaite. 2009. “Chapter Three Structure, Zonation and Dynamic Patterns of Coral Reef Communities.” In *Developments in Marine Geology*, 5:67–122. doi:10.1016/S1572-5480(09)05003-9.

Morris-Pocock, James A, David J Anderson, and Vicki L Friesen. 2016. “Biogeographical Barriers to Dispersal and Rare Gene Flow Shape Population Genetic Structure in Red-Footed Boobies (Sula Sula),” 2125–35. doi:10.1111/jbi.12780.

Moy, G W, L M Mendoza, J R Schulz, W J Swanson, C G Glabe, and V D Vacquier. 1996. “The Sea Urchin Sperm Receptor for Egg Jelly Is a Modular Protein with Extensive Homology to the Human Polycystic Kidney Disease Protein, PKD1.” *The Journal of Cell Biology* 133 (4): 809–17.

Nosil, Patrik, and Jeffrey L Feder. 2012. “Genomic Divergence during Speciation: Causes and Consequences.” *Philosophical Transactions of the Royal Society of London. Series B, Biological Sciences* 367 (1587): 332–42. doi:10.1098/rstb.2011.0263.

O’Dea, Aaron, Harilaos A. Lessios, Anthony G. Coates, Ron I. Eytan, Sergio A. Restrepo-

Moreno, Alberto L. Cione, Laurel S. Collins, et al. 2016. “Formation of the Isthmus of Panama.” *Science Advances* 2 (8): 1–12. doi:10.1126/sciadv.1600883.

Østbye, K., T. F. Næsje, L. Bernatchez, O. T. Sandlund, and K. Hindar. 2005. “Morphological Divergence and Origin of Sympatric Populations of European Whitefish (Coregonus Lavaretus L.) in Lake Femund, Norway.” *Journal of Evolutionary Biology* 18 (3): 683–702. doi:10.1111/j.1420-9101.2004.00844.x.

Palmer, Melody R, Margo H McDowall, Lia Stewart, Aleena Ouaddi, Michael J MacCoss, and Willie J Swanson. 2013. “Mass Spectrometry and Next-Generation Sequencing Reveal an Abundant and Rapidly Evolving Abalone Sperm Protein.” *Molecular Reproduction and Development* 80 (6): 460–65. doi:10.1002/mrd.22182.

Palumbi, S.R. 1994. “Genetic Divergence, Reproductive Isolation and Marine Speciation.” *Annual Review of Ecology and Systematics* 25: 547–72.

Palumbi, S R. 1999. “All Males Are Not Created Equal: Fertility Differences Depend on Gamete Recognition Polymorphisms in Sea Urchins.” *Proceedings of the National Academy of Sciences of the United States of America* 96 (22): 12632–37. doi:10.1073/pnas.96.22.12632.

———. 2009. “Speciation and the Evolution of Gamete Recognition Genes: Pattern and Process.” *Heredity* 102 (1): 66–76. doi:10.1038/hdy.2008.104.

Pandolfi, JM, and AF Budd. 2008. “Morphology and Ecological Zonation of Caribbean Reef Corals: The Montastraea ‘Annularis’ Species Complex.” *Marine Ecology Progress Series* 369 (Goreau 1959): 89–102. doi:10.3354/meps07570.

Parkinson, John Everett, Sebastian Baumgarten, Craig T Michell, Iliana B Baums, Todd C

LaJeunesse, and Christian R Voolstra. 2016. “Gene Expression Variation Resolves Species and Individual Strains among Coral-Associated Dinoflagellates within the Genus Symbiodinium.” *Genome Biology and Evolution* 8 (3): evw019-. doi:10.1093/gbe/evw019.

Pinzón, Jorge H, Bishoy Kamel, Colleen A Burge, C Drew Harvell, Mónica Medina, Ernesto Weil, and Laura D Mydlarz. 2015. “Whole Transcriptome Analysis Reveals Changes in Expression of Immune-Related Genes during and after Bleaching in a Reef-Building Coral.” *Royal Society Open Science* 2 (4): 140214. doi:10.1098/rsos.140214.

Prada, Carlos, Bishoy Hanna, Ann F Budd, Cheryl M Woodley, Jeremy Schmutz, Jane Grimwood, Roberto Iglesias-Prieto, et al. 2016. “Empty Niches after Extinctions Increase Population Sizes of Modern Corals.” *Current Biology* 26 (23): 3190–94. doi:10.1016/j.cub.2016.09.039.

Prada, Carlos, and Michael E Hellberg. 2013. “Long Prereproductive Selection and Divergence by Depth in a Caribbean Candelabrum Coral.” *Proceedings of the National Academy of Sciences of the United States of America* 110 (10): 3961–66. doi:10.1073/pnas.1208931110.

———. 2014. “Strong Natural Selection on Juveniles Maintains a Narrow Adult Hybrid Zone in a Broadcast Spawner.” *The American Naturalist* 184 (6): 702–13. doi:10.1086/678403.

Prada, Carlos, Shelby E. McIlroy, Diana M. Beltr??n, Daniel J. Valint, Scott A. Ford, Michael E. Hellberg, and Mary Alice Coffroth. 2014. “Cryptic Diversity Hides Host and Habitat Specialization in a Gorgonian-Algal Symbiosis.” *Molecular Ecology* 23 (13): 3330–40. doi:10.1111/mec.12808.

Prada, Carlos, Nikolaos V Schizas, and Paul M Yoshioka. 2008. “Phenotypic Plasticity or Speciation? A Case from a Clonal Marine Organism.” *BMC Evolutionary Biology* 8 (1961): 47. doi:10.1186/1471-2148-8-47.

Pujolar, J. M., and G. H. Pogson. 2011. “Positive Darwinian Selection in Gamete Recognition Proteins of Strongylocentrotus Sea Urchins.” *Molecular Ecology* 20: 4968–82. doi:10.1111/j.1365-294X.2011.05336.x.

Quattrini, A.M., I.B. Baums, T.M. Shank, C. Morrison, and E.E. Cordes. 2015. “Testing the Depth-Differentiation Hypothesis in a Deepwater Octocoral.” *Proceedings of the Royal Society B: Biological Sciences* 282 (1807): 20150008. doi:10.1098/rspb.2015.0008.

Quattrini, A.M., S.E. Georgian, L. Byrnes, A. Stevens, R. Falco, and E.E. Cordes. 2013. “Niche Divergence by Deep-Sea Octocorals in the Genus Callogorgia across the Continental Slope of the Gulf of Mexico.” *Molecular Ecology* 22 (15): 4123–40. doi:10.1111/mec.12370.

Quesada, Humberto, David Posada, Armando Caballero, Paloma Morán, and Emilio Rolán-Alvarez. 2007. “Phylogenetic Evidence for Multiple Sympatric Ecological Diversification in a Marine Snail.” *Evolution* 61 (7): 1600–1612. doi:10.1111/j.1558-5646.2007.00135.x.

Raina, Jean-Baptiste, Dianne M Tapiolas, Sylvian Foret, Adrian Lutz, David Abrego, Janja Ceh, François O Seneca, et al. 2013. “DMSP Biosynthesis by an Animal and Its Role in Coral Thermal Stress Response.” *Nature* 502: 677–80. doi:10.1038/nature12677.

Reaka-Kudla, Marjorie L. 2005. “Biodiversity of Caribbean Coral Reefs.” In *Caribbean Marine Biodiversity*, 259–76.

Reitzel, Adam M., Lars Behrendt, and Ann M. Tarrant. 2010. “Light Entrained Rhythmic Gene Expression in the Sea Anemone Nematostella Vectensis: The Evolution of the Animal Circadian Clock.” *PLoS ONE* 5 (9): 1–9. doi:10.1371/journal.pone.0012805.

Roberts, Callum M. 1997. “Connectivity and Management of Caribbean Coral Reefs.” *Science* 278 (5342): 1454–57.

Rocha, L A, D R Robertson, J Roman, and B W Bowen. 2005. “Ecological Speciation in Tropical Reef Fishes.” *Proceedings of the Royal Society B* 272 (March): 573–79. doi:10.1098/2004.3005.

Rundle, Howard D., and Patrik Nosil. 2005. “Ecological Speciation.” *Ecology Letters* 8 (3): 336–52. doi:10.1111/j.1461-0248.2004.00715.x.

Sala, Enric, and Nancy Knowlton. 2006. “Global Marine Biodiversity Trends.” *Annual Review of Environment and Resources* 31 (1): 93–122. doi:10.1146/annurev.energy.31.020105.100235.

Savolainen, Vincent, Marie Charlotte Anstett, Christian Lexer, Ian Hutton, James J. Clarkson, Maria V. Norup, Martyn P. Powell, David Springate, Nicolas Salamin, and William J. Baker. 2006. “Sympatric Speciation in Palms on an Oceanic Island.” *Nature* 441 (7090): 210–13. doi:10.1038/nature04566.

Schluter, Dolph. 2009. “Evidence for Ecological Speciation and Its Alternative.” *Science (New York, N.Y.)* 323 (5915): 737–41. doi:10.1126/science.1160006.

Serrano, Xaymara M, Iliana B Baums, Tyler B Smith, Ross J Jones, Tonya L Shearer, and Andrew C Baker. 2016. “Long Distance Dispersal and Vertical Gene Flow in the Caribbean Brooding Coral Porites Astreoides.” *Scientific Reports* 6 (October 2015).

Nature Publishing Group: 21619. doi:10.1038/srep21619.

Shinzato, Chuya, Sutada Mungpakdee, Nana Arakaki, and Noriyuki Satoh. 2015. "Genome-Wide SNP Analysis Explains Coral Diversity and Recovery in the Ryukyu Archipelago." *Scientific Reports* 5. Nature Publishing Group: 1–8. doi:10.1038/srep18211.

Shoguchi, Eiichi, Makiko Tanaka, Chuya Shinzato, Takeshi Kawashima, and Nori Satoh. 2013. "A Genome-Wide Survey of Photoreceptor and Circadian Genes in the Coral, Acropora Digitifera." *Gene* 515 (2). Elsevier B.V.: 426–31. doi:10.1016/j.gene.2012.12.038.

Spalding, Mark, Philip Kramer, Edmund Green, Suzie Greenhalgh, Hillary Nobles, and Jonathan Kool. 2004. "Reefs at Risk in the Caribbean." World Resources Institute.

Swanson, Willie J, and Victor D Vacquier. 2002. "The Rapid Evolution of Reproductive Proteins." *Genetics* 3 (February): 137–44. doi:10.1038/nrg/733.

Szmant, A. M., E. Weil, M. W. Miller, and D. E. Colon. 1997. "Hybridization within the Species Complex of the Scleractinan Coral Montastraea Annularis." *Marine Biology* 129: 561–72.

Thomas, Luke, and Stephen R. Palumbi. 2017. "The Genomics of Recovery from Coral Bleaching." *Proceedings of the Royal Society B: Biological Sciences* 284 (1865): 20171790. doi:10.1098/rspb.2017.1790.

Tomaiuolo, Maurizio, Thomas F Hansen, and Don R Levitan. 2007. "A Theoretical Investigation of Sympatric Evolution of Temporal Reproductive Isolation as Illustrated by Marine Broadcast Spawners." *Evolution; International Journal of Organic Evolution* 61 (11): 2584–95. doi:10.1111/j.1558-5646.2007.00218.x.

Vacquier. 1998. “Evolution of Gamete Recognition Proteins.” *Science (New York, N.Y.)* 281 (5385): 1995–98.

Vacquier, Victor D, and Willie J Swanson. 2011. “Selection in the Rapid Evolution of Gamete Recognition Proteins in Marine Invertebrates.” *Cold Spring Harbor Perspectives in Biology* 3 (11): a002931. doi:10.1101/cshperspect.a002931.

van Doorn, SG, P Edelaar, and F J Weissing. 2009. “On the Origin of Species by Natural and Sexual Selection.” *Science* 326 (5960): 1704–7.

Van Oppen, Madeleine J H, Pim Bongaerts, Jim N. Underwood, Lesa M. Peplow, and Timothy F. Cooper. 2011. “The Role of Deep Reefs in Shallow Reef Recovery: An Assessment of Vertical Connectivity in a Brooding Coral from West and East Australia.” *Molecular Ecology* 20 (8): 1647–60. doi:10.1111/j.1365-294X.2011.05050.x.

Villinski, JT. 2003. “Depth-Independent Reproductive Characteristics for the Caribbean Reef-Building Coral Montastraea Faveolata.” *Marine Biology*, 1043–53. doi:10.1007/s00227-002-0997-0.

Vize, Peter D. 2006. “Deepwater Broadcast Spawning by Montastraea Cavernosa, Montastraea Franksi, and Diploria Strigosa at the Flower Garden Banks, Gulf of Mexico.” *Coral Reefs* 25 (1): 169–71. doi:10.1007/s00338-005-0082-5.

Weil, E, and N Knowlton. 1994. “A Multi-Character Analysis of the Caribbean Coral Montastraea Annularis (Ellis and Solander, 1786) and Its Two Sibling Species, M. Faveolata (Ellis and Solander, 1786).” *Bulletin of Marine Science* 55 (1): 151–75.

Wolstenholme, Jacqueline K. 2004. “Temporal Reproductive Isolation and Gametic Compatibility Are Evolutionary Mechanisms in the Acropora Humilis Species Group

(Cnidaria; Scleractinia).” *Marine Biology* 144: 567–82. doi:10.1007/s00227-003-1209-2.

Chapter 2

Gene Orthology assessment in *Orbicella* coral

Abstract

Here we studied the genomes of the three Caribbean coral species in the genus *Orbicella* (*O. franksi, O. faveolata*, *O. annularis*) to assess if different gene family expansions/reductions took place after speciation events within each lineage. We compared the genomes of the three sister *Orbicella* species to one another to assess protein-coding gene ortholog differences and used the Caribbean corals *Acropora palmata* and *A. cervicornis* as outgroups. Data suggests there are over forty thousand gene ortholog groups most of which overlap among these corals. Only 14, 4, and 21 ortholog groups were found to be species-specific to *O. franksi, O. faveolata* and *O. annularis*, suggesting high level genome similarity. This observation suggests that the presence of exclusive species-specific ortholog groups may not be responsible for the ecological segregation of *Orbicella*; and, instead, this could be an effect of other factors such as differential gene expression, SNP content, transcription factor activity, or others.

Introduction

The development of modern sequencing technologies and genomics tools has enriched our abilities to learn about organisms in an unprecedented way, particularly non-model organisms that are often difficult to culture, breed or manipulate. Sometimes called “Obscure Model Organisms” taxa with little genomic data available are now being increasingly used to address questions in ecology and evolution (Matz, 2017). Scleractinian corals are a good

example of taxa that have benefitted from the revolution in sequencing technologies and genomic tools that have become available in the last few years. Currently genome assemblies are available for several coral species, such as *Pocillopora damicornis* (Traylor-Knowles, et al.,2018), *Styllopora pistillata* (Voolstra et al.,2017), *Acropora digitifera* (Shinzato et al.,2011), *A. cervicornis, A. palmata* (Kitchen et al.,2019), *Montipora capitata* (Shumaker et al.,2019), *Orbicella faveolata, O. annularis* and *O. franksi* (Prada et al.,2016).

The *Orbicella* species make good study system given the rich amalgam of information available on their ecology, natural history and more recently, genomics (Knowlton, 1993; Knowlton, et al.,1997; Levitan, et al.,2011; Prada et al.,2016). The genus encompasses three sister species (*O. franksi, O. faveolata* and *O. annularis*) that, although currently listed as vulnerable or endangered (Aronson, et al.,2008) are still among the most abundant taxa in the modern Caribbean Sea. Extant *Orbicella* spp. have been abundant in the Caribbean for a couple million years, particularly after the former competitor *O. nancyi* became extinct making the shallow water niches available to *O. faveolata* and *O. annularis* (Prada et al.,2016). Morphological evidence from the fossil record from several locations in the Caribbean suggest all three modern *Orbicella* species appeared within the last 2-4 million years (and the oldest are *O. franksi* and *O. faveolata*, which surely originated 3-4 million years ago) (Budd & Klaus, 2001). Despite the young nature of the genus *Orbicella*, ecological niche preference seems to have been established along a depth gradient. Depth gradients are important in determining other factors like light penetration, temperature fluctuation, wave exposure, and others. Niche preference is therefore important in physically segregating *Orbicella* species, although in a few places they have been reported to occur in overlapping depth ranges (Levitan et al.,2004). In general, *Orbicella franksi* is the deep-water dweller at

approximately 10 to 20 m, *O. annularis* is a shallow water dweller, and finally *O. faveolata* has an intermediate range in between both. Although these species are reproductively isolated and are recognized as distinct evolutionary lineages, previously molecular evidence using ITS region and cytochrome oxidase (COI) failed to discriminate between morphotypes and concluded they were the same species (Medina, et al.,1999).

Currently, genome availability permits the exploration of an organism's biology in a comprehensive way, which can be particularly enriching when inquiring about the ecology and evolution of an obscure model organism. Here, we explore the genomes of the *Orbicella* sister species to establish if there are genomic signatures that reflect their ecology (e.g., niche preference). We used comparative genomics to determine the gene family commonalities between species as well as the and uniqueness of each species by studying orthologous families of *Orbicella* and compared against two other Caribbean corals, *Acropora palmata* and *A. cervicornis*, which are distantly related to *Orbicella* but provide an ecological comparison since they both are shallow water species like *O. faveolata* and *O. annularis*.

Methods

In this analysis we focus on the genomes of *O. annularis, O. franksi, O. faveolata* (Prada et al., 2016), and used *Acropora cervicornis* and *A. palmata* (Kitchen et al., 2019) as outgroups. gVolante version 1.2.1 was used to compare these genomes with CEGMA (with non-vertebrate, CEG database settings) and BUSCO programs (v2/v3, Metazoa database) (Nishimura, et al., 2017). Orthofinder version 2.1.2 (Emms & Kelly, 2015) was used to infer orthologous groups in the input genomes. ClusterProfiler version 3.0.5 with the enrichGO was used to generate gene ontology groupings within the orthologous groups and REVIGO was

then used with default settings to create the summaries (Supek, et al., 2011). NCBI BLAST P and UNIPROT (UniProt Consortium, 2018) were used to identify the protein descriptions by using default search settings. Finally, the online version of HMMER was used with default settings to assess the putative identity of the domains present in each of the species-specific proteins here found (Potter et al., 2018).

Results

Overall, Orthofinder found 287,203 genes within the five genomes used. Eighty one percent of them (234,484 genes) were assigned to 42,612 orthogroups, 12,125 of which are present in the five coral species, 3,720 are present in *Acropora* only, and 9,183 are present in *Orbicella* only. Few orthogroups are species-specific totaling only 62 orthogroups which encompass 230 genes suggesting high overlap in genome content (Table 2-1).

Table 2-1: Metrics of ortholog groups found in the genomes of *A. cervicornis, A. palmata, O. annularis, O. faveolata,* and *O. franksi.*

	A. cervicornis	***A. palmata***	***O. annularis***	***O. faveolata***	***O. franksi***
Number of genes	33322	30899	93491	38734	90757
Number of genes in orthogroups	29620	27752	70425	34605	72082
Number of unassigned genes	3702	3147	23066	4129	18675
Percentage of genes in orthogroups	88.9	89.8	75.3	89.3	79.4
Percentage of unassigned genes	11.1	10.2	24.7	10.7	20.6
Number of species-specific orthogroups	18	5	21	4	14
Number of genes in species-specific orthogroups	64	39	74	10	43
Percentage of genes in species-specific orthogroups	0.2	0.1	0.1	0	0

Commonalities within Orbicella

There are large overlapping numbers of orthologous groups in these species as seen in Table 2-5. These ortholog groups can be summarized in Gene Ontology (GO) categories and

when the three species of *Orbicella* are grouped they all encompass 334, 71, 25 Biological Process, Molecular Function and Cellular Component categories with $p_{adj}<0.05$. These GO can be summarized in 127, 45, and 10 non redundant groups for each category respectively. These are summarized in Appendix Tables 2-1, 2-2 and 2-3 for the Biological Process, Cellular Component and Molecular Function categories.

Table 2-2: Quantitative overlap in orthologous groups.

	A. cervicornis	***A. palmata***	***O. annularis***	***O. faveolata***	***O. franksi***
A. cervicornis	19860	-	-	-	-
A. palmata	17731	19491	-	-	-
O. annularis	15343	15043	36200	-	-
O. faveolata	14052	13826	24091	26320	-
O. franksi	15363	15061	35310	25715	37756

2

Species Specific Ortholog groups in Orbicella

Given that *Acropora* species are used as an outgroup, we choose to focus on the species-specific orthogroups only belonging to *Orbicella* spp. Finding identities to these groups can be difficult given that many produced hits in NCBI BLASTP nr database that were hypothetical, predicted or uncharacterized proteins. Some hits came from *Orbicella faveolata*, reflecting how similar the genomes are between these species and perhaps too that this is the best annotated *Orbicella* genome. However, annotating against the same database but excluding the cnidarian taxa (Taxa ID 6073) confirmed the identity of numerous hits either when a functional annotation was available or when it was just a hypothetical annotation. Note that these are annotations based on sequence homology and few of them have backing empirical evidence.

O. franksi has 14 species-specific orthogroups that encompass 43 genes (Table 2-3). Three orthogroups are related to E3 ubiquitin-protein ligase proteins, which are associated with the modification of other proteins. The other orthogroups with most confident annotation are Dehydrogenase/reductase SDR family (oxygen reduction), Dynein heavy chain 8 axoneal-like (ATP binding and associated with sperm motility), the PREDICTED: zinc finger protein 862-like (transcription factor), DUF 885 protein (DUF stands for Domain of Unknown Function, hence the activity or role of this protein the cell is unknown)(Table 2-6). On the other hand, out of the 14 orthologous groups found in *O. franksi* 10 did not match any known domain architectures. One orthogroup, Ofra.g20388.t1, containing three proteins and uncharacterized annotation, shows inconsistent domain presence since only one of the genes matches a known domain. This is the case of Scavenger receptor cysteine-rich domain, which is in the cell membrane and presumably has immune function. Three other cases are worth discussing. First, there is one group in which the annotation obtained by NCBI (protein identity) and HMMER (domain identity) match suggesting an adequate annotation. This is the orthologue group Ofra.g10124.t1 Dynein heavy chain 8, axonemal-like domain, which are involved in microtubule movement in the cell (cytokinesis). Finally, there were two orthogroups in this species that had uncharacterized proteins whose annotation changed when excluding Cnidarians from the database. These are the Ofra.g10124.t1 Methyltransferase domain-containing protein which are proteins associated with methylation of other proteins in the cell; and, DUF885 domain-containing protein, which is a "bacterial" domain that as it name states is of Unknown Function.

O. faveolata has only four orthologs groups that are species specific (Table 2-4). Of these, DBH-like monooxygenase proteins is the one with the most reliable annotation. These

proteins are involved in metal binding and oxygen reduction of other proteins. The annotation of the domains present in this protein indicates the proteins in this orthogroup (namely, Ofav.g27384.t1) contain at least one DOMON (dopamine beta-monooxygenase N-terminal) and Copper type II ascorbate-dependent monooxygenase, N-terminal and Copper type II ascorbate-dependent monooxygenase, C-terminal domains both of which are involved in enzymatic activity either in redox reactions (DOMON) or in copper-based reactions. All other proteins found as species-specific in *O. faveolata* were uncharacterized.

O. annularis is the species with the largest number of orthologous groups, 21 that encompass 74 genes (Table 2-5). Two of the orthogroups are associated with Tetratricopeptide repeat proteins, which in humans participate in cell division. Other proteins of interest are: Cadherins (calcium binding proteins), Sushi domain proteins (polysaccharide binding, probably involved in immune response and cell division), Pancreatic lipase-related protein 2 (calcium binding, broad spectrum lipase) (Table 2-11), and tilB proteins, which participate in cell motility can also be involved in the regulation of circadian clocks by temperature in *Drosophila* (temperature compensation of the circadian clock, GO:0010378). Eleven of the 21 species-specific orthologue groups in *O. annularis* lack hits to any known protein domains and of those eight have hypothetical unannotated protein with no hits. A few proteins have a putative annotation but the protein domains found are unknown. For example, Orthologue group Oann.g36535.t1 has a putative annotation by homology to Mast/stem cell growth factor receptor Kit from *Bactrocera dorsalis* (with accession number XP_029406349.1) but no domain is found and this annotation is only possible when excluding Cnidarians from the search database. Another scenario is the case of orthogroup Oann.g35351.t1, which according to the NCBI database belongs to the Sushi domain-containing protein 2-like found in

Orbicella faveolata and *Strongylocentrotus purpuratus* (with accession numbers XP_020620341.1 and XP_011674351.1, respectively). However, domain was found with HMMER which does an exhaustive protein domain search.

