Evolutionary Genetics of *Pocillopora* Corals

by

David J. Combosch

Diploma Environmental Scientist, University Duisburg-Essen

A dissertation submitted to the
Faculty of the College of Science of Northeastern University
In partial fulfillment of the requirements for the degree of

Doctor of Philosophy

December 2, 2013

Dissertation directed by
Steven V. Vollmer
Associate Professor of Marine & Environmental Sciences
DEDICATION

Für Walter Combosch
I am grateful to my advisor Steven Vollmer for his support, guidance and friendship. I thank all my committee members, Rebeca Rosengaus, Matthew Bracken, Geoff Trussell and especially Gonzalo Giribet for their support and advice and my German mentor Helmut Schuhmacher for many years of support.

In addition, I want to thank many colleagues and friends from the Vollmer Lab, especially Silvia Libro, Liz Hemond, Sarah Gignoux-Wolfsohn, Chris Marks and Stefan Kaluziak as well as the Giribet Lab, especially Vanessa Gonzales and Alicia Perez-Porro and my collaborators Lauren Toth, Rich Aronson and Ian McIntyre. I benefitted greatly from the Three Sea Program, particularly due to the help of Sal Genovese, Pete Edmunds, Ruth Gates, Brian Helmuth, Josh Idjahdi, Jim Leichter and Bill Precht.

Numerous friends in America, Germany, Panamá, France, French Polynesia and especially in 123 River Street contributed a lot, both directly and indirectly, to this dissertation, especially my girlfriend Sarah Lemer, my brother Florian and his family Fiona, Liva, Clara und Bruno as well as my parents Lene and Walter Combosch.

This work was supported by a German Academic Exchange Service Pre-doctoral fellowship, Teaching assistantship awards from Northeastern University and the Three Seas Program and the National Science Foundation to Steven Vollmer.
ABSTRACT OF DISSERTATION

My dissertation research addresses fundamental questions about the evolution of tropical reef corals in isolated biogeographic regions. The main focus is on the effect of intra- and interspecific gene flow within and among peripheral coral populations. Since introgression and genetic connectivity are keys factor for evolution and adaptation as well as for population resilience, recovery and persistence, the identification and quantification of past and present gene flow is of major interest, not only for basic evolutionary research but also for applied conservation.

The central question of my research is the how the widespread, but generally subordinate reef-building coral *Pocillopora damicornis* became the ecologically dominant reef-builder in the isolated, environmentally challenging Tropical Eastern Pacific (TEP). The dominance of TEP *Pocillopora damicornis* populations is accompanied by strong differences in life-history traits, including a shift in reproductive strategy from brooding predominantly parthenogenetic larvae in the Indo-West Pacific to broadcast spawning in the TEP.

Using a combination of genetic tools ranging from multi-locus Sanger sequencing to RAD-Seq phylogenomic analyses, I show that three genetic lineages that are found in TEP and Central Pacific *P. damicornis* populations are characterized by strong genome-wide differences, which suggests that they belong to separate species. However, a subset of loci was shared among *P. damicornis* colonies across lineages and across large geographic distances. The allelic composition of putative hybrid samples indicate repeated uniparental backcrossing and substantial genomic introgression among TEP *Pocillopora* species.

The population genetic structure of TEP *P. damicornis* indicates limited gene flow among and within populations, which enables effective adaptations to the environmentally
heterogeneous TEP environment. Predominantly sexual reproduction, significant inbreeding and small-scale spatial genetic structures over meter scales within populations further sustain divergent adaptations. This small-scale population genetic structure allows *P. damicornis* to adapt locally to the harsh environmental condition of the Tropical Eastern Pacific and presumably facilitated its regional dominance.

In addition, I show that Central Pacific *P. damicornis* display a unique mixed reproductive strategy that includes the simultaneous production of sexual and asexual larvae. While the majority of larvae are produced parthenogenetically, most colonies also produce a subset of their larvae sexually. The proportion of sexually produced larvae decreases with colony size, suggesting that the mixed reproductive strategy is flexible and changes across the life of the coral, leading to an increasingly asexual propagation of locally successful genotypes.

Combined, I have shown that *Pocillopora* corals in the peripheral TEP are characterized by interspecific gene flow between *P. damicornis* and its congeners *P. eydouxi* and *P. elegans* and limited gene flow among and within *P. damicornis* populations. I showed that the morphospecies *Pocillopora damicornis* is composed of multiple genetic lineages, in part due to introgressive hybridization in the TEP. *Pocillopora* populations further differ in their reproductive strategies with central Pacific populations reproducing predominantly parthenogenetic and TEP populations reproducing mostly sexually.
# TABLE OF CONTENTS

**DEDICATION**  
II

**ACKNOWLEDGEMENTS**  
III

**ABSTRACT OF DISSERTATION**  
IV

**TABLE OF CONTENTS**  
VI

**LIST OF FIGURES**  
IX

**LIST OF TABLES**  
XI

**CHAPTER 1 - GENE FLOW, *POCILLOPORA* CORALS AND THE TROPICAL EASTERN PACIFIC**  
1  
Overview  
1  
Gene Flow  
2  
*Pocillopora* corals  
7  
The Tropical Eastern Pacific  
10  
References  
14

**CHAPTER 2 - POPULATION GENETICS OF AN ECOSYSTEM-DEFINING REEF CORAL - *POCILLOPORA DAMICORNIS* IN THE TROPICAL EASTERN PACIFIC**  
25  
Abstract  
25  
Introduction  
26  
Materials and Methods  
30  
Results  
35  
Discussion  
38
<table>
<thead>
<tr>
<th>Section</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>Conclusions</td>
<td>125</td>
</tr>
<tr>
<td>Acknowledgements</td>
<td>126</td>
</tr>
<tr>
<td>References</td>
<td>127</td>
</tr>
<tr>
<td>Tables</td>
<td>133</td>
</tr>
<tr>
<td>Figures</td>
<td>136</td>
</tr>
<tr>
<td>Supplementary Material</td>
<td>143</td>
</tr>
<tr>
<td><strong>FINAL CONCLUSIONS</strong></td>
<td>149</td>
</tr>
</tbody>
</table>
LIST OF FIGURES

CHAPTER 2

FIGURE 1: MAP OF PANAMÁ, SHOWING THE THREE MAIN PACIFIC REGIONS AND THE SAMPLED LOCATIONS.  67

FIGURE 2: PRINCIPAL COMPONENT ANALYSIS OF POCILLOPORA DAMICORNIS POPULATIONS, CONSTRUCTED USING GENALEX 6.1.  68

FIGURE 3: SPATIAL CORRELOGRAMS OF THE SPAGEDI ANALYSIS OF SPATIAL GENETIC STRUCTURE (SGS) AMONG SPECIMEN WITHIN DISCRETE DISTANCE CLASSES.  69

FIGURE S1: RESULTS OF THE BAYESIAN CLUSTERING APPROACH IMPLEMENTED IN STRUCTURE 2.3 THAT WERE USED TO INFER THE NUMBER OF GENETIC CLUSTERS (K) IN THE MICROSATELLITE DATASET  70

CHAPTER 3

FIGURE 1: A POCILLOPORA DAMICORNIS COLONY ON THE REEF IN FRONT OF THE GUMP RESEARCH STATION IN MOOREA, FRENCH POLYNESIA  96

FIGURE 2: THE PROPORTION OF SEXUALLY PRODUCED LARVAE DECREASES AS A FUNCTION OF COLONY SIZE  97

FIGURE 3: THE PROPORTION OF SEXUALLY PRODUCED LARVAE RELEASED DECREASES DURING THE REPRODUCTIVE CYCLES OF P. DAMICORNIS COLONIES  99

CHAPTER 4

FIGURE 1A: POCILLOPORA ELEGANS COLONY IN A P. DAMICORNIS CARPET IN ISLA COIBA, GULF OF CHIRIQUI, PANAMÁ.  136

FIGURE 1B: PHENOTYPIC PLASTICITY OF P. DAMICORNIS/P. VERRUCOSA IN MOOREA, FRENCH POLYNESIA.  137
FIGURE 2: ORF AND ITS PHYLOGENY

FIGURE 3: RAD PHYLOGENOMIC ANALYSES

FIGURE 4: PRINCIPAL COORDINATE ANALYSIS OF PAIRWISE P-DISTANCES BETWEEN RAD-SEQ SAMPLES.

FIGURE 5: ALLELE COMPOSITION OF PUTATIVE HYBRID P. DAMICORNIS SAMPLES COMPARED TO THE EXPECTED ALLELE COMPOSITIONS OF BACKCROSSING HYBRIDS OVER SEVERAL GENERATIONS.

FIGURE S1: BASIC RAD LOCI CORRELATIONS

FIGURE S2: PAIRWISE P-DISTANCE COMPARISONS BETWEEN SAMPLES AT TRANSCRIBED VS. UNTRANSCRIBED LOCI.
## LIST OF TABLES

### CHAPTER 2

<table>
<thead>
<tr>
<th>Table</th>
<th>Title</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>POPULATION GENETIC INDICES FOR THE SIX MICROSATellite LOCI.</td>
<td>62</td>
</tr>
<tr>
<td>2</td>
<td>POPULATION GENETIC INDICES FOR THE NINE POPULATIONS.</td>
<td>63</td>
</tr>
<tr>
<td>3</td>
<td>HIERARCHICAL AMOVA RESULTS</td>
<td>64</td>
</tr>
<tr>
<td>4</td>
<td>PAIRWISE R_{ST} AND F_{ST} BETWEEN TEP P. DAMICORNIS POPULATIONS</td>
<td>65</td>
</tr>
<tr>
<td>5</td>
<td>POCILLOPORA DAMICORNIS POPULATION GENETIC SURVEY (SORTED BY F_{ST}).</td>
<td>66</td>
</tr>
</tbody>
</table>

### CHAPTER 3

<table>
<thead>
<tr>
<th>Table</th>
<th>Title</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>LOCI-SPECIFIC STATISTICS FOR EACH OF THE SIX SAMPLED MICROSATellite LOCI.</td>
<td>93</td>
</tr>
<tr>
<td>2</td>
<td>LIST OF MATERNAL COLONIES</td>
<td>94</td>
</tr>
<tr>
<td>3</td>
<td>RESULTS OF THE GENERALIZED LINEAR MIXED EFFECTS MODEL FOR THE MAIN LOGISTIC REGRESSION ANALYSIS</td>
<td>95</td>
</tr>
</tbody>
</table>

### CHAPTER 4

<table>
<thead>
<tr>
<th>Table</th>
<th>Title</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>OVERVIEW OF THE RAD-SEQ SAMPLE COMPOSITION AND BASIC POPULATION GENETIC INDICES.</td>
<td>133</td>
</tr>
<tr>
<td>2</td>
<td>PAIRWISE COMPARISONS BETWEEN SAMPLES USING P-DISTANCE AND IBS-PROPORTIONS.</td>
<td>134</td>
</tr>
<tr>
<td>3</td>
<td>GENETIC MAKE-UP OF PUTATIVE HYBRID P. DAMICORNIS SAMPLES</td>
<td>135</td>
</tr>
<tr>
<td>51</td>
<td>READ DEPTH PER SAMPLE AND PER LOCUS</td>
<td>143</td>
</tr>
</tbody>
</table>
TABLE S2: TRANSCRIPTOME MATCHES/ALIGNMENTS. 144

TABLE S3: IBS & NUMBER OF DIFFERENTLY FIXED LOCI 145

TABLE S4: VICE VERSA COMPARISON OF ALLELIC IDENTITY 146
Evolutionary Genetics of Pocillopora Corals

CHAPTER 1
GENE FLOW, POCILLOPORA CORALS AND THE TROPICAL EASTERN PACIFIC

Overview

This dissertation explores the patterns and consequences of gene flow for the evolution of reef-building corals. Each chapter highlights a different aspect of gene flow, within species, between generation and across species. Gene flow acts on many different levels and in different, sometimes contradictive ways, to shape the genetic identity and delimitations of species and populations and the genetic structures between them.

Most commonly, gene flow refers to the movement of alleles from one population to another, i.e. inter-population gene flow or population connectivity. Gene flow within populations, intra-population gene flow, is usually of little concern for mobile species but in sessile marine taxa, populations can be structured internally. Both inter- and intra-population gene flow are the focus of Chapter 2. Since evolutionary consequences of gene flow require genetic mixing, evolutionary meaningful gene flow in eukaryotic taxa is tied to reproduction, which enables the observation of individual gene flow events, from parent to offspring. Chapter 3 explores these individual steps of reproductive, vertical gene flow in a predominantly asexual P. damicornis population from the Central Pacific Islands of French Polynesia. Interspecific gene flow via introgressive hybridization is the focus of Chapter 4.
**Gene Flow**

Gene flow is one of four evolutionary forces that change allele frequencies within and among populations over time. Mutations and genetic drift alter allele frequencies randomly across the genome by generating new alleles and eliminating or fixing alleles in populations, respectively. Natural selection acts only on select loci, favoring beneficial alleles and eliminating detrimental ones. Gene flow disperses genetic variation between populations, reintroducing alleles lost to genetic drift, spreading new mutations and fueling natural selection. However, high rates of gene flow homogenize genetic differences between populations, counter divergent selection (Slatkin 1987; Slatkin & Barton 1989; Garcia-Ramos & Kirkpatrick 1997) and prevent populations from separating into different species (Dobzhansky 1937; Mayr 1963). Gene flow thus generates interactions between mutation, drift and selection and shapes the genetic identity of populations and species, their delimitations and the genetic structure between them (Dobzhansky 1937; Huxley 1942; Wright 1943).

In addition, gene flow is also effective over ecological time scales and has a strong impact on population demography, dynamics and diversity (Hellberg *et al.* 2002; Gaines *et al.* 2007). Demography and diversity are in turn key factors for a population’s resilience, recovery and persistence and thus ultimately drive species distribution dynamics and extinction probabilities (Hellberg *et al.* 2002; Jones *et al.* 2007; Willi & Hoffmann 2009). Gene flow data are therefore increasingly used to inform the design of conservation areas like marine protected areas (Palumbi 2003; Jones *et al.* 2007). Other important ecological aspects of gene flow are species invasions (Lavergne & Molofsky 2007; Dlugosch & Parker 2008) and introgressive gene flow from genetically modified organisms (e.g. Wolfenbarger & Phifer 2000; Lu & Snow 2005).
The identification and quantification of gene flow patterns is a main focus of conservation genetics.

Tropical reef corals are important taxa for evolutionary and population genetic research because they are the architects of the planet’s most diverse ecosystem and due to their recent declines due to a wide range of mostly anthropogenic disturbances (Wilkinson 1998, 2000, 2004), such as rising sea surface temperatures and ocean acidification (Hughes et al. 2003; Pandolfi et al. 2003; Hoegh-Guldberg et al. 2007). Genetic research over the last 15 years has revolutionized our understanding of scleractinian evolution (Romano & Palumbi 1996; Fukami et al. 2004; Fukami et al. 2008; Budd et al. 2010). For example, DNA sequence data has revealed that the order Scleractinia consists of two basal lineages, the ‘robust’ and ‘complex’ clade, that evolve separately (Romano & Palumbi 1996; Chen et al. 2002; Fukami et al. 2008) and are inconsistent with the morphology-based taxonomy (Wells 1956; Veron 2000). However, several major issues are still controversial, for example the role of introgressive hybridization in coral evolution and speciation (Veron 1995; Vollmer & Palumbi 2002; Willis et al. 2006; Richards et al. 2008).

Interspecific Gene Flow

Interspecific gene flow via introgressive hybridization is increasingly recognized as a major source of evolutionary novelty (Arnold 1997; Mallet 2005; Hedrick 2013). The evolutionary significance of hybridization depends on the fitness of hybrids, i.e. on their ability to contribute to subsequent generations. Backcrossing of hybrids with parental species can lead to the incorporation of genetic material from one taxon into the genome of another (Anderson 1949). Introgressed alleles can provide new evolutionary raw material (e.g. Kim & Rieseberg
Prominent examples for adaptive introgression include bidirectional introgression of beak traits between Darwin finch species (Grant & Grant 2008), the transition of black fur colors from dogs to wolves (Anderson et al. 2009) and the transfer of mimicry wing patterns among Heliconius butterflies (Pardo-Diaz et al. 2012; Heliconius Genome Consortium 2012). Possible marine examples include two genomic hotspots of elevated unidirectional introgression between cryptic Cliona intestinalis species (Roux et al. 2013) and the presence of Mytilus edulis alleles at higher frequencies in distant M. galloprovencaulis populations than in neighboring populations (Gosset & Bierne 2012). However, in both cases the adaptive significance of introgressed genes has not been determined and differential introgression might be due to intrinsic factors (Gosset & Bierne 2012; Roux et al. 2013). So far, adaptive introgression has not been documented in corals.

Importantly, introgression does usually not occur at equal rates across the genome of hybridizing species (Martinsen et al. 2001; Hohenlohe et al. 2013). Genes and surrounding gene regions that are involved in hybrid sterility and inviability (Muller–Dobzhansky genes) are unable to cross species boundaries while major parts of the genome introgress unaffected (e.g. Ting et al. 1998; Ting et al. 2001; Barbash et al. 2003; Wu & Ting 2004; Brideau et al. 2006). These speciation genes (sensu Wu 2001; but see Nosil & Schluter 2011 for a discussion of speciation genes) or genomic islands/continents of speciation (Turner et al. 2005; Michel et al. 2010) have so far been characterized mostly for plants (Rieseberg & Blackman 2010) and Drosophila species (e.g. Ting et al. 2001; Barbash et al. 2003; Presgraves et al. 2003). In the marine realm, gamete adhesion molecules of free-spawning invertebrates (sea urchins & abalone) have been characterized as speciation genes of sorts since it is unclear if their
divergence is cause or consequence of speciation (Vacquier 1998; Palumbi 2009; Lessios 2011). The best candidate speciation gene in corals is the Mini-Collagen locus in Caribbean Acropora corals, which is blocked from introgression into A. cervicornis while several other genes introgress from A. palmata (Vollmer & Palumbi 2002). Introgressive hybridization has been shown to occur in Pocillopora corals as well and likely plays an important role in the extraordinary success of Pocillopora corals in the remote Tropical Eastern Pacific (Combosch et al. 2008).

In chapter 4, I analyze the patterns of inter-specific gene flow among Tropical Eastern Pacific Pocillopora species using genome-wide RAD-Seq analyses.

Inter-population Gene Flow

Our understanding of intra-specific gene flow among reefs has changed significantly as well over the last decades. In contrast to the long-held paradigm that long-distance dispersal is pervasive among marine populations (Scheltema 1971; Heck & McCoy 1978; Veron 1995), increasing evidence from coral population genetic analyses, suggest widely disparate levels of population connectivity (Warner & Cowen 2002; Ayre & Hughes 2004; Jones et al. 2009). Several studies have recently addressed fine-scale spatial genetic structures over meter scales within coral populations to assess local gene dispersal, reproductive patterns (e.g. parentage, selfing and inbreeding) and selection pressures (Underwood et al. 2007; Miller & Ayre 2008a; van Oppen et al. 2008).

The potential for local adaptation in corals has become a recent focus of coral population genetic studies. Several studies show that populations adapted to extreme environments might contain stress-resistant genotypes (West & Salm 2003; McClanahan et al. 2007; Miller & Ayre
In addition, corals cope with thermal extremes by shifting between different types of their main algal symbionts (Baker et al. 2004; Rowan 2004; Jones 2008) or host differently adapted symbiont populations (Howells et al. 2011). Divergent adaptations between populations, mostly to different temperature regimes, have been documented over various spatial scales from hundreds of meters to hundreds of kilometers (e.g. D'Croz & Mate 2004; Smith-Keune & van Oppen 2006; Ulstrup et al. 2006; Smith et al. 2007; Barshis et al. 2010). Adaptations to different environments are facilitated by limited gene flow between populations to increase the frequency of differentially beneficial alleles.

In chapter 2, I analyze patterns of gene flow within and among Tropical Eastern Pacific Pocillopora damicornis populations and assess their potential for local population to significantly different environmental conditions.

Reproduction

In eukaryotic taxa, evolutionarily meaningful gene flow requires sexual reproduction. Evolutionary theory implies that sex is important to decrease genetic load by exposing alleles to selection (Fisher 1941; Maynard Smith 1978; Charlesworth & Charlesworth 1987), to increase adaptation potential by generating new allele combinations (Stebbins 1957; Maynard Smith 1978; Bell 1982; Morran et al. 2009) and to escape parasitism (Lively 1987; Howard & Lively 1994). This is particularly important in unstable, heterogeneous habitats with high selection pressure and to colonize new environments (Williams 1975; Maynard Smith 1978). Parthenogenesis, in contrast, is more efficient than sexual reproduction since it is independent of mates. It is particularly suitable for stable environments with low selection pressures and as a means to reproduce quickly to colonize open space (e.g. Williams 1975; Holman 1987; Radtkey
et al. 1995; Weider et al. 1999). Both reproductive strategies clearly have advantages and several taxa combine the two (Bell 1982; Hughes 1989). Instead of sex or no sex, the question is then how much sex is ideal? Theory predicts that a little sex goes a long way and the bulk of reproductive output should be asexual (Hurst & Peck 1996).

Corals predominantly reproduce via broadcast spawning of gametes that fertilize in the water column or by releasing fully developed internally brooded larvae. Both strategies had been considered to be obligatory sexual but a recent study indicated that two coral colonies \textit{(Pocillopora damicornis)} use a mixed reproductive strategy with mostly parthenogenetic and some sexually produced larvae (Yeoh & Dai 2010). However, very little is known about this unique mixed reproductive strategy and its evolutionary significance remains unclear.

In chapter 3, I explore the evolutionary significance of simultaneously sexual and asexual reproductive gene flow in \textit{P. damicornis} from Moorea, French Polynesia.

\textbf{Pocillopora corals}

The genus \textit{Pocillopora} (Lamarck 1816) belongs to the family Pocilloporidae, order Scleractinia, in the phylum Cnidaria. Phylogenetically, \textit{Pocillopora} are monophyletic and belong to the robust clade (Romano & Palumbi 1996; Chen \textit{et al.} 2002).

The family \textit{Pocilloporidae} (Gray 1842) is one of the youngest scleractinian families (Wells, 1956) and appeared for the first time in the Upper Cretaceous (approximately 65mya). It is composed of three extant genera, \textit{Stylophora} Schweigger 1819, \textit{Seriatopora} Lamarck 1816 and \textit{Pocillopora} Lamarck 1818 but recent phylogenetic evidence indicates that the genera \textit{Madracis} Milne Edwards & Haime 1849 and \textit{Stylocoeniella} Ehrenberg 1834 belong into this family as well (Fukami \textit{et al.} 2008). Most \textit{Pocilloporidae} are phenotypically plastic and
phylogenetic studies have revealed major discrepancies between morphospecies and phylogenetic lineages in *Pocillopora* (Combosch et al. 2008; Pinzón et al. 2013), *Stylophora* (Flot et al. 2011; Stefani et al. 2011), *Seriatopora* (Flot et al. 2008) and *Madracis* (Frade et al. 2010).

The genus *Pocillopora* is the most abundant and widespread pocilloporiid genus. Two species, *P. damicornis* and *P. eydouxi*, can be found on virtually any coral reef in the Indo-Pacific, from South Africa and the Red Sea to the Tropical Eastern Pacific and from the southernmost coral reefs in Lord Howe Island to the northernmost reefs in Japan and Israel (Veron 2000). While extant pocilloporids are restricted to the Indo-Pacific, fossil evidence suggests that *Pocillopora cf. palmata* (similar to *P. eydouxi*; Veron 1995) was present in the Caribbean as recently as 82,000 years ago (Geister 1977; Pandolfi et al. 2001).

The taxonomy of *Pocillopora* species is complicated by ecomorphs and widespread phenotypic plasticity in response to light intensity, wave exposure, space availability, sedimentation, crab induction, grazing and cropping (Wells 1956; Pichon & Veron 1976; Veron 2000). Accordingly, well over 50 species have been described (Wells 1956; Pichon & Veron 1976; Knowlton & Jackson 1994; Veron 1995; Glynn 1999; Veron 2000) although others have suggested that all growth forms represent a single polymorphic species (Gardiner 1897; Wood-Jones 1910). For example, Gardiner (1897 p. 942) stated “In the collections there are over 50 specimens, (...) and I am doubtful whether all these so-called species should not rather be described as varieties of one species, the characters of which would be the characters of the whole genus.” However, *Pocillopora* species are amongst the most intensely studied coral taxa (Veron 1995 & 2000) and a robust, well-supported systematic resolution of this genus is missing.
*Pocillopora damicornis* is likely the most studied coral species worldwide (Veron 1995). However, it seems like the more it is studied the more questions arise. For example, although *P. damicornis* is amongst the widest distributed coral species, which testifies to its dispersal potential, population genetic studies have reported significant population structure over kilometer-scales among neighboring reefs (e.g. Benzie *et al.* 1995; Ayre & Miller 2004; Miller & Ayre 2004). Although the genetics of *P. damicornis* larvae brooding have been investigated repeatedly for 30 years, major basic questions about the asexual and/or sexual origin of its planula larvae remain (Stoddart 1983a, 1986; Ayre & Miller 2004; Sherman *et al.* 2006; Yeoh & Dai 2010). Moreover, after I reported in 2008 that *P. damicornis* in the Tropical Eastern Pacific consists of two different genetic types (Combosch *et al.* 2008), others have described up to six different genetic lineages in *P. damicornis* populations throughout its Indo-Pacific distribution range (Souter 2010; Pinzón & LaJeunesse 2011; Pinzón *et al.* 2013; Schmidt-Roach *et al.* 2013).

One of the most intriguing characteristics of *P. damicornis* is its regional dominance in the isolated Tropical Eastern Pacific (TEP) while *P. damicornis* is elsewhere considered a minor component of the hermatypic coral taxa. Glynn *et al.* (1972) first described extensive, monospecific *P. damicornis* stands in the TEP and suspected its inter-specific competitiveness and fast-growing ramose growth form as the main reasons for its regional success. Highsmith (1982) speculated that the local abundance in the TEP is largely due to vegetative fragmentation, i.e. clonal propagation. Richmond (1985) found differences between TEP and Indo-West Pacific populations in fecundity, reproductive allocation, growth rates and age specific mortality and interpreted them as successful adaptations of *P. damicornis* to the TEP. Richmond (1987) later suggested that TEP populations might represent a different species, based on these life-history differences. Glynn *et al.* (1991) first noticed that TEP populations reproduce via gamete-
spawning in contrast to Indo-West Pacific populations that brood mostly parthenogenetic larvae (Stoddart 1983b; Stoddart 1986; Sherman et al. 2006; Yeoh & Dai 2010). In 2008, I provided genetic evidence for inter-specific hybridization among TEP Pocillopora corals and introgressive gene flow into P. damicornis but found no indications for hybridization among Indo-West Pacific populations (Combosch et al. 2008). Pinzon and LaJeunesse (2011) recently suggested that TEP Pocillopora species, including P. damicornis, are comprised of three distinct genetic clusters (Type 1 - 3) that are independent of species identity and morphology. The evolutionary history of TEP P. damicornis population is thus controversial and nothing is known about the genetic basis of this most successful peripheral coral adaptation.

The Tropical Eastern Pacific

The TEP is best known for being one of the most isolated biogeographic regions of the world’s oceans (Darwin 1859; Chavez & Brusca 1991; Grigg & Hey 1992). It is enclosed by land and cold water and separated from the tropical central Pacific by a vast expense of deep island-free ocean, stretching over 5000km between Clipperton and the Line Islands. For continental shelf fauna, this endless open ocean constitutes a virtually “impassable barrier” (Darwin 1859). Yet, trans-Pacific species occur in several taxa, but trans-Pacific gene flow is mostly sporadic (e.g. Rosenblatt & Waples 1986; Lessios et al. 1998; Lessios et al. 2003; Lessios & Robertson 2006) and often too low to effectively connect TEP populations with their Central Pacific counterparts, particularly among corals (Forsman 2003; Combosch et al. 2008; Baums et al. 2012).

The isolation of the TEP makes its populations particularly vulnerable to population collapses and genetic pauperization. Moreover, TEP coral reefs are exposed to strong seasonal
upwelling and periodic El-Niño events. While seasonal upwelling generally limits coral growth (Glynn & Stewart 1973), El-Niño events often cause massive coral mortality by increasing seawater temperature beyond tolerable thresholds for reef corals. Several major El-Niño events causing severe coral mortality have been documented over the last three decades (Glynn 1985, 1988, 1990; Glynn & Colgan 1992; Eakin & Glynn 1996; Glynn et al. 2001; reviewed in Glynn & Enochs 2011). Moreover, suitable reef environments in the TEP are fragmented (Glynn & Ault 2000), subject to elevated nutrient levels (Glynn & Mate 1997), increased terrestrial runoff (Glynn 1976) and toxic algal blooms (Guzman et al. 1990). Some of these challenges are likely to become more frequent for corals worldwide due to environmental degradation and global climate change.

A recent paleontological study further showed a massive population collapse of TEP corals due to increased El-Niño variability, with subsequent recovery ~2600 years ago (Toth et al. 2012) with potentially significant evolutionary consequences. Prolonged significant population bottlenecks cause a dramatic loss of genetic diversity, which eliminates evolutionary raw material and reduces the ability of populations to react to environmental change (Gilpin & Soule 1986). On the other hand, bottlenecks can also purge deleterious alleles and facilitate local selection and adaptation (Carson 1990; Keller et al. 1994). The isolation of the TEP and the fragmentation among TEP coral reefs likely enabled selection to effectively eliminate deleterious and mal-adapted alleles, creating a potentially rapidly evolving small and isolated population that was forced to adapt effectively to difficult local environmental conditions. The ability of TEP \textit{P. damicornis} specimens to adapt to local conditions has been shown experimentally: specimens from the non-upwelling Gulf of Chiriquí are less sensitive to higher temperatures than those from the upwelling Gulf of Panamá (D'Croz & Mate 2004). No population genetic analyses of any
TEP coral had been done before, to assess levels of gene flow and genetic structures among TEP reefs.

In Chapter 2, I use multi-locus microsatellite data to describe the population genetic structure among and within nine Tropical Eastern Pacific *P. damicornis* populations across three environmentally and geographically distinct regions in Panamá. I assess the corals potential to adapt to local environmental conditions and show significant population genetic structure, both among regions and among populations indicating limited gene flow, i.e. larval dispersal. Populations in environmentally distinct regions that are connected by only limited levels of gene flow have a significant potential for differential adaptations. Within reefs, I found high levels of genotypic diversity, significant inbreeding and local spatial genetic structures over 10m distances. These results suggest that TEP *P. damicornis* populations reproduce predominantly sexually, in contrast to previous speculations (Highsmith 1982), and that gene flow is severely limited within reef, even over meter scales.

In Chapter 3, I provide detailed genetic analyses of the unusual, simultaneously mixed reproductive strategy in Indo-West Pacific *P. damicornis* populations. Using multi-locus microsatellite data, I compare the genotypes of over 600 coral larvae with the genotypes of their brood parents. My results show that although larvae are predominantly parthenogenetic, a small subset of larvae from most colonies was produced sexually. I use logistic regression analyses to assess the impact of several external factors on the ratio of parthenogenetic and sexually produced larvae and found significantly variation in the proportion of sexual larvae with colony size, reproductive cycle day and general calendar day. Particularly interesting is the significant decrease in the proportion of sexual larvae with increasing colony size, which suggests that the
mixed reproductive strategy changes during the life cycle of the coral. This has important ecological and evolutionary consequences that I discuss in detail in chapter 3.

In chapter 4, I re-analyzed new and published sequence data from the nuclear rDNA (ITS2) and mitochondrial open reading frame (ORF) with a particular focus on rooting the phylogenetic tree to understand the evolutionary trajectory among *Pocillopora* lineages. Both phylogenetic analyses identified previously defined genetic types that are incongruent with traditional species delimitations. The improved rooting suggests that type 3 is the most basal *Pocillopora* lineage, which might indicate an Indian Ocean origin of the genus *Pocillopora*. I then used genome-wide Restriction-side Associated DNA sequencing (RAD-Seq) to conduct a phylogenomic comparison among three TEP *Pocillopora* species, two types of putative TEP hybrids and four Pacific-wide *P. damicornis* populations. The phylogenomic analyses revealed three distinct genomic lineages that support previous ITS and ORF-based lineages. The two hybrid pools were genetically similar to one parental lineage but also contained numerous alleles from the other parental lineage, which strongly suggest their hybrid origin and subsequent uni-parental backcrossing.
References


Forsman Z. 2003 *Phylogeny and phylogeography of Porites & Siderastrea (Scleractinia: Cnidaria) species in the caribbean and eastern pacific; based on the nuclear ribosomal ITS region*, University of Houston.


Fukami H., C. Chen, A. Budd, et al. 2008 Mitochondrial and nuclear genes suggest that stony corals are monophyletic but most families of stony corals are not (Order Scleractinia, Class Anthozoa, Phylum Cnidaria). *PLoS One* 3.


Richmond R. 1985 Variations in the population biology of *Pocillopora damicornis* across the Pacific 6, 101-106.


Williams G. 1975 *Sex and evolution* Princeton University Press, Princeton, NJ.


CHAPTER 2

POPULATION GENETICS OF AN ECOSYSTEM-DEFINING REEF CORAL -
POCILLOPORA DAMICORNIS IN THE TROPICAL EASTERN PACIFIC¹

Abstract

Background: Coral reefs in the Tropical Eastern Pacific (TEP) are amongst the most peripheral and geographically isolated in the world. This isolation has shaped the biology of TEP organisms and lead to the formation of numerous endemic species. For example, the coral *Pocillopora damicornis* is a minor reef-builder elsewhere in the Indo-West Pacific, but is the dominant reef-building coral in the TEP, where it forms large, mono-specific stands, covering many hectares of reef. Moreover, TEP *P. damicornis* reproduces by broadcast spawning, while it broods mostly parthenogenetic larvae throughout the rest of the Indo-West Pacific. Population genetic surveys for *P. damicornis* from across its Indo-Pacific range indicate that gene flow (i.e. larval dispersal) is generally limited over hundreds of kilometers or less. Little is known about the population genetic structure and the dispersal potential of *P. damicornis* in the TEP.

Methodology: Using multilocus microsatellite data, we analyzed the population structure of TEP *P. damicornis* among and within nine reefs and test for significant genetic structure across three geographically and ecologically distinct regions in Panamá.

Principal Findings/Conclusions: We detected significant levels of population genetic structure (global $R_{ST} = 0.162$), indicating restricted gene flow (i.e. larvae dispersal), both among the three

regions ($R_{RT} = 0.081$) as well as within regions ($R_{SR} = 0.089$). Limited gene flow across a distinct environmental cline, like the regional upwelling gradient in Panamá, indicates a significant potential for differential adaptation and population differentiation. Individual reefs were characterized by unexpectedly high genet diversity (avg. 94%), relatively high inbreeding coefficients (global $F_{IS} = 0.183$), and localized spatial genetic structure among individuals (i.e. unique genets) over 10m intervals. These findings suggest that gene flow is limited in TEP $P. damicornis$ populations, particularly among regions, but even over meter scales within populations.

**Introduction**

Populations at the periphery of species’ distribution ranges often exist in suboptimal and/or unstable conditions at the organism’s physiological or ecological limits (Soule 1973). They tend to be small as well as geographically and genetically isolated from central core populations (Brown *et al.* 1996; Vucetich & Waite 2003). Their small population sizes and genetic isolation increase genetic drift and inbreeding and limit genetic diversity and cohesiveness among populations (Wright 1931; Slatkin 1987; Garcia-Ramos & Kirkpatrick 1997). As a result, peripheral populations tend to have reduced genetic diversity and are often heavily reliant on local sources of recruitment rather than long-distance recruitment from core populations (e.g. Arnaud-Haond *et al.* 2006; Bergl & Vigilant 2007; Beatty *et al.* 2008; Noreen *et al.* 2009). Their isolation and low levels of genetic diversity may also hamper their ability to cope with environmental changes (e.g. global climate change) and to recover from disturbances, making them more vulnerable to extinction (Lawton 1993; Hoffmann & Blows 1994). However, small population sizes in genetic isolation can also facilitate local selection and adaptation, give
rise to evolutionary innovation (Budd & Pandolfi 2010) and allow peripheral populations to exploit new ecological niches (Templeton 1980; Carson & Templeton 1984).

One of the most isolated and peripheral biogeographic regions of the world’s oceans is the Tropical Eastern Pacific (TEP). It is separated from the Indo-West Pacific by 5000km of open-ocean that Darwin (1859) described as “impassable” for shallow water marine species (Ekman 1935; Mayr 1954). While not completely impassable, the marine fauna of the TEP consists of a limited subset of trans-Pacific species (i.e. species distributed across the Pacific), which are typically interconnected by low levels of gene flow (e.g. Lessios et al. 1999; McCartney et al. 2000; Colborn et al. 2001; Lessios et al. 2001; Lessios & Robertson 2006; Combosch et al. 2008; Duda & Lessios 2009), as well as numerous endemic species (Glynn & Ault 2000; Colborn et al. 2001; Hughes et al. 2002). For example, 34 species of reef-building corals are recognized in the TEP (compared to 581 in the Western Pacific), of which 27 are trans-Pacific (80%) and seven (20%) are regional endemics (Veron 2000). Major environmental challenges for reef corals in the TEP include significant freshwater discharges and seasonal upwelling, leading to high turbidity, elevated nutrient levels and reduced seawater temperatures. Reef habitats in the TEP are confined mostly to a narrow continental shelf and often fragmented by long stretches of estuaries, mangroves and sandy beaches (Briggs 1974; Glynn & Ault 2000).

The isolation of the TEP from the Indo-West Pacific and major differences in environmental factors and species compositions between these regions enabled significant ecological and life-history shifts in many TEP species. For example, the scleractinian coral *Pocillopora damicornis* is a minor reef-building coral in the Indo-West Pacific, but the primary reef-builder in the TEP (Glynn & Macintyre 1977; Glynn & Wellington 1983; Guzman & Cortes 1993). *Pocillopora damicornis* in the TEP forms large, mono-specific carpets measuring tens to
hundreds of hectares (Glynn et al. 1996) that are thought to be maintained predominantly by asexual (i.e. vegetative) fragmentation (Highsmith 1982; Richmond 1987). Moreover, while Indo-West Pacific *P. damicornis* populations brood mostly asexual (parthenogenetic) larvae (Stoddart 1983; Yeoh & Dai 2010), populations in the TEP reproduce by broadcast-spawning its gametes (Glynn et al. 1991). This reproductive shift could have important genetic consequences since a broadcast spawning coral should not be able to produce parthenogenetic larvae (Harrison & Wallace 1990; Richmond & Hunter 1990) and because broadcast spawning corals are often considered to have higher dispersal potentials than brooding corals (e.g. Hellberg 1995; Hellberg 1996; Ayre & Hughes 2000; Nishikawa et al. 2003; Nishikawa 2008), particularly over small spatial scales (Stoddart 1988b; McFadden & Aydin 1996; Miller 1998; Costantini et al. 2007 but see Underwood et al. 2007; Miller & Ayre 2008a). Hermaphroditic broadcast-spawning corals might also have a higher potential for selfing and inbreeding resulting from external fertilization (Heyward & Babcock 1986; Hedgecock 1994; Willis et al. 1997).

Population genetic surveys among larvae-brooding *P. damicornis* populations across the Indo-West Pacific indicate that larval dispersal (i.e. gene flow) is generally limited over hundreds of kilometers (e.g. Ayre & Hughes 2004; Miller & Ayre 2004, 2008b) and in some cases even among neighboring reefs (e.g. Benzie et al. 1995; Ayre & Miller 2004; Miller & Ayre 2004). Some of the highest levels of genetic structure have been described at the margins of *P. damicornis* Indo-Pacific range (Stoddart 1984; Ayre & Hughes 2004; Miller & Ayre 2004, 2008b). For example, populations in Southeast Australia are more genetically distinct over 700km (FST = 0.32; Miller & Ayre 2008b) than more central Great Barrier Reef populations over 1200km (FST = 0.06; Ayre et al. 1997). Most studied populations of *P. damicornis* in the Indo-West Pacific also showed significant levels of local inbreeding and contained some level of
clonality (Ayre et al. 1997; Adjeroud & Tsuchiya 1999; Sherman et al. 2006; Whitaker 2006; Souter et al. 2009).

Little is known about the population structure of broadcast-spawning *P. damicornis* populations in the TEP. Combosch et al. (2008) documented evidence for restricted transpacific dispersal between Central and Eastern Pacific populations of *P. damicornis*, using ribosomal DNA sequence data ($F_{ST} = 0.419$, $p < 0.001$). Using multilocus genetic data, Pinzon & LaJeunesse (2011) examined the genetic structure among multiple TEP Pocillopora species, including 7 well-recognized Indo-Pacific species (*P. damicornis*, *P. eydouxi*, *P. meandrina*, *P. verrucosa*, *P. capitata*, *P. elegans*, *P. woodjonesi*). They identified three genetic clusters or types (Type I, II & III) in their multi-species TEP sample of Pocillopora. All three of these types were found in *P. damicornis* from across the TEP. Type I was found throughout the TEP, including Panamá, Type II was found only in Clipperton Atoll and Type III was found in Panamá and Galápagos. Yet, no genetic structure was detected within the broadly distributed Type I group over 3500km of the TEP.

One of the best-studied areas in the TEP is the Pacific coast of Panamá, which is dominated by two major gulfs, the Gulf of Chiriquí in the west and the Gulf of Panamá in the east. The two gulfs differ in several important ecological factors (Glynn et al. 1972; Glynn & Stewart 1973; Glynn 1976, 1977; Highsmith 1980; Glynn & Leyte Morales 1997; Glynn & Ault 2000), most notably in the strength of seasonal upwelling, which is strong in the Gulf of Panamá, but virtually absent in the Gulf of Chiriquí (Glynn & Mate 1997; Glynn & Ault 2000). The Azuero Peninsula, which separates the two gulfs, is a transition zone with moderate upwelling (D'Croz & Mate 2004). Corals from the two gulfs differ significantly in their thermal tolerances.
(Glynn & D'Croz 1990; D'Croz & Mate 2004) and ability to recover from major disturbances like El-Niño induced bleaching events (Glynn 2001).

Using multilocus microsatellite data, we examined the population genetic structure of TEP *P. damicornis* within and among nine Panamanian populations across the three geographically and ecologically distinct regions (Figure 1). Based on the strong geographic and ecologic differences between the three regions, we predicted significant regional differences in *P. damicornis* population genetic structure. Within reefs, we analyzed spatial patterns of genetic diversity to estimate the relative contributions of clones versus unique genets to localized spatial genetic structure (SGS). We predict that high rates of vegetative fragmentation should lead to significant SGS due to clonal aggregations. In contrast, high rates of sexual reproduction by broadcast spawning should inhibit SGS due to increased outcrossing and the mixing of genotypes at the sea surface during larval development.

**Materials and Methods**

Collection of coral samples for this project was approved by the Autoridad Nacional del Ambiente and was conducted under permit number ANAM SEA7108.

**Sampling Sites**

*Pocillopora damicornis* was sampled from nine reefs out of three regions (Gulf of Chiriquí, Azuero Peninsula, Gulf of Panamá) across a well-documented upwelling gradient (e.g. Glynn & Mate 1997) along 1400km of coastline in Panamá (Figure 1). Four reefs (Uva Island, Granito de Oro, Canal de Afuera Island and Bahia de Damas on Coiba Island) were sampled in the Gulf of Chiriquí in the west, where upwelling is absent and sea surface temperatures rarely
drop below 25°C (Glynn & Mate 1997). Coastal coral reefs are poorly developed in the Gulf of Chiriquí due to extensive freshwater runoff, but substantial reef development has occurred around its numerous islands, including the largest (136ha), thickest (13m) and oldest (5600a) TEP reefs (Glynn & Macintyre 1977; Glynn & Mate 1997). Two reefs (Achotines Bay and Iguana Island) were sampled in the centrally located Azuero Peninsula where occasional moderate upwelling occurs (D'Croz & Mate 2004). Coastal reefs are more common in the Azuero Peninsula than in the gulfs, most likely due to reduced freshwater runoff. Three reefs (Saboga and Contadora Islands in the Perlas Archipelago and Taboga Island) were sampled in the eastern Gulf of Panamá where strong seasonal upwelling leads to declines in sea surface temperature to ~15°C (Glynn & Mate 1997). Coastal reefs are uncommon in the Gulf of Panamá and reef development occurs primarily around near-shore islands and in the Perlas Archipelago.

Sampling and Genetic analyses

On each reef, 18 to 36 nubbins (2cm) of *P. damicornis* were sampled every ten meters along reef sites composed of dense stands of *P. damicornis* to analyze its small-scale spatial genetic structure. Samples were preserved in Guanidinium-Isothiocyanate (GITC) DNA buffer for genetic analysis (Fukami et al. 2004). DNA was extracted using standard Phenol-Chloroform extraction and ethanol precipitation.

Ten microsatellite loci were amplified using primers developed by Magalon et al. (2004) (Pv2 & Pv6) and Starger et al. (2007) (Pd2, Pd4, Pd5, Pd6, Pd7, Pd8, Pd9 & Pd10). Six of these loci (Pd2, Pd4, Pd5, Pd6, Pv2 & Pv6) amplified consistently using modified PCR protocols. PCR amplifications were carried out separately for each locus, following a nested PCR protocol for labeling, including forward primers with M13 tails and M13 primers with IR-labels (Schuelke
Each PCR mix contained 2 pmol forward and 8 pmol of each reverse and m13 primer in a 10μl reaction with 1μl 10× PCR buffer, 0.2 mM dNTPs, 0.25 U AmpliTaq DNA polymerase (Applied Biosystems) and 0.5μl template DNA. PCR profiles varied slightly by locus: For Pd2, Pd4, Pd5 and Pd6, the PCR profile consisted of 30 cycles at 94°C for 45s, 55°C for 35s and 72°C for 35s followed by 8 cycles with a lower annealing temperature (53°C) for the M13 labeling. For Pv2 and Pv6, the profile was modified to 30 cycles at 94°C for 45s, 56°C for 45s and 72°C for 30s followed by 8 labeling cycles with 53°C annealing temperature. Both protocols included an initial 3min denaturing step at 94°C and a final extension step at 72°C for 7min. The PCR products with incorporated infrared dyes were then sized and scored on a LICOR 4300 Genetic Analyzer.

Population genetics

Population genetic statistics, including allele frequencies, clonality, probability of identity, F-statistics and Analyses of Molecular Variance (AMOVA) were calculated using GenAlEx 6.1 (Peakall & Smouse 2006). The program Micro-Checker 2.3 (van Oosterhout et al. 2004) was used to test for stutter bands, large allele dropout and null alleles. Linkage disequilibrium among loci was assessed using a Markov chain method (Guo & Thompson 1992) in GENEPOP 4.0.10 (Raymond & Rousset 1995), followed by sequential Bonferroni correction to account for multiple comparisons (Rice 1989). Clones were identified in GenAlEx and because they introduce biases into population genetic analyses, only unique genets were used unless otherwise noted.

Population genetic structure was calculated as F<sub>ST</sub> (Wright 1951), which assumes an infinite allele model (Kimura & Crow 1964), and R<sub>ST</sub> (Slatkin 1995), which is based on a
stepwise mutation model of allele evolution (Kimura & Ohta 1978). Hierarchical AMOVA were used to estimate genetic structure among regions, among populations and among populations within regions (Excoffier et al. 1992). Statistical significance of $R_{ST}$ and $F_{ST}$ is based on 999 permutations. $R_{ST}$ recovered more genetic structure, especially among regions, whereas $F_{ST}$ had a lower resolution and higher intra-regional variation ($F_{RT} = 15\%$ of $F_{ST}$ compared to $R_{RT} = 50\%$ of $R_{ST}$). As a result, we preferentially focus on $R_{ST}$ (e.g. Tables 3 & 4). Principal Coordinate Analyses (Orloci 1978), implemented in GenAlEx, were used to detect and visualize the major patterns of pairwise $R_{ST}$ population comparisons. Isolation by distance among populations was tested using a Mantel test, implemented in GenAlEx 6.1, to test for an association between pairwise population genetic differences and geographic distances.