On the other hand, the ones that had identifiable domains, four are assuring of the annottions suggested by NCBI for example Cadherin (Oann.g33461.t1), the Tetracopeptide (Oann.g40961.t, Oann.g32219.t1), and the Lipase (Oann.g40605.t1) orthogroups have Cadherin, Tetracopeptide, and Lipase domains in them, respectively. The remaining orthologue groups have a combination of annotations with and without domains.

Table 2-3: Species-specific ortholog groups in *O. franksi.* Annotations were obtained by aligning to NCBI, BLAST P program, non-redundant database. The bold sequences are the description sequence of the orthologous group and it is also the one that corresponds to the annotation boxes.

(Next page)

Species-specific Orthologues	Gene ID	Gene count	NCBI first match in BlastP (Name, Best hit organism, accession number)	Same thing but excluding Cnidarians (taxaid 6073)	Protein domain
OG0020029	**Ofra.g44854.t1** Ofra.g63497.t1 Ofra.g6605.t1 Ofra.g79735.t1 Ofra.g90236.t1	5	Uncharacterized protein LOC110065426 [Orbicella faveolata] XP_020628220.1	Uncharacterized protein LOC107453613 [Parasteatoda tepidariorum] XP_015925979.1	No domain hits found in any of these gene IDs.
OG0024446	**Ofra.g22288.t1** Ofra.g22289.t1 Ofra.g36983.t1 Ofra.g36984.t1	4	Dehydrogenase/reductase SDR family member 7-like [Orbicella faveolata] XP_020614368.1	Dehydrogenase/reductase SDR family member 7-like [Crassostrea virginica] XP_022329120.1	These match the Short-chain dehydrogenase family but no domain found.
OG0024447	**Ofra.g50368.t1** Ofra.g50369.t1 Ofra.g50370.t1 Ofra.g50371.t1	4	Uncharacterized protein LOC110063124 [Orbicella faveolata] XP_020625731.1	Sensor domain-containing diguanylate cyclase [Pseudomonas alkylphenolica] WP_128324300.1	No domain hits found in any of these gene IDs.
OG0031985	**Ofra.g2340.t1** Ofra.g33883.t1 Ofra.g33884.t1	3	Trichohyalin-like [Orbicella faveolata] XP_020621800.1	NA	No domain hits found in any of these gene IDs.
OG0031986	**Ofra.g20388.t1** Ofra.g4998.t1 Ofra.g9413.t1	3	Uncharacterized protein LOC110044851 [Orbicella faveolata] XP_020606089.1	Helix-turn-helix domain-containing protein [Marinilactibacillus psychrotolerans] WP_091759638.1	No domain hits found for first or third but the second one has a hit to Scavenger receptor cysteine-rich , which is in the cell membran and has immune function.
OG0031987	**Ofra.g5009.t1** Ofra.g66974.t1 Ofra.g76593.t1	3	Uncharacterized protein LOC114964364 [Acropora millepora] XP_029199535.1	Hypothetical protein DSY43_06820 [Gammaproteobacteria bacterium] RUA04226.1	No domain hits found in any of these gene IDs.
OG0031988	**Ofra.g10124.t1** Ofra.g52665.t1 Ofra.g52667.t1	3	Uncharacterized protein LOC110065822 [Orbicella faveolata] XP_020628647.1	Methyltransferase domain-containing protein [Candidatus Kentron sp. DK] VFJ51043.1	Methyltransferase domain is found in all these gene Ids.
OG0031989	**Ofra.g18156.t1** Ofra.g21749.t1 Ofra.g27310.t1	3	E3 ubiquitin-protein ligase RNF213-like isoform X2 [Orbicella faveolata] XP_020616491.1	PREDICTED: E3 ubiquitin-protein ligase RNF213-like [Saccoglossus kowalevskii] XP_006812911.1	No domain hits found in any of these gene IDs.
OG0031990	**Ofra.g22274.t1** Ofra.g30850.t1 Ofra.g30851.t1	3	Uncharacterized protein LOC110052631 [Orbicella faveolata] XP_020614434.1	PREDICTED: angiopoietin-4-like isoform X1 [Haplochromis burtoni] XP_014196390.1	No domain hits found in any of these gene IDs.
OG0031991	**Ofra.g34979.t1** Ofra.g36787.t1 Ofra.g36788.t1	3	E3 ubiquitin-protein ligase UBR3-like [Stylophora pistillata] XP_022782680.1	PREDICTED: E3 ubiquitin-protein ligase UBR3 [Nanorana parkeri] XP_018410064.1	No domain hits found in any of these gene IDs.
OG0031992	**Ofra.g52391.t1** Ofra.g52392.t1 Ofra.g52393.t1	3	Dynein heavy chain 8, axonemal-like [Acropora millepora] XP_029206097.1	PREDICTED: dynein heavy chain 8, axonemal-like [Priapulus caudatus] XP_014666592.1	These are in the Dynein heavy chain, N-terminal region 2 family, which are involved in microtubule movement in the cell (cytokinesis).
OG0042609	**Ofra.g5218.t1** Ofra.g69705.t1	2	E3 ubiquitin-protein ligase rnf213-alpha-like [Pocillopora damicornis] XP_027055358.1	PREDICTED: E3 ubiquitin-protein ligase RNF213-like [Saccoglossus kowalevskii] XP_006822899.1	No domain hits found in any of these gene IDs.
OG0042610	**Ofra.g31836.t1** Ofra.g55317.t1	2	PREDICTED: zinc finger protein 862-like, partial [Acropora digitifera] XP_015769849.1	PREDICTED: zinc finger protein 862-like [Saccoglossus kowalevskii] XP_006822333.1	No domain hits found in any of these gene IDs.
OG0042611	**Ofra.g35128.t1** Ofra.g35129.t1	2	Uncharacterized protein LOC110068834 [Orbicella faveolata] XP_020631909.1	DUF885 domain-containing protein [Sphingomonas sp. BK235] WP_132911143.1	Bacterial protein of unknown function (DUF885)

Table 2-4: Species-specific ortholog groups in *O. faveolata.* Annotations were obtained by aligning to NCBI, BLAST P program, non-redundant database. The bold sequences are the description sequence of the orthologous group and it is also the one that corresponds to the annotation boxes.

(Next page)

Species-specific Orthologues	Gene ID	Gene count	NCBI first match in BlastP (Name, Best hit organism, accession number)	Same thing but excluding Cnidarians (taxaid 6073)	Protein domain
OG0024410	**Ofav.g27384.t1** Ofav.g27386.t1 Ofav.g27389.t1 Ofav.g32227.t1	4	DBH-like monooxygenase protein 1 [Orbicella faveolata] XP_020615038.1	DBH-like monooxygenase protein 1 [Trichoplax sp. H2] RDD43174.1	These genes have combiantions of one or two of the DOMON (dopamine beta-monooxygenase N-terminal), Copper type II ascorbate-dependent monooxygenase, N-terminal and Copper type II ascorbate-dependent monooxygenase, C-terminal domains. The DOMON is involved in "enzymatic activity involved in redox reactions". The copper one is also ensymatic and uses copper as a cofactor in electron transfers too. These are related. DOMON is an enzyme contained in the Copper family above.
OG0041404	**Ofav.g2083.t1** Ofav.g2084.t1	2	Uncharacterized protein LOC110063006 [Orbicella faveolata] XP_022806602.1	PREDICTED: periphilin-1-like isoform X1 [Pygocentrus nattereri] XP_017539414.1	No domain hits found in any of these gene IDs.
OG0041526	**Ofav.g23305.t1** Ofav.g6482.t1	2	Uncharacterized protein LOC110065773 [Orbicella faveolata] XP_020628601.1	PREDICTED: sec1 family domain-containing protein 2-like [Priapulus caudatus] XP_014676542.1	No domain hits found in any of these gene IDs.
OG0041606	**Ofav.g8313.t1** Ofav.g8314.t1	2	Uncharacterized protein LOC110067462 [Orbicella faveolata] XP_020630448.1	Nephrocystin-3 [Trichoplax sp. H2] RDD40144.1	No domain hits found in any of these gene IDs.

Table 2-5: Species-specific ortholog groups in *O. annularis.* Annotations were obtained by aligning to NCBI, BLAST P program, non-redundant database. The bold sequences are the description sequence of the orthologous group and it is also the one that corresponds to the annotation boxes.

(Next page)

Species-specific Orthologues	Gene ID	Gene count	NCBI first match in BlastP (Name, Best hit organism, accession number)	Same thing but excluding Cnidarians (taxaid 6073)	Protein domain
OG0004428	**Oann.g26636.t1** Oann.g2788.t1 Oann.g61942.t1 Oann.g63172.t1 Oann.g64842.t1 Oann.g66375.t1 Oann.g70393.t1 Oann.g86971.t1 Oann.g9402.t1 Oann.g9506.t1	10	Uncharacterized protein LOC110066631 isoform X3 [Orbicella faveolata] XP_020629515.1	Hypothetical protein AK88_02034 [Plasmodium fragile] XP_012335091.1	No domain hits found in any of these gene IDs.
OG0018576	**Oann.g16543.t1** Oann.g31735.t1 Oann.g62605.t1 Oann.g86014.t1 Oann.g87632.t1	5	Uncharacterized protein LOC110043244 [Orbicella faveolata] XP_020604346.1	Unnamed protein product [Vitrella brassicaformis CCMP3155] CEM13723.1	No domain hits found in any of these gene IDs.
OG0021530	**Oann.g11814.t1** Oann.g11815.t1 Oann.g1960.t1 Oann.g7608.t1	4	Centrosomal protein of 290 kDa-like [Orbicella faveolata] XP_020614739.1	General vesicular transport factor p115 isoform X2 [Mastacembelus armatus] XP_026159118.1	No domain hits found in any of these gene IDs.
OG0021682	**Oann.g15867.t1** Oann.g5093.t1 Oann.g56177.t1 Oann.g91361.t1	4	Uncharacterized protein LOC110045099, partial [Orbicella faveolata] XP_020606364.1	NA	No domain hits found in any of these gene IDs.
OG0022786	**Oann.g33461.t1** Oann.g33462.t1 Oann.g37018.t1 Oann.g39654.t1	4	Cadherin-23-like [Orbicella faveolata] XP_020625379.1	Cadherin-23-like [Branchiostoma belcheri] XP_019635319.1	Five Cadherin domains are present in these. Cadherin are transmembrane proteins, they require Calcium to operate and they "bind cells together"
OG0022997	**Oann.g40961.t1** Oann.g40964.t1 Oann.g40966.t1 Oann.g90446.t1	4	Tetratricopeptide repeat protein 28-like [Orbicella faveolata] XP_020607438.1	Tetratricopeptide repeat protein [Microcystis aeruginosa] WP_084990044.1	Six Tetratricopeptide repeat domains and one CHAT (Caspase HetF Associated with Tprs), which is a peptidase domain. No info on the Tetra… in EMBL
OG0023133	**Oann.g45776.t1** Oann.g46010.t1 Oann.g46012.t1 Oann.g59853.t1	4	Uncharacterized protein LOC110046774 [Orbicella faveolata] XP_020608146.1	PREDICTED: uncharacterized protein LOC106811052 [Priapulus caudatus] XP_014670060.1	No domain hits found in any of these gene IDs.
OG0025733	**Oann.g12463.t1** Oann.g17443.t1 Oann.g2563.t1	3	Protein tilB homolog [Orbicella faveolata] XP_020604301.1	PREDICTED: protein tilB homolog [Xenopus laevis] XP_018123755.1	No domain hits found in any of these gene IDs.
OG0025931	**Oann.g29647.t1** Oann.g29648.t1 Oann.g5421.t1	3	Hypothetical protein pdam_00016508 [Pocillopora damicornis] RMX45141.1	NA	No domain hits found in any of these gene IDs.
OG0026214	**Oann.g53453.t1** Oann.g9870.t1 Oann.g9872.t1	3	Hypothetical protein AWC38_SpisGene5436 [Stylophora pistillata] PFX29821.1	DNA, W-Samurai RAPD marker in retrotranposable element (reverse transcriptase), strain p50 [Operophtera brumata] KOB79292.1	No domain hits found in any of these gene IDs.
OG0026519	**Oann.g14116.t1** Oann.g78288.t1 Oann.g79295.t1	3	Uncharacterized protein LOC107344505 [Acropora digitifera] XP_015765663.1	Hypothetical protein [Bathymodiolus brooksi thiotrophic gill symbiont] WP_139476187.1	No domain hits found in any of these gene IDs.
OG0027272	**Oann.g25604.t1** Oann.g39492.t1 Oann.g56518.t1	3	Uncharacterized protein LOC107338671 [Acropora digitifera] XP_015759392.1	NA	No domain hits found in any of these gene IDs.
OG0027614	**Oann.g32219.t1** Oann.g32546.t1 Oann.g50083.t1	3	Tetratricopeptide repeat protein 28-like, partial [Orbicella faveolata] XP_020627500.1	Tetratricopeptide repeat protein [Microcystis aeruginosa] WP_084990044.1	6 Tetratricopeptide repeat domains 4 Tetratricopeptide repeat domains 4 Tetratricopeptide repeat domains
OG0027662	**Oann.g33304.t1** Oann.g48674.t1 Oann.g74234.t1	3	Peroxidasin homolog [Orbicella faveolata] XP_020608164.1	Roundabout homolog 1-like [Parasteatoda tepidariorum] XP_021003522.1	3 Immunoglobulin I-set domain (cell adhesion) and one Immunoglobulin domain (hundreds of putative functions) in the first two proteins and the last one only has one member of each domain.
OG0027756	**Oann.g35351.t1** Oann.g35352.t1 Oann.g87440.t1	3	Sushi domain-containing protein 2-like [Orbicella faveolata] XP_020620341.1	Sushi, nidogen and EGF-like domain-containing protein 1 [Strongylocentrotus purpuratus] XP_011674351.1	No domain hits found in any of these gene IDs.
OG0027829	**Oann.g37115.t1** Oann.g42008.t1 Oann.g46889.t1	3	RNA-directed DNA polymerase from mobile element jockey-like [Acropora digitifera] XP_015769126.1	Hypothetical protein [Flavobacteriaceae bacterium] MAG86089.1	Phosphatidylinositol-specific phospholipase C, X domain (signal transdution function, lipid signaling (lots of info on this one))
OG0027992	**Oann.g40605.t1** Oann.g56639.t1 Oann.g56640.t1	3	Pancreatic lipase-related protein 2-like [Orbicella faveolata] XP_020607544.1	Pancreatic lipase-related protein 2 [Cryptotermes secundus] PNF41742.1	Lipase domain
OG0028427	**Oann.g50271.t1** Oann.g50272.t1 Oann.g50273.t1	3	Uncharacterized protein LOC110051384 [Orbicella faveolata] XP_020613082.1	Odorant/gustatory chemosensory receptor-like 122 [Saccoglossus kowalevskii] ALR88638.1	7tm Chemosensory receptor domain (these are G-protein-coupled receptors (GPCRs) and have gustatory and olfactory sensor activity in insects).
OG0037472	**Oann.g36535.t1** Oann.g57776.t1	2	Uncharacterized protein LOC110062511 [Orbicella faveolata] XP_020625096.1	Mast/stem cell growth factor receptor Kit [Bactrocera dorsalis] XP_029406349.1	No domain hits found in any of these gene IDs.
OG0037768	**Oann.g45090.t1** Oann.g45091.t1	2	Basigin-like [Orbicella faveolata] XP_020610690.1	NA	Immunoglobulin domain
OG0038164	**Oann.g55530.t1** Oann.g58877.t1	2	Uncharacterized protein LOC110045097 [Orbicella faveolata] XP_020606361.1	Agrin-like [Parasteatoda tepidariorum] XP_021000730.1	Kazal-type serine protease inhibitor domain (serine protease inhibitors)

Discussion

The analyses for the dataset that included five Caribbean corals recovered over 42,000 ortholog groups. Note that the genomes of *O. annularis* and *O. franksi* are approximately three times larger than the other three species. This could be due to contig fragmentation because contigs tend to be very short (see Appendix A, Table A-2) and may be expanding the ortholog list. In the future, it would be interesting to see how these numbers vary if the genomes of *O. annularis* and *O. franksi* had better assemblies.

Unfortunately, protein domain characterization of the orthogroups unique to each species yielded a minority of hits. In a few cases, these hits corroborate the annotations suggesting an adequate homology-based characterization. Identifying the molecular properties of the proteins here reported will require more work to verify protein identity and function, which hopefully can address the molecular basis for the ecological and evolutionary history of *Orbicella*. Future work could incorporate a detailed protein structure characterization in addition to identifying if there is transcriptomic evidence of the presence of the species-specific orthologues at different coral life stages and conditions (for example, in gametes, larva and adults as well as in healthy and bleached corals). Having a functional description of these species-specific orthologue groups will aid in unraveling the complex ecology of *Orbicella* sister species.

There is substantial overlap in the ortholog groups as shown in Table 2-2. This suggests these corals have similar genome content and not surprisingly, species within the same genus share more orthologs than when compared to species of another genus. This could be reflecting either the shared evolutionary history of species within the genus (their ancestor gained some ortholog groups overtime as it diverged from other taxa) or it could be reflecting

the evolutionary history of species within the Order Scleractinia (and so A*cropora* could have lost some ancestral orthologs that remain in *Orbicella*). In order to assess which case is true future research needs to be done in order to evaluate the expansions/reductions in gene families among these corals.

Large overlap within orthogroups is not surprising between clades. For example, it was reported that *Pocillopora damicornis* shares 46% of its ortholog groups with ten other cnidarians, including *O. faveolata* (Traylor-Knowles et al., 2018), and *Montipora capitata* shared almost 90% with at least another of eleven analyzed cnidarian genomes (Shumaker et al., 2019). Another study comparing eight cnidarian species that included corals in the Robust and Complex clade found that not only there is quantitative similarity within ortholog groups but also according to synteny analyses, there is a high level of conservation in gene order (Ying et al., 2018). None of these studies, however, had been looking through the lens of ecology and reproductive isolation in closely related species.

Orbicella is a relatively young genus and it is not surprising that the three species share a large number of ortholog groups. These groups, when clustered in Gene Ontology (GO) terms are particularly enriched for the Biological Process category, which highlights homeostasis processes like G protein couple receptor signaling pathways, behavior processes, molecule transport, and circadian regulation (See Appendix B).

The number of species-specific orthologs among the *Orbicella* spp. is low (14, 4, and 21 out of ≈42000 for *O. franksi, O. faveolata* and *O. annularis*, respectively). This is not surprising since *Orbicella*'s presence in the fossil record is relatively short (less than 5 million years, see (Prada et al.,2016)) suggesting short divergence time. Given the small amount of differences in the ortholog groups here found, it is likely that the ecological and reproductive

differences of these coral species are more susceptible to differential gene expression rather than changes in gene content. It is worth noting that other features in addition to genome content can also be partially shaping the natural history of these corals, hence, the presence of SNPs, microRNA regulators, and differences within the proteins, and others can all be operating in the evolutionary process leading to the distinction of these lineages.

Conclusion

In this study we evaluated the genomes of the sister *Orbicella* species and used the Caribbean *Acropora* as an outgroup with the aim of investigating whether there have been gene family expansions/reductions within each extanct *Orbicella* lineage. High levels of ortholog groups overlapping between *Acropora* and *Orbicella* indicate a presence of "core" scleractinian genes. Furthermore, there is a high similarity in ortholog groups across *Orbicella* accompanied by a slim number of species-specific ortholog groups. It is likely that species-specific genomic signatures discriminators have not had enough time to evolve and rather species differentiation is due to differential gene expression within each of these similar genomes is responsible for the ecology of these species. Further work could explore furthermore how these apparently similar genomes separate these three species in other ways (e.g., assessing differences in transcription factor, SNPs, miRNA-based genome regulation, and others).

References

Aronson, R., Bruckner, A., Moore, J., Precht, B. & E. Weil 2008. *Montastraea annularis*. *The IUCN Red List of Threatened Species* 2008: e.T133134A3592972. http://dx.doi.org/10.2305/IUCN.UK.2008.RLTS.T133134A3592972.en

Budd, A. F., & Klaus, J. S. (2001). The origin and early evolution of the Montastraea annularis species complex (Anthozoa: Scleractinia). *Journal of Paleontology*, *75*(3), 527–545. https://doi.org/10.1666/0022-3360(2001)075<0527:TOAEEO>2.0.CO;2

Emms, D. M., & Kelly, S. (2015). OrthoFinder: solving fundamental biases in whole genome comparisons dramatically improves orthogroup inference accuracy. *Genome Biology*, *16*(1), 1–14. https://doi.org/10.1186/s13059-015-0721-2

Kitchen, S. A., Ratan, A., Bedoya-Reina, O. C., Burhans, R., Fogarty, N. D., Miller, W., & Baums, I. B. (2019). Genomic Variants Among Threatened Acropora Corals. *Genes|Genomes|Genetics*, *9*(5), 1633–1646. https://doi.org/10.1534/g3.119.400125

Knowlton, N. (1993). SIBLING SPECIES IN THE SEA. *Annual Review of Ecology and Systematics*, *24*(1993), 189–216. https://doi.org/10.1146/annurev.es.24.110193.001201

Knowlton, N., Mate, J., Guzman, H., Rowan, R., & Jara, J. (1997). Direct evidence for reproductive isolation among the three species of the Montastraea annularis complex in Central America (Panama and Honduras). *Marine Biology*, *127*, 705–711. Retrieved from http://link.springer.com/article/10.1007/s002270050061

Levitan, D. R., Fogarty, N. D., Jara, J., Lotterhos, K. E. K., & Knowlton, N. (2011). Genetic, spatial, and temporal components of precise spawning synchrony in reef building corals

of the Montastraea annularis species complex. *Evolution*, *65*(5), 1254–1270. https://doi.org/10.1111/j.1558-5646.2011.01235.x

Levitan, D. R., Fukami, H., Jara, J., Kline, D., McGovern, T. M., McGhee, K. E., … Knowlton, N. (2004). Mechanisms of reproductive isolation among sympatric broadcast-spawning corals of the Montastraea annularis species complex. *Evolution; International Journal of Organic Evolution*, *58*(2), 308–323. https://doi.org/10.1111/j.0014-3820.2004.tb01647.x

Matz, M. V. (2017). Fantastic Beasts and How To Sequence Them: Ecological Genomics for Obscure Model Organisms. *Trends in Genetics*, *xx*(2), 1–12. https://doi.org/10.1016/j.tig.2017.11.002

Medina, M., Weil, E., & Szmant, A. (1999). Examination of the Montastraea annularis Species Complex (Cnidaria: Scleractinia) Using ITS and COI Sequences. *Marine Biotechnology (New York, N.Y.)*, *1*(1), 89–97. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/10373615

Nishimura, O., Hara, Y., & Kuraku, S. (2017). GVolante for standardizing completeness assessment of genome and transcriptome assemblies. *Bioinformatics*, *33*(22), 3635–3637. https://doi.org/10.1093/bioinformatics/btx445

Potter, S. C., Luciani, A., Eddy, S. R., Park, Y., Lopez, R., & Finn, R. D. (2018). HMMER web server: 2018 update. *Nucleic Acids Research*, *46*(W1), W200–W204. https://doi.org/10.1093/nar/gky448

Prada, C., Hanna, B., Budd, A. F., Woodley, C. M., Schmutz, J., Grimwood, J., … Medina, M. (2016). Empty Niches after Extinctions Increase Population Sizes of Modern Corals. *Current Biology*, *26*(23), 3190–3194. https://doi.org/10.1016/j.cub.2016.09.039

Shinzato, C., Shoguchi, E., Kawashima, T., Hamada, M., Hisata, K., Tanaka, M., … Satoh, N. (2011). Using the Acropora digitifera genome to understand coral responses to environmental change. *Nature*, *476*(7360), 320–323. https://doi.org/10.1038/nature10249

Shumaker, A., Putnam, H. M., Qiu, H., Price, D. C., Zelzion, E., Harel, A., … Bhattacharya, D. (2019). Genome analysis of the rice coral Montipora capitata. *Scientific Reports*, *9*(1), 2571. https://doi.org/10.1038/s41598-019-39274-3

Supek, F., Bošnjak, M., Škunca, N., & Šmuc, T. (2011). Revigo summarizes and visualizes long lists of gene ontology terms. *PLoS ONE*, *6*(7). https://doi.org/10.1371/journal.pone.0021800

Traylor-Knowles, N., Cunning, R., Baker, A. C., Gillette, P., & Bay, R. A. (2018). Comparative analysis of the Pocillopora damicornis genome highlights role of immune system in coral evolution. *Scientific Reports*, *8*(1), 1–10. https://doi.org/10.1038/s41598-018-34459-8

UniProt Consortium, T. (2018). Erratum: UniProt: the universal protein knowledgebase (Nucleic acids research (2017) 45 D1 (D158-D169)). *Nucleic Acids Research*, *46*(5), 2699. https://doi.org/10.1093/nar/gky092

Voolstra, C. R., Li, Y., Liew, Y. J., Baumgarten, S., Zoccola, D., Flot, J. F., … Aranda, M. (2017). Comparative analysis of the genomes of Stylophora pistillata and Acropora digitifera provides evidence for extensive differences between species of corals. *Scientific Reports*, *7*(1), 1–14. https://doi.org/10.1038/s41598-017-17484-x

Ying, H., Cooke, I., Sprungala, S., Wang, W., Hayward, D. C., Tang, Y., … Miller, D. J. (2018). Comparative genomics reveals the distinct evolutionary trajectories of the robust and complex coral lineages. *Genome Biology*, *19*(1), 175.

https://doi.org/10.1186/s13059-018-1552-8

Chapter 3

Transcriptional insights of temporal isolation in broadcast spawning corals

Abstract

Broadcasting spawning is the most common gamete release strategy in corals. Gamete release occurs during massive spawning events in which localized conspecifics simultaneously release their gametes into the water column. *Orbicella annularis* and *Orbicella franksi* are sister taxa that typically inhabit different reef depths (<10m, and >20m, respectively); and, spawn at precise species-specific time frames in the same evening once per year. This temporal segregation appears to have contributed to reproductive isolation between them. An analysis of transcriptome of these two species was done to test whether different spawning times is associated with differentially expressed genes. We compared by RNAseq the transcriptional profiles of both species at the time of spawning and after spawning. The data suggest these species have different gene expression profiles, which overlap minimally. *O. franksi* expresses more similar transcriptional profiles to *Acropora millepora,* a phylogenetically distant species, than to the congeneric *O. annularis*. These corals use their genetic tool kit differently resulting in the same behavior (gamete release).