The Bayesian clustering method implemented in STRUCTURE 2.3 (Pritchard et al. 2000) was used to infer the likely number of genetic clusters (K) in the dataset. STRUCTURE estimates the posterior probability $P(X|K)$ that the data fit the hypothesis of K clusters using a Markov chain Monte Carlo approach to optimize genotypic equilibrium (linkage equilibrium and HWE) within each cluster. Five independent runs for each K from 2 to 19 were conducted using the admixture model and independent allele frequencies with a burn-in period of $10^5$ steps and $10^6$ Markov chain Monte Carlo replications for data collection (Supplementary Figure S1). STRUCTURE runs were aligned using Clumpp1.1 (Jakobsson & Rosenberg 2007) and the bar plot was generated with Distruct1.1 (Rosenberg 2004). To verify the optimal value for K, we compared the $P(X|K)$ results with $\Delta K$, the second order rate of change in the likelihood of K (Evanno et al. 2005), which corresponds to the strength of genetic subdivision among clusters in the data (Figure S1b).
Fine-scale Spatial Genetic Structure (SGS) within populations was analyzed using the program SPAGeDi 1.2 (Hardy & Vekemans 2002). To distinguish between spatial genetic structure (SGS) due to clonal aggregations and SGS among unique genets, we analyzed and compared two different dataset: one dataset consisted of all 207 samples including clones (i.e. the ramet dataset) and one dataset contained only the 194 unique genotypes excluding clones (i.e. the genet dataset). SGS was estimated for each of the three regions separately and for all populations combined using the pairwise kinship coefficient $F_{ij}$ (Loiselle et al. 1995; Vekemans & Hardy 2004). Other kinship coefficients including Ritland (1996) and Moran’s $I$ (Hardy & Vekemans 1999) produced similar results (not shown). 95% confidence intervals and standard errors were estimated by 10,000 permutations of the genetic and the spatial datasets. Kinship values outside the 95% confidence intervals were interpreted as significant SGS at the spatial distance. Three parameters were used to describe and compare SGS among regions and studies: $F_{(1)}$, genetic patch size and $Sp$. $F_{(1)}$ is the average kinship among individuals within the smallest distance interval (10m). Its statistical significance ($p$) was obtained by comparing the slope $b(F)$ of $F(r)$ on $\ln(r)$ to 9999 random permutations of individuals among locations within populations using a Mantel test. The genetic patch size is the distance that corresponds to the first x-intercept of the kinship correlogram (cf. Sokal & Wartenberg 1983). Within the genetic patch, individuals are more closely related than the population average, i.e. the association between genotypes is positive, while individuals outside of the patch are genetically independent, i.e. their correlation is negative. The $Sp$ statistic (Vekemans & Hardy 2004) is based on $F_{(1)}$ and the decrease of SGS with distance ($b(F)$).
Results

**Genetic diversity, heterozygosity and clonality**

Multilocus microsatellite data ($n \geq 5$ loci) was obtained for 207 *Pocillopora damicornis* colonies from nine reef populations (Tables 1 & 2). The six amplified loci yielded between four and ten alleles each (mean $N_A = 7$). Three pairs of loci showed indications of significant linkage disequilibria, but none showed consistent significant linkage disequilibria in more than two populations. Consequently, all six loci were considered as unlinked. Significant heterozygote deficits were detected in five loci (Table 1), but only Pv2 showed significant deficits in all populations, which is indicative of null alleles. The program Micro-Checker confirmed the presence of null alleles in Pv2, which was then excluded from further analyses.

Using the five remaining loci, 194 unique genets were identified with a high probability of identity ($p < 3 \times 10^{-5}$; Table 2) among the 207 genotyped samples. All nine populations were comprised predominantly of unique, i.e. sexually derived genets (93.8 ± 5.5%). Clonal genotypes were detected in all but two populations (Canal & Iguana), but only Saboga (Gulf of Panamá) contained more than two clones ($n = 4$, i.e. 18%). No clonal genotypes were shared among populations, in contrast to larvae-brooding Indo-West Pacific populations (Souter *et al.* 2009; Starger *et al.* 2010). Since clones introduce biases into population genetic analyses, only unique genets were used in subsequent analyses, unless otherwise noted.

Using the genet only dataset, significant heterozygote deficits were detected in six out of the nine populations (Table 2), indicating widespread local inbreeding. Inbreeding was significant in the Gulf of Panamá ($F_{IS} = 0.255 \pm 0.122$) and in the Gulf of Chiriquí ($F_{IS} = 0.196 \pm 0.171$), but not in the Azuero Peninsula ($F_{IS} = 0.048 \pm 0.135$) due to an excess of heterozygotes in Achotines. Each population contained on average 4.7 alleles per locus. Slightly more alleles
were detected in the Gulf of Chiriquí (avg. $N_A = 5.1 \pm 0.26$) than in the Azuero Peninsula (4.5 ± 0.20) and in the Gulf of Panamá (4.3 ± 0.31). Interestingly, private alleles were only found in Gulf of Chiriquí populations, which contained the only regional-specific allele (Pd2, 223bp) as well.

**Genetic Structure among reefs and regions**

Hierarchical AMOVA revealed significant levels of population genetic structure ($p < 0.005$) among populations, among regions, and among populations within regions (Table 3). Genetic structure among populations was highly significant for both global $R_{ST}$ (0.162; $p < 0.001$) and global $F_{ST}$ (0.053; $p < 0.001$). For the R-statistics, population structure was equivalent among regions ($R_{RT} = 0.081$) and among populations within regions ($R_{SR} = 0.089$). All three regions were significantly different from each other ($p < 0.001$), but the two gulf regions were most distinct (pairwise $R_{ST} = 0.138$), whereas the Azuero Peninsula, which lies between the two gulfs, showed lower pairwise genetic structure (Gulf of Chiriquí $R_{ST} = 0.070$; Gulf of Panamá $R_{ST} = 0.101$). For the $F$-statistics, population structure was predominantly due to differences among populations within regions ($F_{SR} = 0.045$; Table 3).

Fourteen out of 36 pairwise $R_{ST}$ comparisons between populations were significant after sequential Bonferroni adjustments (Table 4). Twelve of the significant pairwise $R_{ST}$s were between populations from different regions, whereas only two comparisons within regions were significant, both involving the Gulf of Chiriquí population Granito. These strong regional differences can be seen as well in the principal coordinate analyses, where populations cluster into distinct regional groupings (Figure 2). However, no significant isolation-by-distance pattern was observed among populations ($p = 0.013$).
Bayesian clustering implemented in STRUCTURE detected five genetic clusters in the dataset (Figure S1). All five clusters were found in each population, but their distribution differed significantly among populations and particularly among regions (Figure 1). The genetic clusters 1 and 2 were most common in the Gulf of Chiriquí (28% & 25%, respectively), clusters 4 and 5 were most prevalent in the Gulf of Panamá (22% & 39%) and cluster 3 was dominant in the Azuero Peninsula (32%). Cluster 1 and 2 were also common in the Azuero Peninsula (23% and 24%), which indicates that its *P. damicornis* populations are more similar to the Gulf of Chiriquí than to the Gulf of Panamá - as indicated by the pairwise R$_{ST}$ values.

*Fine-scale spatial genetic structure (SGS) on reefs*

In addition to the genetic structure among regions and populations, strong fine-scale spatial genetic structure (SGS) was detected among individuals within reefs. SGS was significant over 10m intervals in both the ramet dataset (i.e. including clones; F$_{ramets}$ = 0.102; p <0.001) and the genet dataset (i.e. excluding clones; F$_{genets}$ = 0.086, p <0.001). This indicates that corals 10 meters apart are significantly more closely related to each other than to the rest of the population. Kinship among individuals (F$_{ij}$) was only slightly elevated when clones were included, reflecting the low frequency of clones in the dataset (Figure 3). While SGS is to be expected in coral populations due to clonal aggregations as a consequence of asexual (vegetative) fragmentation, significant SGS in the genet dataset indicates that the observed structure is due to non-random spatial relatedness among distinct, i.e. sexually produced, genets. Across all nine populations, the average genetic patch size (i.e. the distance at which individuals on average are as related as across the entire population) was 50m and the Sp statistic was 0.055 (i.e. rather high compared to plant literature data; Vekemans & Hardy 2004).
Significant levels of SGS out to 10m were detected in all three regions (p < 0.05; Figure 3b-d). Average $F_{10m}$-kinship in the Azuero Peninsula and Gulf of Panamá was similarly high ($F_{10m} \sim 0.109$), almost equivalent to a first-cousin relatedness ($F_{IJ} = 0.125$) whereas $F_{10m}$-kinship in the Gulf of Chiriquí was two times lower (0.059; p < 0.05). The genetic patch size was almost twice as large in the Gulf of Panamá (70m) than in the Gulf of Chiriquí and Azuero Peninsula (40m). These differences in the strength and extent of SGS are reflected in the Sp statistic, which indicates that SGS is strongest in the western Gulf of Panamá (0.062), intermediate in the central Azuero Peninsula (0.059) and lowest in the eastern Gulf of Chiriquí (0.049).

**Discussion**

The Tropical Eastern Pacific (TEP) is one of the most geographically isolated biogeographic regions in the world’s oceans (Darwin 1859). Its isolation has profound consequences for the diversity and ecology of TEP ecosystems, most notably coral reefs (Glynn & Ault 2000). TEP coral reefs are remarkable since they are built predominantly by *Pocillopora damicornis* (Glynn & Macintyre 1977; Glynn & Wellington 1983; Guzman & Cortes 1993). TEP *P. damicornis* also distinguish by its broadcast gamete-spawning reproductive strategy. Multilocus microsatellite data from nine Panamanian populations demonstrate that TEP *P. damicornis* populations from Panamá are characterized by strong population genetic structure ($R_{ST} = 0.162$), within and particularly among its three major regions, the Gulf of Chiriquí, the Azuero Peninsula and the Gulf of Panamá. Moreover, the large and mono-specific stands of *P. damicornis* that characterize TEP reefs are not dominated by a few, highly clonal genotypes as previously thought (Highsmith 1982; Richmond 1987), but instead are comprised of numerous distinct genets, indicating frequent sexual reproduction. Significant fine-scale spatial genetic
structure (SGS) among individual genets and highly significant levels of heterozygote deficits further indicate widespread inbreeding and limited gene dispersal over meter scales, which is unusual for free-spawners. These tiered layers of significant population genetic structure indicate that gene flow (i.e. larval dispersal) is restricted among and within reefs, particularly between the environmentally distinct regions, which indicates a significant potential for regional adaptations to the different ecological conditions in each region.

*Population Structure*

As hypothesized, we detected strong genetic differences among the three geographically and ecologically distinct regions in Panamanian *P. damicornis* populations ($R_{RT} = 0.081$). Panamá’s two major gulfs – the Gulf of Chiriquí and the Gulf of Panamá - which are geographically most distant, were most genetically distinct ($R_{ST} = 0.138$). These genetic differences correspond with pronounced ecological differences between these two regions, e.g. in species compositions and environmental factors (e.g. Glynn 1976; Glynn et al. 1991; Glynn & Mate 1997). The most significant environmental difference is the pattern of seasonal upwelling, which is strong in the Gulf of Panamá dropping temperatures down to 15°C, but is virtually absent in the Gulf of Chiriquí (Glynn & Mate 1997; Glynn & Ault 2000). The Azuero Peninsula, which bisects the two regions, is a transition zone with intermediate upwelling conditions (D'Croz & Mate 2004). It also has intermediate levels of genetic structure compared to the other regions ($R_{ST} = 0.070$ & 0.101).

Upwelling delivers cool, nutrient-rich waters to the near-shore environment. In tropical habitats like the TEP, they have significant impacts on nutrient and temperature-sensitive reef fauna (Glynn 1977; Glynn & Leyte Morales 1997). In the TEP, variations in the severity of
upwelling restrict the regional and local distributions of coral reefs (Glynn & Stewart 1973; Glynn & Ault 2000) and of major reef taxa, including corals (Glynn & Ault 2000) and key corallivores like the sea star *Acanthaster planci* (Glynn 1976). Lower temperatures associated with seasonal upwelling reduce reef development (Glynn *et al.* 1972), coral growth rates (Glynn 1977) and coral reproduction (Glynn *et al.* 1991), while elevated nutrients increase competition with algae (Glynn & Stewart 1973; Birkeland 1977), bioerosion (Highsmith 1980) and water turbidity (Glynn & Ault 2000). Corals from regions with different upwelling regimes often display different thermal tolerances, indicating local adaptations. For example, *P. damicornis* from the non-upwelling Gulf of Chiriquí are less sensitive to higher temperatures than those from the upwelling Gulf of Panamá (Glynn & D'Croz 1990; D'Croz & Mate 2004). Similar differences in thermal adaptation among *P. damicornis* populations have also been reported from the Great Barrier Reef (*Ulstrup et al.* 2006). Yet, upwelling can also buffer against critically high sea surface temperatures and reduce coral bleaching and mortality. For example, during the 1997/98 El-Niño, seasonal upwelling in the Gulf of Panamá reduced rising seawater temperatures (mean 27°C), preventing significant coral bleaching and mortality. In contrast, at the same time in the non-upwelling Gulf of Chiriquí, mean seawater temperature rose to 30°C causing severe coral bleaching and mortality (Glynn 2001). While reefs in upwelling environments may thus be better able to weather predicted increases in sea surface temperatures and El Niño frequencies resulting from global climate change (e.g. Trenberth & Hoar 1997), coral populations in non-upwelling habitats seem to be better adapted to elevated temperatures (Glynn & D'Croz 1990; D'Croz & Mate 2004). The genetic structure of Panamanian *P. damicornis* indicates that even among broadcast-spawning coral populations, adaptations to different thermal regimes can occur over distances of ~100km.
Limited inter-regional gene flow in Panamanian *P. damicornis* may also explain prolonged differences in post-disturbance reef recovery between the Gulf of Chiriquí and the Gulf of Panamá. For example, the 1982/83 El Niño caused mass coral mortalities and reduced average coral cover on most Panamanian reefs by 80%, but even though mortality rates were similar in both gulfs, reefs in the Gulf of Panamá had recovered by 1992, whereas reefs in the Gulf of Chiriquí recovered only 3-6% of their populations by 1997 (Glynn 2001). The significant population genetic structure between regions indicates that larval dispersal from recovered Gulf of Panamá populations to the still depauperate Gulf of Chiriquí populations was very limited and this may have prolonged the recovery of reefs in the Gulf of Chiriquí. Moreover, due to the limited gene flow (i.e. larval dispersal) between individual reefs, damaged populations depended strongly on surviving colonies for repopulation. *Acanthaster* predation on surviving colonies in the Gulf of Chiriquí potentially hampered initial recovery significantly, while corallivorous sea stars are absent in the Gulf of Panamá. Another important factor for the differences in *P. damicornis* recovery between Gulfs may be geographic differences in *Symbiodinium* identity (or type). Baker *et al.* (2004) showed that in the Gulf of Panamá, *Pocillopora* corals with clade C bleached severely in 1997 while colonies with clade D were unaffected by bleaching, which lead to a relative increase in colonies with clade D from 43% in 1995 to 63% in 2001. Higher levels of clonality in the Gulf of Panamá (e.g. Saboga) further indicate that asexual fragmentation might have played a more prominent role in the recovery of Gulf of Panamá populations compared to their Gulf of Chiriquí counterparts.

Using multiple genetic markers, Pinzon & LaJeunesse (2011) recently suggested that TEP *Pocillopora* are comprised of three distinct genetic types (Type I - III), independent of species identity and morphology. This is controversial because they included at least seven well-
recognized Indo-Pacific species that must then either be reclassified or synonymized by their genetic types. Their dataset included a limited number of *P. damicornis* samples from the Gulf of Panamá (n = 26), which were comprised of two of the three genetic types (Type I & II). Using a similar set of microsatellites, our STRUCTURE analyses do not support two distinct genetic types, but instead identified five genetic clusters (Figure 1 & Supplementary Figure S1). In addition, Pinzon & LaJeunesse (2011) failed to detect significant genetic structures within their *Pocillopora* type I over 3500km from Mexico to Panamá and the Galápagos. Given that we observe strong genetic structure in *P. damicornis* within Panamá, it seems likely that the absence of geographic structure in Pinzon & LaJeunesse’s (2011) analyses are an artifact of pooling multiple species (i.e. gene pools) for population genetic analyses.

*Comparisons with brooding Indo-West Pacific P. damicornis populations*

Similar levels of population genetic structure over scales of hundreds of kilometers have also been observed in other brooding and broadcast-spawning coral species in the Indo-Pacific and the Caribbean (van Oppen & Gates 2006; Vollmer & Palumbi 2007). In *P. damicornis*, allozyme studies from across the Indo-West Pacific indicate that populations exhibit a range of significant population genetic structures with *F* _ST_ values ranging from 0.06 to as high as 0.39 (Table 5). Some of the highest levels of population structure have been detected in isolated or peripheral populations. For example, *F* _ST_ values among high-latitude *P. damicornis* populations in southeast Australia averaged 0.37 over 700km (Miller & Ayre 2008b) compared to 0.06 among Great Barrier Reef populations over 1200km (Ayre *et al.* 1997). Similarly, *F* _ST_ values among the southernmost *P. damicornis* populations around Lord Howe Island (Miller & Ayre
were twice as high as among similarly spaced populations on the Great Barrier Reef ($F_{ST} = 0.102$ vs. $F_{ST} = 0.044$; Ayre & Miller 2004).

Direct comparisons of population genetic data are possible with two *P. damicornis* microsatellite datasets from peripheral populations in East-African (Souter *et al.* 2009) and core populations in the Indonesian Archipelago (Starger *et al.* 2010) since they used the same or similar subsets of the microsatellite loci used here. Both studies found highly significant population genetic structure. Levels of genetic differentiation among the peripheral East-African populations were lower ($F_{ST} = 0.02$ over 860km) than among TEP populations ($F_{ST} = 0.05$ over 400km). Indonesian populations show the same degree of genetic differentiation as TEP populations ($F_{ST} = 0.05$) over a 10-fold wider geographic distance (3000km vs. 400km). We observed a much lower allelic diversity in the TEP ($N_A/locus = 7$) compared to *P. damicornis* in Indonesia and East-Africa ($N_A/locus = 14.0$ & 13.4, respectively). Lower allelic diversities are expected for peripheral populations and the lower diversity in TEP populations, compared to the similarly peripheral East-African populations, is likely due to the extreme isolation of the TEP and very limited transpacific gene flow ($F_{ST} = 0.419$; Combosch *et al.* 2008). As expected, inbreeding was found to be highly significant in both peripheral populations, in the TEP and in Eastern Africa (Souter *et al.* 2009). Comparisons among *P. damicornis* allozyme studies confirm all three population genetic trends in peripheral populations: 1) higher genetic structure (e.g. Ayre *et al.* 1997 vs. Miller & Ayre 2008b) 2) decreased allelic diversity (Miller & Ayre 2008b) and 3) higher inbreeding (Miller & Ayre 2008b). These patterns have also been observed in peripheral populations of other coral species (Stoddart 1988a; Ayre & Hughes 2004; Noreen *et al.* 2009; Nunes *et al.* 2009).
Our results indicate lower genetic diversity and higher levels of population structure among broadcast-spawning populations (TEP) than among larvae brooding populations (Indo-West Pacific), which may reflect intrinsic differences between the two reproductive strategies. In Indo-West Pacific populations, parthenogenetic larvae brooding preserves genotypes, which reduces the rate of genetic drift and allelic loss, while allelic diversity in TEP populations are likely reduced due to reproductive sweepstakes, promoted by its broadcast-spawning reproductive strategy (Hedgecock 1994; Ruzzante et al. 1996; Christie et al. 2010). Contrary to our results, inter-specific comparisons mostly indicate that brooding corals have reduced dispersal potentials compared to broadcast spawners (Hellberg 1996; Carlon 1999; Ayre & Hughes 2000; Nishikawa et al. 2003). This indicates that population structure in reef corals is not predominantly driven by reproductive strategies. Other factors are likely as important, for example effective population sizes and species-specific larval behaviors (e.g. Palumbi 1995; Miller & Ayre 2008a).

Fine-scale Population Genetic Structure

The observed fine-scale spatial genetic structure (SGS) and the high inbreeding coefficients (average $F_{IS} = 0.18$) indicate that gene dispersal is limited in Panamanian *P. damicornis*, not only between reefs, but within reefs as well. Fine-scale SGS arises from a variety of historic, demographic, and evolutionary processes (Loveless & Hamrick 1984; Epperson 2003; Vekemans & Hardy 2004). Over small spatial scales, SGS is most often the result of limited gene dispersal and can be influenced by a broad array of life-history and reproductive traits including clonality, selfing, inbreeding, fertilization, fecundity, larval life history, and recruitment (Epperson 2003; Vekemans & Hardy 2004; Born et al. 2008). Clonality
can be an important determinant of SGS because clonal aggregations increase the average kinship over short distance intervals (e.g. Stoddart 1988b). We observed fine-scale SGS out to 10m in both, the genet-only and the genet-plus-clone dataset, but since clones were rare (6.2% of samples), clonality had only a small effect on the observed SGS (Figure 3). Instead, the fine-scale SGS observed in TEP *P. damicornis* was predominantly driven by significant non-random relatedness among individual genets out to 10 meters, which is unusual for free-spawning benthic marine populations.

The two most likely explanations for the observed SGS among genets in TEP *P. damicornis* are cohort recruitment (i.e. SGS among larval recruits, independent of parents) or proximity recruitment (i.e. SGS among parents and recruits). In brooding coral species, proximity recruitment is facilitated by the internal (i.e. stationary) larval development and the frequent settlement of larvae within meters of the maternal colony (Gerrodette 1981; Best & Resing 1987; Carlon & Olson 1993). Although broadcast spawning was never directly observed, several independent studies provide convincing evidence that TEP *P. damicornis* do not release brooded larvae but instead free-spawn gametes (Richmond 1985; Glynn *et al.* 1991; Carpizo-Ituarte *et al.* 2010). In broadcast-spawning corals, like TEP *P. damicornis*, larval development occurs externally and larval settlement after gamete release takes at least 1-3 days (Miller & Mundy 2003). Although corals generally spawn during times of minimal water movement (Babcock *et al.* 1986; van Woesik 2009), it is unlikely that drifting coral larvae would stay so close together (cohort recruitment) or remain within meters of the maternal colony (inhibiting proximity recruitment) over several days during their dispersive phase to create such a fine-scale spatial genetic structure (10m).
One potential scenario that could significantly reduce passive dispersal would be if gametes and/or larvae sink down and escape the flow of the water column. Although coral gametes typically float on the sea surface during fertilization and development (Willis & Oliver 1990), gametes of two Pocillopora (P. eydouxi & P. verrucosa; Kinzie 1993) and several other coral species (e.g. Fungia scutaria, F. fungites, Goniastrea favulus) are known to be negatively buoyant (Krupp 1983; Harrison & Wallace 1990; Gilmour 2002; Miller & Ayre 2008a). Negatively buoyant gametes sink down after spawning and fertilize close to the substrate. Fertilization at the substrate also favors mating among nearby colonies, which limits gene dispersal, and should facilitate SGS. Even if fertilization occurs at the sea surface but the zygotes sink down promptly, larval dispersal would still be reduced considerably. A minority of propagules could then remain in the water column to account for long-distance dispersal. Whether TEP P. damicornis gametes or larvae are positively or negatively buoyant is currently unknown, since gamete spawning has not yet been observed directly.²

A second indication for limited gene dispersal within populations is the significant inbreeding coefficients. Inbreeding is a consequence of non-random mating among a limited number of individuals, which limits gene dispersal and promotes SGS (Zhao et al. 2009). The broadcast-spawning reproductive strategy of TEP P. damicornis might facilitate inbreeding in several ways. Firstly, the external fertilization of broadcast-spawners lacks the potential parental control of internal fertilization in brooders (Heyward & Babcock 1986; Willis et al. 1997). Secondly, the high fecundity of broadcast-spawners leads to reproductive sweepstakes, which can promote inbreeding (Heyward & Babcock 1986; Willis et al. 1997). However, heterozygote deficits were as pronounced in larvae-brooding East-African populations (F$_{IS}$ = 0.26), and

² P. damicornis has meanwhile been observed to release negatively buoyant eggs at the GBR by Schmidt et al. (2012)
significant inbreeding is common in *P. damicornis* populations throughout the Indo-Pacific, regardless of reproductive strategy, local environment, population history, and analyzed genetic marker (Table 5). Inbreeding, and potentially SGS, may thus be inherent features of *P. damicornis*.

While SGS is common in plants (reviews by Heywood 1991; Hardy & Vekemans 1999; Vekemans & Hardy 2004), it has only been documented in four other coral species. Underwood *et al.* (2007) described a similar pattern of SGS in the brooding coral *Seriatopora hystrix*, including a significantly elevated kinship among genets in the smallest distance interval (F$_1$ = 20m) and a roughly comparable genetic patch size (80m). Among broadcast-spawning corals, Stoddart (1988b) and Miller and Ayre (2008a) found no significant genet SGS in *Acropora digitifera* and *Goniastrea favulus*, respectively. However, Miller and Ayre (2008a) found significant genet SGS in the broadcast-spawning *Platygyra daedalea*, including significantly positive kinship in the smallest distance interval (F$_1$ = 5m) and a Genetic Patch Size of only 17m. This is the only other documented broadcast-spawning coral with genet SGS. To account for this pattern, Miller and Ayre (2008a) suspected “an unknown element of larval behavior or development” - e.g. that larvae might be negatively buoyant.

Small-scale SGS over tens of meters has been documented for at least eight other sessile marine invertebrates, including two sponges (Calderon *et al.* 2007; Blanquer *et al.* 2009), a temperate soft coral (McFadden & Aydin 1996), a black coral (Miller 1998), a red coral (Costantini *et al.* 2007), a bryozoan (Pemberton *et al.* 2007) and two tunicates (Yund & O'Neil 2000; David *et al.* 2010). For six out of these eight taxa, SGS seems to be driven by limited dispersal of brooded, philopatric larvae. Both broadcast-spawners with significant SGS, the
Anthipatharian coral *Antipathes fiordensis* (Miller 1998) and the solitary tunicate *Styela plicata* (David *et al.* 2010) have negatively buoyant larvae (*A. fiordensis*) or eggs (*S. plicata*).

**Clonal structure**

Because of its large mono-specific stands and little to no detectable larvae recruitment (Richmond 1985), TEP *P. damicornis* reefs were long thought to be highly clonal, persisting locally through vegetative asexual fragmentation (Highsmith 1982; Richmond 1987). Our data indicate that only 6% of corals sampled every 10m apart were clones, i.e. big clones (>10m) were uncommon. In contrast, 94% of samples consist of unique, sexually-derived genets. This indicates that sexual recruitment is sufficient to maintain high levels of genet diversity in TEP *P. damicornis*. Simulation models by Neigel & Avise (1983) indicate that in populations of the Caribbean coral *Acropora cervicornis*, which depends heavily on vegetative fragmentation, 5% sexual recruitment (i.e. 95% asexual) is enough to generate high genotypic diversities (~60%), independent of initial genet diversity.

Compared to other *P. damicornis* studies, levels of clonality in the TEP (6%) seem low. However, direct comparisons among studies are biased by different sampling strategies and genetic markers. Lower levels of clonality tend to be detected with microsatellites compared to allozyme markers (avg. 22% vs. 43%; Table 5) and with large sampling intervals compared to exhaustive samplings (e.g. Souter *et al.* 2009, Starger *et al.* 2010 and the present study vs. Yeoh & Dai 2010). Comparisons among studies using the same marker and similar sampling indicate higher levels of clonality in peripheral populations (TEP, East Africa & Lord Howe Island) compared to more central populations (Indonesia & Great Barrier Reef; Table 5). Higher clonality in peripheral populations has been observed in other corals (Baums *et al.* 2006; Noreen
et al. 2009) and terrestrial plants (Eckert & Barrett 1993; Beatty et al. 2008). Clonal reproduction is considered to enable persistence in the commonly adverse environmental conditions towards range margins, at the expense of sexual reproduction benefits (Herlihy & Eckert 2002). One example of reduced sexual benefits due to asexual reproduction is limited gene flow among and within populations as described here for the peripheral TEP P. damicornis populations.

Conclusion

Regional structure and localized SGS detected in Panamanian Pocillopora damicornis has important consequences for the conservation management of TEP coral reefs. Recent devastating mass mortalities, for example due to El Niño, demonstrate the fragility of TEP coral reefs (Glynn 1985, 1990; Glynn & Colgan 1992; Eakin & Glynn 1996; LaJeunesse et al. 2007). However, regional populations reacted very differently to these large-scale disturbances (Glynn 2001) and are further differently affected by regional disturbances including Acanthaster predation (Glynn 1973, 1988), toxic algal blooms (Guzman et al. 1990) and seasonal upwelling (Glynn & Stewart 1973). In combination with the significant genetic differences among regions, these variations highlight the importance of environmentally, ecologically and genetically defined regional management and conservation units in Panamá and possibly throughout the TEP. Limited gene flow and significant genetic differences between reefs further indicate that local population dynamics on individual reefs are vital for reef resilience and adaptation in Panamá. For example, although the limited gene flow potentially hampers population recovery, it also allows for local and regional adaptation, which should be preserved and accounted for in conservation management.
Acknowledgements

Field and logistic support was provided by Rich Aronson, Ian McIntyre and Carmen Schloeder. Thank you to Sarah Lemer, Liz Hemond, Sean Kent, Silvia Libro, Flavia Nunes and two anonymous reviewers for their comments on the manuscript.
References


Hardy, O.J. & X. Vekemans 2002 SPAGeDi: a versatile computer program to analyse spatial genetic structure at the individual or population levels. *Molecular Ecology Notes* 2: 618-620.


Tables

Table 1: Population genetic indices for the six microsatellite loci.

Indices are based on genets, i.e. unique genotypes, per locus and over all loci (= Total).

N = Number of samples; N_A = Number of alleles per locus; Ho = Heterozygosity observed; He = Heterozygosity expected; F_IS = (He-Ho)/He = Inbreeding coefficient; SE = Standard Error; *p<0.05, **p<0.005 per loci

<table>
<thead>
<tr>
<th>Marker</th>
<th>N</th>
<th>N_A</th>
<th>Ho ± SE</th>
<th>He ± SE</th>
<th>F_IS</th>
</tr>
</thead>
<tbody>
<tr>
<td>Pd2</td>
<td>183</td>
<td>8</td>
<td>0.551 ± 0.035</td>
<td>0.633 ± 0.019</td>
<td>0.239**</td>
</tr>
<tr>
<td>Pd4</td>
<td>161</td>
<td>5</td>
<td>0.292 ± 0.056</td>
<td>0.489 ± 0.044</td>
<td>0.606**</td>
</tr>
<tr>
<td>Pd5</td>
<td>188</td>
<td>9</td>
<td>0.509 ± 0.031</td>
<td>0.641 ± 0.014</td>
<td>0.267**</td>
</tr>
<tr>
<td>Pd6</td>
<td>187</td>
<td>6</td>
<td>0.625 ± 0.077</td>
<td>0.629 ± 0.021</td>
<td>0.077</td>
</tr>
<tr>
<td>Pv2</td>
<td>173</td>
<td>4</td>
<td>0.109 ± 0.033</td>
<td>0.426 ± 0.062</td>
<td>0.831**</td>
</tr>
<tr>
<td>Pv6</td>
<td>172</td>
<td>10</td>
<td>0.597 ± 0.057</td>
<td>0.738 ± 0.020</td>
<td>0.298**</td>
</tr>
<tr>
<td><strong>Total</strong></td>
<td>194</td>
<td>42</td>
<td><strong>0.447 ± 0.032</strong></td>
<td><strong>0.593 ± 0.020</strong></td>
<td><strong>0.367</strong>**</td>
</tr>
</tbody>
</table>
Table 2: Population genetic indices for the nine populations.

Indices are based on genets, i.e. unique genotypes, except for number of samples.
Pop = Population: U = Uva, G = Granito, C = Canal de Afuera, B = Bahia de Damas, A = Achotines, I = Iguana, S = Saboga, Co = Contadora, T = Taboga;
N = Number of samples; N_G = Number of unique genotypes; N_A = Average number of alleles per population; SE = Standard Error; Ho = Heterozygosity observed; He = Heterozygosity expected; F IS = (He-Ho)/He = Inbreeding coefficient; *p<0.005 per pop

<table>
<thead>
<tr>
<th>Region</th>
<th>Pop</th>
<th>N</th>
<th>N_G/N</th>
<th>N_A ± SE</th>
<th>Ho ± SE</th>
<th>He ± SE</th>
<th>F IS ± SE</th>
</tr>
</thead>
<tbody>
<tr>
<td>Gulf of Chiriqui</td>
<td>U</td>
<td>23</td>
<td>0.91</td>
<td>5.2 ± 0.8</td>
<td>0.589 ± 0.069</td>
<td>0.624 ± 0.049</td>
<td>0.062 ± 0.065</td>
</tr>
<tr>
<td></td>
<td>G</td>
<td>21</td>
<td>0.95</td>
<td>5.4 ± 0.9</td>
<td>0.531 ± 0.110</td>
<td>0.687 ± 0.018</td>
<td>0.236 ± 0.154*</td>
</tr>
<tr>
<td></td>
<td>C</td>
<td>18</td>
<td>1.00</td>
<td>5.0 ± 0.9</td>
<td>0.378 ± 0.081</td>
<td>0.640 ± 0.039</td>
<td>0.421 ± 0.103*</td>
</tr>
<tr>
<td></td>
<td>B</td>
<td>30</td>
<td>0.93</td>
<td>4.8 ± 0.5</td>
<td>0.580 ± 0.081</td>
<td>0.613 ± 0.075</td>
<td>0.065 ± 0.026</td>
</tr>
<tr>
<td>Azuero Peninsula</td>
<td>A</td>
<td>23</td>
<td>0.91</td>
<td>4.4 ± 0.7</td>
<td>0.615 ± 0.079</td>
<td>0.607 ± 0.044</td>
<td>-0.048 ± 0.168</td>
</tr>
<tr>
<td></td>
<td>I</td>
<td>24</td>
<td>1.00</td>
<td>4.6 ± 1.0</td>
<td>0.544 ± 0.120</td>
<td>0.605 ± 0.080</td>
<td>0.143 ± 0.159*</td>
</tr>
<tr>
<td>Gulf of Panama</td>
<td>S</td>
<td>22</td>
<td>0.82</td>
<td>4.0 ± 0.3</td>
<td>0.390 ± 0.072</td>
<td>0.575 ± 0.059</td>
<td>0.336 ± 0.087*</td>
</tr>
<tr>
<td></td>
<td>Co</td>
<td>25</td>
<td>0.96</td>
<td>4.6 ± 0.7</td>
<td>0.445 ± 0.093</td>
<td>0.648 ± 0.047</td>
<td>0.314 ± 0.128*</td>
</tr>
<tr>
<td></td>
<td>T</td>
<td>21</td>
<td>0.95</td>
<td>4.2 ± 0.4</td>
<td>0.562 ± 0.057</td>
<td>0.634 ± 0.019</td>
<td>0.114 ± 0.084*</td>
</tr>
<tr>
<td>Total</td>
<td></td>
<td>207</td>
<td>0.94</td>
<td>4.7 ± 0.2</td>
<td>0.515 ± 0.029</td>
<td>0.626 ± 0.016</td>
<td>0.183 ± 0.042*</td>
</tr>
</tbody>
</table>
Table 3: Hierarchical AMOVA results

Hierarchical AMOVA results showing levels of genetic structure among regions (R\textsubscript{RT}/F\textsubscript{RT}), among populations within regions (R\textsubscript{SR}/F\textsubscript{SR}) and among populations (R\textsubscript{ST}/F\textsubscript{ST}).

<table>
<thead>
<tr>
<th>Source of Variation</th>
<th>df</th>
<th>Variance</th>
<th>Variation</th>
<th>Differentiation</th>
<th>p</th>
</tr>
</thead>
<tbody>
<tr>
<td>R-statistics</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Among Regions</td>
<td>2</td>
<td>1331</td>
<td>8.1%</td>
<td>R\textsubscript{RT} = 0.081</td>
<td>&lt; 0.001</td>
</tr>
<tr>
<td>Among Populations</td>
<td>6</td>
<td>1336</td>
<td>8.1%</td>
<td>R\textsubscript{SR} = 0.089</td>
<td>&lt; 0.001</td>
</tr>
<tr>
<td>Within Populations</td>
<td>379</td>
<td>13755</td>
<td>83.8%</td>
<td>R\textsubscript{ST} = 0.162</td>
<td>&lt; 0.001</td>
</tr>
<tr>
<td>F-statistics</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Among Regions</td>
<td>2</td>
<td>0.014</td>
<td>0.8%</td>
<td>F\textsubscript{RT} = 0.008</td>
<td>0.004</td>
</tr>
<tr>
<td>Among Populations</td>
<td>6</td>
<td>0.081</td>
<td>4.5%</td>
<td>F\textsubscript{SR} = 0.045</td>
<td>&lt; 0.001</td>
</tr>
<tr>
<td>Within Populations</td>
<td>379</td>
<td>1.700</td>
<td>94.7%</td>
<td>F\textsubscript{ST} = 0.053</td>
<td>&lt; 0.001</td>
</tr>
</tbody>
</table>
Table 4: Pairwise $R_{ST}$ and $F_{ST}$ between TEP *P. damicornis* populations

Pairwise $R_{ST}$ (above diagonal) and $F_{ST}$ (below diagonal) between TEP *P. damicornis* populations.

*p<0.05; bold = significant after Sequential Bonferroni correction

GC = Gulf of Chiriquí, AP = Azuero Peninsula, GP = Gulf of Panamá; U = Uva, G = Granito, C = Canal de Afuera, B = Bahía de Damas, A = Achetines, I = Iguana, S = Saboga, Co = Contadora, T = Taboga;

<table>
<thead>
<tr>
<th></th>
<th>Gulf of Chiriquí</th>
<th>Azuero Peninsula</th>
<th>Gulf of Panama</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>$R_{ST}$</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>U</td>
<td>G</td>
<td>C</td>
</tr>
<tr>
<td>GC</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>U</td>
<td>0.114*</td>
<td>0.025</td>
<td>0.025</td>
</tr>
<tr>
<td>G</td>
<td>0.035*</td>
<td>0.229*</td>
<td>0.124*</td>
</tr>
<tr>
<td>C</td>
<td>0.033*</td>
<td>0.229*</td>
<td>0.029</td>
</tr>
<tr>
<td>B</td>
<td>0.065*</td>
<td>0.073*</td>
<td>0.047*</td>
</tr>
<tr>
<td>AP</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>A</td>
<td>0.015</td>
<td>0.001</td>
<td>0.037*</td>
</tr>
<tr>
<td>I</td>
<td>0.039*</td>
<td>0.048*</td>
<td>0.025*</td>
</tr>
<tr>
<td>GP</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>S</td>
<td>0.019</td>
<td>0.061*</td>
<td>0.015</td>
</tr>
<tr>
<td>Co</td>
<td>0.025*</td>
<td>0.039*</td>
<td>0.040*</td>
</tr>
<tr>
<td>T</td>
<td>0.087*</td>
<td>0.072*</td>
<td>0.064*</td>
</tr>
</tbody>
</table>
Table 5: *Pocillopora damicornis* population genetic survey (sorted by $F_{ST}$).

1= consistent heterozygote deficits, reported as differences between He & Ho, i.e. inconvertible;
2= two different types of *P. damicornis*;
* = significant; - = not significant;

<table>
<thead>
<tr>
<th>Locations</th>
<th>Scale (km)</th>
<th>$F_{ST}$</th>
<th>$F_{IS}$</th>
<th>Clones</th>
<th>Marker</th>
<th>Reference</th>
</tr>
</thead>
<tbody>
<tr>
<td>GBR (One Tree Island)</td>
<td>5</td>
<td>0.04 *</td>
<td>0.23</td>
<td>18% Allozymes (7 loci)</td>
<td>Ayre &amp; Miller 2004</td>
<td></td>
</tr>
<tr>
<td>GBR (One Tree Island)</td>
<td>1</td>
<td>0.05 *</td>
<td>0.03</td>
<td>-</td>
<td>Allozymes (5 loci)</td>
<td>Benzie et al. 1995</td>
</tr>
<tr>
<td>GBR (One Tree Island)</td>
<td>5</td>
<td>0.06 *</td>
<td>0.11 *</td>
<td>17% Allozymes (8 loci)</td>
<td>Sherman et al. 2006</td>
<td></td>
</tr>
<tr>
<td>GBR</td>
<td>1200</td>
<td>0.06 *</td>
<td>0.28 *</td>
<td>20% Allozymes (7 loci)</td>
<td>Ayre et al. 1997</td>
<td></td>
</tr>
<tr>
<td>Japan (Okinawa)</td>
<td>650</td>
<td>0.06 *</td>
<td>0.18 *</td>
<td>39% Allozymes (7 loci)</td>
<td>Adjeroud &amp; Tsuchiya 1999</td>
<td></td>
</tr>
<tr>
<td>Northwest Australia</td>
<td>900</td>
<td>0.08 *</td>
<td>0.18</td>
<td>85% Allozymes (6 loci)</td>
<td>Whitaker 2006</td>
<td></td>
</tr>
<tr>
<td>Lord Howe Island</td>
<td>10</td>
<td>0.10 *</td>
<td>0.17</td>
<td>28% Allozymes (7 loci)</td>
<td>Miller &amp; Ayre 2004</td>
<td></td>
</tr>
<tr>
<td>Lord Howe Island to GBR</td>
<td>2500</td>
<td>0.15 *</td>
<td>N/A</td>
<td>N/A Allozymes (7 loci)</td>
<td>Ayre &amp; Hughes 2004</td>
<td></td>
</tr>
<tr>
<td>Southeast Australia to GBR</td>
<td>1200</td>
<td>0.24 *</td>
<td>0.44</td>
<td>49% Allozymes (8 loci)</td>
<td>Miller &amp; Ayre 2008b</td>
<td></td>
</tr>
<tr>
<td>Southeast Australia</td>
<td>700</td>
<td>0.32 *</td>
<td>0.47</td>
<td>42% Allozymes (8 loci)</td>
<td>Miller &amp; Ayre 2008b</td>
<td></td>
</tr>
<tr>
<td>Southwest Australia</td>
<td>400</td>
<td>0.39 *</td>
<td>N/A</td>
<td>82% Allozymes (4 loci)</td>
<td>Stoddart 1984</td>
<td></td>
</tr>
<tr>
<td>Hawaii (Kaneohe Bay)</td>
<td>10</td>
<td>N/A</td>
<td>N/A</td>
<td>71% Allozymes (4 loci)</td>
<td>Stoddart 1986</td>
<td></td>
</tr>
<tr>
<td>East Africa</td>
<td>860</td>
<td>0.02 *</td>
<td>0.26 *</td>
<td>13% Microsats (6 loci)</td>
<td>Souter et al. 2009</td>
<td></td>
</tr>
<tr>
<td>Indonesia</td>
<td>3300</td>
<td>0.05</td>
<td>N/A</td>
<td>2% Microsats (9 loci)</td>
<td>Starger et al. 2010</td>
<td></td>
</tr>
<tr>
<td><strong>Tropical Eastern Pacific</strong></td>
<td><strong>400</strong></td>
<td>**0.05 ***</td>
<td>**0.18 ***</td>
<td><strong>6% Microsats (5 loci)</strong>*</td>
<td><strong>THIS STUDY</strong></td>
<td></td>
</tr>
<tr>
<td>Taiwan (Nanwan Bay)</td>
<td>0.5</td>
<td>N/A</td>
<td>N/A</td>
<td>66% Microsats (7 loci)</td>
<td>Yeoh &amp; Dai 2010</td>
<td></td>
</tr>
<tr>
<td>TEP to Central-West Pacific</td>
<td>8000</td>
<td>0.42 *</td>
<td>N/A</td>
<td>N/A ITS2-5.8S</td>
<td>Combosch et al. 2008</td>
<td></td>
</tr>
</tbody>
</table>
Figure 1: Map of Panamá, showing the three main Pacific regions and the sampled locations.

The pie charts show the distribution of the five genetic clusters identified by STRUCTURE 2.3 in the three regional and nine reef populations. The bar plot shows each individual sample as a vertical bar with colors indicating the relative proportions of each genetic cluster. Map created on www.aquarius.ifm-geomar.de.

U = Uva Island; C = Canal de Afuera Island; G = Granito de Oro Island; B = Bahia de Damas, Coiba Island; A = Achotines Bay; I = Iguana Island; S = Saboga Island; Co = Contadora Island; T = Taboga Island.
Figure 2: Principal component analysis of *Pocillopora damicornis* populations, constructed using GenAlEx 6.1.

Distances between populations were analyzed as RST pairwise genetic distances. The two axes explained 90.4% of the total variation. Blue circles = Populations in the Gulf of Chiriquí; Red triangles = Azuero Peninsula; Green squares = Gulf of Panamá.
Figure 3: Spatial correlograms of the SpaGeDi analysis of Spatial Genetic Structure (SGS) among specimen within discrete distance classes.

**Figure 3a)** Spatial correlogram of the average pairwise kinship $F_{ij}$ as a function of distance over all populations.

**Figure 3b-d)** Spatial correlograms over the populations in each of the three regions, b) the Gulf of Chiriquí, c) the Azuero Peninsula and d) the Gulf of Panamá.

Black lines represent the genet dataset, excluding clones and grey lines show the results for the ramet dataset, including clones. Dashed lines are the permuted 95% confidence intervals. Error bars are 95% bootstrap errors of the genet datasets.
Supplementary Material

Figure S1: Results of the Bayesian clustering approach implemented in STRUCTURE 2.3 that were used to infer the number of genetic clusters (K) in the microsatellite dataset.

The graph shows the mean log likelihood L(K) for K = 1-19 (±SD) and ΔK, the second order rate of change of L(K) for K = 2–18. The most likely number of cluster is K = 5, since it has the highest likelihood value (L(K) = -2059) and a local maximum for ΔK.
Chapter 3
MIXED ASEXUAL AND SEXUAL REPRODUCTION IN THE INDO-PACIFIC REEF CORAL POCILLOPORA DAMICORNIS

Abstract

Pocillopora damicornis is one of the best-studied reef-building corals, yet its somewhat unique reproductive strategy remains poorly understood. Genetic studies indicate that P. damicornis larvae are produced almost exclusively parthenogenetically, and yet population genetic surveys suggest frequent sexual reproduction. Using microsatellite data from over 580 larvae from 13 colonies, we demonstrate that P. damicornis displays a mixed reproductive strategy where sexual and asexual larvae are produced simultaneously within the same colony. The majority of larvae were parthenogenetic (94%), but most colonies (10 out of the 13) produced a subset of their larvae sexually. Logistic regression indicates that the proportion of sexual larvae varied significantly with colony size, cycle day and calendar day. In particular, the decrease in sexual larvae with colony size suggests that the mixed reproductive strategy changes across the life of the coral. This unique shift in reproductive strategy leads to increasingly asexual replications of successful genotypes, which (in contrast to exclusive parthenogens) have already contributed to the recombinant gene pool.