Introduction

Reproductive isolation (RI) among species requires the presence of pre-zygotic and/or post-zygotic mechanisms to operate and is fundamental in the speciation processes. Pre- and post-zygotic barriers are often well documented in terrestrial systems in which physical barriers provide the means for RI (Coyne and Orr 2004). However, in marine systems physical barriers are scarce and other isolating mechanisms operate (Palumbi, 1994; Prada and Hellberg 2013). In fact, physical barriers rarely segregate populations in the ocean (except across the Isthmus of Panama and along the Sunda Shelf) and most marine species generate planktonic larvae that are capable of dispersing hundreds to thousands of kilometers connecting populations across large ocean basins such as the entire Pacific Ocean (Lessios & Robertson, 2006). The lack of physical barriers and the extreme dispersal capabilities of marine larvae suggest that other mechanisms of isolation other than vicariance may be operating to establish RI (Palumbi, 1994). Ecological factors such as differential niche occupancy may operate in conjunction with physical barriers to generate RI (Mayr, 1947). Hence, ecologically driven speciation can help explain biodiversity in hotspots like coral reefs ecosystems.

Coral reefs are the most speciose ecosystems in the ocean, yet we have little knowledge of how biodiversity is generated there. Ecological factors like depth, light availability, topography, and competition can contribute to segregation of species (reviewed in González et al, 2018). Population segregation (the differentiation of populations is the precursor to RI) is apparent in Caribbean corals where genetic differences have been detected between ecomorphs or habitat-segregated populations with overlapping ranges (Brazeau &

Harvell, 1994; Levitan et al., 2004; Prada & Hellberg, 2013). A classic example of depth segregation in the Caribbean is the *Orbicella* species complex, which includes three species *O. franksi*, *O. faveolata* and *O. annularis* (Knowlton, et al., 1992). *O. franksi* are found in deep waters (>20m), *O. faveolata* is found at intermediate depths and *O. annularis*, is located in shallow waters (<10m), although the distribution of these species can overlap, particularly at intermediate depths (Weil & Knowlton, 1994).

Coral can reproduce either sexually or asexually (Jackson & Hughes, 1985; Jackson & Coates, 1986). Asexual reproduction may occur by budding, fragmentation or coral polyp expulsion; and, ultimately favors the persistence of long-lived genets (genetically distinct individuals) by the multiplication of ramet colonies or clones (Foster, et al., 2007; Jackson & Hughes, 1985; Jackson & Coates, 1986; Kramarsky-Winter, et al., 1996). Sexual reproduction is important for generating genetic diversity. Coral species display a variety of sexual reproduction strategies. Species may be gonochoric or hermaphroditic, and brooders or broadcast spawners (Szmant, 1986). The most common strategy is hermaphroditic broadcast spawning in which each polyp releases eggs and sperm into the water column and fertilization is external (Szmant, 1986). Gamete release, known as spawning, typically occurs within a species-specific time range presumably optimizing the chances of successful fertilization (Levitan, 2010; Levitan et al., 2004). Spawning at species-specific time frames remains punctual even when populations are in different environments. For instance, corals of three Caribbean species (*Montastraea cavernosa*, *Diploria strigosa* and *Orbicella franksi*) found between 33 and 45 m in depth spawn within the same time frame as their shallow water conspecifics (Vize, 2006).

Several environmental cues have been correlated with this sophisticated synchronicity (Baird, et al., 2009) including solar irradiance, seawater temperature (Richmond & Hunter, 1990; van Veghel, 1994), sunset time (Knowlton, et al., 1997; Levitan, et al., 2011), and lunar phase (Levy et al., 2007; van Veghel, 1994). Synchronous mass spawning events in which multiple species spawn simultaneously are common, and suggest coral species respond similarly to physical factors (Babcock et al., 1986). Synchronous mass spawning, however, may also facilitate hybridization among closely related species (Harrison, et al., 1984) and evidence for hybridization has been reported in both Caribbean and Indo-Pacific corals in the genus *Acropora* (Fogarty, 2012; Fogarty, et al., 2012; Isomura, et al., 2013; Kenyon, 1997; van Oppen, et al., 2001; van Oppen, et al., 2000). Thus, except for acroporid corals there is little evidence of hybridization in the field as a consequence of mass spawning events.

Ecological evidence suggests that there are reproductive barriers operating among closely related corals despite the fact that many of them engage in mass spawning events. Perhaps the best documented case is the *O. annularis* complex. Formerly, three discrete morphotypes used to be recognized as either columnar, massive, or bumpy *Orbicella* (=*Montrastraea*) *annularis* (van Veghel, et al., 1993). Morphotype differences were initially ascribed to phenotypic plasticity to environmental conditions (Foster, 1979) and several lines of evidence suggested morphotypes were all the same species. Competitive behavior studies reported morphotypes did not appear aggressive to each other but were to other species (van Veghel, et al., 1996) and no differences in gametogenesis, spawning behavior or timing were reported in *Orbicella* corals from Curaçao (van Veghel, 1994). Columnar and massive morphs, (later renamed *Montastraea annularis* and *M. faveolata*) were reported to spawn in synchrony, potentially facilitating hybridization in Rosario Islands (Sanchez et al.,1999; van

Veghel, 1994). Successful hybrid larvae production in laboratory trials suggested a lack of pre-zygotic barriers to cross-fertilization although this could be attributed to colony misidentification at the time (Szmant, et al., 1997).

Studies focused on these species' reproductive strategies have shown evidence of intrinsic reproductive isolation. Differences in fecundity, number of gonads per polyp, and fertility suggested varying investment towards sexual reproduction (van Veghel & Kahmann, 1994). It has now been established that gamete bundles are released into the water column within narrow species-specific time windows (allochronic assortative mating)(Levitan et al., 2004). All members of the *Orbicella* genus are broadcast spawning corals and release their gametes during their annual mass-spawning event, which typically occurs between August and October and four to seven evenings after the full moon (Levitan et al., 2011, 2004). Spawning times vary across the Caribbean depending on latitude. For example, a 7.5° difference between Honduras and Panama causes a slight spawn delay (due to the ≈1 hour in sunset delay) in Panama relative to Honduras (Knowlton et al., 1997). *Orbicella* species spawn gametes at different times. Data from Panama, Bahamas and Curaçao reported that on average, *O. franski*, *O. annularis* and *O. faveolata* spawn 2 hours, 3:43 hours, 3:56 hours after sunset (Levitan et al., 2004; Levitan, et al., 2011). Although *O. annularis* and O. *faveolata* spawn at very close time frames, experimental interspecific crosses between them are unsuccessful. Interspecific crosses between *O. faveolata* and *O. franksi* are also unsuccessful.

Crosses between *O. annularis* and *O. franksi* can be successful with high fertilization rates, particularly when combining *O. annularis* sperm and *O. franksi* egg, which yield 50 % fertilization rates and which are almost as high as the intra-specific crosses that have minimally 60% fertilization rates in this site (although when combining *O. franksi* sperm with

O. annularis egg fertilization rates are much lower, 10%) (Levitan et al., 2004) Although gametes of these two species are not likely to meet in the water column since they are spawned within approximately 100 minutes of each other (Knowlton et al., 1997; Levitan et al., 2004). *O. franksi* spawns one to two hours earlier than *O. annularis,* and by the time *O. annularis* spawns *O. franksi* eggs have drifted long distances and sperm cells are diluted and old reducing their fertilization potential (Levitan et al., 2004). *Orbicella* corals do not self-fertilize (Levitan et al., 2004). Overall, these factors suggest that timely spawning performance is the most effective way to enhance fertilization rates in the field.

Corals use environmental cues to synchronize spawning. Spawning behavior (and gamete maturation) correlates with increase (Babcock et al., 1986; Soong, 1991). Lunar and diel light cycles have been proposed to act as zeitgebers in coral spawning (Willis, et al., 1985). Although the reproductive behavior of *Orbicella* is well described, it is not clear what molecular mechanisms control the spawning behavior and whether molecular differences in proteins or gene expression are responsible for RI among these species. Given that these mechanisms underlay how reef building coral species are generated, understanding them could have key insights on biodiversity, and coral reef conservation. This study is the first one to comparatively assess the transcriptional profiles of two co-occurring coral species *during* and after spawning with the goal of studying the molecular differences responsible for temporal reproductive isolation generating RI in reef building corals

Methods

Sample collections and processing

We collected coral samples (either scrapped tissue or coral cores) on September 14, 2014 during spawning (approximately 20:30 and 22:10 for *O. franksi* and *O. annularis*, respectively) and post-spawning (~23:30) (Figure 3.1). A total of 16 samples were collected (4 samples x 2 time points x 2 species). Note that the samples collected during the time of spawning may have had gamete bundles whereas the postspawning samples only had parental coral (by then gamete bundles had been released). Given that samples for spawning may have had gamete bundles still in them, these gamete remnants could have had an impact in the gene expression profiles here reported. Comparing the data from spawning to the data from gametes transcriptomes allows it to discern the expression of the parent coral and the gametes. Transcriptome data from gametes of each species (*O. annularis* and *O. franksi*) can hence be useful to determine these differences. For this, frozen gametes bundles were used to extract RNA with the mirVana kit by Ambion, following the manufacturer's protocol. To reduce sequencing costs three samples from different colonies per species were pooled them into one sample and Library preparation followed the Illumina True Seq protocol. Each pooled sample was sequenced with pair-ended reads of 100 nucleotides on Illumina HiSeq400 in the University of Chicago Genomics Facility. 36,612,122 and 34,744,812 reads were obtained from *O. annularis* and *O. franksi* respectively and they were processed using the Trinity v2.4.0 pipeline to assemble a *de novo* transcriptome.

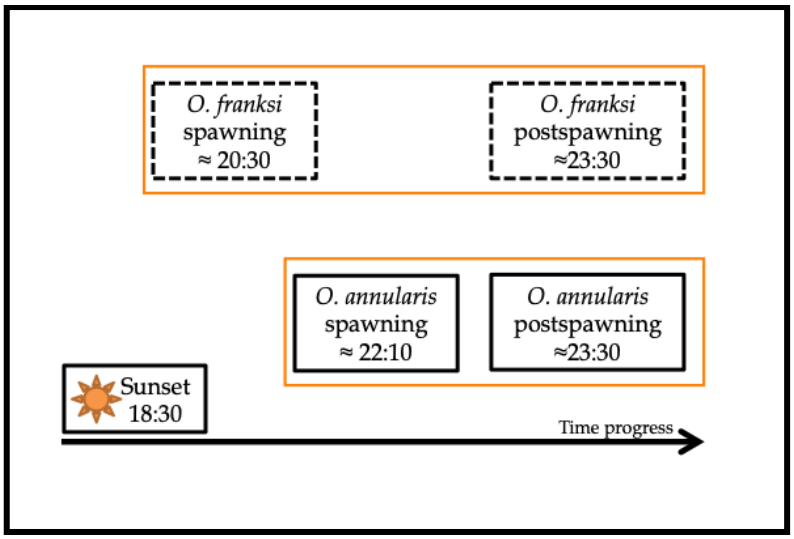


Figure 3.1: Sample collection time points sequenced in this study and experimental design. The orange boxes around the sampling time points for each species show the transcriptome comparisons we present here. The horizontal arrow at the bottom signals time progress and spawning occurred on September 14 2014.

Adult coral samples were immediately flash frozen aboard the boat, stored, and transported to the laboratory in dry shippers for further processing. Samples were powdered in liquid nitrogen and kept frozen at all times until RNA extraction. Approximately 0.2 grams of coral powder were used in the total RNA extraction protocol, which followed the mirVana™ miRNA Isolation Kit by Life Technologies (Carlsbad, California). RNA was purified and concentrated using the Zymo Research RNA Clean & Concentrator™ kit (Irvine, California). Sample quantity was measured using the Qubit 2.0 RNA Broad Range Assay Kit, Invitrogen (Carlsbad, California). NanoDrop Spectrophotometer by Thermo Scientific (Wilmington, Delaware) was used to assess protein contamination and sample quality and the Bionalyzer 2100 instrument by Agilent (Santa Clara, California) was used to measure RNA integrity. All

samples had RIN values greater than 7.1 with one exception, which had a RIN score of 6.8. Nonetheless, all samples had distinct bands typical from the 18S and 28S RNA subunits indicating good RNA integrity (Schroeder et al.,2006). Libraries were prepared following the Illumina TruSeq stranded mRNA protocol (San Diego, California). The Illumina HiSeq 2500 sequencer was used to generate 100bp single-end reads. Library preparation and sequencing were done at Pennsylvania State University Genomics Core Facility.

Data processing

Seventeen to 25 million raw reads were obtained per sample, most of which had high quality Phred scores and very low adaptor contamination (data not shown). Transcripts were mapped against the transcriptome for *Orbicella faveolata* (Anderson et al. 2016) and abundances were quantified using Kallisto (Bray, et al., 2015). At the time there were no published transcriptomes for *O. annularis* or *O. franksi*. Draft genomes for both species were available but due to the high contig numbers, genome guided assembly with Trinity pipeline and quantification with Cuffdiff proved difficult (István Alvert, personal communication). Therefore, the published transcriptome from *Orbicella faveolata* (Anderson et al. 2016) was used as a reference, given that it was a well annotated source with environmental samples (Table 3.1).

Table 3.1: Compiled data from genomes available at the time of the analyses.

Genome file	version	Format	Type	Number of contigs	Total length	Contig minimum length	Contig average length	Contig maximum length
annularis.fasta	1	FASTA	DNA	1,462,333	826,496,413	77	565.2	202,029
franksi.fa	1	FASTA	DNA	1,508,759	810,129,945	75	537	167,780
faveoFlorida.fa	1	FASTA	DNA	6,076,806	1,236,531,036	71	203.5	495,114
Faveolata_Dovetal.fna	2	FASTA	DNA	72,291	512,429,270	200	7,088.40	4,771,691

Given that the reference transcriptome was made from a series of environmentally collected adult samples, presence of non-coral eukaryotes in the data was inevitable (i.e., corals host many other organisms in and on their tissues). To focus exclusively on coral host gene expression, the analysis included a step in which the quantified transcript abundances were filtered using the genomes of *O. franksi* and *O. annularis* to separate the reads that came from corals from the ones that are not coral before quantifying the read abundances with Kallisto. Then, differential expression analyses were conducted with DESEq using its default normalization method, which accounts for the number of raw transcripts counts and the size of the library where they are found (Anders & Huber, 2010). DESEq was used to compare spawning and postspawning samples of each species. Heatmaps and hierarchical clustering were generated using MeV version 4.8.1 (Saeed et al., 2003) or the command line in *Useful R scripts* (Albert, 2017). Revigo (Supek et al. 2011) and AgriGO (version 2.0, default settings, (Tian et al, 2017)) were used to assess the gene ontologies maps.

Reads that lacked an annotation in the reference transcriptome (shown as UNIPROT ID) were putatively annotated by homology by blasting against the NCBI non-redundant database and using the BLASTN program. The top hit was used to identify the gene name. UNIPROT IDs were obtained from the reference transcriptome annotations in Anderson et al. (2016). Venn diagrams were generated using Venny (Oliveros, 2007).

Data from both species were compared against the list of differentially expressed genes of a similar study in *Acropora millepora* (Kaniewska, et al., 2015) by using their reported UNIPROT Id and the UNIPROT Ids available for this study.

Results

Information from the gamete bundles

Gamete transcriptomes yielded 87702 transcripts and 71611 transcripts from *Orbicella annularis* and *O. franksi* respectively. However, since these samples were pooled and there is only an n=1 of gametes it is impossible to assess transcript abundances in comparison to the adult samples by using differential gene expression analyses. With this data it is only possible to assess transcript presence or absence. It is apparent that there are overlapping transcripts in the gamete and adult samples since 30 transcripts from the 48 DEG in *O. annularis* are also present in the gamete transcriptome. Of these, 14 and 16 are upregulated and downregulated, respectively in the adult *O. annularis*. Similarly, out of 117 DEG found in *O. franksi* during spawning, 82 of them are also present in the gametes (out of which 63 and 19 are upregulated and downregulated, respectively in the adults of *O. franksi*). A summary of these results is Appendix #3.

Information from the adult samples

The transcriptome of the *O. faveolata* reference has approximately 33000 transcripts. Datasets from *O. franksi* and *O. annularis* aligned against 18925, and 18910 of those respectively with a small number of them showing significant differential expression. For *O. franksi,* 117 out of 18925 aligned transcripts were differentially expressed ($p_{adj} < 0.05$) (See Appendix table 4-1) whereas *O. annularis* had only 48 differentially expressed genes (See Appendix Table 4-2) (Figure 4-2). Five genes were commonly found to be differentially expressed in both species. One gene lacked annotation and the others are 1-barH-like 2 homeobox protein (Uniprot ID P48031), Krueppel-like factor 12 (Uniprot ID Q9EPW2), Zinc

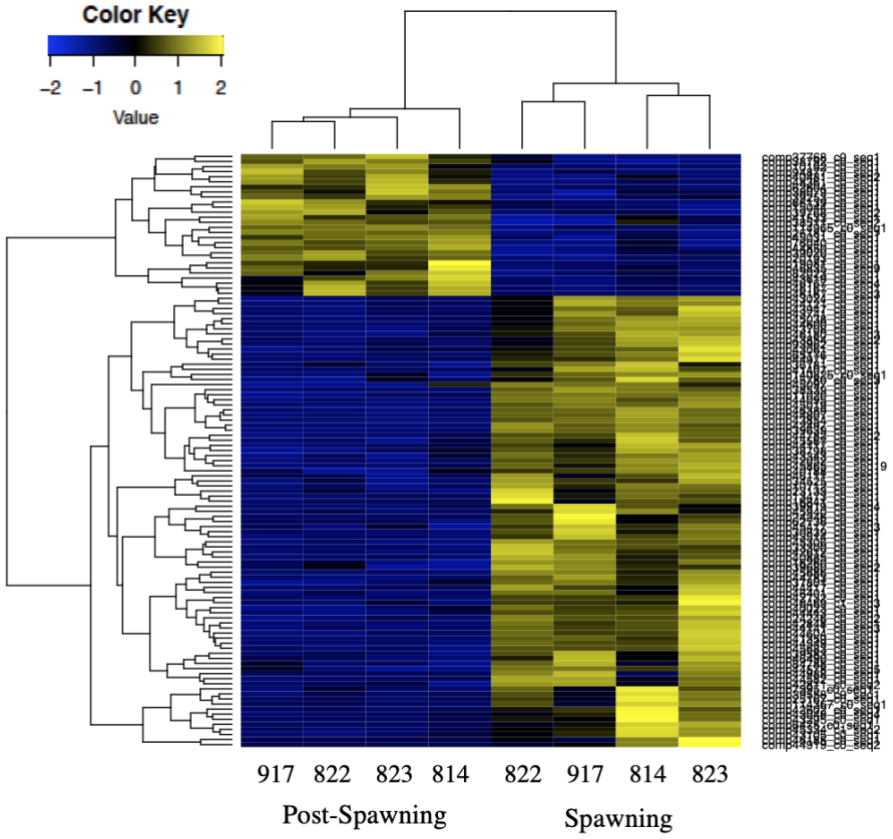
finger CCHC (Uniprot ID B2RVL6), and Patched domain-containing protein 3 (Uniprot ID O15118). These last four genes represent 11, 10, 4 and 7 Gene Ontology groupings respectively, which taken altogether mostly represent binding, intracellular and metabolic processes. One of them, Krueppel-like factor 12 is up regulated in spawning compared to postspawning in both coral species whereas the other three are downregulated in spawning compared to post-spawning.

Both species had many uncharacterized differentially expressed genes in their transcriptomes, making it difficult to assess the complexity of the physiological state during spawning: 75 out of 117 and 24 put of 48 were uncharacterized for *O. franksi* and *O. annularis*.