Introduction

Evolutionary theory predicts that sex is beneficial because recombination generates new, potentially advantageous gene combinations (Weismann 1886; Fisher 1930; Muller 1932) and helps to purge deleterious alleles (Muller 1964; Kondrashov 1988) by exposing them to natural selection (Hill & Robertson 1966; Felsenstein 1974). In contrast, asexual reproduction propagates successful genotypes and does not require mates (e.g. Weismann 1886; Williams 1975; Holman 1987; Radtkey et al. 1995; Weider et al. 1999). Most asexual lineages depend on occasional sexual recombination to add variation to their genomes, purge deleterious mutations, and adapt to changing environments. A mixed strategy of asexual propagation with low levels of sex combines the best of both strategies (Hurst & Peck 1996). For example, sperm-dependent parthenogenesis allows for recombination through occasional leakage of parental DNA from cons- or hetero-specific sexual progenitors (Beukeboom & Vrijenhoek 1998). Sexual and parthenogenetic reproduction are usually well separated by time (e.g. in cyclic parthenogens) or population (geographic parthenogens) or produce different offspring (e.g. in many insects).

Most reef-building corals reproduce sexually and asexually. Sexual reproduction is achieved by either free spawning eggs and sperm, or internally brooding larvae inside the coral polyp (Harrison 2011). The resulting planula larvae then disperse, settle, and metamorphose to form a new coral polyp. In colonial corals, the coral colony then grows by asexual fission of individual polyps. Some corals can reproduce asexually via vegetative fragmentation (Highsmith 1982). In a handful of coral species, genetically identical coral colonies can also be formed through the production of parthenogenetic larvae (Harrison 2011). Parthenogenesis was first documented in the Indo-Pacific coral *Pocillopora damicornis* (Stoddart 1983), where it has been confirmed repeatedly using allozyme data (Stoddart 1986; Ayre & Miller 2004; Sherman et al. 2007).
2006). None of these studies detected any sign of sexual reproduction, even though population genetic surveys suggest frequent sexual recombination amidst predominantly asexual reproduction (e.g. Ayre & Miller 2004). Recent microsatellite analyses indicate that two *P. damicornis* colonies in Taiwan produced a subset (6 and 29%) of their larvae sexually (Yeoh & Dai 2010), suggesting that *P. damicornis* can simultaneously produce asexual and sexual larvae. The exact origin of the larvae remains unclear. Several authors hypothesized that asexual larvae originate via budding, analogous to polyp budding (Muir 1984; Stoddart & Black 1985; Ward 1992), however, histological observations of gametogenesis and embryogenesis suggest that parthenogenetic larvae are the result of vegetative embryogenesis in diploid oocytes (Harriott 1983; Martin-Chavez 1986; Diah-Permata *et al.* 2000).

The evolutionary significance of parthenogenetic reproduction in *P. damicornis* remains unknown. The extent and patterns of parthenogenetic larvae production among colonies are poorly defined and nothing is known about the potentially different roles and fates of sexual vs. parthenogenetic larvae. Using microsatellite genotyping, we quantified the extent of sexual and asexual reproduction across multiple colonies of *P. damicornis*. We examined whether colony genotype, size or habitat are associated with the production of parthenogenetic and/or sexual larvae. In addition, we asked if parthenogenetic and/or sexual larvae are preferentially released at particular times during the lunar and/or reproductive cycle, to characterize potential differences (e.g. dispersion vs. population maintenance - sensu Stoddart 1988).
Materials and Methods

Organisms and larvae collection methods

In January 2010, 23 *P. damicornis* colonies (Figure 1) were collected from fringe and back reef sites on the north shore of Moorea, French Polynesia. Colonies were transported to the Gump Research Station and kept in large seawater tanks (~5m$^3$). A small nubbin (1cm) was sampled from each colony and preserved in GITC-buffer for genetic analysis. Three days before the new moon (January 15$^{\text{th}}$ and February 13$^{\text{th}}$), colonies were transferred to individual containers with continuous seawater inflow. The outflow of each container was diverted into a short PVC tube with a mesh filter (100µm) where the larvae accumulated during planulation (Figure 1). Larvae were collected using a 1000µl micropipette and preserved individually in 200µl Guanidinium-Isothiocyanate (GITC) buffer (Fukami et al. 2004). Between January 17 and 20, colonies were monitored throughout the night from 1900h to 0600h and larvae were collected and preserved as frequently as possible (i.e. every two to three hours). In February 2010, planulation lasted from February 14 to 26, i.e. until two nights before the full moon, and larvae were collected only in the morning.

Parent colonies

*Pocillopora damicornis* were collected from fringing and back reef habitats to assess potential environmental influences on reproductive strategies. *Pocillopora damicornis* is very abundant on fringe reefs around Moorea while on back reefs it is restricted to wave-protected grooves and crevices.
Colony size is often related to colony age and fecundity (Harrison 2011). Its impact on reproduction was measured in two ways. Colony size (as plane surface area, based on the longest and shortest colony diameter) was included in the main logistic regression analyses as a continuous variable. In addition, a separate comparison of larvae production between large and small colonies was conducted. Small colonies (n=7) were defined as having a maximum diameter of less than 10 cm and a surface area under 75 cm$^2$, while large colonies (n=4) had minimum diameter over 10 cm and a surface area over 130 cm$^2$ (Table 2).

Eighteen out of the 23 P. damicornis colonies collected released each between 1 and over 1000 larvae during our observations following the new moon in January and February 2010. In order to ensure reliable characterization of each colony’s reproductive patterns, only colonies that released more than 10 larvae (n = 13) were included in the analyses presented here.

Genetic analyses

DNA was extracted with Agencourt DNAdvance® isolation kits (Beckman Coulter Inc.). Ten microsatellite loci were amplified with primers developed by Magalon et al. (2004) and Starger et al. (2007). PCR products included infrared labels on forward primers, and were manually sized and scored on a LICOR 4300 Genetic Analyzer. Six loci amplified consistently using modified PCR protocols (Combosch & Vollmer 2011) and were used in all subsequent analyses (Table 1). Only larvae that were genotyped successfully at five or more loci were included in the analyses. Linkage disequilibrium among loci was assessed using a Markov Chain Monte Carlo resampling method (Guo & Thompson 1992) in GENEPOP (Raymond & Rousset 1995), followed by sequential Bonferroni correction to account for multiple comparisons.
Probability of Identity (PI) and the inbreeding coefficient $F_{IS}$ were calculated using GenAlEx 6.4.1 (Peakall & Smouse 2006).

Larvae with genotypes identical to their parent were considered parthenogenetically produced. Larvae with genotypes that differed from the maternal genotype were considered sexually produced. Sexually produced larval genotypes were examined for signs of selfing (i.e. absence of alleles not already present in the maternal genotype).

**Data analyses**

Logistic regression analyses were used because the data are dichotomous (i.e., sexual vs. asexual) and binomially distributed (Wilson & Hardy 2002; Warton & Hui 2011). The main logistic regression analysis was conducted using a generalized linear mixed effects model with the proportion of sexually produced larvae per colony per day as the response variable. Six different predictor variables were included as fixed effects in the model. Three factors were colony-related, time-independent factors: colony genotype, colony size, and colony habitat. Three factors were time-related, colony-independent factors: calendar day, lunar day, and release day. Lunar day and release day were used to assess daily patterns of larval release across reproductive cycles (i.e. January and February 2010). Lunar day refers to days after the new moon while release day refers to the day after a specific colony started to release larvae (i.e. days into the larval-release cycle of a colony). “Colony ID” was treated as a random effect to determine whether the results were consistent across colonies.

A second logistic regression analysis was used to compare the proportion of sexually produced larvae between small and large colonies. The same type of generalized linear mixed effects model was used with the two size categories as predictor variables. All statistical analyses
were conducted in the program R version 2.14.0 (R Core Team 2012) using the glmer function in the lme4 package (Bates et al. 2009).

Results

Over 3000 larvae were released by the 18 Pocillopora damicornis colonies. One colony (P10), released more than 1000 larvae over just three nights in January 2010. In total, 960 larvae were collected and preserved in GITC buffer for microsatellite genotyping; 597 larvae were successfully genotyped at 5 to 6 microsatellite loci (Table 1). Data analyses were conducted on 583 larvae that were released by 13 colonies, which each released 10 or more larvae (Table 2).

The six microsatellite loci had between 2 and 6 alleles (Table 1). No significant linkage disequilibria (LD) were detected among colony genotypes (p > 0.05 after Sequential Bonferroni correction). The probability of identity (PI) to identify unique genotypes using all six loci was 3.9x10^{-4} (Table 1).

Out of the 583 genotyped larvae, 546 larvae (93.7%) had identical genotypes to their parent and were scored as parthenogenetic larvae. The genotypes of the remaining 37 larvae (6.3%) differed from the parental genotype at one allele over one or more loci and were scored as sexual larvae. Eight of the 37 sexual larvae could have been produced via meiotic selfing; seven of these larvae had identical siblings, but differed from their brood parent’s genotypes.

Ten out of the 13 genotyped P. damicornis colonies released both sexually and asexually produced larvae (Table 2). All ten produced mostly parthenogenetic larvae (>70%). The other three colonies produced only parthenogenetic larvae (see Table 2). Two out of these three coral colonies had identical genotypes (P8 & P9). However, a third colony had the identical genotype
as well (P6) and produced some larvae sexually. This suggests that parthenogenesis is influenced by environmental, demographic and genetic factors.

Reproduction changes with colony size and over time

Logistic regression analyses revealed significant effects of Colony Size, Cycle Day and Calendar Day on the proportions of sexually produced larvae (Table 3). Colony Genotype, Colony Habitat, and Lunar Day were not statistically significant (p > 0.05).

The proportion of sexually produced larvae declined significantly with colony size (LRT, p = 0.029; Figure 2). Moreover, small colonies produced significantly more larvae sexually (~10.5%) than large colonies (~2%; Table 2 and Figure 2; single factor LRT: p = 0.02). The proportion of sexually produced larvae also declined significantly over the course of each colony’s larval release cycle (Cycle Day; Figure 3; LRT, p = 0.004). However, the proportion of sexually produced larvae increased significantly over the two month monitoring period (Calendar Day; LRT, p = 0.040), e.g. from January (7/223; 3.1%) to February (30/360; 8.3%).

Discussion

Microsatellite data from 583 coral larvae and their 13 brood-parents indicate that *Pocillopora damicornis* in Moorea use a simultaneously mixed sexual and asexual reproductive strategy. Over two reproductive cycles, ten out of the thirteen surveyed *P. damicornis* colonies produced both sexual and parthenogenetic larvae. All colonies produced on average 87% of their larvae parthenogenetically; three colonies produced only parthenogenetic larvae. The proportion of sexual reproduction decreased significantly with increasing colony size and during each colony’s individual larval release cycle. Yet, over the bimonthly monitoring period, sexual
reproduction increased. Parental genotype, habitat and lunar day did not have significant effects on the proportion of sexual larvae. While the exact mechanisms of this strategy remain unknown, it appears that *P. damicornis* applies a minimal sex strategy, dominated by asexual parthenogenesis.

Previous genetic analyses of *P. damicornis* did not detect any sexually produced larvae (Stoddart 1983; Stoddart 1986; Ayre & Miller 2004; Sherman *et al.* 2006). Sexually produced larvae might have been missed due to their low frequency (~6%), or mis-identified as parthenogenetic due to the low resolution of allozyme markers. In contrast, allozyme-based population genetic surveys of *P. damicornis* provided evidence for sexual recombination (i.e. high genotypic diversity) in *P. damicornis* populations across the Central-West Pacific (Western Australia: Stoddart 1984a, b and Whitaker 2006; Great Barrier Reef: Benzie *et al.* 1995, Ayre *et al.* 1997, Ayre & Miller 2004 and Sherman *et al.* 2006; Lord Howe Island: Miller & Ayre 2004; Okinawa: Adjeroud & Tsuchiya 1999; and potentially in Hawaii: Stoddart 1986). Histological studies of *P. damicornis* in Taiwan and Hawaii also suggested regular embryogenesis and found no indications of asexual budding (Martin-Chavez 1986; Diah-Permata *et al.* 2000). Recently, Yeoh and Dai (2010) used multi-locus microsatellite markers to compare larvae of two *P. damicornis* colonies in Taiwan and documented asexual and sexual larvae release. Both histologic and indirect and direct genetic studies thus indicate that *P. damicornis* reproduces mixed, sexually and asexually, throughout the Central-West Pacific.

*Mixed reproduction*

One of the most intriguing characteristics of the mixed reproductive strategy of *P. damicornis* is the observed simultaneous release of parthenogenetic and sexual larvae.
Simultaneous mixed reproduction is rare in animals. In most animals that display mixed reproduction, sexual and asexual reproduction are separated by different conditions. In cyclical parthenogens such as monogonont rotifers, aphids and *Daphnia*, sexual reproduction is limited to particular seasons, environmental conditions and/or population densities (e.g. Hebert 1974; Gómez & Carvalho 2001; Simon *et al.* 2002). In geographic parthenogens like squamate lizards and grasshoppers, asexual reproduction occurs in isolated populations, at particular high altitudes and latitudes or at the margins of species ranges (Vrijenhoek & Parker 2009). Animals that reproduce simultaneously sexually and asexually generally produce parthenogenetic offspring that are either haploid (e.g. hymenoptera) or predominantly homozygous (e.g. cockroaches and termites). *Pocillopora damicornis* are somewhat distinct in that sexual and apomictic parthenogenetic reproduction occurs simultaneously within the same coral colony.

Recent histological examinations clearly indicate that *P. damicornis* larvae result from oocytes following regular embryogenesis (Diah-Permata *et al.* 2000). Since parthenogenetic larvae have identical genotypes, including several heterozygotic loci, they must stem from apomictic diploid eggs. But haploid eggs are required for sexually produced larvae. This suggests that both haploid and diploid eggs would have to be produced; yet no indications of distinct types of oocytes and/or eggs have ever been reported in *P. damicornis*.

Alternatively, eggs might be generally diploid and the maternal chromosome constitution is retained or restored post-fertilization in parthenogenetic larvae. For example, in pseudogametic taxa, parthenogenetic reproduction requires sperm to stimulate embryogenesis without incorporating the parental DNA. Frequently, however, sperm nuclei evade complete degeneration and chromosome fragments, complete chromosome or entire sperm genomes ‘leak’ into the parthenogenetic eggs (Schartl *et al.* 1995; D'Souza *et al.* 2006; D’Souza & Michiels
The paternal chromosomes are then either added to the genome (chromosome addition) or the maternal ploidy level is restored by removing a maternal chromosome set. Thus, a predominantly apomorphic parthenogenetic clutch would contain a few genetically diverse propagules with small to moderate genotypic differences (D’Souza & Michiels 2009), which is exactly what we observed in *P. damicornis*.

Pseudogamy is rare, but occurs in a wide variety of taxa. It is more common among hybrid lineages, but non-hybrid pseudogametic taxa are also known (e.g. platyhelminthes, Beukeboom & Vrijenhoek 1998). Since parental leakage occurs regularly in pseudogametic taxa, it has been termed “minimal sex” (Beukeboom & Vrijenhoek 1998) or “cryptic sex” (Schlupp 2005).

There are several striking similarities between the genetic patterns of mixed reproduction in *P. damicornis* and in the pseudogametic turbellarian *Schmidtea polychroa* (Platyhelminthes). Two microsatellite studies of *S. polychroa* (D’Souza et al. 2004; D’Souza et al. 2006) report similar proportions of sexual offspring per parent (2-25% compared to 2-30% here) and across parents (6% each). Additionally, the number of loci affected by parental leakage (1-4 out of 4) and the relative prevalence of parental leakage among fertile parents (6/11) is very similar to our results (1-5 out of 6 loci and 10/13 fertile parents). Pseudogamy also explains the remarkable synchrony of asexual and sexual reproduction.

A common consequence of parental leakage in pseudogametic taxa is temporary polyploidy (Beukeboom & Vrijenhoek 1998). In *P. damicornis* larvae, polyploid allozyme genotypes have been described previously (Stoddart 1986) and several sexually produced larvae had multiple alleles at individual loci in this study (i.e. two maternal and one paternal). However, in other cases sexual reproduction led to a decrease in the number of observed alleles, from
heterozygous in the colony to homozygous in the larvae. The reproductive patterns observed in *P. damicornis* thus match patterns described in pseudogametic taxa. Because most genotypic distances between maternal and larval genotypes were small (one or two loci), it is possible that somatic mutations and/or chimeric colonies (reviewed in van Oppen *et al.* 2011) may have contributed to the signature of sexual reproduction.

**Evolutionary significance**

Mixed reproduction is often described as a best-of-both-worlds scenario (Hurst & Peck 1996; Peck & Waxman 2000). It enables effective propagation of successful genotypes via parthenogenesis while sexual recombination constantly generates offspring with new allele combinations for selection. The extent of these benefits depends on the frequency of sex. In many cases a predominantly asexual strategy with a low frequency of sex seems to be most effective (Hurst & Peck 1996; Peck & Waxman 2000).

Mixed reproduction conveys many evolutionary advantages. In organisms using mixed reproductive strategies, natural selection acts not only on the level of the entire genome, as in self-replicating asexual taxa, but also on the level of individual genes, which are exposed to natural selection due to occasional sexual recombination. Moreover, the extent of asexual and sexual reproduction could be fine-tuned among genotypes so that successful genotypes are replicated mostly intact while less successful genotypes are more involved in sexual recombination.

Mixed reproduction can help organisms adapt to changing environments. The extent of asexual and sexual reproduction may vary across environments. If environmental conditions deteriorate and organisms become stressed, mixed reproducers might be capable of increasing
the proportion of sexual reproduction. Increased sexual recombination would generate a higher proportion of novel genotypes and a more diverse cohort of offspring to deal with the new condition.

Another benefit of mixed reproduction is reproductive assurance. Reproductive failure because of sperm limitation is a serious problem for many broadcast-sperm dependent marine invertebrates (Yund 2000). Parthenogenesis assures reproduction independent of fertilization success. Moreover, if unfertilized eggs of pseudogametic taxa are capable of eventually developing into parthenogenetic larvae without pseudo-fertilization, a second tier of parthenogenetic larvae would start to develop and consequently be released slightly delayed. This is consistent with the observed decrease in the proportion of sexual larvae during the reproductive cycle of individual colonies.

Sexual reproduction decreases with increasing colony size

Among the most remarkable characteristics of *P. damicornis* mixed reproduction is that bigger colonies reproduce more asexually than smaller colonies (Figure 2). Because the proportion of sexual reproduction was genotype and habitat independent, the decrease of sexual reproduction with colony size appears to be a general life-history strategy. Colony size is an important proxy for a colony’s fecundity and evolutionary fitness (Hughes *et al*. 1992). Correlations between colony size and reproductive strategy therefore have important evolutionary consequences. Large colonies affirm the success of their genotypes in local environments. The shift towards increasingly asexual reproduction by larger colonies could thus lead to increased recruitment and survival of these successful genotypes in larval cohorts. However, in contrast to fully parthenogenetic taxa, successful brood parents will have already
contributed “good alleles” to the gene pool when they were small. Over time, this reproductive strategy should lead to a highly inbred population structure with moderately high genotypic diversity (due to ongoing sexual recombination). The genotype distribution should consist of a few successful genotypes, present in multiple colonies (but particularly dominant among big colonies) and many subordinate genotypes. This is exactly what has been observed in population genetic and demographic studies, not only in French Polynesia (Marks et al., in prep), but also in other *P. damicornis* populations across the Central-West Pacific (e.g. Stoddart 1984b; Whitaker 2006).

*Sexual reproduction decreases during colonies’ reproductive cycle*

Sexually produced larvae were released significantly earlier during the reproductive cycle of individual colonies than asexually produced larvae. One possible reason is that unfertilized larvae eventually develop into parthenogenetic larvae, as outlined above (reproductive assurance hypothesis). Alternatively, sexually produced larvae might simply develop faster than their asexual counterparts, as is commonly observed in other taxa, for example in termites (Matsuura 2010) and cockroaches (Corley & Moore 1999). Parental factors might stimulate the development of sexual offspring, as in *Drosophila melanogaster* (Yasuda *et al.* 1995) and *Caenorhabditis elegans* (Browning & Strome 1996). Faster development of sexual offspring could also be mediated genetically, e.g. via heterosis. Ultimately, sexually produced larvae might simply be preferred by the mother, as a means to allocate resources in a way that maximizes overall offspring fitness (Michiels *et al.* 1999).
Poly-embryos

Three different colonies (P3, P6 & P17) released sexually produced, but genetically identical siblings (two pairs of twins and one triplet). One pair of genetically identical, sexually produced *P. damicornis* larvae had also been observed in Taiwan (Yeoh & Dai 2010). Genetically identical, sexually produced twins result from a single zygote that eventually divides into two or more clonal poly-embryos, akin to monozygotic twins in humans. This form of limited poly-embryony is considered accidental and does not seem to bear much evolutionary significance (Craig *et al.* 1997). While most *P. damicornis* polyps tend to produce only one larva at the time, multiple larvae per polyp are not uncommon (e.g. Stoddart & Black 1985; Martin-Chavez 1986). Moreover, viable settlement-competent poly-embryos are known from other coral species, where free-swimming embryos break apart and multiply (Heyward & Negri 2012).

Conclusion

The results of this study in combination with Yeoh and Dai (2010), and several population genetics studies throughout the Central-West Pacific, showed that asexual and sexual larvae brooding is the predominant reproductive strategy of *P. damicornis* populations in the Central-West Pacific. Populations at the margins of the species distribution range reproduce by broadcast spawning, *e.g.* at the Southern GBR (Schmidt-Roach *et al.* 2012), in the Tropical Eastern Pacific (Glynn *et al.* 1991; Chávez-Romo & Reyes-Bonilla 2007) and in Western Australia (Stoddart & Black 1985; Ward 1992), although taxonomic uncertainties exist for some of these peripheral populations (e.g. Schmidt-Roach *et al.* 2013). Multiple *P. damicornis* types with reproductive differences have been confirmed in Hawaii (Richmond & Jokiel 1984) and inferred in Western Australia (Ward 1995) and the Western Indian Ocean (Souter 2010). In
addition, inter-specific and inter-generic hybridization was inferred for peripheral *P. damicornis* populations in the Tropical Eastern Pacific (Combosch *et al.* 2008) and at Lord Howe Island (Miller & Ayre 2004). In light of the confusing, sometimes contradictory results of allozyme-based studies, it is now clear that more high-resolution data (microsatellites, SNPs) are needed to assess the spatial and temporal variability of sexual reproduction and the impact of environmental variation on the reproductive strategy of *P. damicornis*. Eighty years after the first description of the reproductive strategy of *P. damicornis* (Marshall & Stephenson 1933) a lot of basic questions remain to be answered.

Acknowledgements

Field and logistics support was provided by Pete Edmunds, Gonzalo Giribet, Sal Genovese and the Three Seas Program. Statistical advice was provided by Sean Kent and Tarik Gouhier. DJC was supported by a Three Seas Teaching Fellowship and a German Academic Exchange Service (DAAD) Pre-doctoral Research Fellowship. Thank you to Sarah Lemer, Sarah Gignoux-Wolfsohn, Christopher Marks, and two anonymous reviewers for their comments on the manuscript.
References


Muir P. 1984 *Periodicity and asexual planulae production in Pocillopora damicornis (Linnaeus) at Magnetic Island* BSc Honours thesis, James Cook University of North Queensland, Townsville.


Stoddart J. 1984a Genetic differentiation amongst populations of the coral *Pocillopora damicornis* off Southwest Australia. *Coral Reefs* **3**: 149-156.


Weismann A. 1886 *Die Bedeutung der sexuellen Fortpflanzung für die Selektions-Theorie* Gustav Fischer Verlag, Jena, Germany.


Williams G. 1975 *Sex and evolution* Princeton University Press, Princeton, NJ.


### Tables

**Table 1: Loci-specific statistics for each of the six sampled microsatellite loci.**


\[ n = \text{Number of samples}; \quad N(A) = \text{Number of Alleles}; \quad \text{PI} = \text{Probability of Identity}; \quad n, N(A) \text{ and PI are based on the full dataset}; \quad \text{Sex} = \text{Number of times the locus varied in sexual larvae}; \quad F_{IS} = \text{Inbreeding coefficient among unique colony genotypes}; \quad * = \text{significant deviation from HWE (} p < 0.05) ; \]

<table>
<thead>
<tr>
<th>Locus</th>
<th>n</th>
<th>N(A)</th>
<th>PI</th>
<th>Sex</th>
<th>F_{IS}</th>
</tr>
</thead>
<tbody>
<tr>
<td>Pd1</td>
<td>585</td>
<td>2</td>
<td>0.38</td>
<td>5</td>
<td>-0.54 *</td>
</tr>
<tr>
<td>Pd2</td>
<td>591</td>
<td>3</td>
<td>0.21</td>
<td>14</td>
<td>-0.42 *</td>
</tr>
<tr>
<td>Pd4</td>
<td>558</td>
<td>3</td>
<td>0.20</td>
<td>12</td>
<td>-0.29 *</td>
</tr>
<tr>
<td>Pd5</td>
<td>592</td>
<td>3</td>
<td>0.34</td>
<td>9</td>
<td>-0.36</td>
</tr>
<tr>
<td>Pd6</td>
<td>593</td>
<td>3</td>
<td>0.47</td>
<td>16</td>
<td>0.10</td>
</tr>
<tr>
<td>Pv6</td>
<td>545</td>
<td>6</td>
<td>0.15</td>
<td>11</td>
<td>-0.16</td>
</tr>
<tr>
<td>Total</td>
<td>596</td>
<td>20</td>
<td>3.9 \times 10^{-4}</td>
<td>67</td>
<td>-0.28</td>
</tr>
</tbody>
</table>
Table 2: List of maternal colonies

List of maternal colonies, including colony size, the total number of genotyped larvae per colony (# Larvae), the number and the percentage of sexually produced larvae per colony (% and # Sex, respectively). Colonies with identical genotypes have the same grey background (P3 & P5; P6, P8 & P9).

<table>
<thead>
<tr>
<th>Parent</th>
<th>Size [cm²]</th>
<th># Larvae</th>
<th># Sex</th>
<th>% Sex</th>
</tr>
</thead>
<tbody>
<tr>
<td>P3</td>
<td>67</td>
<td>27</td>
<td>4</td>
<td>14.8%</td>
</tr>
<tr>
<td>P5</td>
<td>53</td>
<td>23</td>
<td>1</td>
<td>4.3%</td>
</tr>
<tr>
<td>P6</td>
<td>74</td>
<td>10</td>
<td>2</td>
<td>20.0%</td>
</tr>
<tr>
<td>P8</td>
<td>139</td>
<td>55</td>
<td>0</td>
<td>0%</td>
</tr>
<tr>
<td>P9</td>
<td>69</td>
<td>45</td>
<td>0</td>
<td>0%</td>
</tr>
<tr>
<td>P10</td>
<td>163</td>
<td>165</td>
<td>4</td>
<td>2.4%</td>
</tr>
<tr>
<td>P11</td>
<td>130</td>
<td>71</td>
<td>2</td>
<td>2.8%</td>
</tr>
<tr>
<td>P12</td>
<td>72</td>
<td>56</td>
<td>1</td>
<td>1.8%</td>
</tr>
<tr>
<td>P17</td>
<td>60</td>
<td>46</td>
<td>12</td>
<td>26.1%</td>
</tr>
<tr>
<td>P18</td>
<td>25</td>
<td>41</td>
<td>6</td>
<td>14.6%</td>
</tr>
<tr>
<td>P20</td>
<td>174</td>
<td>14</td>
<td>0</td>
<td>0%</td>
</tr>
<tr>
<td>P23</td>
<td>n/a</td>
<td>10</td>
<td>3</td>
<td>30.0%</td>
</tr>
<tr>
<td>P28</td>
<td>n/a</td>
<td>20</td>
<td>2</td>
<td>10.0%</td>
</tr>
<tr>
<td><strong>Total</strong></td>
<td><strong>60</strong></td>
<td><strong>583</strong></td>
<td><strong>37</strong></td>
<td><strong>6.3%</strong></td>
</tr>
</tbody>
</table>
Table 3: Results of the generalized linear mixed effects model for the main logistic regression analysis

Significance was determined using a Likelihood Ratio Test (LRT) $\chi^2$

** = $p < 0.005$ = highly significant effects; * = $p < 0.05$ = significant effect.

<table>
<thead>
<tr>
<th>Predictor variable</th>
<th>$\chi^2$</th>
<th>df</th>
<th>p</th>
</tr>
</thead>
<tbody>
<tr>
<td>Cycle Day</td>
<td>8.36</td>
<td>1</td>
<td>0.004 **</td>
</tr>
<tr>
<td>Colony Size</td>
<td>4.74</td>
<td>1</td>
<td>0.029 *</td>
</tr>
<tr>
<td>Calendar Day</td>
<td>4.21</td>
<td>1</td>
<td>0.040 *</td>
</tr>
<tr>
<td>Colony Habitat</td>
<td>0.73</td>
<td>1</td>
<td>0.392</td>
</tr>
<tr>
<td>Colony Genotype</td>
<td>0.44</td>
<td>1</td>
<td>0.505</td>
</tr>
<tr>
<td>Lunar Day</td>
<td>0.03</td>
<td>1</td>
<td>0.873</td>
</tr>
</tbody>
</table>
Figure 1: *Pocillopora damicornis* colony on the reef in front of the Gump Research Station in Moorea, French Polynesia

The small insert shows a *P. damicornis* larvae on a mesh collection filter (100µm).
Figure 2: The proportion of sexually produced larvae decreases as a function of colony size

The proportion of sexually produced larvae decreases as a function of colony size (solid line). The average proportions of sexually produced larvae per colony are shown as black diamonds (◆), the average proportions per colony per cycle day are shown as white circles (○). Two different colony size classes are distinguished in this graph. Small colonies have a planar surface area under 75cm$^2$ and a maximum diameter of less than 10cm (n=7). Large colonies have a surface area over 130cm$^2$ (n=4) a minimum diameter of more than 10cm. The cross on the left
shows the average and the standard deviation for sexually produced larvae (and colony size) by small colonies. The cross on the right shows the significantly different results for big colonies (single factor LRT: $p = 0.02$).
Figure 3: The proportion of sexually produced larvae released decreases during the reproductive cycles of *P. damicornis* colonies

The proportion of sexually produced larvae released decreases during the reproductive cycles of *P. damicornis* colonies (solid line). The average proportions of sexually produced larvae per cycle day are shown as black diamonds (w), the average proportions per cycle day per colony are shown as white circles (o). The cumulative percentage of sexual larvae across colonies is indicated by a dashed line (---).
CHAPTER 4

EVOLUTIONARY GENOMICS OF POCILLOPORA CORALS

Abstract

Background: Discrepancies between morphology-based taxonomy and phylogenetic systematics are common in corals. In the reef coral genus Pocillopora, nine recently described genetic lineages disagree fundamentally with the 16 currently recognized species. A coherent systematic classification, however, depends on an evolutionary understanding of the observed phenomenon.

Method: To clarify the phylogeographic relationships within this genus, we re-analyzed new and published sequence data from the rDNA marker ITS2 and the mitochondrial open reading frame (ORF). We then used genome-wide Restriction-side Associated DNA sequencing (RAD-Seq) to conduct a phylogenomic comparison among three Tropical Eastern Pacific (TEP) Pocillopora species, three genetic lineages (types 1, 3 & 5) and two putative TEP hybrids.

Results/Conclusion: Our phylogenetic analyses indicated four distinct groups of genotypes. Phylogenomic analyses revealed that genetic lineages from three of these groups are consistent with significant genome-wide differences, which suggests that they belong to separate species although they were all derived from colonies of the morphospecies Pocillopora damicornis. Putative hybrids were composed of a mix of alleles from both parental lineages and allelic ratios indicate repeated uni-parental backcrossing and genome-wide introgression. However, a subset of loci distinguished P. damicornis colonies from colonies with other morphologies, across
lineages and across large geographic distances, which indicates a genetic component of species-specific colony morphology.

**Introduction**

The primary goal of taxonomy is to arrange taxa in hierarchical categories that reflect their systematic or evolutionary relationships. Evolutionary relationships are commonly described using phylogenetic data while taxonomic classifications are still largely based on morphology. Disagreements between morphology-based taxonomic descriptions and phylogenetic systematics can result from a variety of factors ranging from phenotypic plasticity to introgressive hybridization. Phylogenetic data often provides better resolution than morphological characters and can be used to identify cryptic species that are morphologically indistinguishable. Cryptic species may be the result of morphological stasis or unexpected convergent evolution of morphological traits (i.e. homoplasy). Incomplete lineage sorting and introgressive hybridization are the most common sources for misleading molecular phylogenies.

Disagreements between morphology- and genetics-based classifications are common in scleractinian corals (e.g. Fukami et al. 2008). Morphological characters traditionally used to delineate coral taxonomic groups have been shown to be relatively poor predictors of evolutionary relationships, even at the family level (Romano & Palumbi 1996; Fukami et al. 2004b; Fukami et al. 2008), and molecular data revealed cryptic species in many genera (Stobart & Benzie 1994; Fukami et al. 2004a; Keshavmurthy et al. 2013). Commonly used genetic markers evolve slowly in corals, limiting their phylogenetic resolution (e.g. van Oppen et al. 1999; Shearer et al. 2002; Hellberg 2006), and genetic mixing via introgressive hybridization appears to be common in many coral genera (reviewed by Willis et al. 2006), which is
particularly problematic for analyses that rely on relatively few molecular markers, as is often the case for corals (reviewed by Fukami 2008),

Introgressive hybridization confounds species identities by passing genetic material from one taxon into the genome of another. Rates of introgression (i.e. genetic mixing) are not equal across the genomes of hybridizing species (e.g. Martinsen et al. 2001; Hohenlohe et al. 2013). Selection can limit interspecific gene flow of genes involved in hybrid inviability and sterility while allowing introgression in other parts of the genome (e.g. Martinsen et al. 2001; Ting et al. 2001; Barbash et al. 2003; Wu & Ting 2004; Brideau et al. 2006). These “speciation” genes define species differences (i.e. genic view of speciation, Wu 2001) and can be used to study the evolutionary forces that shaped the species (e.g. Rieseberg et al. 1999; Presgraves et al. 2003; Rieseberg et al. 2003). For example, regulatory elements that diverged under positive selection are frequently implicated to cause hybrid incompatibilities among closely related Drosophila species (e.g. Ting et al. 1998; Wu & Ting 2004; Brideau et al. 2006; Ortíz-Barrientos et al. 2007). In other cases, introgression can provide evolutionary novelty by passing potentially adaptive genetic variation between species (e.g. Grant & Grant 1994; Kim & Rieseberg 1999; Mallet 2005; Whitney et al. 2010; Kunte et al. 2011; Hedrick 2013).

Conflicts between molecular phylogenetics and species descriptions based on morphology are common in the ecological important and widespread coral genus Pocillopora. Lack of genetic monophyly in Pocillopora species was first described in the Tropical Eastern Pacific (TEP) and attributed to introgressive hybridization (Combosch et al. 2008). Most TEP Pocillopora species share the same rDNA and mtDNA lineages (type 1 after Pinzón et al. 2008).
However, the most common TEP coral, *P. damicornis*, contains two distinct lineages, the shared type 1 and the private type 3, which in the TEP is exclusive to *P. damicornis* (Combosch et al. 2008; Pinzón & LaJeunesse 2011). In addition, several *P. damicornis* colonies contain alleles from both lineages, i.e. they are heterozygous for type 1 and 3 (1/3). This indicates a pattern of introgressive hybridization between *P. damicornis* and its TEP congeneric *P. eydouxi* and *P. elegans* that allows for inter-specific gene flow of type 1 alleles into *P. damicornis* (*damicornis* (Combosch et al. 2008). Subsequent broad-scale phylogenetic analyses of *Pocillopora* sp. by Pinzón et al. (2013) indicate that widespread and well-defined *Pocillopora* morphospecies often contain multiple distinct genetic lineages that are shared by other morphospecies across broad swaths of their Indo-Pacific ranges. In particular the widely distributed genetic lineages type 1 and 3 are shared among several morphospecies while other types (2, 4, 5, 6, 7 and 8) seem to be mono-specific.

In order to assess the evolutionary history of the coral genus *Pocillopora*, we first re-analyze the phylogenetic relationships among *Pocillopora* lineages using new and published sequence data from nuclear internal transcribed spacer 2 (ITS) and the mitochondrial open reading frame (ORF). We then use genome-wide Restriction-site Associated DNA sequencing (RAD-Seq) to conduct a phylogenomic comparison among three TEP *Pocillopora* species (*P. damicornis*, *P. eydouxi* and *P. elegans*), four trans-Pacific *P. damicornis* populations (from Japan, French Polynesia and Panamá) and two types of putative TEP hybrids (sensu Combsch et al. 2008).

---

4 This is the same lineage that had been termed Type III by Combsch et al. (2008) and species A by Flot et al. (2010). To avoid further confusion, I am subsequently adopting the names of Pinzon et al. (2013), which includes most described lineages.

5 based on taxonomic re-assessments of Figure 5 in Pinzon & LaJeunesse (2011).
Material and Methods

Sample collection & DNA extraction

Coral nubbins were collected from *Pocillopora damicornis*, *P. eydouxi* and *P. elegans* in the Gulf of Chiriquí and the Gulf of Panamá (both Panamá) as well as from *P. damicornis* populations in Okinawa (Japan) and the Tuamotus (French Polynesia). Samples were preserved in Guanidinium-Isothiocyanate (GITC) DNA buffer and genomic DNA was extracted using standard Phenol-Chloroform and Agencourt DNAdvance isolation kits (Beckman Coulter Inc.). In total, more than 300 coral nubbins were collected and processed. ITS2 and ORF markers were sequenced in 231 and 165 specimens, respectively and 147 specimens were RAD sequenced in eight pooled samples (Table 1).

ITS & ORF genotyping and phylogenetics

The ribosomal internal transcribed spacer 2 (ITS) and a mitochondrial open reading frame (ORF) were PCR amplified and sequenced using published primers (Flot & Tillier 2007; Combosch *et al.* 2008). Additional sequence data for major *Pocillopora* lineages (Pinzón *et al.* 2013) was obtained from GeneBank (accession numbers HQ378559–HQ378561, JX994072–JX994088, HQ378552–HQ378557 and KC015015–KC015037). ORF outgroup sequences were obtained from GeneBank for the pocilloporiids *Seriatopora hystrix* (ORF: HQ878586; ITS: AY722794) and *Stylophora* sp. (ORF: JN558909, JN558908, JN558882; ITS: JN559111, JN559109, JN559080). Sequences were aligned with MUSCLE as implemented in Geneious® and adjusted manually.

---

6 *Stylophora* lineages A, B and C (Flot *et al.* 2011)
ITS sequence data from 38 samples contained double peaks, indicating two distinct and heterozygote sequences. Heterozygote sequences were resolved manually using Clark’s method (Clark 1990) and the program Indelligent (Dmitriev & Rakitov 2008) for length-variant heterozygotes. ITS sequence data were aligned as described above, but a complex microsatellite region in ITS was not alignable in a reliable, evolutionarily consistent way (Figure S4a) and between 15 and 34bp were therefore removed from each sequence.

Maximum Likelihood phylogenetic analyses were conducted with RAxML 7.4.4 (Stamatakis 2006; Stamatakis et al. 2008) on the CIPRES Web Portal (Miller et al. 2010). We used the GTR-GAMMA model, which includes a parameter for site heterogeneity among sites, and 1000 rapid bootstrap replicates to estimate clade confidence. Additional Maximum Parsimony analyses were conducted with MEGA (Tamura et al. 2011), including 1000 Bootstrap replicates.

**RAD-Seq**

Genomic DNA quality and quantity was checked on agarose gels after calibration with a subset of samples with DNA 1000 chips on an Agilent BioAnalyzer 2100 (Agilent Technologies). Approximately 50ng of genomic DNA from each specimen was pooled to yield 1µg of pooled genomic DNA per sample. The RAD-Seq pools were composed as described in Table 1. Briefly, samples Pdam_Ja and Pdam_Tu contained *P. damicornis* specimens from Japan and the Tuamotus Islands in French Polynesia, respectively, with ITS type 5 genotypes. The samples Pdam_GC and Pdam_GP contained *P. damicornis* specimens from populations in the Gulf of Chiriquí and Gulf of Panamá, respectively, with ITS type 3 genotypes. The samples Peyd and Pele contained *P. eydouxi* and *P. elegans* specimens from Panamá with ITS type 1 genotypes.
The sample Pdam_T1 contained *P. damicornis* specimens from Panamá with ITS type 1 genotypes and the sample Pdam_T1/3 contained *P. damicornis* specimens from Panamá with heterozygous ITS type 1 and type 3 genotypes.

RAD-Seq libraries were prepared according to modified protocols (Etter *et al.* 2012). Genomic DNA was digested, using a high-fidelity *PstI* enzyme (New England Biolabs). The resulting fragments were ligated to P1 adapters that consist of the corresponding restriction-site overhang, and a sample-specific 4bp barcode sequence. Resulting DNA constructs were randomly fragmented to an average size of 500bp, using a Covaris S220 Ultrasonicator (Covaris Inc.), and 300-500bp fragments were size-selected using agarose gel electrophoresis. DNA fragments were then ligated to a second P2 adapter, with the reverse amplification primer sequence and a divergent Y end that enables PCR amplification only if the corresponding P1 adapter is present within the same DNA fragment. PCR amplifications were carried out with 10µl 5x Phusion Buffer, 0.5µl dNTPs (25mM), 2.5µl of forward and reverse primer, 0.5µl Phusion Polymerase, 5µl DNA template and water to 50µl (all NEB). The PCR profile consisted of 14-18 cycles at 98°C for 10s, 65°C for 30s and 72°C for 30s with an initial denaturation step at 98°C for 30s and a final extension step at 72°C for 5min. Resulting libraries were size-selected again (350-550bp) to remove remaining adapters and primers and verified electrophoretically on an Agilent BioAnalyzer 2100 using DNA High Sensitivity chips (Agilent Technologies) and fluorometrically using a Qubit Fluorometer (Invitrogen). Multiplexed samples were single-end sequenced (1x 50bp, 4 samples per lane) on an Illumina HiSeq 2000 platform (Illumina, Inc, San Diego, California, USA) at the FAS Center for Systems Biology (Harvard University).

Quality filtering and loci assembly was conducted with the STACKS software pipeline (version 0.999; Catchen *et al.* 2011; 2013). Samples were de-multiplexed and raw reads were
quality trimmed to remove low-quality reads (Phred score <10) and reads without complete barcode or restriction cut site. Barcode sequences were removed and remaining reads (47bp) were assembled into ‘stacks’, the RAD equivalent of alleles, if five or more identical reads were found within a sample. These invariable stacks were then compiled into sample-specific loci if they were less than 2 nucleotides (= 5%) different from each other. Sample-specific loci were then assembled into homologous loci if they were less than 3 nucleotides (= 7%) different between samples. A range of assembly parameters was tested against the *P. damicornis* transcriptome (-m: 3-20; -M: 1-3; -n: 1-3) and the final parameter setting (-m 3; -M 2; -n 3) resulted in the highest number and the highest proportion of loci aligning to unique locations within the TEP *P. damicornis* transcriptome (Table S2).

Maximum likelihood models, implemented in STACKS, were used to identify sequence polymorphisms and filter out sequencing errors. Three restrictive filters were applied to increase the proportion of unique genomic loci: 1) all loci with more than 1 SNP were removed from both datasets, 2) the STACKS deleveraging algorithm was used to identify and remove significantly highly repetitive stacks that likely represent sequencing errors and/or repetitive genomic regions, and 3) the STACKS *populations* program was used to filter out loci with invalid SNP calls. The remaining loci consist of 47bp of sequence data per locus, either with zero or one variable nucleotide position within and across samples.

*RNA-Seq*

A transcriptome was assembled from RNA-Seq data in order to identify the proportion of RAD loci that were derived from the coral versus its algal symbionts by mapping the RAD loci against the transcriptome. Total RNA was extracted from *P. damicornis* samples collected at
Taboga Island (Panamá Province) and in Achotines Bay (Los Santos Province) in the Panamanian Pacific. Coral tissues were preserved in TriReagent (Molecular Research Center, Inc.), flash frozen in liquid nitrogen and stored at -80°C. Colonies were genotyped as described above and the transcriptomes of four hybrid colonies were sequenced, two *P. damicornis* specimen heterozygous for type 1 and type 3 and two specimens homozygous for type 1. Total RNA was extracted following the manufacturer's protocol, RNA quality was assessed on an Agilent Bioanalyzer 2100 and extractions with RIN values of 6 or higher were used for cDNA library preparations. RNA sequencing was performed using a modified RNA-Seq protocol (Libro *et al.*, in press) and multiplexed samples were sequenced as described above (1x 50bp, 4 samples per lane, Illumina HiSeq 2000 at Harvard University). Barcoded samples were de-multiplexed and raw sequencing reads were quality trimmed to remove sequences with a Phred score of less than 30 and a read length of less than 15bp using custom Perl Scripts in the FASTX-Toolkit (http://hannonlab.cshl.edu). De novo transcriptome assembly was carried out in Trinity (Grabherr *et al*. 2011) using the combined reads of all 4 samples.

For each RAD locus, strict majority consensus sequences were generated and aligned to the transcriptome using Bowtie version 0.12.9 (Langmead *et al*. 2009) with up to one mismatch over the 47bp of RAD locus length. In addition, RAD loci consensus sequences were aligned to a publicly available *P. damicornis* transcriptome (Traylor-Knowles *et al*. 2011) and transcriptomes of two different *Symbiodinium* types that are commonly found in *Pocillopora* species (Ladner *et al*. 2012). By mapping RAD loci against these transcriptomes, we are able to identify transcribed symbiont and coral loci and estimate the overall proportion of coral versus symbiont loci.

Transcriptome-mapped RAD loci were also used to annotate protein-coding RAD loci. Putative gene identities were identified using homology searches with BLASTx (Altschul *et al*. 2000).
1990) against the Swiss-Prot protein database. Blast matches with an e-value of less than $10^{-5}$ were deemed homologous and their GO terms and gene functions were obtained with Blast2Go (Conesa et al. 2005).

**RAD-Seq Analyses:**

Assembled RAD loci were organized in two different datasets: The main dataset contained all loci that were recovered from at least two samples (referred to as the ‘P2’ dataset) and was used for all analyses, unless stated otherwise. A second dataset, containing loci found across all 8 samples (‘P8’ dataset) was used to calculate pairwise population genetic metrics (e.g. heterozygosity, private alleles, pairwise distances). For the phylogenomic analyses, heterozygous SNPs were coded following IUPAC guidelines and a single consensus sequence was generated for each pooled sample by concatenating variable RAD loci. Phylogenomic analyses were conducted with two complementary methods for P2, P8 and two reduced P2/P8 subsets that contained only differently fixed loci (i.e. homozygous within samples but polymorphic between).