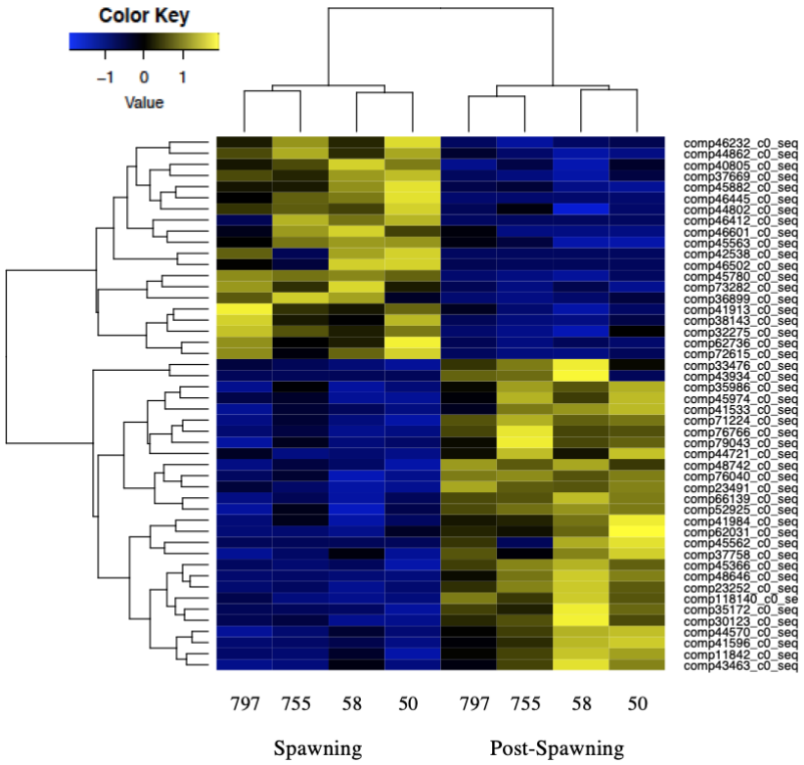
In *O. franksi*, out of the 117 DEG, 88 were upregulated in spawning compared to post-spawning, and 29 were downregulated in spawning compared to post-spawning (see Table 3-2). According to gene ontology (GO) groupings *O. franksi* is enriched for 77 unique GO terms but only three of them within the Molecular Function division are significantly expressed with $p < 0.05$ (Transcription factor activity, sequence-specific DNA binding (GO:0003700), Nucleic acid binding transcription factor activity (GO:0001071), DNA binding (GO:0003677). These GO terms are mapped in figure 3-4. In *O. annularis* 20 of the 48 DEGs were up-regulated during spawning compared to post-spawning, and the rest were down-regulated. When summarized in gene ontology terms, *O. annularis* differentially expressed genes are grouped in only 25 gene ontologies.

Figure **3-2**: Heatmaps showing the differentially expressed genes in *O. annularis* and *Orbicella franksi. A) Orbicella franksi data B) O. annularis data.* The samples in batches of four correspond to the biological replicates during either spawning or post-spawning. There are 117 in *O. franksi* and 48 DEG in *O. annularis*. Colors indicate expression such that yellow represents upregulation respect to blue hues and log2 fold change of Z scores.

a)

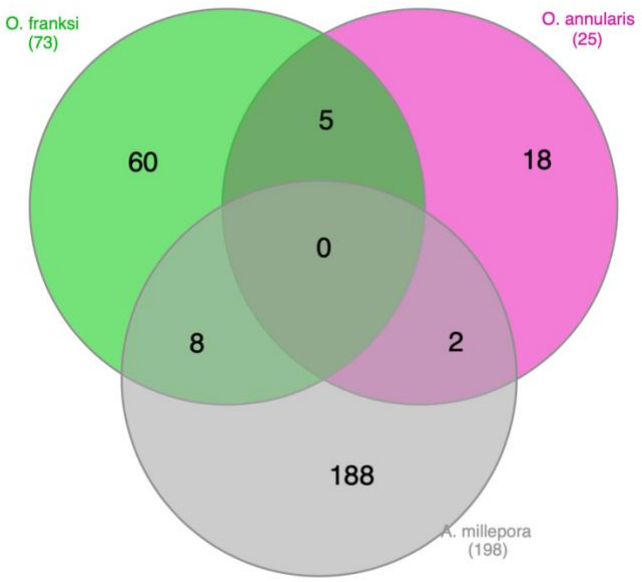


b)



When comparing the differential gene expression profiles of both *Orbicella* species against *Acropora millepora* there is a small number of shared genes. Note that UNIPROT/SWISSPROT identifier codes were used to summarize the identity and function of these transcripts to compare the data from this dissertation and the *A. millepora* data (Kaniewska et al.,2015). It is remarkable that there are no sequences that are differentially expressed during spawning and after spawning in all three species (see Figure 3-3). *O. franksi* however has more commonly expressed genes to *A. millepora* than to its sister species, *O. annularis* suggesting species-specific gene expression profiles.

Figure 3-3: Venn Diagrams showing the overlap in differentially gene expression profiles in three species during spawning and after spawning. Comparisons were based on UNIPROT/SWISSPROT identifier codes. Note than only the transcripts that had an annotation are plotted here (73/117 known to be DEG in *O. franksi*, 25/48 known to be DEG in *O. annularis*, 198/206 known to be DEG in *Acropora millepora*).



Discussion

This is the first study that simultaneously evaluates the gene expression profiles for two coral species during and after spawning. Timing of spawning is fundamental component for the reproductive isolation of *Orbicella* spp., which are sometimes found in sympatry and

annually mass-spawn in the same night. The physiology of gamete release is a complex process that involves the transport of the gamete bundle from the mesenterial tissue to the coral polyp's mouth for a timely release of it. Given the complexity of this process, some degree of overlap was expected in the *O. franksi* and *O. annularis* spawning transcriptomes. The low degree of transcriptome similarity suggests these species use their genetic toolkit differently when they undergo their gamete release and afterwards, which may be playing a role in reproductive isolation. One of the four the DEGs commonly found in both coral profiles, the Krueppel-like factor was upregulated during spawning, and it is an interesting transcriptional factor involved in the regulation of the clock genes in the heart. Another gene was Zinc finger CCHC domain-containing protein 24 (nucleic acid binding gene), which was down regulated during spawning in both species and in both analyses. Homeobox protein GBX-2 (also known as Gastrulation and brain-specific homeobox protein) is another transcription factor and it may be involved in cell pluripotency. Finally, Pregnancy zone protein or PZP is another commonly expressed gene, which functions as a protease and can be involved in embryo implantation.

Though it is hard to assess the rhythmicity of the gene profiles with 2 timepoints, the fact that some of these genes play a role in biological clocks shows the promise of a transcriptome approach for RI research in corals In the particular case of *Orbicella annularis* the list of differentially expressed transcripts was short, 48. In twelve cases the UNIPROT and the manual NCBI BLAST identification coincided, suggesting great confidence in their identification and function.

Acropora millepora and *Orbicella* spp. belong to different coral clades that separated at least 240 million years ago (Romano & Palumbi, 1996). Comparative genomics support the

presence of both clades and very high conservation of gene arrangements within corals (e.g., HOX genes arrangements and symbiosis genes share great similarity among corals (Bhattacharya et al.,2016; Ying et al.,2018).

Given that *Acropora millepora* is a broadcast spawner and engages in mass spawning events, we expected to find some overlap between DEGs of Acropora and the *Orbicella* species. In *Acropora millepora* 177 genes were differentially upregulated during spawning and 29 were differentially downregulated; most of which are associated with functions in cell cycle, GTPase activity and signal transduction during the day of spawning (Kaniewska et al., 2015).

In this study, there were no common DEG during and after spawning among both *Orbicella* species and *Acropora millepora* (note that in both studies the time frame between spawning and post-spawning was two to three hours). Hence, the data reveal no core set of ancestral genes responsible for spawning behavior in scleractinian corals. Determining if such a core of genes exists remains to be shown and will require the examination of additional species across the coral tree. It would be interesting to incorporate brooding species into this analyses and assess the different gene expression profiles of each reproductive strategy because cycles and timing of reproduction peaks vary greatly depending on the reproductive strategy.

It is clear that the gene expression profiles in the sister species *Orbicella annularis* and *O. franksi* are distinct during the time of spawning and within the next few hours. Ecological evidence suggests that in these species the zeitgeber to this sophisticated process is the onset of darkness during sunset (Levitan et al., 2011). Having a better understanding of how the onset of darkness triggers the chain reaction that leads to species-specific spawning time will

be crucial in providing the physiological mechanisms responsible for this temporal isolation. Simulating a premature sunset in the lab a few days prior to the predicted day of spawning by covering experimental coral colonies, induces early onset of gamete release in *O. annularis* for a couple of hours (Weil & Knowlton, 1994). However, several days are needed to do so indicating the circadian clock associated with coral spawning is entrained and will go on for a few a few days in the absence of the zeitgeber.

Coral spawning is a complex and underexplored biological process. The implications executing a timely gamete release are dramatically important considering that not only sexual reproduction occurs once per year but also gametes get preyed on, diluted and old within a short time window of time (Levitan et al., 2011, 2004). Corals exhibit circadian activity and respond to light stimuli from solar and lunar cycles (Reitzel, et al., 2013). Corals have molecular mechanisms to respond to light such as the presence of light sensitive proteins (Levy et al., 2007) and distinct gene expression profiles that associate with the moon cycle (Kaniewska et al., 2015). Change in light availability are known to disrupt entrained circadian rhythms in corals (Brady, et al., 2009). Corals in these situations spawn at unnatural times or fail to spawn at all.

Here we show that there are different genes operating during the time of spawning and after spawning when comparing two sister species of broadcasting corals. Confidence in this data is supported by the fact that the gene expression profiles were very consistent over the four genetically distinct biological replicates here used. It is worth noting, however, that given the nature of the sample elements in the coral themselves such as the gamete bundles or the algal can be present the data. Hence, in the case of gamete transcripts for which we can show expression, these transcripts should be considered with caution. This reduces even more

the list of transcripts that confidently were pertaining to the spawning process itself. Having additional timepoints and perhaps a seasonal/diurnal expression study to compare against could be helpful in finding the nature of the expression of these genes (i.e. if their expression is common or if it mostly conducive to spawning events).

Conclusion

This study highlights the importance of how even closely-related species evolved slightly different genomic mechanisms to produce similar behaviors and how this divergence already provided a robust mechanism for reproductive isolation and coexistence on coral reefs.

Funding

This work was funded by NSF OCE 1442206 and OCE 1642311 and the findings and conclusions do not necessarily reflect the view of the funding agency.

References

Albert, I. (2017). *The Biostar Handbook*. https://doi.org/10.1002/ejoc.201200111

Babcock, R. C., Bull, G. D., Harrison, P. L., Heyward, A. J., Oliver, J. K., Wallace, C. C., & Willis, B. L. (1986). Synchronous spawning of 105 scleractinian coral species on the Great Barrier Reef. *Marine Biology*, *90*(3), 379–394. https://doi.org/10.1007/BF00428562

Baird, A. H., Guest, J. R., & Willis, B. L. (2009). Systematic and Biogeographical Patterns in the Reproductive Biology of Scleractinian Corals. *Annual Review of Ecology, Evolution,*

and Systematics, *40*(1), 551–571. https://doi.org/10.1146/annurev.ecolsys.110308.120220

Bhattacharya, D., Agrawal, S., Aranda, M., Baumgarten, S., Belcaid, M., Drake, J. L., … Falkowski, P. G. (2016). Comparative genomics explains the evolutionary success of reef-forming corals. *ELife*, *5*, 1–26. https://doi.org/10.7554/eLife.13288

Brady, A., Hilton, J., & Vize, P. (2009). Coral spawn timing is a direct response to solar light cycles and is not an entrained circadian response. *Coral Reefs*, *28*(3), 677–680. https://doi.org/10.1007/s00338-009-0498-4

Fogarty, N. D. (2012). Caribbean acroporid coral hybrids are viable across life history stages. *MEPS*, *446*, 145–159.

Fogarty, N. D., Vollmer, S. V, & Levitan, D. R. (2012). Weak Prezygotic Isolating Mechanisms in Threatened Caribbean Acropora Corals. *PloS One*, *7*(2), e30486. Retrieved from file:///Users/electrik/Documents/Publicaciones/2011 prezygotic isolation?.pdf

Foster, A. (1979). Phenotypic plasticity in the reef corals Montastraea annularis (Ellis & Solander) and Siderastrea siderea (Ellis & Solander). *Journal of Experimental Marine Biology and Ecology*, *39*, 25–54. Retrieved from http://www.sciencedirect.com/science/article/pii/0022098179900030

Foster, N. L., Baums, I. B., & Mumby, P. J. (2007). Sexual vs. asexual reproduction in an ecosystem engineer: the massive coral Montastraea annularis. *The Journal of Animal Ecology*, *76*(2), 384–391. https://doi.org/10.1111/j.1365-2656.2006.01207.x

González, A. M., Prada, C. A., Ávila, V., & Medina, M. (2018). Ecological speciation in corals. In *Population Genomics: Marine Organisms* (pp. 2056–2074). Springer.

https://doi.org/10.1111/j.1558-5646.2007.00171.x

Harrison, P., Babcock, R., & Bull, G. (1984). Mass spawning in tropical reef corals. *Science*, *223*(4641), 1186–1189. Retrieved from http://www.sciencemag.org/content/223/4641/1186.short

Isomura, N., Iwao, K., & Fukami, H. (2013). Possible natural hybridization of two morphologically distinct species of Acropora (Cnidaria, Scleractinia) in the Pacific: fertilization and larval survival rates. *PloS One*, *8*(2), e56701. https://doi.org/10.1371/journal.pone.0056701

Jackson, J. B. C., & Hughes, T. P. (1985). Adaptive Strategies of Coral-Reef Invertebrates. *American Scientist*, *73*(3), 265–274.

Jackson, J., & Coates, A. (1986). Life cycles and evolution of clonal (modular) animals. *Philosophical Transactions of the Royal Society of London. Series B, Biological Sciences*, *313*(1159), 7–22. Retrieved from http://rstb.royalsocietypublishing.org/content/313/1159/7.short

Kaniewska, P., Alon, S., Karako-Lampert, S., Hoegh-Guldberg, O., & Levy, O. (2015). Signaling cascades and the importance of moonlight in coral broadcast mass spawning. *ELife*, *4*, 1–14. https://doi.org/10.7554/eLife.09991

Kenyon, J. (1997). Models of reticulate evolution in the coral genus Acropora based on chromosome numbers: parallels with plants. *Evolution*, *51*(3), 756–767. Retrieved from http://www.jstor.org/stable/10.2307/2411152

Knowlton, N., Mate, J., Guzman, H., Rowan, R., & Jara, J. (1997). Direct evidence for reproductive isolation among the three species of the Montastraea annularis complex in Central America (Panama and Honduras). *Marine Biology*, *127*, 705–711. Retrieved

from http://link.springer.com/article/10.1007/s002270050061

Kramarsky-Winter, E. ., Fine, M. ., & Loya, Y. (1996). Coral polyp expulsion. *Nature*, *379*(6561), 126–126. https://doi.org/10.1038/379126b0

Levitan, D. R. (2010). Sexual Selection in External Fertilizers. In *Evolutionary behavioral ecology* (pp. 365–379).

Levitan, D. R., Fogarty, N. D., Jara, J., Lotterhos, K. E. K., & Knowlton, N. (2011). Genetic, spatial, and temporal components of precise spawning synchrony in reef building corals of the Montastraea annularis species complex. *Evolution*, *65*(5), 1254–1270. https://doi.org/10.1111/j.1558-5646.2011.01235.x

Levitan, D. R., Fukami, H., Jara, J., Kline, D., McGovern, T. M., McGhee, K. E., … Knowlton, N. (2004). Mechanisms of reproductive isolation among sympatric broadcast-spawning corals of the Montastraea annularis species complex. *Evolution; International Journal of Organic Evolution*, *58*(2), 308–323. https://doi.org/10.1111/j.0014-3820.2004.tb01647.x

Levy, O., Appelbaum, L., Leggat, W., Gothlif, Y., Hayward, D. ., Miller, D. ., & Hoegh-Guldberg, O. (2007). Light-Responsive Cryptochromes from a Simple Multicellular Animal, the Coral Acropora millepora. *Science*, *318*(5849), 467–470. Retrieved from http://www.sciencemag.org/content/318/5849/467.short

Mayr, E. (1947). Ecological Factors in Speciation. *Evolution*, *1*, 263–288. https://doi.org/10.2307/2405335

Oliveros, J.C. (2007-2015) Venny. An interactive tool for comparing lists with Venn's diagrams. http://bioinfogp.cnb.csic.es/tools/venny/index.html

Palumbi, S. R. (1994). Genetic divergence, reproductive isolation and marine speciation. *Annual Review of Ecology and Systematics*, *25*, 547–572.

Reitzel, A. M., Tarrant, A. M., & Levy, O. (2013). Circadian clocks in the cnidaria: environmental entrainment, molecular regulation, and organismal outputs. *Integrative and Comparative Biology*, *53*(1), 118–130. https://doi.org/10.1093/icb/ict024

Richmond, R., & Hunter, C. (1990). Reproduction and recruitment of corals: comparisons among the Caribbean, the Tropical Pacific, and the Red Sea. *Marine Ecology Progress Series*, *60*, 185–203. https://doi.org/10.3354/meps060185

Sanchez, J., Alvarado, E., Gil, M., Charry, H., Arenas, O., Chasqui, L., & García, R. (1999). Synchronous mass spawning of Montastraea annularis (Ellis & Solander) and Montastraea faveolata (Ellis & Solander) (Faviidae: Scleractinia) at Rosario Islands, Caribbean Coast of Colombia. *Bulletin of Marine Science*, *65*(3), 873–879. Retrieved from http://www.ingentaconnect.com/content/umrsmas/bullmar/1999/00000065/00000003/art00026

Supek, F., Bošnjak, M., Škunca, N., & Šmuc, T. (2011). Revigo summarizes and visualizes long lists of gene ontology terms. *PLoS ONE*, *6*(7). https://doi.org/10.1371/journal.pone.0021800

Szmant, A. (1986). Coral Reefs Reproductive ecology of Caribbean reef corals. *Coral Reefs*, *5*, 43–53.

Szmant, A. M., Weil, E., Miller, M. W., & Colon, D. E. (1997). Hybridization within the species complex of the scleractinan coral Montastraea annularis. *Marine Biology*, *129*, 561–572. Retrieved from http://link.springer.com/article/10.1007/s002270050197

Tian, T., Liu, Y., Yan, H., Yi, X., Du, Z., Xu, W., Su, Z.(2017) agriGO v2.0: a GO analysis toolkit for the agricultural community, 2017 update. Nucleic Acids Res 2017 gkx382. doi: 10.1093/nar/gkx382

van Oppen, M. J., McDonald, B. J., Willis, B., & Miller, D. J. (2001). The evolutionary history of the coral genus Acropora (Scleractinia, Cnidaria) based on a mitochondrial and a nuclear marker: reticulation, incomplete lineage sorting, or morphological convergence? *Molecular Biology and Evolution*, *18*(7), 1315–1329. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/11420370

van Oppen, M., Willis, B., van Vugt, W., & Miller, D. (2000). Examination of species boundaries in the Acropora cervicornis group (Scleractinia, Cnidaria) using nuclear DNA sequence analyses. *Molecular Ecology*, *9*, 1363–1373. Retrieved from http://onlinelibrary.wiley.com/doi/10.1046/j.1365-294x.2000.01010.x/full

van Veghel, M. (1994). Reproductive characteristics of the polymorphic Caribbean reef building coral Montastrea annularis. I. Gametogenesis and spawning behavior. *Marine Ecology Progress Series*, *109*, 209–219. Retrieved from http://www.onedaymanagement.nl/mediapool/62/624323/data/m109p209.pdf

van Veghel, M., Bak, R., Van, M. L. J., & Rolf, P. M. (1993). Intraspecific variation of a dominant Caribbean reef building coral, Montastrea annularis: genetic, behavioral and morphometric aspects. *Mar Ecol Prog Ser*, *92*(Stoddart 1983), 255–265.

van Veghel, M., Cleary, D., & Bak, R. (1996). Interspecific interactions and competititve ability of the polumorphic reef-building coral Montastrea annularis. *Bulletin of Marine …*, *58*(3), 792–803. Retrieved from

http://www.ingentaconnect.com/content/umrsmas/bullmar/1996/00000058/00000003/art00014

van Veghel, M. L. J., & Kahmann, M. E. H. (1994). Reproductive characteristics of the polymorphic Caribbean reef building coral Montastrea annularis. II. Fecundity and colony structure. *Marine Ecology Progress Series*, *109*(2–3), 221–227. https://doi.org/10.3354/meps109209

Vize, P. D. (2006). Deepwater broadcast spawning by Montastraea cavernosa, Montastraea franksi, and Diploria strigosa at the Flower Garden Banks, Gulf of Mexico. *Coral Reefs*, *25*(1), 169–171. https://doi.org/10.1007/s00338-005-0082-5

Weil, E., & Knowlton, N. (1994). A multi-character analysis of the Caribbean coral Montastraea annularis (Ellis and Solander, 1786) and its two sibling species, M. faveolata (Ellis and Solander, 1786). *Bulletin of Marine Science*, *55*(1), 151–175. Retrieved from http://www.ingentaconnect.com/content/umrsmas/bullmar/1994/00000055/00000001/art00008

Ying, H., Cooke, I., Sprungala, S., Wang, W., Hayward, D. C., Tang, Y., … Miller, D. J. (2018). Comparative genomics reveals the distinct evolutionary trajectories of the robust and complex coral lineages. *Genome Biology*, *19*(1), 175. https://doi.org/10.1186/s13059-018-1552-8

Chapter 4

Conclusion

During my doctoral studies, I became interested in learning about speciation, particularly in marine taxa, which are often not restricted by physical barriers, and have long lived and dispersive larvae. This leads to a potential for maintenance of high gene flow between large and distant populations and suggests the ocean could be very homogenized and inhabited by few well-spread taxa with little divergence. There is, however, a great deal of biodiversity in the ocean, particularly in coral reef ecosystems that concentrate in small areas. The central theme of this dissertation was to explore some genetic and molecular mechanisms might be responsible for differentiation among recently divergent taxa.

First, I explored the factors that operate in the ecological speciation in corals. Ecological factors can drive enough segregation in populations to eventually lead them to be reproductive isolated. The ocean is a very stratified environment and physical factors like depth, light, temperature, salinity and others vary along gradients. When populations are segregated by ecological factors, they will be genetically similar to other populations in their ecological niche despite being separated by vast distances, yet they will be divergent from sympatric populations that are in different ecological niches. For example, shallow populations of a species tend to be more alike to other shallow populations than to local deep inhabiting

populations. In corals, depth is perhaps the best documented segregating ecological factor. Other factors such as light and temperature inherently vary along the depth gradient.

Additionally, biological factors also vary by depth, such as prevalence of the mutualistic endosymbiont algae that inhabit corals. I discuss numerous examples of ecological speciation in corals but the best documented cases are the Octocorals *Eunicea* and the Scleractinian corals *Orbicella.* In fact, *Orbicella* became my study system given that it is a group with well-studied natural history, fossil record, ecology, reproductive behavior, and their genomes are available. *Orbicella* species segregated by depth have at least two mechanisms conducive to reproductive isolation: the timing of spawning and the ability of gametes to detect each other as compatible by gamete recognition proteins. Low fertilization rates among some inter-specific crosses in the lab suggest gamete recognition proteins operate in these corals but their description is still pending and warrants future work.

Second, I compared the genome content of *Orbicella* species by assessing the protein-coding ortholog groups to explore whether there have been gene family expansions within each extant *Orbicella* lineage. Here, I first made genome assemblies of *Orbicella*'s close relatives *Cyphastrea serailia* and *C. microphtalma* to use for comparison. At the moment the assemblies are preliminary drafts since both genomes are lacking most core eukaryote and metazoan genes, and have very low coverage making an accurate proteome inference unlikely. Hence, I used the genomes of the Caribbean coral *A. palmata* and *A. cervicornis*, as outgroups. Overall, I found a very large extent of ortholog overlap in these corals, particularly in the *Orbicella* genus. In fact, very few ortholog groups were found to be species-specific and this suggests that modulations in gene expression may be driving speciation within *Orbicella*, rather than changes in genome architecture. It remains to be seen whether gene expression, SNP variation, protein

modification, miRNA expression and ecological factors, among others could be the catalyst for speciation in these corals.

Lastly, I explored the other mechanism of reproductive isolation in *Orbicella*: timing of spawning also known as allochronic assortative mating. While there is ecological evidence of allochronic assortative mating, the molecular mechanisms that lead to shifts in spawning time had not yet been studied. This is the first comparative study that explores this aspect in two young sister broadcasting species. Transcriptome evidence from *O. franksi* (early spawner) and *O. annularis* (late spawner) during and after spawning suggests these sister species vary in number of differentially expressed genes, but also in the identity of these genes (i.e., only 4 genes were found to be differentially expressed between the two taxa). When compared to *Acropora millepora*, we find that *O. franksi* shares more DEG with A. *millepora* than *O. annularis*. Overall, it seems *Orbicella* species use their genetic toolkit differently to regulate the spawning timing. In the future, incorporating timepoints of setting (minutes before spawning, when gamete bundles are transported to the mouth of the polyp) and sunset (an ecological zeitgeber for spawning) alone would enhance our ability to detect gene expression fluctuations and provide further insight into the mechanisms operating on temporal isolation in *Orbicella*.