Maximum Likelihood analyses were conducted with RAxML 7.4.4 as described above. Maximum parsimony (MP) analyses are well suited for phylogenetic analyses of SNP data (e.g. Decker et al. 2009; Yoder et al. 2013). MP trees were identified with the program dnapars, as included in the Phylip package (Version 3.695; Felsenstein 1989, 2005). Parsimony tree searches were conducted with 111 randomized sample input orders (using the “jumble” option) and the default settings (‘More thorough search’; ‘Ordinary Parsimony’; ‘retaining 10,000 trees while searching’). Bootstrap support was assessed with 100 (P2) and 1000 (P8) bootstrap replicate datasets, respectively, that were generated with the program seqboot (Phylip 3.695).
Heterozygote loci, private alleles and pairwise comparisons were calculated using parsing scripts in R (R Core Team 2012). Pairwise genetic distances between samples (Table 3) were assessed in three different ways. The most conservative approach (P-distance) is based on the proportion of differentially fixed loci over all fixed loci in a specific pairwise comparison between samples (Nei & Kumar 2000). To include heterozygous loci, two approaches developed for microsatellite (Bowcock et al. 1994) and SNP data (Isolation-by-state, IbS) were tested. The Bowcock method is based on the proportion of shared alleles in each pairwise comparison (e.g. A/A vs. A/C = ½). The IbS method is based on the proportion of identical comparisons (i.e. A/A vs. A/C = ¾). Since both methods resulted in virtually identical pairwise distances, only the IbS results are presented here. Both P-distances and IbS-distances are reported as proportional differences (Table 3) and as the total number of differently fixed loci/different comparisons (Table S3). In addition, a Principal Coordinate Analysis (PCoA) was conducted based on pairwise P-distances using GenAlEx (Version 6.4, Peakall & Smouse 2006). To assess genome-wide levels of selection among samples, pairwise distances between samples were calculated separately for transcribed and un-transcribed loci (Figure S2).

Since *P. damicornis* is known to share alleles with its Tropical Eastern Pacific congeners *P. eydouxi* and *P. elegans*, patterns of allele sharing in the putative hybrid samples (Pdam_T1/3 & Pdam_T1) were analyzed at 4333 loci segregating between the type 1 samples *P. eydouxi* and *P. elegans* versus the type 3 sample *P. damicornis* from the Gulf of Chiriquí to determine the degree of genetic admixture in the two putative hybrid samples.
Results

*ITS & ORF Phylogenies*

Phylogenetic analyses of mitochondrial open reading frame (ORF) sequence data recovered four distinct genetic groups among the eight previously identified types (Figure 2). The new outgroups indicated that type 3 sequences are basal and closely related to type 7. The remaining three genetic groups cluster types 1 and 8 sequences, types 2 and 6 sequences, and types 4 and 5 sequences. This phylogenetic order separates genetic types found only in larvae-brooding *P. damicornis* (types 4 & 5) from types that have been found in broadcast spawning *Pocillopora* species (types 1 and 3). Types 2 and 3, which were only found in locally endemic *P. effusus* (type 2, from Clipperton & Galápagos) and *P. ligulata* (type 6, from Clipperton), seem to be closer related to the larvae-brooding *P. damicornis* clade (types 4 & 5).

The ITS phylogeny was largely consistent with the ORF phylogeny (Figure 2) but the bootstrap support for the four major genetic groups was low. Type 1, 3 and 5 were clearly separated on the ITS tree and Type 1 sequences differed from most other *Pocillopora* lineages by seven fixed nucleotide substitutions. The two local endemics, types 2 and 6, shared some of these fixed differences, but type 1 genotypes were still separated by at least 3 unique fixed nucleotide changes (Figure S3).

*RAD-Seq loci*

In total, over 154 million high-quality RAD-Seq reads were obtained with an average of 12.4 million reads per pooled sample. STACKS assembled 452,367 RAD loci; 207,228 of which were shared between two or more samples and 33,710 (7.5%) occurred in all eight. Further quality filters left 112,644 RAD loci, from which 39,581 single nucleotide polymorphisms
(SNPs) were recovered from at least two samples (P2) with an average cover of 89 reads per locus. 15,511 loci with 6,769 SNPs were recovered from all 8 samples (P8) with an average locus cover of 133 reads (Table S1).

RAD phylogenomic analyses consistently returned three well-supported genetic clades, regardless of the applied phylogenetic method and analyzed data subset (Figure 3). The three clades matched the three major genetic groups identified with ITS and ORF and were labeled accordingly as clade 1, 3 and 5 based on their match to the existing ORF and ITS genetic types. Clade 1 contained all four TEP samples with type 1 genotypes, i.e. the two morphospecies *P. eydouxi* (Peyd) and *P. elegans* (Pele) and both hybrid pools (Pdam_T1 & Pdam_T1/3). Clade 3 contained the TEP *P. damicornis* populations with type 3 genotypes (Pdam_GC & Pdam_GP). Clade 5 contained the central-west Pacific *P. damicornis* populations from Japan (Pdam_Ja) and the Tuamotus (Pdam_Tu) with type 5 genotypes. *Pocillopora damicornis* populations in the central-west Pacific (clade 5) and in the TEP (clade 3) were clearly separated by significant phylogenetic distances with 100% bootstrap support (Figures 3 & S2). Samples within clade 1 were not well separated.

Pairwise comparisons supported the large genetic differences among the three major clades (Figure 4, Tables 2 & S3). For example, less than 5% of all fixed SNPs differed within clades while over 20% usually differed between clades. Both pairwise comparisons, however, indicated that *P. damicornis* from the Gulf of Panamá (Pdam_GP) shares more alleles with type 1 samples than does its *P. damicornis* counterpart from the Gulf of Chiriquí (Pdam_GC; Figure 4). Pairwise comparisons indicated significant levels of population differentiation between type 5 populations over a large geographic distance (10,000km). TEP type 3 populations showed comparable levels of population structure over a much shorter geographic
distance (~400km), which is congruent with population genetic analyses of TEP *P. damicornis* populations (Combosch & Vollmer 2011). However, the intermediate position of the Pdam_GP sample (Figure 4) indicates that pairwise comparisons were affected by its unexpectedly mixed composition, which cannot be accurately separated from the population structure. Pairwise comparisons showed little evidence for genetic differentiation within the type 1 genetic lineage and the mixed *P. damicornis* sample. The morphospecies *P. eydouxi* and *P. elegans* differed by 12 fixed differences (P-distance: 0.003; the largest within this group) and the two hybrid pools did not differ by a single fixed locus. Remarkably, the heterozygous *P. damicornis* sample (Pdam_T1/3) was less different from all four type 3 and type 5 samples than any other type 1 sample (followed by Pdam_T1; Figure 4, Table 2).

**Evidence for Introgressive Hybridization**

The genetic make-up of the two putative *P. damicornis* hybrid pools – heterozygous ITS type 1 and type 3 *P. damicornis* (Pdam_T1/3) and homozygous type 1 *P. damicornis* (Pdam_T1) - was analyzed in detail to detect signatures of genetic admixture at the 4,333 segregating loci between the parental lineages type 3 (Pdam_GC) and type 1 (Peyd & Pele; Table 3). Both *P. damicornis* hybrid types were dominated by type 1 alleles shared with *P. eydouxi* and *P. elegans*. The *P. damicornis* hybrids that were heterozygous for ITS type 1 and 3 had 87% type 1 alleles and 13% type 3 alleles. Type 1 alleles were fixed at 78% of all segregating loci, both alleles were present at 18% and fixed type 3 alleles were found at 4% of segregating loci. The *P. damicornis* hybrids that were homozygous at ITS for type 1 (the *P. eydouxi* and *P. elegans* type) had 94% type 1 alleles and 6% type 3 alleles at segregating sites. They were fixed for type 1 alleles at 89% of all segregating loci, had both allele types at 10% and fixed type 3 alleles at 1% of these
Remarkably, 21 out of the 34 loci that were fixed for type 3 alleles in hybrids with ITS type 1 (Pdam_T1) were also fixed for type 3 alleles in the other hybrid (62%). This indicates that these fixed loci are not a random sample of the 4,333 segregating loci but are selected.

Across all six TEP samples, 105 loci segregated between samples with *P. damicornis* morphologies and samples with *P. eydouxi* and/or *P. elegans* morphologies. A trans-Pacific comparison of these loci with the *P. damicornis* samples from Japan and the Tuamotus identified 33 loci that were fixed for the *P. damicornis* allele (if present). The other 72 segregating loci were either absent (n = 56) or fixed for the Peyd/Pele alleles (n = 16); only 5% were mixed within or between samples (compared to 35% genome-wide). These loci are suitable candidates to further explore genetic differences between morpho-species, within and across genetic lineages.

**RAD-Seq transcriptome matches**

Approximately one quarter of all RAD P8 loci (3,831 out of 15,511) mapped to either our newly assembled transcriptome or the published *P. damicornis* transcriptome (Traylor-Knowles *et al.* 2011) and 10% of transcriptome contigs were directly sampled by a RAD locus. In contrast, less than 0.001% of all RAD loci (35/42 out of 112,644) mapped to the publicly available transcriptomes of *Symbiodinium* type C and type D (Ladner *et al.* 2012) that are commonly found in TEP and central-west Pacific *Pocillopora* species and lineages (e.g. Cunning *et al.* 2013). This indicates that the RAD-Seq data set is dominated by genetic variation in the coral host and not its algal symbionts.

---

*Only loci are considered that were present in one type 3 sample (Pdam_GC and/or Pdam_GP), one hybrid sample (Pdam_T1 and/or Pdam_T1/3) and *P. eydouxi* or *P. elegans.*
Discussion

The goal of this study was to describe the evolutionary relationships in the coral genus *Pocillopora* by re-examining published and new sequence data and to analyze genome-wide allelic patterns among common morphospecies (*P. damicornis*, *P. eydouxi* & *P. elegans*) and their putative TEP hybrids using RAD-Seq. We confirmed the presence of several distinct genetic groups in *Pocillopora* using mitochondrial ORF sequence data and re-rooted the phylogenetic tree using new outgroups. We found that mtDNA and rDNA-based genetic types are consistent with genome-wide differences in the RAD-Seq data, which suggests that some of the described genetic lineages correspond to different species, despite their incongruence with taxonomically recognized morphospecies. Allelic comparisons between putative hybrid and pure-breed samples confirmed introgressive hybridization between genetic lineages in Panamá (Combosch *et al.* 2008), but identified a subset of species-specific RAD loci that were shared among all *P. damicornis* specimens regardless of their introgression history and their geographic origin.

Phylogenetic analyzes of nuclear rDNA (ITS), mitochondrial ORF, and now genome-wide RAD-Seq data demonstrate that well-recognized *Pocillopora* morphospecies are not monophyletic, but instead are comprised of multiple distinct genetic lineages that are shared with other morphospecies. This disagreement between morphospecies and genetics lineages in *Pocillopora* corals was first reported as multiple shared ITS genotypes in Tropical Eastern Pacific (TEP) species due to introgressive hybridization (Combosch *et al.* 2008). *Pocillopora* species with multiple genotypes have subsequently been reported from elsewhere in the TEP (Flot *et al.* 2010; Pinzón & LaJeunesse 2011), East Africa (Souter 2010) and the Great Barrier Reef (Schmidt-Roach *et al.* 2012). A recent Indo-Pacific wide phylogeographic study identified
nine different genetic types, including the three types that were used in the present study (types 1, 3 & 5 sensu Pinzón et al. 2013). The disagreement between morphology and phylogenetics has caused some to argue for a lineage-based description of *Pocillopora* species by genetic type (e.g. type 1 versus type 3; Pinzón et al. 2013). However, this does little to resolve the conflict since these genetic types cannot be distinguished visually in the field and because typing does not reconcile how or why the disagreement between morphology and genetics exists.

*Phylogenetic analyses*

Our re-analyses of published ORF data with the inclusion of new data from 231 *P. damicornis* samples from Japan (n = 17), French Polynesia (23) and Panamá (191) and a focus on outgroups to root the tree identified three main genetic groups in *Pocillopora*. The basal *Pocillopora* type 3 lineage is most genetically diverse (11 ORF genotypes), speciose (7 morphospecies) and geographically widespread (TEP to the Red Sea; Figure 2; Pinzón et al. 2013). Even though type 3 contains multiple species, they all appear to be broadcast spawning. The center of type 3 diversity is in the Indian Ocean and the most basal genotypes, 3e and 3h, were found only in the western Indian Ocean and the Red Sea (Pinzón et al. 2013). This might indicate an Indian Ocean origin of the genus, like its closest sister genus *Stylophora* (Flot et al. 2011; Keshavmurthy et al. 2013). The other two *Pocillopora* lineages that we sequenced with RAD-Seq (types 1 and 5) were significantly different from the type 3 lineage (Figure 2). Type 1 genotypes are geographically widespread, have been found across multiple morphospecies and appear to be broadcast spawners (like type 3). Type 5 genotypes have so far only been found in *P. damicornis* colonies in the Indo-West Pacific and are notably absent from the TEP. Type 5 and type 4 lineages include the brooding *Pocillopora* corals, for example the predominantly
asexually brooding *P. damicornis* (Combosch & Vollmer 2013; Torda *et al.* 2013). All lineages that are closely related to type 5 (i.e. types 2, 4 and 6) appear to be restricted to single morphospecies (*P. effusus*, *P. damicornis* & *P. ligulata*, respectively).

Even though the three lineages are not monophyletic with respect to the current species classification scheme in *Pocillopora*, the newly rooted tree indicates that reproductive strategy may be a key character in the group with the basal genetic groups containing types 3 being broadcast spawners and all brooding corals containing type 4 or type 5 genotypes. Brooding and broadcast spawning were thought to be plastic within *Pocillopora* species, while the phylogeny suggest they are a key evolutionary trait, with brooding probably originating only once within this predominantly broadcast-spawning genus. This result disagrees with the published phylogeny by Pinzon *et al.* (2013) that suggested that brooding was ancestral to broadcast spawning because of their rooting scheme.

Our RAD-based phylogenomic analyses included almost 40,000 variable SNPs from three *Pocillopora* morphospecies and three genetic lineages (types 1, 3 and 5). The resulting RAD phylogeny (Figure 3) is consistent with the phylogenetic patterns observed with rDNA ITS and mitochondrial ORF trees (Figure 2). The two sympatric, but genetically distinct types of TEP *P. damicornis*, described based on ITS (Combosch *et al.* 2008), ORF and microsatellite loci (Pinzon & LaJeunesse 2011), were confirmed by the RAD-Seq data. Genome-wide genetic differences between TEP *P. damicornis* types far exceeded levels of population differentiation within types, even over large biogeographic distances (e.g. over 10,000km between Japan and the Tuamotus; Table 2). *Pocillopora eydouxi* and *P. elegans* from the TEP belonged to the same RAD-Seq clade, shared the same ITS2 and ORF haplotypes (Combosch *et al.* 2008) and were only 0.3% different based on pairwise comparisons (P-distance, Table 2), indicating that they are
closely related and or possibly interbreeding. The RAD-Seq phylogeny further showed that
*P. damicornis* from the central-west Pacific, which had previously been described as a distinct
type based on ITS sequence data (Type I in Combosch *et al.* 2008), indeed belongs to a different
genomic lineage.

*Hybridization in the TEP*

Introgressive hybridization between Tropical Eastern Pacific (TEP) *Pocillopora* had
previously been inferred based on the presence of heterozygous ITS type 1/type 3 genotypes in
Panamanian *P. damicornis* specimens (Combosch *et al.* 2008), based on admixtures signatures in
multi-locus microsatellite data (Pinzón & LaJeunesse 2011; Pinzón *et al.* 2013) and due to
mismatching ITS type 3 and ORF type 1 genotypes (Combosch, unpublished results; Pinzón *et al.*
2013). RAD-Seq phylogenomic analyzes confirm this pattern of admixture across the
genomes of these species in the TEP. In the RAD-Seq data, *P. damicornis* with mixed type 1 and
3 ITS lineages (Pdam_T1/3) was genetically much closer to the type 1 lineages of their TEP
congeners *P. elegans* and *P. eydouxi* than to the pure *P. damicornis* type 3 sample (Figure 4).
This is expected if specimens heterozygous for ITS type 1/3 are the result of hybridization
between types 1 and 3, but cannot be explained as a result of ITS incomplete lineage sorting in
type 3. The heterozygous *P. damicornis* type 1 and 3 sample (Pdam_T1/3) contained a mix of
type 1 (87%) and type 3 alleles (13%) in both homozygous and heterozygous states (Table 3).
This suggests that these heterozygous Pdam_T1/3 specimens were at least a few generations
removed from their most recent type 3 ancestors (Figure 5). Hybridization and backcrossing

---

8 Pinzon *et al.* (2013) label these ITS types 1D and 1F but they differ by 3 fixed SNPs from type
1 while they are only 1 SNP different from type 3.
seem limited to the type 1 maternal lineage, i.e. type 1 eggs\(^9\), because virtually all heterozygotes had mitochondrial type 1 genotypes. Since these colonies have the morphology of \(P. damicornis\), colony morphology is likely dictated by the paternal lineage even though their genomes are largely the other type. The allele composition of the Pdam_T1 sample (94% type 1 vs. 6% type 3) indicates that \(P. damicornis\) specimens homozygous for ITS type 1 are further removed from their most recent type 3 ancestors (Figure 5), as indicated by their ITS genotype.

The dominant species of TEP coral reefs, \(P. damicornis\), is the central component of the local hybridization system. It is the only species in the TEP that frequently contains both genetic lineages and its TEP population is the only one that contains genotypes heterozygous for ITS type 1 and 3. Based on population genetic surveys in Panamá, 14% of \(P. damicornis\) specimens are heterozygous for ITS type 1 and 3 (n = 26 out of 191 genotyped specimens), 46% are homozygous for type 1 and 40% are homozygous for type 3. Thus, so far, we have sampled more specimens with a hybrid history than pure breed \(P. damicornis\) (114 vs. 77). Despite population genetic evidence for significant regional differentiation of \(P. damicornis\) populations between the upwelling Gulf of Chiriquí and the non-upwelling Gulf of Panamá (Combosch & Vollmer 2011), we detected no significant difference in the distribution of hybrids versus pure-breeds. This might indicate a mosaic hybrid zone (sensu Rand & Harrison 1989) between type 1 and type 3 specimens with a patchy distribution of hybrid and pure breed \(P. damicornis\) in Panamá.

Despite extensive genotyping (>500 genotyped samples) elsewhere in the TEP, not a single ITS type 3 genotype has so far been found in Mexico (Pinzón & LaJeunesse 2011). Since type 3 genotypes are widespread throughout the Indo-Pacific, selection likely accounts for the absence of type 3 genotypes in Mexico. Since none of the surveyed Mexican reefs is located in

---

\(^9\) Reports of type 3 genotypes in \(P. capitata\) and \(P. elegans\) are likely due to taxonomic mis-identification (e.g. in Figure 5b Pinzon & LaJeunesse 2011) or indicate backcrossing.
an upwelling zone, the more heterogeneous environment in Panamá may benefit the sympatric coexistence of both types. Moreover, differential thermal adaptations of local populations in two gulfs have been shown experimentally (D'Croz & Mate 2004) and can generate adaptation clines that can stabilize or trap hybrid zones locally (Bierne et al. 2011), which might contribute to the occurrence of local hybridization among *Pocillopora* types in Panamá. Alternatively, type 3 genotypes might have arrived in Panamá only ‘recently’, e.g. with human facilitation or during a strong El-Niño event (Glynn & Ault 2000), or type 1 is replacing type 3 specimens in Panamá just like it might have in Mexico.

In the TEP, 105 loci segregated between samples with *P. damicornis* morphologies (Pdam_GC, Pdam_GP, Pdam_T1 and/or Pdam_T1/3) and samples with *P. eydouxi* and/or *P. elegans* morphologies. Preliminary analyses against our unpublished transcriptome indicate that 10 out of these 105 loci map against contigs in the transcriptome and three contigs have well-defined protein annotations (e-value 10-5). All three corresponding proteins are involved in gene regulation, two are involved in transcription regulation, and one is a RNA-binding protein. Regulatory genes are prime candidates for hybrid sterility and inviability, i.e. speciation genes (e.g. Ortíz-Barrientos et al. 2007) and several speciation genes in other taxa were found to have regulatory functions (e.g. Ods & Hmr; Ting et al. 1998; Brideau et al. 2006). Other segregating RAD loci are possibly linked to additional candidate speciation genes as indicated by their low heterozygosity (5% compared to 23% genome-wide) and high fixation within lineages (90%). In particular the 33 loci that segregate for *P. damicornis* samples across all three genetic lineages and over 20,000 km of open ocean are suitable candidates for further investigations.
Hybridization outside the TEP

Indications for hybridization among Pocillopora species and genetic lineages have been reported from many other locations throughout the Indo-Pacific (e.g. Miller & Ayre 2004; Souter 2010; Schmidt-Roach et al. 2013). For example, Pinzon et al. (2013) identified admixture between types (presumably 3 and 5) in Hawaii using multi-locus microsatellite data and found one specimen with a heterozygous mitochondrial ORF type 5 and 6. This is interesting because it would involve hybridization among presumed broadcasting and brooding Pocillopora lineages. In addition, we re-analyzed published sequence data from Hawaii (Flot et al. 2008) and found heterozygote ITS type 3 and 5 genotypes in 9 out of 16 P. damicornis specimen, mismatching ITS type 3 and ORF type 5 genotypes in two additional P. damicornis specimen and one heterozygote ITS type 3 and 5 with ORF type 4, i.e. three distinct genetic lineages. Moreover, mismatching ORF type 1 with ITS type 3 (mislabeled as 1E akin type 1D and 1F discussed above) were detected in 7 P. meandrina and one P. eydouxi (Flot et al. 2008). These mixed typed individuals suggest that introgressive hybridization is more common among Pocillopora types than previously thought.

Colony morphology, phylogenetics and the evolution of Pocillopora corals

The main result of this study, that rDNA- and mtDNA-based genetic types are consistent with genome-wide RAD-Seq lineages, strongly supports the hypothesis that some of the described genetic types correspond to interbreeding and reproductively isolated units, i.e. biological species (sensu Mayr 1963). Additional morphological, ecological and reproductive studies will now be necessary to identify and delimitate valid Pocillopora species within and among these distinct genetic lineages. Importantly, the analyzed types 1, 3 and 5 correspond to
the three main *Pocillopora* lineages (Figure 2). Most other *Pocillopora* types are in fact poorly delimitated and either share the same ITS type (3 and 7), share both ITS and ORF types (1 & 9), are based on less than a handful of samples (type 8) or can simultaneously be shared with two other types (type 4).

If biological species in *Pocillopora* are defined solely by genetic lineages, it is unexpected that species share multiple distinct colony morphologies, often all across their distribution range. Colony morphologies can be shared because they are ancestral (i.e. reticulate evolution), plastic (i.e. taxonomy is based on unsuitable characters), as a result of independent convergent evolution or because the underlying morphology genes have been exchanged via introgressive hybridization. Highly variable, plastic morphologies are widespread among *Pocillopora* corals and have lead to numerous new species descriptions (e.g. Pichon & Veron 1976; Figure 1b). However, it is unlikely that phenotypic plasticity alone accounts for the full range of discrepancies observed here, including eight distinct, recurrent colony morphologies in two distinct lineages. First of all, studies have shown consistent links between *Pocillopora* morphospecies and specific genotypes (e.g. Schmidt-Roach et al. 2013). In addition, phenotypic plasticity depends on environmental differences to drive morphologic variation (West-Eberhard 1989) and many of these morphospecies occur sympatrically together on reefs. Within the homogeneous TEP *P. damicornis* carpets (Figure 1a), it is hard to imagine that environmental differences could account for the stark morphological differences between morphospecies.

Divergent positive selection can generate and/or maintain genetically determined recurrent colony morphologies within distinct genetic lineages via convergent evolution, morphological stasis or introgressive hybridization. Convergent morphological evolution could be present to some extent, but it seems very unlikely that it can account for all eight observed
morphospecies shared between type 1 and type 3. The exchange of morphology-driving genes via introgressive hybridization cannot fully account for the observed number of shared morphologies either because the number of distinct morphologies (at least 8) is much bigger than the number of genetic lineages that share them (2; types 1 & 3). The presence of shared ancestral morphologies is somewhat supported by the observation that the phylogenetically basal types 1 and 3 shared most morphospecies while more derived lineages are monomorphic (e.g. types 2, 4, 5 & 6), which might indicate a loss of ancestral morphologies, e.g. due to founder effects. Importantly, none of these different scenarios are mutually exclusive and some of them are in fact more likely to occur together than separately, for example due to mutually favorable conditions (e.g. strong selection). It is therefore most likely that a combination of multiple factors generates the observed discrepancies in *Pocillopora*.

In the TEP, we have shown that in some cases disagreements can be explained by introgression hybridization. It is unlikely, however, that introgressive hybridization explains all disagreements between morphological taxonomy and phylogenetic lineages through the Indo-Pacific. It is a first step, however, and we have shown that the isolated Tropical Eastern Pacific with its reduced complexity of *Pocillopora* corals is a good place to continue disentangling this evolutionary puzzle. Alternatively, Hawaii potentially offers the unique opportunity to study the evolution and interaction of both major coral reproductive strategies. Since all Hawaiian *P. damicornis* colonies that were heterozygous for ITS type 3 and 5 contained either ORF type 4 or 5, broadcast-spawned sperm from type 3 might be capable of fertilizing the eggs of larval-brooding type 4 and 5 colonies internally. This type of hybrid fertilization seems to be unique to Hawaii and would allow the identification of genes involved in different coral reproductive strategies and possibly in the origin of some of these different *Pocillopora* lineages.
Significant discrepancies between morphology-based taxonomy and phylogenetic systematics are common in pocilloporiid genera and have been described in *Stylophora* (Flot *et al.* 2011; Stefani *et al.* 2011), *Seriatopora* (Flot *et al.* 2008) and *Madracis* (Frade *et al.* 2010). Different evolutionary scenarios and combinations of mechanisms have been proposed for each genus. For example, while introgressive hybridization has been documented in *Madracis* (Frade *et al.* 2010) and *Pocillopora* (this study), it has been ruled out in *Stylophora* based on the congruence of nuclear and mitochondrial markers (Flot *et al.* 2011). However, studies in all four genera have benefitted from common transferable technical advances. For example the description of the mitochondrial genome in *Pocillopora* (Flot & Tillier 2007) enabled the development of mitochondrial sequence markers in *Seriatopora* (Flot *et al.* 2008) and *Stylophora* (Flot *et al.* 2011). We hope that his study testifies to the usefulness and power of genome-wide RAD sequencing to disentangle the complex evolutionary history of corals in general and pocilloporiids in particular.

**Conclusions**

In summary, we conclude that genome-wide differences between genetically defined lineages likely correspond to interbreeding, reproductively isolated units, and hence different species. To define, delimit and recognize these new species, additional genetic, morphological, ecological, and reproductive studies will be necessary (e.g. Combosch & Vollmer 2013; Torda *et al.* 2013). In addition, we strongly support the re-analysis of the morphology-based taxonomic descriptions of corals in general and *Pocillopora* in particular in light of recent genetic findings. Several recent publications have made significant progress (e.g. Stefani *et al.* 2008; Benzoni *et al.*
2010; Arrigoni et al. 2012; Schmidt-Roach et al. 2013). However, coral phylogeneticists must provide taxonomists with sound evolutionary explanations in addition to their list of cryptic genetic types. The evolution of many coral taxa is certainly complex and multi-facetted and genomic approaches might be necessary to resolve their evolutionary histories.

Acknowledgements

Support for fieldwork, laboratory work, bioinformatics and genetic analyses was provided by Sarah Lemer, Rich Aronson, Pete Edmunds, Rosa Fernández, Sal Genovese, Vanessa Gonzales, Gonzalo Giribet, Stefan Kaluziak, Ian McIntyre, Alicia Pérez-Porro, Carmen Schloeder and Prashant Sharma.
References


Fukami H., C. Chen, A. Budd, et al. 2008 Mitochondrial and nuclear genes suggest that stony corals are monophyletic but most families of stony corals are not (Order Scleractinia, Class Anthozoa, Phylum Cnidaria). *PLoS One* **3**: e3222.


Nei M. & S. Kumar 2000 Molecular evolution and phylogenetics Oxford University Press, USA.


Central-west Pacific *P. damicornis* populations were characterized by high levels of genetic polymorphism (9-12%), many private alleles (>800) and high proportions of fixed private alleles (5-10%). Most TEP samples (Peyd, Pele, Pdam_T1 & Pdam_T1/3) had low levels of polymorphisms (7-8%), few private alleles (<250) and low proportions of fixed private alleles (<2%). No significant correlations were found between the number of raw reads and the number of pooled individuals per sample (Figure S1) with the number of heterozygous loci and/or the number of private alleles.

Region: CWP = central-west Pacific; TEP = Tropical Eastern Pacific; Location: Tuamotus = French Polynesia; Gulf of Chiriquí & Gulf of Panamá = Panamá; N\textsubscript{pool} = Number of pooled specimens; IPA = Private Alleles; Reads [M] = number of raw reads [in million reads] after quality control; Hets = heterozygous loci = loci that are variable within each sample.

### Tables

**Table 1: Overview of the RAD-Seq sample composition and basic population genetic indices.**

<table>
<thead>
<tr>
<th>Sample</th>
<th>Species</th>
<th>Region</th>
<th>Location</th>
<th>Genotypes</th>
<th>N\textsubscript{pool}</th>
<th>Reads [M]</th>
<th>Variable loci</th>
<th>Hets</th>
<th>PA'</th>
</tr>
</thead>
<tbody>
<tr>
<td>Pdam_Ja</td>
<td><em>P. damicornis</em></td>
<td>CWP</td>
<td>Japan</td>
<td>Type 5</td>
<td>17</td>
<td>3.5</td>
<td>18,350</td>
<td>1,405</td>
<td>842</td>
</tr>
<tr>
<td>Pdam_Tu</td>
<td><em>P. damicornis</em></td>
<td>CWP</td>
<td>Tuamotus</td>
<td>Type 5</td>
<td>23</td>
<td>9.3</td>
<td>26,210</td>
<td>1,793</td>
<td>993</td>
</tr>
<tr>
<td>Pdam_GC</td>
<td>*Pocillopora</td>
<td>Novel</td>
<td>Gulf of Chiriquí</td>
<td>Type 3</td>
<td>21</td>
<td>17.6</td>
<td>27,037</td>
<td>1,151</td>
<td>539</td>
</tr>
<tr>
<td>Pdam_GP</td>
<td>damicornis</td>
<td>Novel</td>
<td>Gulf of Panamá</td>
<td>Type 3</td>
<td>18</td>
<td>7.6</td>
<td>21,533</td>
<td>1,561</td>
<td>240</td>
</tr>
<tr>
<td>Pdam_T1/3</td>
<td><em>P. damicornis</em></td>
<td>TEP</td>
<td>Gulf of Chiriquí</td>
<td>Type 1/3</td>
<td>21</td>
<td>26.3</td>
<td>31,586</td>
<td>1,217</td>
<td>135</td>
</tr>
<tr>
<td>Pdam_T1</td>
<td><em>P. damicornis</em></td>
<td>TEP</td>
<td>Gulf of Panamá &amp; Chiriquí</td>
<td>Type 1</td>
<td>13</td>
<td>9.5</td>
<td>15,275</td>
<td>1,201</td>
<td>116</td>
</tr>
<tr>
<td>Peyd</td>
<td><em>P. eydouxi</em></td>
<td></td>
<td></td>
<td>Type 1</td>
<td>16</td>
<td>22.5</td>
<td>33,301</td>
<td>1,224</td>
<td>225</td>
</tr>
<tr>
<td>Pele</td>
<td><em>P. elegans</em></td>
<td></td>
<td></td>
<td>Type 1</td>
<td>18</td>
<td>2.8</td>
<td>29,293</td>
<td>1,074</td>
<td>167</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Total</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>147</td>
<td>99.1</td>
<td>39,581</td>
<td>6,128</td>
<td>3,242</td>
</tr>
</tbody>
</table>
Table 2: Pairwise comparisons between samples using P-distance and IbS-proportions.

The average amount of differently fixed SNPs between sample pairs was 588; since 4,187 loci were on average fixed between samples, the average P-Distance was 0.14. Allele-specific pairwise comparisons (IBS) indicated on average 7,073 alleles differences in 27,076 allelic comparisons, which correspond to an average IbS of 1,780 and an average IbS-proportion of 0.26. The average sample pair was thus fixed for the same allele at 12,901 P8 loci, for different alleles at 588 loci and contained 2,000 heterozygous loci.

Above: IbS-proportions = Proportion of allele comparisons that are different;

Average = 0.261 ± 0.094

Below: P-distance = Proportion of fixed loci that are different; Average = 0.140 (± 0.084)

Intra-clade comparisons are boxed in

Grey background highlights the increase in pairwise distances between type 1 samples and

*P. damicornis* samples with ITS type 3 and type 5 (Pdam_T1/3 < Pdam_T1 < Pele/Peyd).
Table 3: Genetic make-up of putative hybrid *P. damicornis* samples

Genetic make-up of putative hybrid *P. damicornis* samples (Pdam_T1/3 & Pdam_T1) at 4333 loci that are segregating between type 3 (Pdam_GC) and type 1 (Peyd/Pele). Percentages are calculated based on the number of alleles/loci present in each sample. 581 segregating loci were not recovered from Pdam_T1/3, 882 loci were not recovered from Pdam_T1 and 1065 were not recovered from either one (i.e. “both”).

To evaluate levels of random ‘background’ variation of SNPs between lineages, the type 1 samples Peyd and Pele were analyzed vice versa at all 3,723 loci segregating between type 3 (Pdam_GC) and Pdam_T1-Pdam_T1/3 (Table S4).

<table>
<thead>
<tr>
<th>Alleles</th>
<th>Loci</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Type 3 (Pdam_GC)</strong></td>
<td><strong>Type 1 (Peyd/Pele)</strong></td>
</tr>
<tr>
<td>Pdam_T1/3</td>
<td>967 (13%)</td>
</tr>
<tr>
<td>Pdam_T1</td>
<td>403 (6%)</td>
</tr>
<tr>
<td>Both</td>
<td>298 (5%)</td>
</tr>
<tr>
<td>Type 1 (Peyd/Pele)</td>
<td>8666 (100%)</td>
</tr>
</tbody>
</table>
Figure 1a: *Pocillopora elegans* colony in a *P. damicornis* carpet in Isla Coiba, Gulf of Chiriquí, Panamá.
Figure 1b: Phenotypic plasticity of *P. damicornis/P. verrucosa* in Moorea, French Polynesia.
Figure 2: ORF and ITS Phylogeny

Phylogenetic tree and 1000 Bootstrap replicates from RAxML. The final ORF alignment was 632 bp long and contained on average 565 bp of *Pocillopora* sequence data. The final ITS alignment was 643 bp long and contained on average 389 bp of *Pocillopora* sequence data.
Outgroup sequences were obtained from GeneBank for pocilloporiid corals *Seriatopora* sp. and *Stylophora* sp. A, B & C (sensu Flot et al. 2011).
Figure 3: RAD phylogenomic analyses

Tree topology and branch lengths [substitutions/site] are based on Maximum likelihood reconstructions (RAxML). The tree is rooted according to the ORF phylogeny (Figure 2). Bootstrap support (in bold) is based on 1000 pseudo-replicates for RAxML and 100 pseudo-replicates for Maximum Parsimony (DNApars) and reported for analyses based on all loci and fixed SNPs loci only (RAxML / fix RAxML / DNApars / fix DNApars). The tree topology was identical among all four analyses with the exception of one node (intra type 1) in the fixed Maximum Parsimony analyses. The tree topology for P8 loci was identical and the bootstrap support was virtually identical (results not shown).
Figure 4: Principal Coordinate Analysis of pairwise P-distances between RAD-Seq samples.

PCA is drawn proportionally (i.e. according to percentages per axis). The first principal component explains 73% of the variance (53% for IBS, Figure S3) and separated the pure *P. damicornis* samples from the hybridizing samples in clade 1. The second principal component explained 26% of the genetic variance and separates the clade 3 from the others.

Among type 1 sample, the two hybrids (Pdam_T1/3 & Pdam_T1) are more similar to other *P. damicornis* samples than the other two type 1 samples (Peyd & Pele). The unexpected intermediate identity of Pdam_GP is also evident.
Figure 5: Allele composition of putative hybrid *P. damicornis* samples compared to the expected allele compositions of backcrossing hybrids over several generations.

The heterozygote sample Pdam_T1/3 is on average probably only 3-4 generations removed from the initial hybridization event while Pdam_T1 is on average probably 4-5 generations removed.
Supplementary Material

Table S1: Read Depth per sample and per locus

Table S1a shows the minimum (Min), Average (Avg) and Maximum (Max) read depth per sample across loci and the total number of reads per sample.

<table>
<thead>
<tr>
<th>Reads per sample</th>
<th>Pdam Ja</th>
<th>Pdam Tu</th>
<th>Pdam GC</th>
<th>Pdam GP</th>
<th>Peyd</th>
<th>Pele</th>
<th>Pdam T1/3</th>
<th>Pdam T1</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>Min</td>
<td>5</td>
<td>5</td>
<td>5</td>
<td>5</td>
<td>5</td>
<td>5</td>
<td>10</td>
<td>5</td>
<td>5</td>
</tr>
<tr>
<td>Avg</td>
<td>37</td>
<td>54</td>
<td>164</td>
<td>77</td>
<td>175</td>
<td>25</td>
<td>220</td>
<td>64</td>
<td>133</td>
</tr>
<tr>
<td>Max</td>
<td>381</td>
<td>866</td>
<td>1264</td>
<td>609</td>
<td>1129</td>
<td>230</td>
<td>1220</td>
<td>394</td>
<td>1264</td>
</tr>
<tr>
<td>Sum</td>
<td>255,555</td>
<td>375,134</td>
<td>1,134,319</td>
<td>531,066</td>
<td>1,211,264</td>
<td>175,617</td>
<td>1,517,396</td>
<td>442,122</td>
<td>8,288,661</td>
</tr>
</tbody>
</table>
Table S2: Transcriptome matches/alignments.

RAD loci are considered to match transcriptome contigs if their consensus sequence aligns with 0 or 1 mismatch over the 47bp sequence (2%).

**Matches #:** The number of RAD loci consensus sequences that match each transcriptome.

**Matches %:** The proportion of RAD loci that match to a specific transcriptome.

**Multiple matches:** Percentage of loci that matched to multiple sites within a specific transcriptome (mostly due to multiple perfect alignments across the 47bp sequences).

**Transcriptome coverage:** Percentage of transcriptome contigs that had at least one RAD locus

**Transcriptome contigs:** The number of contigs in each of the 4 transcriptomes.

1New = new transcriptome for TEP Pdam_T1 and Pdam_T1/3 (Combosch, unpublished);

2Public = Traylor-Knowles et al. 2012 = likely type 3 and/or type 5.

3Combined = loci that match to either transcriptomes were considered to be transcribed loci.

4Symbiodinium Clade C & D = Ladner et al. 2012

<table>
<thead>
<tr>
<th></th>
<th>Transcripts</th>
<th>Matches #</th>
<th>Matches %</th>
<th>Multiple matches</th>
<th>Transcriptome coverage</th>
<th>Transcriptome contigs</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>P. damicornis</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1New</td>
<td>P8</td>
<td>3,383</td>
<td>21.8%</td>
<td>1.7%</td>
<td>3.7%</td>
<td>73,522</td>
</tr>
<tr>
<td></td>
<td>P2</td>
<td>10,673</td>
<td>9.5%</td>
<td>0.8%</td>
<td>10.1%</td>
<td></td>
</tr>
<tr>
<td>2Public</td>
<td>P8</td>
<td>1,961</td>
<td>12.6%</td>
<td>2.9%</td>
<td>2.4%</td>
<td>70,786</td>
</tr>
<tr>
<td></td>
<td>P2</td>
<td>6,518</td>
<td>5.8%</td>
<td>1.2%</td>
<td>7.2%</td>
<td></td>
</tr>
<tr>
<td>3Combined</td>
<td>P8</td>
<td>3,831</td>
<td>24.7%</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>P2</td>
<td>12,772</td>
<td>11.3%</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Symbiodinium</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Clade C</td>
<td>P8</td>
<td>1</td>
<td>&lt;0.1%</td>
<td></td>
<td></td>
<td>26,947</td>
</tr>
<tr>
<td></td>
<td>P2</td>
<td>35</td>
<td>&lt;0.1%</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Clade D</td>
<td>P8</td>
<td>0</td>
<td>0.0%</td>
<td></td>
<td></td>
<td>23,657</td>
</tr>
<tr>
<td></td>
<td>P2</td>
<td>42</td>
<td>&lt;0.1%</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
Table S3: IbS & Number of differently fixed loci

**Above:** = IbS distances between sample pairs; Average = 1775 (+/- 574)

**Below:** = Differently fixed loci among sample pairs; Average = 588 (+/- 395)

Intra-clade comparisons are boxed in

Grey background highlights the increase in pairwise distances between type 1 samples and *P. damicornis* samples with ITS type 3 and type 5 (Pdam_T1/3 < Pdam_T1 < Pele/Peyd).

<table>
<thead>
<tr>
<th></th>
<th>Pdam_Ja</th>
<th>Pdam_Tu</th>
<th>Pdam_GC</th>
<th>Pdam_GP</th>
<th>Pdam_T1/3</th>
<th>Pdam_T1</th>
<th>Pele</th>
<th>Peyd</th>
</tr>
</thead>
<tbody>
<tr>
<td>Pdam_Ja</td>
<td>1544</td>
<td>2126</td>
<td>2199</td>
<td>2262</td>
<td>2282</td>
<td>2273</td>
<td>2321</td>
<td></td>
</tr>
<tr>
<td>Pdam_Tu</td>
<td>181</td>
<td>2139</td>
<td>2223</td>
<td>2280</td>
<td>2299</td>
<td>2289</td>
<td>2339</td>
<td></td>
</tr>
<tr>
<td>Pdam_GC</td>
<td>887</td>
<td>722</td>
<td>1332</td>
<td>2055</td>
<td>2083</td>
<td>2070</td>
<td>2124</td>
<td></td>
</tr>
<tr>
<td>Pdam_GP</td>
<td>778</td>
<td>633</td>
<td>206</td>
<td>1609</td>
<td>1622</td>
<td>1621</td>
<td>1660</td>
<td></td>
</tr>
<tr>
<td>Pdam_T1/3</td>
<td>983</td>
<td>820</td>
<td>917</td>
<td>397</td>
<td>792</td>
<td>817</td>
<td>862</td>
<td></td>
</tr>
<tr>
<td>Pdam_T1</td>
<td>1009</td>
<td>845</td>
<td>947</td>
<td>408</td>
<td>0</td>
<td>798</td>
<td>844</td>
<td></td>
</tr>
<tr>
<td>Pele</td>
<td>1063</td>
<td>895</td>
<td>988</td>
<td>443</td>
<td>5</td>
<td>6</td>
<td>852</td>
<td></td>
</tr>
<tr>
<td>Peyd</td>
<td>1037</td>
<td>874</td>
<td>979</td>
<td>427</td>
<td>2</td>
<td>3</td>
<td>14</td>
<td></td>
</tr>
</tbody>
</table>
Table S4: Vice versa comparison of allelic identity

Vice versa comparison of allelic identity revealed minor allele sharing between type 3 and Peyd/Pele, most likely due to backcrossing. To evaluate the level of random ‘background’ variation of SNPs between lineages, the type 1 samples Peyd and Pele were analyzed vice versa at all 3,723 loci segregating between type 3 (Pdam_GC) and Pdam_T1-Pdam_T1/3 (Table S4a). Vice-versa comparisons revealed minor background allele sharing between type 3 and Peyd/Pele. Interestingly Peyd contained more type 3 alleles than Pele, which could be due to backcrossing and introgression of type 3 alleles into Peyd and would indicates preferred backcrossing of hybrids with *P. eydouxi*.

Hybrids comparison- Type 3 (Pdam_GC) vs. Type 1(Peyd &\ Pele): 4333 segregating loci

<table>
<thead>
<tr>
<th></th>
<th>Type 3 fix</th>
<th>Hets</th>
<th>Peyd-Pele fixed</th>
</tr>
</thead>
<tbody>
<tr>
<td>Both hybrids</td>
<td>21 0.5%</td>
<td>230 5%</td>
<td>2,638 61%</td>
</tr>
<tr>
<td>Pdam_T1</td>
<td>34 0.8%</td>
<td>335 8%</td>
<td>3,076 71%</td>
</tr>
<tr>
<td>Pdam_T1/3</td>
<td>134 3.1%</td>
<td>699 16%</td>
<td>2,961 68%</td>
</tr>
<tr>
<td>Pdam_GP</td>
<td>1,153 44%</td>
<td>897 35%</td>
<td>543 21%</td>
</tr>
</tbody>
</table>

Vice Versa comparison – Type 3 (Pdam_GC) vs. (Pdam_T1/Pdam_T1/3): 3723 segregating loci

<table>
<thead>
<tr>
<th></th>
<th>Type 3 fix</th>
<th>Hets</th>
<th>T1-T1/3 fixed</th>
</tr>
</thead>
<tbody>
<tr>
<td>Peyd &amp; Pele</td>
<td>2 0.1%</td>
<td>13 0%</td>
<td>1,572 42%</td>
</tr>
<tr>
<td>Pele</td>
<td>7 0.2%</td>
<td>47 1%</td>
<td>1,685 45%</td>
</tr>
<tr>
<td>Peyd</td>
<td>29 0.8%</td>
<td>150 4%</td>
<td>3,022 81%</td>
</tr>
<tr>
<td>Pdam_GP</td>
<td>858 23%</td>
<td>869 23%</td>
<td>493 13%</td>
</tr>
</tbody>
</table>
Figure S1: Basic RAD loci correlations

No significant correlation between the number of heterozygous loci and the number of specimens per sampling pool ($R^2 = 0.27$), raw reads per sample ($R^2 = 0.07$) and the number of reads per locus ($R^2 = 0.11$) nor between private alleles and the number of raw reads per sample ($R^2 = 0.10$), the number of pool specimen ($R^2 = 0.25$) and the number of raw reads ($R^2 = 0.11$).
Figure S2: Pairwise P-distance comparisons between samples at transcribed vs. untranscribed loci.

Green line: expected neutral 1:1 ratio. Dotted black line: average trend among samples.

Average All = 0.140; Average Transcribed = 0.133; Average Untranscribed = 0.142;

Pairwise comparisons over transcribed versus un-transcribed loci indicated higher levels of differentiation among un-transcribed loci in 24 out of 28 pairwise comparisons, indicating genome-wide mild balancing selection. Further indications for balancing selection are the increasing gap between transcribed vs. transcribed loci with larger genetic differences between samples and higher proportions of private alleles at untranscribed loci (not shown).

Three pairwise comparisons (in red) are more differentiated at transcribed loci, which could indicate positive selection between the two geographic *P. damicornis* populations pairs Gulf of Chiriquí vs. Gulf of Panamá & Tuamotus vs. Japan.
FINAL CONCLUSIONS

I have shown that even in stable, oceanic environments TEP *P. damicornis* populations do have sex, which is well organized and guarantees that every genotype contributes into the common gene pool while successful genotypes are allowed to produce more offspring on their own, i.e. parthenogenetically (Chapter 3). In contrast, TEP *P. damicornis* populations have more sex, are genetically fine-structured (Chapter 2) and exchange large parts of their genome with local congenerics (Chapter 4). This combination facilitates adaptations and is most likely the key to the extraordinary success of *P. damicornis* in the environmentally challenging TEP. In addition, I showed that distinct genetic lineages within TEP *Pocillopora damicornis* correspond to genome-wide differences but colonies with *P. damicornis* morphologies share a core set of genes (candidate speciation genes) across genomic lineages (Chapter 4). Moreover, regional *P. damicornis* populations are composed of different genomic lineages and exchange genes between lineages and with local congenerics via introgressive hybridization (Chapter 4).

However, it feels like with every question I answered, five new questions arose: What are the molecular adaptations that enable corals to survive in the harsh TEP environments and how do they differ among environments? What are the ecological functions of different larval types and what are the molecular mechanisms that enable their simultaneous production. What evolutionary mechanisms keep the distinct genomic lineages separate within morphospecies and why is the same morphospecies composed of different genomic lineages in different locations? But I will leave that for my next research project…