While further work will be necessary to add more detail to the speciation story of the *Orbicella* lineage this dissertation takes us a couple of steps forward. Now we know that *Orbicella* species have very similar genomes and that gene regulation seems to be a more attributable reason for species differentiation than gene content alone. A list of potential factors both biological and environmental could aid in the separation of these lineages. In the case of gene expression profiling this suggests each species uses a singular gene set during the time of spawning, with implication in their timing. Future work could expand the datasets to enhance

our ability to study gene expression fluctuations over time in this broadcast spawning species. Other aspects of *Orbicella*'s biology also require further studies, like the describing the gamete recognition proteins responsible for mating (or lack thereof), and the assessment of the postzygotic mechanisms that could also be operating when prezygotic mechanisms fail. It would also be worth exploring what Symbiodinaceae inhabit these corals since modern genotyping techniques and recent taxonomy revisions suggest previous genotyping results could be unaccurate, providing another layer to studying the complexity of ecological speciation in these corals. Overall, the broadcast spawning *Orbicella* species offer a great study system for a range of ecologically and evolutionary relevant studies with implications in ecological speciation.

Appendix A

Cyphastrea genome assembly supplementary information

Abstract

Here we report preliminary genome assemblies for *Cyphastrea serailia* and C. *microphtalma,* both of which are Indo Pacific corals. The genome versions currently available are rough drafts given low coverage (less than 20x at best). Further sequencing from more samples including high density DNA samples such as sperm is recommended given that improving the genome assemblies of *Cyphastrea* will be helpful to provide a more closely related outgroup to *Orbicella* and presumably enhance the resolution of comparative genomic Cnidarian studies.

Introduction

We report a preliminary genome assembly for members of the closest living relative to *Orbicella,* namely *Cyphastrea serailia* and *C. microphtalma,* which live in the Indo-Pacific, are the closest living relative to modern *Orbicella* corals (Huang et al.,2014) and share some ecological traits with *Orbicella.* For example, *Cyphastrea* species are common reef corals (Baird, Hoogenboom, & Huang, 2017). They can also be found in deeper waters like *O. franksi* and they are broadcast spawners (DeVantier, et al.,2014,(Richmond & Hunter, 1990; Wilson & Harrison, 2003)

Methods

One tissue sample and its DNA were processed from *C. serailia* and *C. microphtalma* both of which were collected from Lord Howe Island in Australia following the Global Coral Microbiome Project methods (Pollock et al.,2018). In brief, coral fragments were flash frozen in liquid nitrogen and subsequently air blasted to collect only tissue samples. Sequencing of the *Cyphastrea* metagenomes was done at the Joint Genome Institute (Sequencing project codes 1107374 and 1107375 for *C. serailia* and *C. microphtalma*, respectively). In both cases the Illumina HiSeq 2500 was used to obtain 151 pair ended reads and 21,006,930 and 210,069,306 reads were produced initially. Then, according to the JGI BBDuk was used to trim adapters and filter reads of poor quality; BBMap (Bushnell, 2014) was used to map reads to the Human Genome version 19 (reads with more than 93% similarity were removed); and Megahit was used to assemble an initial metagenome (Li, Liu, Luo, Sadakane, & Lam, 2015).

Corals are an ecosystem onto themselves and host many other organisms such as many dinoflagellate algae (i.e., Symbiodinaceae) (LaJeunesse et al.,2018), green algae (i.e., *Ostreobium)* (Del Campo, Pombert, Šlapeta, Larkum, & Keeling, 2017), bacteria (Pollock et al.,2018), virus (Thurber & Correa, 2011), and others. Hence, a series of steps were taken to remove as many reads from organisms other than *Cyphastrea* as possible, since JGI produced a metagenome and the objective herein was to assemble the *Cyphastrea* genomes. First, a Pan-cnidarian genome was concatenated using the genomes of *Orbicella faveolata, Orbicella annularis, Orbicella franksi* (Prada et al.,2016), *Acropora cervicornis*, *Acropora palmata* (Kitchen et al.,2019), *Acropora digitifera* (Shinzato et al.,2011), *Exaiptasia pallida* (Baumgarten et al.,2015), *Montastraea cavernosa* (Fuller et al.,2018), *Porites rus* (Celis et al.,2018), *Styllophora pistillata* (Voolstra et al.,2017), *Hydra magnipapillata* (Chapman et

al.,2010), and *Nematostella vectensis* (Putnam et al.,2007). Both *Cyphastrea* metagenomes were mapped to the Pan-cnidarian genome using BBMap with default settings to remove all reads that were non-cnidarian. Similarly, a second step included the removal of dinoflagellate reads from the Pan-cnidarian mapped data from the previous step using BBMap with default settings. For this a Pan-Symbiodinaceae genome was concatenated with the genomes of *Cladocopium goreaui* (Liu et al.,2018), *Fugacium kawaguti* (Lin et al.,2015), *Symbiodinium microadriaticum* (Aranda et al.,2016), and *Brevolium minutum* (Shoguchi et al.,2013). The output of this mapping still may include non-cnidarian reads so two identical steps were added to remove bacteria and green algae (Chlorophyta). The bacterial database was obtained from NCBI and green algae were obtained by concatenating the JGI genome databases for *Chlorella, Coccomyxia, Micromonas pusilla, Ostreococcus, Volvox carterii,* and *Bathycoccus* genomes.

Then, using BBMap as described previously reads mapped to these databases were removed leaving only coral. The remaining reads were presumably only from coral and used for genome assembly. After clean up, FastQC reported that only 9,769,296 sequences and 71,499,774 sequences remained in *C. serailia* and *C. microphtalma* to proceed with genome assembly. Additionally, FastQC also reported approximately 30% duplication in both data sets (33% for *C. serailia* and 37% for *C. microphtalma*). To assess if there was an ideal KMER to assemble these sequences into a genome for each *Cyphastrea*, Kmergenie (version 1.6982 (Chikhi & Medvedev, 2014)) analyzed the *Cyphastrea* remnant data that was left from the "cleaned" metagenome. Default settings were used. No best kmer was found in either case. Megahit v.1.1.2 was ran to assemble these reads into contigs and it yielded 156,057 contigs and 457,319 contigs for *C. serailia* and *C. microphtalma*, respectively.

Given that the genome size of either *Cyphastrea* species are unknown a range of known coral genomes were used to calculate an estimated genome size including the *Orbicella* genomes, the genome of *Montipora capitata* (this is the largest coral genome known so far) and the genome of *Pocillopora damicornis* (which is the smallest coral genome known to date). The equation used to calculate it was the following:

$$Coverage = \frac{(length\ of\ reads)x(number\ of\ reads)}{total\ genome\ size}$$

Additionally, gVolante version 1.2.1 was used to quantify a range of statistics with Core Eukaryotic Genes Mapping Approach or (CEGMA, ran with non-vertebrate, CEG database settings) and Bench- marking Universal Single-copy Orthologs (BUSCO, ran with v2/v3, Metazoa database) programs (Nishimura, Hara, & Kuraku, 2017). Finally, EukRep version 0.6.5 was used to assess the quantities of data that have Eukaryote to Prokaryote origin (Probst, Grigoriev, West, Banfield, & Thomas, 2018).

Results

Coverage of the coral genomes available range from 45x (*Pocillopora damicornis*, which also has the smallest genome size known to date, 349Mb) to 475x (for *Stylophora pistillata*, which has a genome size of 400Mb), and for example the coverages for the first and second genome versions of *Orbicella faveolata* are 100x and 250x, respectively. Using a range of known coral genomes sizes the genomes of *Cyphastrea* have 30x coverage at best (see Table A-1).

Table A-1: Coverage estimates of the *Cyphastrea* genomes based on other known genome sizes. *Pocillopora damicornis* and *Montipora capitata* have the smallest and largest coral genome size known to date. For both *Cyphastrea* species the length of the reads was 151 base pairs. The number of reads for each were 9,769,296 and 71,499,774 for *C. serailia* and *C. microphtalma*, respectively.

Reference genome		**Estimated coverage (x)**	
Species	**Size (bases)**	***C. serailia***	***C. microphtalma***
Orbicella annularis	826,496,413	1.8	13.1
Orbicella franksi	810,129,945	1.8	13.3
Orbicella faveolata v2	512,429,270	2.9	21.1
Pocillopora damicornis	349,000,000	4.2	30.9
Montipora capitata	885,704,498	1.7	12.2

Less than 10% of the eukaryotic core genes were complete in each *Cyphastrea* assembly, and 30% of the core eukaryotic genes are found when partial genes are included in the report by CEGMA through the gVolante server (Table A-2). Similarly, less than 12% of the metazoan core genes were complete in each *Cyphastrea* assembly, although the percentage increases up to 24% in *C. microphtalma* when partial genes are found by BUSCO through the gVolante server (Table A-2). These results suggest the *Cyphastrea* genome assemblies are currently missing most of the core metazoan and eukaryote genes.

Table A-2: Summary statistics for the genomes of *Cyphastrea* and five others for comparison.

Stats		*O. annularis*	*O. franksi*	*O. faveolata*	*A. cervicornis*	*A. palmata*	*C. serailia*	*C. microphtalma*
# of core genes detected by CEGMA	Complete:	121 (48.79%)	122 (49.19%)	151 (60.89%)	165 (66.53%)	154 (62.10%)	6 (2.42%)	20 (8.06%)
	Complete + Partial:	211 (85.08%)	215 (86.69%)	230 (92.74%)	225 (90.73%)	223 (89.92%)	21 (8.47%)	90 (36.29%)
# of core genes detected by BUSCO	Complete:	708 (72.39%)	701 (71.68%)	835 (85.38%)	838 (85.69%)	843 (86.20%)	67 (6.85%)	115 (11.76%)
	Complete + Partial:	857 (87.63%)	853 (87.22%)	883 (90.29%)	873 (89.26%)	872 (89.16%)	135 (13.80%)	236 (24.13%)
# of contig sequences:		1462333	1508759	72291	4383	2048	156057	457319
Total length (nt):		826496413	810129945	512429270	318373619	304059572	94349283	347624413
Longest sequence (nt):		202029	167780	4771691	1044265	1429101	10033	19923
Shortest sequence (nt):		77	75	200	947	5989	200	200
Mean sequence length (nt):		565	537	7088	72638	148830	605	760
Median sequence length (nt):		135	131	263	32278	86433	492	585
N50 sequence length (nt):		5046	5028	1083318	162227	282027	627	865
Base composition (%):	A	29.98	29.98	22.8	30.51	30.49	30.07	30.2
	T	29.99	29.98	22.8	30.49	30.51	29.57	29.85
	G	19.52	19.53	14.55	19.49	19.52	20.13	19.94
	C	19.56	19.55	14.56	19.51	19.48	20.22	20.01
	N	0.95	0.96	25.29	0	0	0	0
	Other	0.01	0.01	0	0	0	0	0
GC-content (%):		39.46	39.46	38.97	39	39	40.46	39.95
# of sequences containing non-ACGTN (nt):		86392	85745	3449	0	0	0	0

Finally, according to EukRep (Probst et al.,2018) the eukaryote sequences left within these original assemblies were not very numerous with 544 (0.3%) and 4235 (0.9%) left for *C. serailia* and *C. microphtalma*, respectively, see Table 3-3. This indicates that despite the data filtration used, there is still prokaryote contamination in these datasets and minimal coral data.

Table A-3: Genome data separation of prokaryote and eukaryote sequences by EukRep.

Cyphastrea	Name	n	L50	N50	sum
C. serailia	contigs from Megahit	156057	25037	818	6.33E+07
	Eukaryotes sequences	544	217	4097	2264056
	Prokaryotes sequences	155513	25880	800	61E+06
C. microphtalma	contigs from Megahit	457319	83640	1044	2.80E+08
	Eukaryotes sequences	4235	1664	3868	1.73E+07
	Prokaryotes sequences	453084	87867	994	2.63E+08

Discussion

The *Cyphastrea* genome assemblies should be considered rough drafts. Currently, most coral genomes have very high coverages of at least 50x and assemblies start with at least 250 million raw reads. In the case of *Cyphastrea* the samples used to sequence the genome were scrapped tissue from adult colonies (we were unfortunately unable to obtain sperm

DNA). Since corals host a multitude of other organisms, using this kind of tissue inevitably incorporates coral cohabitants into the sequencing data. When only adult coral tissue is available, sequencing at high depth and/or sequencing many samples will provide better coverage. For example, the *Acropora cervicornis* genome was obtained from adult coral samples one of which was sequenced at 160x coverage and the assembly was supplemented using 20 low coverage samples allowing to detect for Single Nucleotide Variants (Kitchen et al.,2019).

Improving the coverage of the *Cyphastrea* genome assemblies (currently 1-30x depending on the genome used for calculations) would also aid in recovering more of the core eukaryotic and metazoan genes to a higher percentage (currently 10-20%). In the future it will be ideal to enrich the datasets by sequencing more samples and if possible, sperm. This would provide plenty of high-quality material to assemble a comprehensive genome, infer accurate predicted proteomes and use them as another source of genomic information for comparative genomic studies in Cnidaria and other taxa.

Data availability

The raw data, processed data and genome assembly drafts by Megahit are in the server owned by the Medina Lab called Montastraea. The path to the folder is

/home/anagonzangel/anadata/othergenomes/Cyphastraea

See the README_Cypha_genome_assembly_notes.txt for details.

References

Aranda, M., Li, Y., Liew, Y. J., Baumgarten, S., Simakov, O., Wilson, M. C., … Voolstra, C. R. (2016). Genomes of coral dinoflagellate symbionts highlight evolutionary adaptations conducive to a symbiotic lifestyle. *Scientific Reports*, *6*(August), 1–15. https://doi.org/10.1038/srep39734

Baird, A. H., Hoogenboom, M. O., & Huang, D. (2017). Cyphastrea salae, a new species of hard coral from Lord Howe Island, Australia (Scleractinia, Merulinidae). *ZooKeys*, *2017*(662), 49–66. https://doi.org/10.3897/zookeys.662.11454

Baumgarten, S., Simakov, O., Esherick, L. Y., Liew, Y. J., Lehnert, E. M., Michell, C. T., … Voolstra, C. R. (2015). The genome of *Aiptasia* , a sea anemone model for coral symbiosis. *Proceedings of the National Academy of Sciences*, 201513318. https://doi.org/10.1073/pnas.1513318112

Celis, J. S., Wibberg, D., Ramírez-Portilla, C., Rupp, O., Winkler, A. S. A., Kalinowski, J., & Wilke, T. (2018). Binning enables efficient host genome reconstruction in cnidarian holobionts. *GigaScience*, *7*(7), 1–12. https://doi.org/10.1093/gigascience/giy075

Chapman, J. a, Kirkness, E. F., Simakov, O., Hampson, S. E., Mitros, T., Weinmaier, T., … Steele, R. E. (2010). The dynamic genome of Hydra. *Nature*, *464*(7288), 592–596. https://doi.org/10.1038/nature08830

Chikhi, R., & Medvedev, P. (2014). Informed and automated k-mer size selection for genome assembly. *Bioinformatics*, *30*(1), 31–37. https://doi.org/10.1093/bioinformatics/btt310

Del Campo, J., Pombert, J. F., Šlapeta, J., Larkum, A., & Keeling, P. J. (2017). The “other” coral symbiont: Ostreobium diversity and distribution. *ISME Journal*, *11*(1), 296–299.

https://doi.org/10.1038/ismej.2016.101

Huang, D., Benzoni, F., Fukami, H., Knowlton, N., Smith, N. D., & Budd, A. F. (2014). Taxonomic classification of the reef coral families Merulinidae, Montastraeidae, and Diploastraeidae (Cnidaria: Anthozoa: Scleractinia). *Zoological Journal of the Linnean Society*, *171*(2), 277–355. https://doi.org/10.1111/zoj.12140

Kitchen, S. A., Ratan, A., Bedoya-Reina, O. C., Burhans, R., Fogarty, N. D., Miller, W., & Baums, I. B. (2019). Genomic Variants Among Threatened Acropora Corals. *Genes|Genomes|Genetics*, *9*(5), 1633–1646. https://doi.org/10.1534/g3.119.400125

LaJeunesse, T. C., Parkinson, J. E., Gabrielson, P. W., Jeong, H. J., Reimer, J. D., Voolstra, C. R., & Santos, S. R. (2018). Systematic Revision of Symbiodiniaceae Highlights the Antiquity and Diversity of Coral Endosymbionts. *Current Biology*, *28*(16), 1–11. https://doi.org/10.1016/j.cub.2018.07.008

Li, D., Liu, C. M., Luo, R., Sadakane, K., & Lam, T. W. (2015). MEGAHIT: An ultra-fast single-node solution for large and complex metagenomics assembly via succinct de Bruijn graph. *Bioinformatics*, *31*(10), 1674–1676. https://doi.org/10.1093/bioinformatics/btv033

Lin, S., Liou, H.-L., Zhang, Y., Zhang, H., Ji, Z., Cai, M., … Morse, D. (2015). The Symbiodinium kawagutii genome illuminates dinoflagellate gene expression and coral symbiosis. *Science*, *350*(6261), 691–694. https://doi.org/10.1126/science.aad0408

Liu, H., Stephens, T. G., González-Pech, R. A., Beltran, V. H., Lapeyre, B., Bongaerts, P., … Chan, C. X. (2018). Symbiodinium genomes reveal adaptive evolution of functions related to coral-dinoflagellate symbiosis. *Communications Biology*, *1*(1), 95. https://doi.org/10.1038/s42003-018-0098-3

Nishimura, O., Hara, Y., & Kuraku, S. (2017). GVolante for standardizing completeness assessment of genome and transcriptome assemblies. *Bioinformatics*, *33*(22), 3635–3637. https://doi.org/10.1093/bioinformatics/btx445

Pollock, F. J., McMinds, R., Smith, S., Bourne, D. G., Willis, B. L., Medina, M., … Zaneveld, J. R. (2018). Coral-associated bacteria demonstrate phylosymbiosis and cophylogeny. *Nature Communications*, *9*(1), 1–13. https://doi.org/10.1038/s41467-018-07275-x

Prada, C., Hanna, B., Budd, A. F., Woodley, C. M., Schmutz, J., Grimwood, J., … Medina, M. (2016). Empty Niches after Extinctions Increase Population Sizes of Modern Corals. *Current Biology*, *26*(23), 3190–3194. https://doi.org/10.1016/j.cub.2016.09.039

Probst, A. J., Grigoriev, I. V., West, P. T., Banfield, J. F., & Thomas, B. C. (2018). Genome-reconstruction for eukaryotes from complex natural microbial communities. *Genome Research*, *28*(4), 569–580. https://doi.org/10.1101/gr.228429.117

Putnam, N. H., Srivastava, M., Hellsten, U., Dirks, B., Chapman, J., Salamov, A., … Rokhsar, D. S. (2007). Sea anemone genome reveals ancestral eumetazoan gene repertoire and genomic organization. *Science (New York, N.Y.)*, *317*(5834), 86–94. https://doi.org/10.1126/science.1139158

Richmond, R., & Hunter, C. (1990). Reproduction and recruitment of corals: comparisons among the Caribbean, the Tropical Pacific, and the Red Sea. *Marine Ecology Progress Series*, *60*, 185–203. https://doi.org/10.3354/meps060185

Shinzato, C., Shoguchi, E., Kawashima, T., Hamada, M., Hisata, K., Tanaka, M., … Satoh, N. (2011). Using the Acropora digitifera genome to understand coral responses to environmental change. *Nature*, *476*(7360), 320–323. https://doi.org/10.1038/nature10249

Shoguchi, E., Shinzato, C., Kawashima, T., Gyoja, F., Mungpakdee, S., Koyanagi, R., …

Satoh, N. (2013). Draft assembly of the symbiodinium minutum nuclear genome reveals dinoflagellate gene structure. *Current Biology*, *23*(15), 1399–1408. https://doi.org/10.1016/j.cub.2013.05.062

Thurber, R. L. V., & Correa, A. M. S. (2011). Viruses of reef-building scleractinian corals. *Journal of Experimental Marine Biology and Ecology*, *408*(1–2), 102–113. https://doi.org/10.1016/j.jembe.2011.07.030

Voolstra, C. R., Li, Y., Liew, Y. J., Baumgarten, S., Zoccola, D., Flot, J. F., … Aranda, M. (2017). Comparative analysis of the genomes of Stylophora pistillata and Acropora digitifera provides evidence for extensive differences between species of corals. *Scientific Reports*, *7*(1), 1–14. https://doi.org/10.1038/s41598-017-17484-x

Wilson, J., & Harrison, P. (2003). Spawning patterns of scleractinian corals at the Solitary Islands — a high latitude coral community in eastern Australia. *Marine Ecology Progress Series*, *260*, 115–123.

Appendix B

Lists of summarized (p_{adj} <0.05) gene ontology groupings in the genomes of *Orbicella*

Appendix Table 2-1: List of summarized (p_{adj} <0.05) gene ontology groupings in the genomes of *Orbicella* sp for 127 Biological Process.

Term_ID	Description
GO:0000165	MAPK Cascade
GO:0001501	Skeletal System Development
GO:0001505	Regulation Of Neurotransmitter Levels
GO:0001525	Angiogenesis
GO:0001964	Startle Response
GO:0002021	Response To Dietary Excess
GO:0006790	Sulfur Compound Metabolic Process
GO:0006816	Calcium Ion Transport
GO:0006820	Anion Transport
GO:0006836	Neurotransmitter Transport
GO:0006855	Drug Transmembrane Transport
GO:0006898	Receptor-Mediated Endocytosis
GO:0006939	Smooth Muscle Contraction
GO:0006952	Defense Response
GO:0006954	Inflammatory Response
GO:0007158	Neuron Cell-Cell Adhesion

GO:0007167	Enzyme Linked Receptor Protein Signaling Pathway
GO:0007186	G-Protein Coupled Receptor Signaling Pathway
GO:0007187	G-Protein Coupled Receptor Signaling Pathway, Coupled To Cyclic Nucleotide Second Messenger
GO:0007188	Adenylate Cyclase-Modulating G-Protein Coupled Receptor Signaling Pathway
GO:0007200	Phospholipase C-Activating G-Protein Coupled Receptor Signaling Pathway
GO:0007205	Protein Kinase C-Activating G-Protein Coupled Receptor Signaling Pathway
GO:0007212	Dopamine Receptor Signaling Pathway
GO:0007218	Neuropeptide Signaling Pathway
GO:0007268	Chemical Synaptic Transmission
GO:0007416	Synapse Assembly
GO:0007431	Salivary Gland Development
GO:0007585	Respiratory Gaseous Exchange
GO:0007586	Digestion
GO:0007610	Behavior
GO:0007618	Mating
GO:0007625	Grooming Behavior

GO:0007631	Feeding Behavior
GO:0008202	Steroid Metabolic Process
GO:0008284	Positive Regulation Of Cell Proliferation
GO:0009187	Cyclic Nucleotide Metabolic Process
GO:0009190	Cyclic Nucleotide Biosynthetic Process
GO:0009582	Detection Of Abiotic Stimulus
GO:0009611	Response To Wounding
GO:0009612	Response To Mechanical Stimulus
GO:0010092	Specification Of Animal Organ Identity
GO:0010243	Response To Organonitrogen Compound
GO:0010518	Positive Regulation Of Phospholipase Activity
GO:0010753	Positive Regulation Of Cgmp-Mediated Signaling
GO:0010771	Negative Regulation Of Cell Morphogenesis Involved In Differentiation
GO:0010817	Regulation Of Hormone Levels
GO:0015696	Ammonium Transport
GO:0015837	Amine Transport
GO:0015844	Monoamine Transport
GO:0015849	Organic Acid Transport
GO:0015850	Organic Hydroxy Compound Transport
GO:0018298	Protein-Chromophore Linkage

GO:0018958	Phenol-Containing Compound Metabolic Process
GO:0019233	Sensory Perception Of Pain
GO:0019932	Second-Messenger-Mediated Signaling
GO:0019935	Cyclic-Nucleotide-Mediated Signaling
GO:0019954	Asexual Reproduction
GO:0021884	Forebrain Neuron Development
GO:0022600	Digestive System Process
GO:0023014	Signal Transduction By Protein Phosphorylation
GO:0023058	Adaptation Of Signaling Pathway
GO:0030198	Extracellular Matrix Organization
GO:0030582	Reproductive Fruiting Body Development
GO:0030587	Sorocarp Development
GO:0031128	Developmental Induction
GO:0031279	Regulation Of Cyclase Activity
GO:0031960	Response To Corticosteroid
GO:0032846	Positive Regulation Of Homeostatic Process
GO:0032963	Collagen Metabolic Process
GO:0033555	Multicellular Organismal Response To Stress
GO:0035272	Exocrine System Development
GO:0042133	Neurotransmitter Metabolic Process
GO:0042330	Taxis
GO:0042391	Regulation Of Membrane Potential

GO:0042401	Cellular Biogenic Amine Biosynthetic Process
GO:0042403	Thyroid Hormone Metabolic Process
GO:0042493	Response To Drug
GO:0042744	Hydrogen Peroxide Catabolic Process
GO:0042749	Regulation Of Circadian Sleep/Wake Cycle
GO:0043062	Extracellular Structure Organization
GO:0043410	Positive Regulation Of MAPK Cascade
GO:0044070	Regulation Of Anion Transport
GO:0044236	Multicellular Organism Metabolic Process
GO:0044706	Multi-Multicellular Organism Process
GO:0045744	Negative Regulation Of G-Protein Coupled Receptor Protein Signaling Pathway
GO:0046717	Acid Secretion
GO:0046942	Carboxylic Acid Transport
GO:0048017	Inositol Lipid-Mediated Signaling
GO:0048265	Response To Pain
GO:0048638	Regulation Of Developmental Growth
GO:0048705	Skeletal System Morphogenesis
GO:0050803	Regulation Of Synapse Structure Or Activity
GO:0050808	Synapse Organization
GO:0050817	Coagulation
GO:0050878	Regulation Of Body Fluid Levels

GO:0050880	Regulation Of Blood Vessel Size
GO:0050905	Neuromuscular Process
GO:0050982	Detection Of Mechanical Stimulus
GO:0051241	Negative Regulation Of Multicellular Organismal Process
GO:0051339	Regulation Of Lyase Activity
GO:0051481	Negative Regulation Of Cytosolic Calcium Ion Concentration
GO:0051588	Regulation Of Neurotransmitter Transport
GO:0051606	Detection Of Stimulus
GO:0051923	Sulfation
GO:0051930	Regulation Of Sensory Perception Of Pain
GO:0051937	Catecholamine Transport
GO:0051954	Positive Regulation Of Amine Transport
GO:0060078	Regulation Of Postsynaptic Membrane Potential
GO:0060343	Trabecula Formation
GO:0061383	Trabecula Morphogenesis
GO:0061448	Connective Tissue Development
GO:0070206	Protein Trimerization
GO:0070207	Protein Homotrimerization
GO:0071875	Adrenergic Receptor Signaling Pathway
GO:0072376	Protein Activation Cascade

GO:0075259	Spore-Bearing Organ Development
GO:0090066	Regulation Of Anatomical Structure Size
GO:0098742	Cell-Cell Adhesion Via Plasma-Membrane Adhesion Molecules
GO:0098771	Inorganic Ion Homeostasis
GO:0099504	Synaptic Vesicle Cycle
GO:0099560	Synaptic Membrane Adhesion
GO:1901571	Fatty Acid Derivative Transport
GO:1901615	Organic Hydroxy Compound Metabolic Process
GO:1901698	Response To Nitrogen Compound
GO:1903034	Regulation Of Response To Wounding
GO:1903510	Mucopolysaccharide Metabolic Process
GO:2000479	Regulation Of Camp-Dependent Protein Kinase Activity

Appendix Table 2-2: List of summarized (p_{adj} <0.05) gene ontology groupings in the genomes of *Orbicella* sp for Molecular Function.

Term_ID	Description
GO:0004930	G-Protein Coupled Receptor Activity
GO:0005201	Extracellular Matrix Structural Constituent

GO:0005216	Ion Channel Activity
GO:0008146	Sulfotransferase Activity
GO:0046906	Tetrapyrrole Binding
GO:0004383	Guanylate Cyclase Activity
GO:0004497	Monooxygenase Activity
GO:0008237	Metallopeptidase Activity
GO:0005164	Tumor Necrosis Factor Receptor Binding
GO:0042562	Hormone Binding
GO:0042277	Peptide Binding
GO:0005539	Glycosaminoglycan Binding
GO:0008144	Drug Binding
GO:0005509	Calcium Ion Binding
GO:0051380	Norepinephrine Binding
GO:1901338	Catecholamine Binding
GO:0020037	Heme Binding
GO:0004713	Protein Tyrosine Kinase Activity
GO:0008201	Heparin Binding
GO:0009378	Four-Way Junction Helicase Activity
GO:0016782	Transferase Activity, Transferring Sulfur-Containing Groups
GO:0070405	Ammonium Ion Binding

GO:0008395	Steroid Hydroxylase Activity
GO:0031996	Thioesterase Binding
GO:0003964	RNA-Directed DNA Polymerase Activity
GO:0072349	Modified Amino Acid Transmembrane Transporter Activity
GO:0005102	Receptor Binding
GO:0070330	Aromatase Activity
GO:0008519	Ammonium Transmembrane Transporter Activity
GO:0022803	Passive Transmembrane Transporter Activity
GO:0038024	Cargo Receptor Activity
GO:0009881	Photoreceptor Activity
GO:0042626	Atpase Activity, Coupled To Transmembrane Movement Of Substances
GO:0001609	G-Protein Coupled Adenosine Receptor Activity
GO:0001594	Trace-Amine Receptor Activity
GO:0001517	N-Acetylglucosamine 6-O-Sulfotransferase Activity
GO:0004952	Dopamine Neurotransmitter Receptor Activity
GO:0001653	Peptide Receptor Activity
GO:0030594	Neurotransmitter Receptor Activity
GO:0015280	Ligand-Gated Sodium Channel Activity

GO:0099589	Serotonin Receptor Activity
GO:0035586	Purinergic Receptor Activity
GO:0005126	Cytokine Receptor Binding
GO:0004222	Metalloendopeptidase Activity
GO:0099528	G-Protein Coupled Neurotransmitter Receptor Activity

Appendix Table 2-3: List of summarized (p_{adj} <0.05) gene ontology groupings in the genomes of *Orbicella* sp for Molecular Function.

Term_ID	Description
GO:0005615	Extracellular Space
GO:0009986	Cell Surface
GO:0043235	Receptor Complex
GO:0045121	Membrane Raft
GO:0031253	Cell Projection Membrane
GO:0045177	Apical Part Of Cell
GO:0036477	Somatodendritic Compartment
GO:0031463	Cul3-RING Ubiquitin Ligase Complex
GO:0005581	Collagen Trimer
GO:0044420	Extracellular Matrix Component

Appendix C

Lists of transcripts present in gametes and parental samples

As stated in Chapter 3, the samples analyzed during the timepoint of spawning may have had gamete remnants in them and the following tables show the overlapping transcripts found in both the gamete transcriptomes as well as in the parental samples during the time of spawning. Given that there gametes were pooled into one sample for the sequencing step, we are unable to report differential expression in the transcripts here reported, and can only report their presence or absence in the parental samples for *Orbicella annularis* and *O. franksi.*

Appendix Table 3 -1 Table of Differentially Expressed Genes present in the samples of this study for *Orbicella annularis*. Note that the "Adult" columns correspond to transcripts that are DEG ($p_{adj} < 0.05$) during spawning compared to postspawning, and their expression (up or downregulation) is noted in the Expression column. Genes highlighted in yellow were also DEG at padj<0.001. Additionally, although at this point differential expression calculations are not possible for the gametes note that the transcripts found in the gamete transcriptomes are listed with the numbers of hits here found.

(See the next page.)

"Adult" coral						Gamete bundles	
Sequence name	UNIPROT ID	UNIPROT Protein name	NCBI BLAST N ACCESSION # (NCBI Reference Sequence)	NCBI BLAST N, first match	Expression	Counts	Transcript ID
comp23252_c0_seq1	P54417	Glycine betaine transporter OpuD	XM_020769058.1	Glycine betaine transporter OpuD-like (LOC110062187), mRNA	↓	2	comp23252_c0_seq1
comp35986_c0_seq1	NA	NA	XM_020770138.1	Patched domain-containing protein 3-like (LOC110063182), mRNA	↓	3	comp35986_c0_seq1
comp37669_c0_seq1	NA	NA	XR_002298673.1	Methenyltetrahydrofolate synthase domain-containing protein-like (LOC110067398), transcript variant X2, misc_RNA	↑	4	comp37669_c0_seq1
comp38143_c0_seq1	P97812	Indian hedgehog protein	XM_020776357.1	Indian hedgehog protein-like (LOC110068937), mRNA	↑	4	comp38143_c0_seq1
comp40805_c0_seq2	Q9SZW4	Cadmium/zinc-transporting ATPase HMA2	XM_020748322.1	Cadmium/zinc-transporting ATPase HMA3-like (LOC110042932), transcript variant X5, mRNA	↑	12	comp40805_c0_seq2
comp41533_c0_seq1	O15118	NPC intracellular cholesterol transporter 1	XM_020770138.1	Patched domain-containing protein 3-like (LOC110063182), mRNA	↓	2	comp41533_c0_seq1
comp41596_c0_seq1	NA	NA	XM_020760507.1	Uncharacterized LOC110054171 (LOC110054171), mRNA	↓	1	comp41596_c0_seq1
comp41913_c0_seq1	NA	NA	XM_020760688.1	Uncharacterized LOC110054306 (LOC110054306), mRNA	↑	2	comp41913_c0_seq1
comp41984_c0_seq2	Q02410	Amyloid-beta A4 precursor protein-binding family A member 1	XM_020771129.1	Dentin sialophosphoprotein-like (LOC110064130), mRNA	↓	3	comp41984_c0_seq2
comp42538_c0_seq3	Q9P2F6	Rho GTPase-activating protein 20	op	Uncharacterized LOC110060432 (LOC110060432), transcript variant X4, mRNA	↑	2	comp42538_c0_seq3
comp43934_c0_seq2	Q803Z2	Protein YIPF3	XM_020746535.1	Protein YIPF3-like (LOC110041252), mRNA	↓	4	comp43934_c0_seq2
comp44570_c0_seq1	NA	NA	XM_020769444.1	Uncharacterized LOC110062517 (LOC110062517), transcript variant X1, mRNA	↓	4	comp44570_c0_seq1
comp44721_c0_seq1	Q9HCJ5	Zinc finger SWIM domain-containing protein 6	XM_020769053.1	Zinc finger SWIM domain-containing protein 6-like (LOC110062186), transcript variant X2, mRNA	↓	3	comp44721_c0_seq1
comp44802_c0_seq1	P52962	Moesin	XM_020767196.1	Radixin-like (LOC110060420), mRNA	↑	1	comp44802_c0_seq1
comp44862_c0_seq2	Q6DRG7	Protein phosphatase 1 regulatory subunit 12A	XM_020746303.1	Protein phosphatase 1 regulatory subunit 12A-like (LOC110041021), transcript variant X3, mRNA	↑	10	comp44862_c0_seq2
comp45366_c0_seq1	NA	NA	XM_020773127.1	Uncharacterized LOC110065941 (LOC110065941), mRNA	↓	4	comp45366_c0_seq1
comp45562_c0_seq2	Q76LC6	RNA-binding protein 24	XM_020761791.1	RNA-binding protein 24-A-like (LOC110055404), mRNA	↓	2	comp45562_c0_seq2
comp45563_c0_seq1	Q9UKN7	Unconventional myosin-XV	XM_020755641.1	Uncharacterized LOC110049813 (LOC110049813), mRNA	↑	1	comp45563_c0_seq1
comp45882_c0_seq3	P25291	Pancreatic secretory granule membrane major glycoprotein GP2	XM_020767946.1	Hemicentin-2-like (LOC110061112), mRNA	↑	5	comp45882_c0_seq3
comp45974_c0_seq2	O95084	Serine protease 23	XM_020753854.1	Serine protease 23-like (LOC110048082), transcript variant X1, mRNA	↓	4	comp45974_c0_seq2
comp46232_c0_seq1	NA	NA	XM_020756253.1	Non-specific lipid-transfer protein-like (LOC110050336), mRNA	↑	5	comp46232_c0_seq1
comp46412_c0_seq2	NA	NA	XM_020756106.1	Sporulation-specific protein 15-like (LOC110050217), transcript variant X2, mRNA	↑	2	comp46412_c0_seq2
comp46445_c0_seq1	NA	NA	XM_020760887.1	Zinc finger MIZ domain-containing protein 1-like (LOC110054549), transcript variant X8, mRNA	↑	2	comp46445_c0_seq1
comp46502_c0_seq5	NA	NA	XM_020748922.1	Nuclear factor of activated T-cells 5-like (LOC110043473), transcript variant X4, mRNA	↑	1	comp46502_c0_seq5

"Adult" coral						Gamete bundles	
comp46601_c0_seq3	NA	NA	XM_020771679.1	Pappalysin-1-like (LOC110064612), transcript variant X3, mRNA	↑	1	comp46601_c0_seq3
comp48646_c0_seq1	P45335	Uncharacterized transporter HI_1706	XM_020769058.1	Glycine betaine transporter OpuD-like (LOC110062187), mRNA	↓	3	comp48646_c0_seq1
comp62031_c0_seq1	NA	NA	XM_020763974.1	Monocarboxylate transporter 10-like (LOC110057384), transcript variant X2, mRNA	↓	1	comp62031_c0_seq1
comp66139_c0_seq1	P48031	Homeobox protein GBX-2	XM_020764401.1	BarH-like 2 homeobox protein (LOC110057804), transcript variant X2, mRNA	↓	1	comp66139_c0_seq1
comp76766_c0_seq1	P55013	Solute carrier family 12 member 2	XM_020769451.1	Solute carrier family 12 member 2-like (LOC110062523), mRNA	↓	1	comp76766_c0_seq1
comp79043_c0_seq1	NA	NA	XM_020769451.1	Solute carrier family 12 member 2-like (LOC110062523), mRNA	↓	1	comp79043_c0_seq1
comp73282_c0_seq1	NA	NA	XM_020751859.1	L-gulonolactone oxidase-like (LOC110046178), mRNA	↑	NA	NA
comp72615_c0_seq1	NA	NA	XM_020768129.1	Homeobox protein zampogna-like (LOC110061283), mRNA	↑	NA	NA
comp45780_c0_seq1	NA	NA	XM_020756471.1	Uncharacterized LOC110050541 (LOC110050541), mRNA	↑	NA	NA
comp32275_c0_seq1	NA	NA	XM_020768302.1	Uncharacterized LOC110061466 (LOC110061466), mRNA	↑	NA	NA
comp62736_c0_seq1	Q9EPW2	Krueppel-like factor 15	XM_020750796.1	Krueppel-like factor 12 (LOC110045197), mRNA	↑	NA	NA
comp36899_c0_seq1	NA	NA	NA	NA	↑	NA	NA
comp35172_c0_seq1	A6NMZ7	Collagen alpha-6(VI) chain	XM_020755960.1	Uncharacterized LOC110050092 (LOC110050092), mRNA	↓	NA	NA
comp52925_c0_seq1	NA	NA	XM_020750744.1	T-box transcription factor TBX20-like (LOC110045149), transcript variant X2, mRNA	↓	NA	NA
comp23491_c0_seq1	NA	NA	XM_020768572.1	Uncharacterized LOC110061726 (LOC110061726), mRNA	↓	NA	NA
comp71224_c0_seq1	Q9BDJ5	Pantetheinase	NA	NA	↓	NA	NA
comp76040_c0_seq1	B2RVL6	Zinc finger CCHC domain-containing protein 24	XM_020753214.1	Zinc finger CCHC domain-containing protein 24-like (LOC110047464), mRNA	↓	NA	NA
comp30123_c0_seq2	A6H584	Collagen alpha-5(VI) chain	XM_020755969.1	Uncharacterized LOC110050097 (LOC110050097), transcript variant X3, mRNA	↓	NA	NA
comp118140_c0_seq1	P70551	Type II iodothyronine deiodinase	XM_020747166.1	Thyroxine 5-deiodinase-like (LOC110041843), mRNA	↓	NA	NA
comp48742_c0_seq1	NA	NA	XM_020768753.1	Uncharacterized threonine-rich GPI-anchored glycoprotein PJ4664.02-like (LOC110061891), partial mRNA	↓	NA	NA
comp33476_c0_seq1	NA	NA	XM_020773127.1	Uncharacterized LOC110065941 (LOC110065941), mRNA	↓	NA	NA
comp37758_c0_seq1	Q6DCQ6	von Willebrand factor A domain-containing protein 2	XM_020776739.1	Integrin alpha-D-like (LOC110069246), mRNA	↓	NA	NA
comp11842_c0_seq2	NA	NA	XM_020755974.1	Uncharacterized LOC110050097 (LOC110050097), transcript variant X7, mRNA	↓	NA	NA
comp43463_c0_seq1	NA	NA	XM_020750644.1	Uncharacterized LOC110045051 (LOC110045051), mRNA	↓	NA	NA

Appendix Table 3 – 2: Table of Differentially Expressed Genes present in the samples of this study for *Orbicella franksi*. Note that the "Adult" columns correspond to transcripts that are DEG ($p_{adj} < 0.05$) during spawning compared to postspawning, and their expression (up or downregulation) is noted in the Expression column. Genes highlighted in yellow were also DEG at padj<0.001. Additionally, although at this point differential expression calculations are not possible for the gametes note that the transcripts found in the gamete transcriptomes are listed with the numbers of hits here found.

(Next page)

"Adult" coral						Gamete bundles	
Sequence name	UNIPROT ID	UNIPROT Protein name	NCBI BLAST N ACCESSION # (NCBI Reference Sequence)	NCBI BLAST N, first match	Expression	Counts	Transcript ID
comp10712_c0_seq1	Q9U3S9	Zinc metalloproteinase nas-6	XM_020747405.1	Zinc metalloproteinase nas-4-like (LOC110042061), transcript variant X2, mRNA	↑	4	comp10712_c0_seq1
comp11020_c0_seq1	Q07352	mRNA decay activator protein ZFP36L1	XM_020748063.1	mRNA decay activator protein ZFP36L1-like (LOC110042696), mRNA	↑	1	comp11020_c0_seq1
comp114367_c0_seq1	NA	NA	XM_020775084.1	Protein FAM124A-like (LOC110067759), transcript variant X4, mRNA	↑	4	comp114367_c0_seq1
comp15044_c0_seq1	NA	NA	XM_020772305.1	Bifunctional TH2 protein, mitochondrial-like (LOC110065196), mRNA	↓	1	comp15044_c0_seq1
comp19081_c0_seq1	NA	NA	XM_020763974.1	Monocarboxylate transporter 10-like (LOC110057384), transcript variant X2, mRNA	↓	1	comp19081_c0_seq1
comp20193_c0_seq1	NA	NA	XM_020761342.1	Bifunctional TH2 protein, mitochondrial-like (LOC110065196), mRNA	↓	1	comp20193_c0_seq1
comp24246_c0_seq1	P50616	Protein Tob1	XM_020770942.1	Protein Tob2-like (LOC110063947), mRNA	↑	1	comp24246_c0_seq1
comp24246_c0_seq2	P50616	Protein Tob1	XM_020770942.1	Protein Tob2-like (LOC110063947), mRNA	↑	1	comp24246_c0_seq2
comp26181_c0_seq1	NA	NA	XM_020764214.1	Transcription factor VBP-like (LOC110057617), mRNA	↓	1	comp26181_c0_seq1
comp30839_c0_seq1	NA	NA	XM_020777062.1	Neurogenic locus notch homolog protein 1-like (LOC110069530), mRNA	↑	1	comp30839_c0_seq1
comp33300_c0_seq1	Q6UB98	Ankyrin repeat domain-containing protein 12	XM_020763516.1	Cortactin-binding protein 2-like (LOC110056953), mRNA	↑	1	comp33300_c0_seq1
comp36781_c0_seq1	P15105	Glutamine synthetase	XM_020752370.1	Glutamine synthetase-like (LOC110046681), mRNA	↑	1	comp36781_c0_seq1
comp37801_c0_seq1	Q6NTY6	Early growth response protein 1-B	XM_020751274.1	Early growth response protein 1-like (LOC110045671), transcript variant X1, mRNA	↑	3	comp37801_c0_seq1
comp38079_c0_seq1	Q4A3R3	Deleted in malignant brain tumors 1 protein	XR_002295978.1	Uncharacterized LOC110043178 (LOC110043178), transcript variant X2, ncRNA	↓	1	comp38079_c0_seq1
comp38511_c0_seq2	Q08CS6	DBH-like monooxygenase protein 2 homolog	XM_020770343.1	Uncharacterized LOC110063370 (LOC110063370), mRNA	↓	1	comp38511_c0_seq2
comp38539_c0_seq1	Q9VVY3	Glycogen-binding subunit 76A	XM_020775678.1	Glycogen-binding subunit 76A-like (LOC110068299), mRNA	↑	2	comp38539_c0_seq1
comp38792_c0_seq1	O35738	Krueppel-like factor 12	XM_020769222.1	Krueppel-like factor 6 (LOC110062347), mRNA	↓	1	comp38792_c0_seq1
comp38798_c0_seq1	P34743	Protein BTG1	XM_020772618.1	Protein BTG2-like (LOC110065472), mRNA	↑	2	comp38798_c0_seq1

"Adult" coral						Gamete bundles	
comp39020_c0_seq1	NA	NA	XR_002296924.1	Uncharacterized LOC110051260 (LOC110051260), ncRNA	↓	1	comp39020_c0_seq1
comp39253_c0_seq1	NA	NA	XM_020765268.1	Roundabout homolog 2-like (LOC110058616), partial mRNA	↓	2	comp39253_c0_seq1
comp39639_c0_seq1	NA	NA	XM_020772814.1	Uncharacterized LOC110065656 (LOC110065656), mRNA	↓	1	comp39639_c0_seq1
comp39680_c0_seq2	Q91VS7	Microsomal glutathione S-transferase 1	XM_020775453.1	Microsomal glutathione S-transferase 1-like (LOC110068084), mRNA	↑	1	comp39680_c0_seq2
comp39919_c0_seq4	P26652	Metalloproteinase inhibitor 3	XM_020752328.1	Metalloproteinase inhibitor 3-like (LOC110046634), mRNA	↑	3	comp39919_c0_seq4
comp40023_c0_seq1	NA	NA	NA	NA	↑	1	comp40023_c0_seq1
comp40188_c0_seq1	NA	NA	XM_020749343.1	Uncharacterized LOC110043845 (LOC110043845), transcript variant X1, mRNA	↑	2	comp40188_c0_seq1
comp40481_c0_seq2	NA	NA	XM_020758728.1	Uncharacterized LOC110052587 (LOC110052587), mRNA	↓	2	comp40481_c0_seq2
comp40885_c0_seq1	Q99581	Protein FEV	XM_020775929.1	ETS translocation variant 4-like (LOC110068544), mRNA	↑	3	comp40885_c0_seq1
comp41359_c0_seq1	Q9H0K1	Serine/threonine-protein kinase SIK2	XM_020765851.1	Serine/threonine-protein kinase SIK2-like (LOC110059158), transcript variant X4, mRNA	↑	3	comp41359_c0_seq1
comp41407_c0_seq1	NA	NA	XM_020752427.1	Uncharacterized LOC110046733 (LOC110046733), mRNA	↑	3	comp41407_c0_seq1
comp41533_c0_seq1	O15118	NPC intracellular cholesterol transporter 1	XM_020770138.1	Patched domain-containing protein 3-like (LOC110063182), mRNA	↓	1	comp41533_c0_seq1
comp42602_c0_seq1	Q7KM13	Hairy/enhancer-of-split related with YRPW motif protein	XM_020760876.1	Transcription factor cwo-like (LOC110054546), transcript variant X1, mRNA	↑	2	comp42602_c0_seq1
comp42612_c0_seq1	NA	NA	XM_020767223.1	Uncharacterized LOC110060441 (LOC110060441), transcript variant X1, mRNA	↑	1	comp42612_c0_seq1
comp42811_c0_seq2	Q9Z2H5	Band 4.1-like protein 1	XM_020768122.1	Band 4.1-like protein 1 (LOC110061280), mRNA	↑	1	comp42811_c0_seq2
comp43024_c0_seq1	NA	NA	XM_020761342.1	Uncharacterized LOC110054961 (LOC110054961), mRNA	↑	1	comp43024_c0_seq1
comp43061_c0_seq1	NA	NA	XR_002297290.1	Uncharacterized LOC110055117 (LOC110055117), ncRNA	↑	5	comp43061_c0_seq1
comp43105_c0_seq1	NA	NA	XM_020771973.1	Uncharacterized LOC110064869 (LOC110064869), mRNA	↑	2	comp43105_c0_seq1
comp43295_c0_seq1	Q9NPA2	Matrix metalloproteinase-25	XM_020762837.1	Matrix metalloproteinase-25-like (LOC110056369), mRNA	↑	3	comp43295_c0_seq1
comp43343_c0_seq1	Q99542	Matrix metalloproteinase-19	XM_020762821.1	Matrix metalloproteinase-24-like (LOC110056355), mRNA	↑	4	comp43343_c0_seq1
comp43514_c0_seq1	O42342	Transcription factor Sox-7	XM_020764830.1	Transcription factor Sox-14-like (LOC110058232), mRNA	↑	1	comp43514_c0_seq1
comp43619_c0_seq1	P0C6B8	Sushi, von Willebrand factor type A, EGF and pentraxin domain-containing protein 1	XM_020755666.1	Uncharacterized LOC110049839 (LOC110049839), mRNA	↓	1	comp43619_c0_seq1
comp43708_c0_seq1	A1Z6E0	Protein gustavus	XM_020768301.1	Protein gustavus-like (LOC110061465), mRNA	↑	5	comp43708_c0_seq1
comp43741_c0_seq1	Q9D0B5	Thiosulfate sulfurtransferase /rhodanese-like domain-containing protein 3	XM_020756079.1	Thiosulfate sulfurtransferase/rhodanese-like domain-containing protein 1 (LOC110050189), mRNA	↑	2	comp43741_c0_seq1
comp43854_c0_seq1	Q76I25	HIG1 domain family member 1C	NA	NA	↑	1	comp43854_c0_seq1
comp43854_c0_seq2	Q76I25	HIG1 domain family member 1C	NA	NA	↑	1	comp43854_c0_seq2

"Adult" coral						Gamete bundles	
comp43947_c0_seq1	Q9R0S2	Matrix metalloproteinase-24	XM_020762836.1	72 kDa type IV collagenase-like (LOC110056368), mRNA	↑	3	comp43947_c0_seq1
comp44095_c0_seq1	Q91827	Apoptosis regulator R1	XM_020746225.1	Apoptosis regulator R1-like (LOC110040960), mRNA	↑	1	comp44095_c0_seq1
comp44117_c0_seq1	Q9EP86	Neuropeptide FF receptor 1	XM_020748456.1	Neuropeptide FF receptor 1-like (LOC110043018), transcript variant X1, mRNA	↑	5	comp44117_c0_seq1
comp44141_c0_seq3	Q3MHQ7	Lysoplasmalogenase-like protein TMEM86A	XM_020768719.1	Lysoplasmalogenase-like protein TMEM86A (LOC110061859), mRNA	↑	7	comp44141_c0_seq3
comp44417_c0_seq1	NA	NA	XM_020774855.1	Uncharacterized LOC110067515 (LOC110067515), mRNA	↑	1	comp44417_c0_seq1
comp44443_c0_seq1	Q0KIA2	PP2C-like domain-containing protein CG9801	XM_020751862.1	PP2C-like domain-containing protein CG9801 (LOC110046180), transcript variant X1, mRNA	↑	3	comp44443_c0_seq1
comp44578_c0_seq5	NA	NA	XM_020751975.1	Uncharacterized LOC110046287 (LOC110046287), mRNA	↑	1	comp44578_c0_seq5
comp44582_c0_seq1	Q80ZQ5	Juxtaposed with another zinc finger protein 1	XM_020745898.1	Juxtaposed with another zinc finger protein 1-like (LOC110040630), mRNA	↑	1	comp44582_c0_seq1
comp44600_c0_seq1	NA	NA	XM_020752518.1	Uncharacterized LOC110046805 (LOC110046805), mRNA	↑	1	comp44600_c0_seq1
comp44604_c0_seq1	Q5FVC7	Arf-GAP with coiled-coil, ANK repeat and PH domain-containing protein 2	XM_020751023.1	Arf-GAP with coiled-coil, ANK repeat and PH domain-containing protein 2-like (LOC110045421), mRNA	↑	1	comp44604_c0_seq1
comp44782_c0_seq2	Q03629	Uncharacterized protein YML079W	XM_020773262.1	Uncharacterized LOC110066062 (LOC110066062), mRNA	↑	1	comp44782_c0_seq2
comp44783_c0_seq1	Q8K3X4	Probable E3 ubiquitin-protein ligase IRF2BPL	XM_020777109.1	Interferon regulatory factor 2-binding protein-like (LOC110069586), mRNA	↑	6	comp44783_c0_seq1
comp44807_c0_seq1	NA	NA	XM_020774952.1	Melanocortin receptor 5-like (LOC110067607), mRNA	↑	5	comp44807_c0_seq1
comp44850_c0_seq1	Q9Y5Z4	Heme-binding protein 2	XM_020771695.1	Heme-binding protein 2-like (LOC110064621), transcript variant X2, mRNA	↓	7	comp44850_c0_seq1
comp44919_c0_seq2	NA	NA	NA	NA	↑	3	comp44919_c0_seq2
comp44971_c0_seq1	P58545	BTB/POZ domain-containing protein 3	XM_020764957.1	BTB/POZ domain-containing protein 3-like (LOC110058317), mRNA	↑	3	comp44971_c0_seq1
comp44988_c0_seq1	Q92626	Peroxidasin homolog	XM_020749475.1	Papilin-like (LOC110043970), mRNA	↑	1	comp44988_c0_seq1
comp45019_c0_seq1	P79145	cAMP-responsive element modulator	XM_020752048.1	Cyclic AMP-responsive element-binding protein 1-like (LOC110046357), mRNA	↑	2	comp45019_c0_seq1
comp45031_c0_seq1	NA	NA	NA	NA	↓	1	comp45031_c0_seq1
comp45038_c0_seq1	Q6PJ21	SPRY domain-containing SOCS box protein 3	XM_020772297.1	SPRY domain-containing SOCS box protein 3-like (LOC110065191), transcript variant X1, mRNA	↑	2	comp45038_c0_seq1
comp45097_c0_seq1	NA	NA	XM_020768661.1	Uncharacterized LOC110061806 (LOC110061806), transcript variant X1, mRNA	↑	2	comp45097_c0_seq1
comp45334_c1_seq2	P07152	Stromelysin-2	XM_020762835.1	Stromelysin-1-like (LOC110056367), mRNA	↑	3	comp45334_c1_seq2
comp45567_c0_seq1	Q9CYL5	Golgi-associated plant pathogenesis-related protein 1	XM_020763535.1	Golgi-associated plant pathogenesis-related protein 1-like (LOC110056967), transcript variant X2, mRNA	↑	6	comp45567_c0_seq1
comp45683_c0_seq1	Q3UFK8	FERM domain-containing protein 8	XM_020752472.1	Putative FERM domain-containing protein FRMD8P1 (LOC110046765), mRNA	↑	13	comp45683_c0_seq1
comp45686_c0_seq1	P28562	Dual specificity protein phosphatase 1	XM_020771923.1	Dual specificity protein phosphatase 1-A-like (LOC110064824), mRNA	↑	1	comp45686_c0_seq1
comp45865_c0_seq19	A3KN95	Transmembrane protein 151B	XM_020771394.1	Uncharacterized LOC110064353 (LOC110064353), mRNA	↑	77	comp45865_c0_seq19

"Adult" coral						Gamete bundles	
comp45940_c0_seq1	NA	NA	XM_020752427.1	Uncharacterized LOC110046733 (LOC110046733), mRNA	↑	2	comp45940_c0_seq1
comp46111_c0_seq1	Q923Q2	StAR-related lipid transfer protein 13	XM_020745441.1	StAR-related lipid transfer protein 13-like (LOC110040233), transcript variant X1, mRNA	↑	5	comp46111_c0_seq1
comp46167_c0_seq4	Q07496	Ephrin type-A receptor 4	XM_020763507.1	Uncharacterized LOC110056945 (LOC110056945), transcript variant X2, mRNA	↓	2	comp46167_c0_seq4
comp46180_c0_seq3	A4IGD2	N-acetylaspartate synthetase	XM_020776004.1	N-acetyltransferase 8-like (LOC110068605), mRNA	↑	2	comp46180_c0_seq3
comp46389_c0_seq1	NA	NA	XM_020761628.1	Uncharacterized LOC110055247 (LOC110055247), transcript variant X1, mRNA	↑	12	comp46389_c0_seq1
comp46469_c1_seq3	Q96NU1	Sterile alpha motif domain-containing protein 11	XM_020770656.1	Inosine-5'-monophosphate dehydrogenase 1-like (LOC110063657), mRNA	↑	15	comp46469_c1_seq3
comp46490_c0_seq1	NA	NA	XM_020747242.1	Uncharacterized LOC110041916 (LOC110041916), transcript variant X1, mRNA	↑	7	comp46490_c0_seq1
comp46835_c0_seq9	Q03141	MAP/microtubule affinity-regulating kinase 3	XM_020762621.1	Serine/threonine-protein kinase MARK2-like (LOC110056174), partial mRNA	↓	17	comp46835_c0_seq9
comp66139_c0_seq1	P48031	Homeobox protein GBX-2	XM_020764401.1	BarH-like 2 homeobox protein (LOC110057804), transcript variant X2, mRNA	↓	1	comp66139_c0_seq1
comp7395_c0_seq1	P24507	Synaptotagmin-C	XM_020770265.1	Synaptotagmin-C-like (LOC110063297), mRNA	↑	4	comp7395_c0_seq1
comp93962_c0_seq1	NA	NA	XM_020753270.1	Meiosis-specific protein MEI4-like (LOC110047516), mRNA	↑	1	comp93962_c0_seq1
comp94877_c0_seq1	Q16534	Hepatic leukemia factor	XM_020748493.1	Thyrotroph embryonic factor-like (LOC110043079), transcript variant X2, mRNA	↓	4	comp94877_c0_seq1
comp42848_c0_seq1	NA	NA	NA	NA	↑	NA	NA
comp42497_c0_seq1	NA	NA	XM_020757372.1	Uncharacterized LOC110051338 (LOC110051338), mRNA	↑	NA	NA
comp48361_c0_seq1	P21956	Lactadherin	XM_020768975.1	Transmembrane protease serine 9-like (LOC110062106), transcript variant X2, mRNA	↓	NA	NA
comp44549_c0_seq1	Q0VCJ7	Ras-related and estrogen-regulated growth inhibitor	XM_020758527.1	Ras-related and estrogen-regulated growth inhibitor-like (LOC110052403), mRNA	↑	NA	NA
comp42732_c0_seq1	P29773	Protein C-ets-2	XM_020764572.1	ETS-related transcription factor Elf-4-like (LOC110057963), mRNA	↑	NA	NA
comp65176_c0_seq1	NA	NA	XM_020761343.1	Uncharacterized LOC110054962 (LOC110054962), mRNA	↑	NA	NA
comp46167_c0_seq3	Q55GQ5	Superoxide dismutase [Cu-Zn] 1	XM_020753493.1	Uncharacterized LOC110047748 (LOC110047748), mRNA	↓	NA	NA
comp42052_c0_seq1	NA	NA	XM_020766955.1	Protein HOS4-like (LOC110060199), transcript variant X2, mRNA	↑	NA	NA
comp10192_c0_seq1	NA	NA	XM_020745206.1	Cuticle collagen 1-like (LOC110040019), mRNA	↓	NA	NA
comp44121_c0_seq1	Q6GQ68	Protein phosphatase 1 regulatory subunit 3B-B	XM_020763439.1	Protein phosphatase 1 regulatory subunit 3B-B-like (LOC110056888), mRNA	↑	NA	NA
comp13473_c0_seq1	Q8K430	Kelch-like protein 17	XM_020768730.1	Kelch-like protein 12 (LOC110061871), mRNA	↑	NA	NA
comp39583_c0_seq1	Q9D119	Protein phosphatase 1 regulatory subunit 27	XM_020766864.1	Protein phosphatase 1 regulatory subunit 27-like (LOC110060114), mRNA	↑	NA	NA
comp39768_c0_seq2	Q8N2G6	Zinc finger CCHC domain-containing protein 24	XM_020753214.1	Zinc finger CCHC domain-containing protein 24-like (LOC110047464), mRNA	↓	NA	NA

"Adult" coral						Gamete bundles	
comp82023_c0_seq1	Q4PZA2	Endothelin-converting enzyme 1	XM_020761456.1	Endothelin-converting enzyme homolog (LOC110055076), mRNA	↓	NA	NA
comp46167_c0_seq1	Q55GQ5	Superoxide dismutase [Cu-Zn] 1	XM_020753493.1	Uncharacterized LOC110047748 (LOC110047748), mRNA	↓	NA	NA
comp62736_c0_seq1	Q9EPW2	Krueppel-like factor 15	XM_020750796.1	Krueppel-like factor 12 (LOC110045197), mRNA	↑	NA	NA
comp46401_c0_seq1	NA	NA	XM_020771276.1	Uncharacterized LOC110064251 (LOC110064251), transcript variant X7, mRNA	↑	NA	NA
comp45306_c0_seq1	Q9Y5X5	Neuropeptide FF receptor 2	XM_020746248.1	Neuropeptide FF receptor 2-like (LOC110040976), mRNA	↑	NA	NA
comp16260_c0_seq1	Q5ZIJ9	E3 ubiquitin-protein ligase MIB2	XM_020761206.1	Uncharacterized LOC110054848 (LOC110054848), mRNA	↑	NA	NA
comp45780_c0_seq3	NA	NA	XM_020756471.1	Uncharacterized LOC110050541 (LOC110050541), mRNA	↑	NA	NA
comp45417_c0_seq3	NA	NA	XM_020754812.1	Uncharacterized LOC110049045 (LOC110049045), mRNA	↑	NA	NA
comp34525_c0_seq1	Q76KB2	Heparan-sulfate 6-O-sulfotransferase 1	NA	NA	↑	NA	NA
comp140825_c0_seq1	NA	NA	XM_020761343.1	Uncharacterized LOC110054962 (LOC110054962), mRNA	↑	NA	NA
comp76040_c0_seq1	B2RVL6	Zinc finger CCHC domain-containing protein 24	XM_020753214.1	Zinc finger CCHC domain-containing protein 24-like (LOC110047464), mRNA	↓	NA	NA
comp6475_c0_seq1	A3RK75	Forkhead box protein O1-B	XM_020758681.1	Forkhead box protein O1-like (LOC110052551), mRNA	↑	NA	NA
comp16841_c0_seq1	P22544	Homeobox protein B-H1	XM_020758353.1	Homeobox protein Hox-C6a-like (LOC110052243), mRNA	↑	NA	NA
comp37768_c0_seq1	NA	NA	XM_020752458.1	Ski oncogene-like (LOC110046743), transcript variant X2, mRNA	↓	NA	NA
comp25167_c0_seq1	A4GT88	Type I iodothyronine deiodinase	XM_020768932.1	Type I iodothyronine deiodinase-like (LOC110062071), mRNA	↑	NA	NA
comp46185_c0_seq1	NA	NA	NA	NA	↑	NA	NA
comp62604_c0_seq1	Q14549	Homeobox protein GBX-1	XM_020764397.1	BarH-like 1 homeobox protein (LOC110057802), mRNA	↓	NA	NA
comp39558_c0_seq1	NA	NA	XM_020758567.1	Uncharacterized LOC110052434 (LOC110052434), mRNA	↑	NA	NA
comp23133_c0_seq1	NA	NA	NA	NA	↑	NA	NA
comp43048_c0_seq4	P47236	Paired box protein Pax-1	XM_020770259.1	Paired box protein Pax-3-B-like (LOC110063291), mRNA	↑	NA	NA
comp114965_c0_seq1	NA	NA	XM_020764397.1	BarH-like 1 homeobox protein (LOC110057802), mRNA	↓	NA	NA
comp84788_c0_seq1	Q5P5G2	Acetophenone carboxylase alpha subunit	XM_020762627.1	Uncharacterized LOC110056180 (LOC110056180), mRNA	↑	NA	NA

Appendix D

Lists of Differentially Expressed Genes in *Orbicella annularis* and *O. franksi*

Appendix Table 4-1: Differentially expressed genes during and after spawning with p adjusted<0.05 in *Orbicella franksi.* Genes highlighted in yellow are expressed to p_{adj}<0.001 and all other genes are DEG to p_{adj}<0.05. The arrows in the Expression tab indicate regulation (up or down regulation) in spawning respective to postspawning.

(Next page)

Sequence name	UNIPROT ID	UNIPROT Protein name	NCBI BLAST N ACCESSION #	NCBI BLAST N, first match	Expression
comp42848_c0_seq1	NA	NA	NA	NA	↑
comp44782_c0_seq2	Q03629	Uncharacterized protein YML079W	XM_020773262.1	Uncharacterized LOC110066062 (LOC110066062), mRNA	↑
comp42497_c0_seq1	NA	NA	XM_020757372.1	Uncharacterized LOC110051338 (LOC110051338), mRNA	↑
comp40885_c0_seq1	Q99581	Protein FEV	XM_020775929.1	ETS translocation variant 4-like (LOC110068544), mRNA	↑
comp41359_c0_seq1	Q9H0K1	Serine/threonine-protein kinase SIK2	XM_020765851.1	Serine/threonine-protein kinase SIK2-like (LOC110059158), transcript variant X4, mRNA	↑
comp33300_c0_seq1	Q6UB98	Ankyrin repeat domain-containing protein 12	XM_020763516.1	Cortactin-binding protein 2-like (LOC110056953), mRNA	↑
comp44783_c0_seq1	Q8K3X4	Probable E3 ubiquitin-protein ligase IRF2BPL	XM_020777109.1	Interferon regulatory factor 2-binding protein-like (LOC110069586), mRNA	↑
comp48361_c0_seq1	P21956	Lactadherin	XM_020768975.1	Transmembrane protease serine 9-like (LOC110062106), transcript variant X2, mRNA	↓
comp44549_c0_seq1	Q0VCJ7	Ras-related and estrogen-regulated growth inhibitor	XM_020758527.1	Ras-related and estrogen-regulated growth inhibitor-like (LOC110052403), mRNA	↑
comp43854_c0_seq1	Q76I25	HIG1 domain family member 1C	NA	NA	↑
comp39639_c0_seq1	NA	NA	XM_020772814.1	Uncharacterized LOC110065656 (LOC110065656), mRNA	↓
comp26181_c0_seq1	NA	NA	XM_020764214.1	Transcription factor VBP-like (LOC110057617), mRNA	↓
comp44807_c0_seq1	NA	NA	XM_020774952.1	Melanocortin receptor 5-like (LOC110067607), mRNA	↑
comp43708_c0_seq1	A1Z6E0	Protein gustavus	XM_020768301.1	Protein gustavus-like (LOC110061465), mRNA	↑
comp42732_c0_seq1	P29773	Protein C-ets-2	XM_020764572.1	ETS-related transcription factor Elf-4-like (LOC110057963), mRNA	↑
comp65176_c0_seq1	NA	NA	XM_020761343.1	Uncharacterized LOC110054962 (LOC110054962), mRNA	↑
comp46167_c0_seq3	Q55GQ5	Superoxide dismutase [Cu-Zn] 1	XM_020753493.1	Uncharacterized LOC110047748 (LOC110047748), mRNA	↓
comp42052_c0_seq1	NA	NA	XM_020766955.1	Protein HOS4-like (LOC110060199), transcript variant X2, mRNA	↑
comp24246_c0_seq2	P50616	Protein Tob1	XM_020770942.1	Protein Tob2-like (LOC110063947), mRNA	↑
comp45019_c0_seq1	P79145	cAMP-responsive element modulator	XM_020752048.1	Cyclic AMP-responsive element-binding protein 1-like (LOC110046357), mRNA	↑
comp45865_c0_seq19	A3KN95	Transmembrane protein 151B	XM_020771394.1	Uncharacterized LOC110064353 (LOC110064353), mRNA	↑
comp38798_c0_seq1	P34743	Protein BTG1	XM_020772618.1	Protein BTG2-like (LOC110065472), mRNA	↑
comp42612_c0_seq1	NA	NA	XM_020767223.1	Uncharacterized LOC110060441 (LOC110060441), transcript variant X1, mRNA	↑
comp30839_c0_seq1	NA	NA	XM_020777062.1	Neurogenic locus notch homolog protein 1-like (LOC110069530), mRNA	↑
comp46167_c0_seq4	Q07496	Ephrin type-A receptor 4	XM_020763507.1	Uncharacterized LOC110056945 (LOC110056945), transcript variant X2, mRNA	↓

Sequence name	UNIPROT ID	UNIPROT Protein name	NCBI BLAST N ACCESSION #	NCBI BLAST N, first match	Expression
comp10192_c0_seq1	NA	NA	XM_020745206.1	Cuticle collagen 1-like (LOC110040019), mRNA	↓
comp44121_c0_seq1	Q6GQ68	Protein phosphatase 1 regulatory subunit 3B-B	XM_020763439.1	Protein phosphatase 1 regulatory subunit 3B-B-like (LOC110056888), mRNA	↑
comp44604_c0_seq1	Q5FVC7	Arf-GAP with coiled-coil, ANK repeat and PH domain-containing protein 2	XM_020751023.1	Arf-GAP with coiled-coil, ANK repeat and PH domain-containing protein 2-like (LOC110045421), mRNA	↑
comp43514_c0_seq1	O42342	Transcription factor Sox-7	XM_020764830.1	Transcription factor Sox-14-like (LOC110058232), mRNA	↑
comp44971_c0_seq1	P58545	BTB/POZ domain-containing protein 3	XM_020764957.1	BTB/POZ domain-containing protein 3-like (LOC110058317), mRNA	↑
comp37801_c0_seq1	Q6NTY6	Early growth response protein 1-B	XM_020751274.1	Early growth response protein 1-like (LOC110045671), transcript variant X1, mRNA	↑
comp15044_c0_seq1	NA	NA	XM_020772305.1	Bifunctional TH2 protein, mitochondrial-like (LOC110065196), mRNA	↓
comp43947_c0_seq1	Q9R0S2	Matrix metalloproteinase-24	XM_020762836.1	72 kDa type IV collagenase-like (LOC110056368), mRNA	↑
comp44988_c0_seq1	Q92626	Peroxidasin homolog	XM_020749475.1	Papilin-like (LOC110043970), mRNA	↑
comp13473_c0_seq1	Q8K430	Kelch-like protein 17	XM_020768730.1	Kelch-like protein 12 (LOC110061871), mRNA	↑
comp44417_c0_seq1	NA	NA	XM_020774855.1	Uncharacterized LOC110067515 (LOC110067515), mRNA	↑
comp39583_c0_seq1	Q9D119	Protein phosphatase 1 regulatory subunit 27	XM_020766864.1	Protein phosphatase 1 regulatory subunit 27-like (LOC110060114), mRNA	↑
comp46389_c0_seq1	NA	NA	XM_020761628.1	Uncharacterized LOC110055247 (LOC110055247), transcript variant X1, mRNA	↑
comp39768_c0_seq2	Q8N2G6	Zinc finger CCHC domain-containing protein 24	XM_020753214.1	Zinc finger CCHC domain-containing protein 24-like (LOC110047464), mRNA	↓
comp46490_c0_seq1	NA	NA	XM_020747242.1	Uncharacterized LOC110041916 (LOC110041916), transcript variant X1, mRNA	↑
comp45686_c0_seq1	P28562	Dual specificity protein phosphatase 1	XM_020771923.1	Dual specificity protein phosphatase 1-A-like (LOC110064824), mRNA	↑
comp45038_c0_seq1	Q6PJ21	SPRY domain-containing SOCS box protein 3	XM_020772297.1	SPRY domain-containing SOCS box protein 3-like (LOC110065191), transcript variant X1, mRNA	↑
comp43741_c0_seq1	Q9D0B5	Thiosulfate sulfurtransferase/rhodanese-like domain-containing protein 3	XM_020756079.1	Thiosulfate sulfurtransferase/rhodanese-like domain-containing protein 1 (LOC110050189), mRNA	↑
comp94877_c0_seq1	Q16534	Hepatic leukemia factor	XM_020748493.1	Thyrotroph embryonic factor-like (LOC110043079), transcript variant X2, mRNA	↓
comp44600_c0_seq1	NA	NA	XM_020752518.1	Uncharacterized LOC110046805 (LOC110046805), mRNA	↑
comp82023_c0_seq1	Q4PZA2	Endothelin-converting enzyme 1	XM_020761456.1	Endothelin-converting enzyme homolog (LOC110055076), mRNA	↓
comp46167_c0_seq1	Q55GQ5	Superoxide dismutase [Cu-Zn] 1	XM_020753493.1	Uncharacterized LOC110047748 (LOC110047748), mRNA	↓
comp66139_c0_seq1	P48031	Homeobox protein GBX-2	XM_020764401.1	BarH-like 2 homeobox protein (LOC110057804), transcript variant X2, mRNA	↓
comp40023_c0_seq1	NA	NA	NA	NA	↑
comp41407_c0_seq1	NA	NA	XM_020752427.1	Uncharacterized LOC110046733 (LOC110046733), mRNA	↑
comp39253_c0_seq1	NA	NA	XM_020765268.1	Roundabout homolog 2-like (LOC110058616), partial mRNA	↓

Sequence name	UNIPROT ID	UNIPROT Protein name	NCBI BLAST N ACCESSION #	NCBI BLAST N, first match	Expression
comp39020_c0_seq1	NA	NA	XR_002296924.1	Uncharacterized LOC110051260 (LOC110051260), ncRNA	↓
comp93962_c0_seq1	NA	NA	XM_020753270.1	Meiosis-specific protein MEI4-like (LOC110047516), mRNA	↑
comp43854_c0_seq2	Q76I25	HIG1 domain family member 1C	NA	NA	↑
comp62736_c0_seq1	Q9EPW2	Krueppel-like factor 15	XM_020750796.1	Krueppel-like factor 12 (LOC110045197), mRNA	↑
comp36781_c0_seq1	P15105	Glutamine synthetase	XM_020752370.1	Glutamine synthetase-like (LOC110046681), mRNA	↑
comp43024_c0_seq1	NA	NA	XM_020761342.1	Uncharacterized LOC110054961 (LOC110054961), mRNA	↑
comp20193_c0_seq1	NA	NA	XM_020761342.1	Bifunctional TH2 protein, mitochondrial-like (LOC110065196), mRNA	↓
comp44095_c0_seq1	Q91827	Apoptosis regulator R1	XM_020746225.1	Apoptosis regulator R1-like (LOC110040960), mRNA	↑
comp45334_c1_seq2	P07152	Stromelysin-2	XM_020762835.1	Stromelysin-1-like (LOC110056367), mRNA	↑
comp42602_c0_seq1	Q7KM13	Hairy/enhancer-of-split related with YRPW motif protein	XM_020760876.1	Transcription factor cwo-like (LOC110054546), transcript variant X1, mRNA	↑
comp46401_c0_seq1	NA	NA	XM_020771276.1	Uncharacterized LOC110064251 (LOC110064251), transcript variant X7, mRNA	↑
comp45306_c0_seq1	Q9Y5X5	Neuropeptide FF receptor 2	XM_020746248.1	Neuropeptide FF receptor 2-like (LOC110040976), mRNA	↑
comp16260_c0_seq1	Q5ZIJ9	E3 ubiquitin-protein ligase MIB2	XM_020761206.1	Uncharacterized LOC110054848 (LOC110054848), mRNA	↑
comp44578_c0_seq5	NA	NA	XM_020751975.1	Uncharacterized LOC110046287 (LOC110046287), mRNA	↑
comp45940_c0_seq1	NA	NA	XM_020752427.1	Uncharacterized LOC110046733 (LOC110046733), mRNA	↑
comp45780_c0_seq3	NA	NA	XM_020756471.1	Uncharacterized LOC110050541 (LOC110050541), mRNA	↑
comp44919_c0_seq2	NA	NA	NA	NA	↑
comp42811_c0_seq2	Q9Z2H5	Band 4.1-like protein 1	XM_020768122.1	Band 4.1-like protein 1 (LOC110061280), mRNA	↑
comp45417_c0_seq3	NA	NA	XM_020754812.1	Uncharacterized LOC110049045 (LOC110049045), mRNA	↑
comp38079_c0_seq1	Q4A3R3	Deleted in malignant brain tumors 1 protein	XR_002295978.1	Uncharacterized LOC110043178 (LOC110043178), transcript variant X2, ncRNA	↓
comp38539_c0_seq1	Q9VVY3	Glycogen-binding subunit 76A	XM_020775678.1	Glycogen-binding subunit 76A-like (LOC110068299), mRNA	↑
comp46111_c0_seq1	Q923Q2	StAR-related lipid transfer protein 13	XM_020745441.1	StAR-related lipid transfer protein 13-like (LOC110040233), transcript variant X1, mRNA	↑
comp19081_c0_seq1	NA	NA	XM_020763974.1	Monocarboxylate transporter 10-like (LOC110057384), transcript variant X2, mRNA	↓
comp34525_c0_seq1	Q76KB2	Heparan-sulfate 6-O-sulfotransferase 1	NA	NA	↑
comp44141_c0_seq3	Q3MHQ7	Lysoplasmalogenase-like protein TMEM86A	XM_020768719.1	Lysoplasmalogenase-like protein TMEM86A (LOC110061859), mRNA	↑

Sequence name	UNIPROT ID	UNIPROT Protein name	NCBI BLAST N ACCESSION #	NCBI BLAST N, first match	Expression
comp140825_c0_seq1	NA	NA	XM_020761343.1	Uncharacterized LOC110054962 (LOC110054962), mRNA	↑
comp10712_c0_seq1	Q9U3S9	Zinc metalloproteinase nas-6	XM_020747405.1	Zinc metalloproteinase nas-4-like (LOC110042061), transcript variant X2, mRNA	↑
comp45031_c0_seq1	NA	NA	NA	NA	↓
comp76040_c0_seq1	B2RVL6	Zinc finger CCHC domain-containing protein 24	XM_020753214.1	Zinc finger CCHC domain-containing protein 24-like (LOC110047464), mRNA	↓
comp45683_c0_seq1	Q3UFK8	FERM domain-containing protein 8	XM_020752472.1	Putative FERM domain-containing protein FRMD8P1 (LOC110046765), mRNA	↑
comp44117_c0_seq1	Q9EP86	Neuropeptide FF receptor 1	XM_020748456.1	Neuropeptide FF receptor 1-like (LOC110043018), transcript variant X1, mRNA	↑
comp6475_c0_seq1	A3RK75	Forkhead box protein O1-B	XM_020758681.1	Forkhead box protein O1-like (LOC110052551), mRNA	↑
comp43295_c0_seq1	Q9NPA2	Matrix metalloproteinase-25	XM_020762837.1	Matrix metalloproteinase-25-like (LOC110056369), mRNA	↑
comp46180_c0_seq3	A4IGD2	N-acetylaspartate synthetase	XM_020776004.1	N-acetyltransferase 8-like (LOC110068605), mRNA	↑
comp46469_c1_seq3	Q96NU1	Sterile alpha motif domain-containing protein 11	XM_020770656.1	Inosine-5'-monophosphate dehydrogenase 1-like (LOC110063657), mRNA	↑
comp44850_c0_seq1	Q9Y5Z4	Heme-binding protein 2	XM_020771695.1	Heme-binding protein 2-like (LOC110064621), transcript variant X2, mRNA	↓
comp44582_c0_seq1	Q80ZQ5	Juxtaposed with another zinc finger protein 1	XM_020745898.1	Juxtaposed with another zinc finger protein 1-like (LOC110040630), mRNA	↑
comp39919_c0_seq4	P26652	Metalloproteinase inhibitor 3	XM_020752328.1	Metalloproteinase inhibitor 3-like (LOC110046634), mRNA	↑
comp44443_c0_seq1	Q0KIA2	PP2C-like domain-containing protein CG9801	XM_020751862.1	PP2C-like domain-containing protein CG9801 (LOC110046180), transcript variant X1, mRNA	↑
comp24246_c0_seq1	P50616	Protein Tob1	XM_020770942.1	Protein Tob2-like (LOC110063947), mRNA	↑
comp16841_c0_seq1	P22544	Homeobox protein B-H1	XM_020758353.1	Homeobox protein Hox-C6a-like (LOC110052243), mRNA	↑
comp43619_c0_seq1	P0C6B8	Sushi, von Willebrand factor type A, EGF and pentraxin domain-containing protein 1	XM_020755666.1	Uncharacterized LOC110049839 (LOC110049839), mRNA	↓
comp37768_c0_seq1	NA	NA	XM_020752458.1	Ski oncogene-like (LOC110046743), transcript variant X2, mRNA	↓
comp43105_c0_seq1	NA	NA	XM_020771973.1	Uncharacterized LOC110064869 (LOC110064869), mRNA	↑
comp25167_c0_seq1	A4GT88	Type I iodothyronine deiodinase	XM_020768932.1	Type I iodothyronine deiodinase-like (LOC110062071), mRNA	↑
comp46185_c0_seq1	NA	NA	NA	NA	↑
comp62604_c0_seq1	Q14549	Homeobox protein GBX-1	XM_020764397.1	BarH-like 1 homeobox protein (LOC110057802), mRNA	↓
comp39558_c0_seq1	NA	NA	XM_020758567.1	Uncharacterized LOC110052434 (LOC110052434), mRNA	↑
comp40481_c0_seq2	NA	NA	XM_020758728.1	Uncharacterized LOC110052587 (LOC110052587), mRNA	↓
comp45567_c0_seq1	Q9CYL5	Golgi-associated plant pathogenesis-related protein 1	XM_020763535.1	Golgi-associated plant pathogenesis-related protein 1-like (LOC110056967), transcript variant X2, mRNA	↑

Sequence name	UNIPROT ID	UNIPROT Protein name	NCBI BLAST N ACCESSION #	NCBI BLAST N, first match	Expression
comp43343_c0_seq1	Q99542	Matrix metalloproteinase-19	XM_020762821.1	Matrix metalloproteinase-24-like (LOC110056355), mRNA	↑
comp7395_c0_seq1	P24507	Synaptotagmin-C	XM_020770265.1	Synaptotagmin-C-like (LOC110063297), mRNA	↑
comp43061_c0_seq1	NA	NA	XR_002297290.1	Uncharacterized LOC110055117 (LOC110055117), ncRNA	↑
comp11020_c0_seq1	Q07352	mRNA decay activator protein ZFP36L1	XM_020748063.1	mRNA decay activator protein ZFP36L1-like (LOC110042696), mRNA	↑
comp39680_c0_seq2	Q91VS7	Microsomal glutathione S-transferase 1	XM_020775453.1	Microsomal glutathione S-transferase 1-like (LOC110068084), mRNA	↑
comp38792_c0_seq1	O35738	Krueppel-like factor 12	XM_020769222.1	Krueppel-like factor 6 (LOC110062347), mRNA	↓
comp23133_c0_seq1	NA	NA	NA	NA	↑
comp41533_c0_seq1	O15118	NPC intracellular cholesterol transporter 1	XM_020770138.1	Patched domain-containing protein 3-like (LOC110063182), mRNA	↓
comp46835_c0_seq9	Q03141	MAP/microtubule affinity-regulating kinase 3	XM_020762621.1	Serine/threonine-protein kinase MARK2-like (LOC110056174), partial mRNA	↓
comp43048_c0_seq4	P47236	Paired box protein Pax-1	XM_020770259.1	Paired box protein Pax-3-B-like (LOC110063291), mRNA	↑
comp38511_c0_seq2	Q08CS6	DBH-like monooxygenase protein 2 homolog	XM_020770343.1	Uncharacterized LOC110063370 (LOC110063370), mRNA	↓
comp40188_c0_seq1	NA	NA	XM_020749343.1	Uncharacterized LOC110043845 (LOC110043845), transcript variant X1, mRNA	↑
comp45097_c0_seq1	NA	NA	XM_020768661.1	Uncharacterized LOC110061806 (LOC110061806), transcript variant X1, mRNA	↑
comp114965_c0_seq1	NA	NA	XM_020764397.1	BarH-like 1 homeobox protein (LOC110057802), mRNA	↓
comp84788_c0_seq1	Q5P5G2	Acetophenone carboxylase alpha subunit	XM_020762627.1	Uncharacterized LOC110056180 (LOC110056180), mRNA	↑
comp114367_c0_seq1	NA	NA	XM_020775084.1	Protein FAM124A-like (LOC110067759), transcript variant X4, mRNA	↑

Appendix Table 4-2: Differentially expressed genes during and after spawning with p adjusted<0.05 in *Orbicella franksi.* Genes highlighted in yellow are expressed to p_{adj}<0.001 and all other genes are DEG to p_{adj}<0.05. The arrows in the Expression tab indicate regulation (up or down regulation) in spawning respective to postspawning.

(Next page)

Sequence name	UNIPROT ID	UNIPROT Protein name	NCBI BLAST N ACCESSION NUMBER	NCBI BLAST N, first match	Expression
comp46445_c0_seq1	NA	NA	XM_020760887.1	Zinc finger MIZ domain-containing protein 1-like (LOC110054549), transcript variant X8, mRNA	↑
comp44862_c0_seq2	Q6DRG7	Protein phosphatase 1 regulatory subunit 12A	XM_020746303.1	Protein phosphatase 1 regulatory subunit 12A-like (LOC110041021), transcript variant X3, mRNA	↑
comp42538_c0_seq3	Q9P2F6	Rho GTPase-activating protein 20	op	Uncharacterized LOC110060432 (LOC110060432), transcript variant X4, mRNA	↑
comp73282_c0_seq1	NA	NA	XM_020751859.1	L-gulonolactone oxidase-like (LOC110046178), mRNA	↑
comp72615_c0_seq1	NA	NA	XM_020768129.1	Homeobox protein zampogna-like (LOC110061283), mRNA	↑
comp46412_c0_seq2	NA	NA	XM_020756106.1	Sporulation-specific protein 15-like (LOC110050217), transcript variant X2, mRNA	↑
comp45780_c0_seq1	NA	NA	XM_020756471.1	Uncharacterized LOC110050541 (LOC110050541), mRNA	↑
comp46232_c0_seq1	NA	NA	XM_020756253.1	Non-specific lipid-transfer protein-like (LOC110050336), mRNA	↑
comp46502_c0_seq5	NA	NA	XM_020748922.1	Nuclear factor of activated T-cells 5-like (LOC110043473), transcript variant X4, mRNA	↑
comp32275_c0_seq1	NA	NA	XM_020768302.1	Uncharacterized LOC110061466 (LOC110061466), mRNA	↑
comp45563_c0_seq1	Q9UKN7	Unconventional myosin-XV	XM_020755641.1	Uncharacterized LOC110049813 (LOC110049813), mRNA	↑
comp62736_c0_seq1	Q9EPW2	Krueppel-like factor 15	XM_020750796.1	Krueppel-like factor 12 (LOC110045197), mRNA	↑
comp41913_c0_seq1	NA	NA	XM_020760688.1	Uncharacterized LOC110054306 (LOC110054306), mRNA	↑
comp40805_c0_seq2	Q9SZW4	Cadmium/zinc-transporting ATPase HMA2	XM_020748322.1	Cadmium/zinc-transporting ATPase HMA3-like (LOC110042932), transcript variant X5, mRNA	↑
comp45882_c0_seq3	P25291	Pancreatic secretory granule membrane major glycoprotein GP2	XM_020767946.1	Hemicentin-2-like (LOC110061112), mRNA	↑
comp46601_c0_seq3	NA	NA	XM_020771679.1	Pappalysin-1-like (LOC110064612), transcript variant X3, mRNA	↑
comp38143_c0_seq1	P97812	Indian hedgehog protein	XM_020776357.1	Indian hedgehog protein-like (LOC110068937), mRNA	↑
comp37669_c0_seq1	NA	NA	XR_002298673.1	Methenyltetrahydrofolate synthase domain-containing protein-like (LOC110067398), transcript variant X2, misc_RNA	↑
comp36899_c0_seq1	NA	NA	NA	NA	↑
comp44802_c0_seq1	P52962	Moesin	XM_020767196.1	Radixin-like (LOC110060420), mRNA	↑
comp66139_c0_seq1	P48031	Homeobox protein GBX-2	XM_020764401.1	BarH-like 2 homeobox protein (LOC110057804), transcript variant X2, mRNA	↓
comp35172_c0_seq1	A6NMZ7	Collagen alpha-6(VI) chain	XM_020755960.1	Uncharacterized LOC110050092 (LOC110050092), mRNA	↓
comp76766_c0_seq1	P55013	Solute carrier family 12 member 2	XM_020769451.1	Solute carrier family 12 member 2-like (LOC110062523), mRNA	↓
comp52925_c0_seq1	NA	NA	XM_020750744.1	T-box transcription factor TBX20-like (LOC110045149), transcript variant X2, mRNA	↓
comp23491_c0_seq1	NA	NA	XM_020768572.1	Uncharacterized LOC110061726 (LOC110061726), mRNA	↓

Sequence name	UNIPROT ID	UNIPROT Protein name	NCBI BLAST N ACCESSION NUMBER	NCBI BLAST N, first match	Expression
comp41533_c0_seq1	O15118	NPC intracellular cholesterol transporter 1	XM_020770138.1	Patched domain-containing protein 3-like (LOC110063182), mRNA	↓
comp71224_c0_seq1	Q9BDJ5	Pantetheinase	NA	NA	↓
comp76040_c0_seq1	B2RVL6	Zinc finger CCHC domain-containing protein 24	XM_020753214.1	Zinc finger CCHC domain-containing protein 24-like (LOC110047464), mRNA	↓
comp35986_c0_seq1	NA	NA	XM_020770138.1	Patched domain-containing protein 3-like (LOC110063182), mRNA	↓
comp30123_c0_seq2	A6H584	Collagen alpha-5(VI) chain	XM_020755969.1	Uncharacterized LOC110050097 (LOC110050097), transcript variant X3, mRNA	↓
comp118140_c0_seq1	P70551	Type II iodothyronine deiodinase	XM_020747166.1	Thyroxine 5-deiodinase-like (LOC110041843), mRNA	↓
comp23252_c0_seq1	P54417	Glycine betaine transporter OpuD	XM_020769058.1	Glycine betaine transporter OpuD-like (LOC110062187), mRNA	↓
comp79043_c0_seq1	NA	NA	XM_020769451.1	Solute carrier family 12 member 2-like (LOC110062523), mRNA	↓
comp48742_c0_seq1	NA	NA	XM_020768753.1	Uncharacterized threonine-rich GPI-anchored glycoprotein PJ4664.02-like (LOC110061891), partial mRNA	↓
comp44570_c0_seq1	NA	NA	XM_020769444.1	Uncharacterized LOC110062517 (LOC110062517), transcript variant X1, mRNA	↓
comp45974_c0_seq2	O95084	Serine protease 23	XM_020753854.1	Serine protease 23-like (LOC110048082), transcript variant X1, mRNA	↓
comp33476_c0_seq1	NA	NA	XM_020773127.1	Uncharacterized LOC110065941 (LOC110065941), mRNA	↓
comp48646_c0_seq1	P45335	Uncharacterized transporter HI_1706	XM_020769058.1	Glycine betaine transporter OpuD-like (LOC110062187), mRNA	↓
comp44721_c0_seq1	Q9HCJ5	Zinc finger SWIM domain-containing protein 6	XM_020769053.1	Zinc finger SWIM domain-containing protein 6-like (LOC110062186), transcript variant X2, mRNA	↓
comp41596_c0_seq1	NA	NA	XM_020760507.1	Uncharacterized LOC110054171 (LOC110054171), mRNA	↓
comp37758_c0_seq1	Q6DCQ6	von Willebrand factor A domain-containing protein 2	XM_020776739.1	Integrin alpha-D-like (LOC110069246), mRNA	↓
comp41984_c0_seq2	Q02410	Amyloid-beta A4 precursor protein-binding family A member 1	XM_020771129.1	Dentin sialophosphoprotein-like (LOC110064130), mRNA	↓
comp45562_c0_seq2	Q76LC6	RNA-binding protein 24	XM_020761791.1	RNA-binding protein 24-A-like (LOC110055404), mRNA	↓
comp62031_c0_seq1	NA	NA	XM_020763974.1	Monocarboxylate transporter 10-like (LOC110057384), transcript variant X2, mRNA	↓
comp45366_c0_seq1	NA	NA	XM_020773127.1	Uncharacterized LOC110065941 (LOC110065941), mRNA	↓
comp43934_c0_seq2	Q803Z2	Protein YIPF3	XM_020746535.1	Protein YIPF3-like (LOC110041252), mRNA	↓
comp11842_c0_seq2	NA	NA	XM_020755974.1	Uncharacterized LOC110050097 (LOC110050097), transcript variant X7, mRNA	↓
comp43463_c0_seq1	NA	NA	XM_020750644.1	Uncharacterized LOC110045051 (LOC110045051), mRNA	↓

VITA
Ana María González Ángel

Address: 1000 W Aaron drive, apt. H9, State College, PA, 16803, USA
Telephone: (+1) 814-880-3529
E-mail: anagonzangel@gmail.com , amg449@psu.edu

EDUCATION

Doctor of Philosophy
The Pennsylvania State University, State College, PA, USA
Biology, 2019
Genomic And Transcriptomic Insights On The Orbicella Species Complex

Master of Science
Nova Southeastern University, Fort Lauderdale, FL, USA
Marine Biology, 2013
"*Porites astreoides* larval response to acute salinity stress"

Bachelor of Science
Universidad de los Andes, Bogotá, Colombia
Biology, 2009
"Evaluación de la red colonial de *Pacifigorgia* spp. (Octocorallia: Cnidaria) mediante Análisis de Fourier"

RESEARCH INTERESTS

I am interested in coral reproduction, ecology and evolution. During my PhD I have focused on speciation in the Caribbean corals *Orbicella* spp. I have studied the reproductive isolation mechanisms that prevent these species from interbreeding. I am also interested in science communication and education.

PUBLICATIONS

González, A.M., Prada C.A., Ávila V., Medina M. (2018) Ecological Speciation in Corals. In: . Population Genomics. Springer, Cham

González, A. M., Sebastian, A., Prada, C., Levy, O., Levitan, D., Knowlton, N., Albert, I., Medina, M. (2019). Reproducing with your kind: the story of temporal prezygotic isolation in *Orbicella* corals. *In prep.*

González, A. M., Medina M. (2020) Gamete recognition proteins in *Orbicella* sister species. *In prep.*

González, A. M., Medina M. (2019) Gene ontology enrichment in *Orbicella* corals. *In prep.*.

TEACHING

Teaching Assistant for BIOL 110 Biology: Basic Concepts and Biodiversity, Department of Biology, The Pennsylvania State University. Falls of 2015, 2017, 2018, spring and fall of 2019

Teaching Assistant for Tropical Field Biology (499), s Department of Biology, The Pennsylvania State University. Springs 2014, 2015, 2016.

ctor of the miniMOOC (Massive Open Online Course) Earth and Atmospheric Sciences course entitled "Ciencias atmósféricas y de la tierra: energía en los océanos vivos" hosted by Clubes de Ciencia, Mexico 2017, 2018, 2019 (Visit: https://www.youtube.com/watch?v=B4ll4HGRnXk)

Instructor of the science club course "Descubriendo el mar y sus tesoros: ecología y evolución" hosted by Clubes de Ciencia, Colombia, 2017, Riohacha, Colombia 26-30 June 2017.