MOLECULAR ECOLOGY

Molecular Ecology (2012) **21**, 2224–2238 doi: 10.1111/j.1365-294X.2012.05528.x

Extensive sympatry, cryptic diversity and introgression throughout the geographic distribution of two coral species complexes

JASON T. LADNER and STEPHEN R. PALUMBI

Department of Biology, Stanford University, Hopkins Marine Station, Pacific Grove, CA 93950, USA

Abstract

The identification of species is one of the most basic, and yet critically important, issues in biology with far-reaching potential implications for fields such as biodiversity conservation, population ecology and epidemiology. Morphology has long been the primary tool biologists have used to categorize life. However, we now know that a significant portion of natural diversity is morphologically hidden, and therefore, we must integrate nonmorphological tools into the description of biodiversity. Here, we demonstrate the utility of multilocus population genetic data for identifying and characterizing cryptic species complexes, even when species share large amounts of genetic variability. Specifically, we have used DNA sequence data from 12 genomic regions to characterize two widespread species complexes in the coral genus *Acropora*: *A. cytherea* and *A. hyacinthus*. These two morphospecies have each been sampled from 5 to 7 locations throughout their Indo-Pacific distributions, and with the use of *structure* and hierarchical clustering, we demonstrate the presence of at least six widespread cryptic species within these two morphospecies complexes. After identifying cryptic lineages, we then utilize the genetic data to examine the history of introgressive hybridization within and between these morphospecies complexes. Our data indicate that these two complexes form a global syngameon with consistent patterns of introgression between species across large geographic distributions.

Keywords: *Acropora*, cryptic species, hybridization, population genetics, population structure

Received 31 October 2011; revision received 20 January 2012; accepted 24 January 2012

Introduction

Morphology is the original tool used by biologists to categorize life (Mayr 1949; p. 115), and morphological traits remain immensely important for taxonomic description and identification of species in the field. However, it has become clear that a large amount of natural diversity is morphologically hidden, with sometimes only subtle, if any, morphological distinctions between species (Bickford *et al.* 2007). Cryptic species complexes have now been described in many major metazoan taxa (Pfenninger & Schwenk 2007) as well as many groups of plants (e.g. Paris *et al.* 1989; Grundt *et al.* 2006). As evidence has grown in support of the extensiveness of cryptic diversity, the data have also highlighted many important distinctions among different cryptic groups. Cryptic species may be recently (Sáez *et al.* 2003) or deeply diverged (Elmer *et al.* 2007), allopatrically (Brown *et al.* 2007; Ross *et al.* 2010) or sympatrically distributed (Stuart *et al.* 2006; Boissin *et al.* 2008). They may represent sister lineages (Yoder *et al.* 2002; Brown *et al.* 2007), or they may have been formed through evolutionary convergence (Goodman *et al.* 2009). Some cryptic species complexes have many members (Hebert *et al.* 2004); others have only a few (Souter 2010; Piggott *et al.* 2011). Identification and characterization of the world's many cryptic species complexes is an ongoing challenge for evolutionary biologists with far-reaching potential implications for fields such as biodiversity conservation (Bickford *et al.* 2007; Trontelj & Fišer 2009), epidemiology (Mayr 1970;

Correspondence: Jason T. Ladner, Fax: 831-655-6215;
E-mail: jtladner@gmail.com

Collins & Paskewitz 1996; Miles *et al.* 2003) and biological control (Hafez & Doutt 1954; DeBach 1969; Rosen 1986; Heraty *et al.* 2007).

Many nonvisual traits (e.g. vocalizations, mating behaviours) have contributed to the identification of cryptic species, but few have been as powerful and as widely applicable as genetics. Although the concept of cryptic species has been articulated for centuries (Winker 2005), the number of publications describing cryptic species has been increasing exponentially ever since the introduction of cost-effective methods to investigate population genetic variation (Bickford *et al.* 2007). Genetic markers can provide critical information regarding evolutionary relationships even in instances of morphological homogeneity, and when groups are strongly diverged, even a few markers can provide clear demarcations of species boundaries (e.g. Boissin *et al.* 2008; Forsman *et al.* 2009; Souter 2010; Piggott *et al.* 2011). Genetic identification of cryptic species can be considerably more complicated when cryptic species are recently diverged (e.g. Sáez *et al.* 2003), when individual species exhibit complex metapopulation structure (e.g. Ross *et al.* 2010; Pinzon & LaJeunesse 2011) and/or when species are connected through introgression (e.g. Forsman *et al.* 2009; Ross *et al.* 2010). This is because in these instances, there is little difference between the levels of genetic differentiation among species and those among populations. However, even in the absence of fixed sequence differences (i.e. reciprocally monophyletic groups), multilocus population genetic data can be extremely powerful for elucidating cryptic taxa living in sympatry.

Using multilocus nucleotide sequence data, Ladner (2012) recently described two, previously unsuspected, cryptic species complexes within the *Acropora* syngameon (i.e. a groups of species connected through genetic exchange). *Acropora* is the most diverse and one of the most abundant genera of reef-building corals (Wallace 1999; Veron 2000). The genus is notorious for high levels of shared polymorphism between species, which is due, in part at least, to widespread genetic exchange through introgression (e.g. van Oppen *et al.* 2001, 2002; Vollmer & Palumbi 2002; Palumbi *et al.* 2011; Ladner 2012). Despite the absence of any fixed sequence differences, Ladner (2012) utilized allele frequency variation at 10 loci to demonstrate that two common coral species in Australia, *A. cytherea* and *A. hyacinthus*, actually represent species complexes with at least two and four cryptic species, respectively. These cryptic species were found in complete sympatry and form part of a large syngameon with a complex network of gene flow among species (Ladner 2012). Cryptic species were identified in accordance with the 'genotypic cluster' species definition, which defines species as distinguishable genotypic groups with few or no intermediates when in contact (Mallet 1995). This species definition is practical and appropriate for corals, and many other taxa, because it is independent from theories of how speciation occurs and it uses relatively easily attainable data to identify the groups of individuals that are interbreeding at levels that maintain cohesion, while allowing for introgression (Mallet 1995).

Acropora cytherea and *A. hyacinthus* each have extensive geographic distributions (Wallace 1999; Veron 2000), but nothing is known about the cryptic species composition of these complexes outside Australia. One possibility is that each morphospecies is composed of several narrowly endemic species, each with a distribution that is a subset of that described for the morphospecies as a whole. Several studies have demonstrated that a number of suspected cosmopolitan species actually consist of several geographically distinct, morphologically cryptic species (e.g. Klautau *et al.* 1999; Barroso *et al.* 2010). Taxonomic 'lumping' across geography has been hypothesized to be particularly widespread in marine invertebrates owing to a disconnect between the visible differences used by taxonomists and the, often chemical, mating cues used by many invertebrates (Knowlton 1993; Klautau *et al.* 1999). Under this hypothesis of narrowly endemic species, the presence of multiple sympatric species groups, like that seen in Australia, might represent independent 'endemic syngameons,' each with its own distinct members. An alternative possibility is that the cryptic species in Australia may themselves exhibit large geographic distributions similar to those reported for the morphospecies as a whole. In this case, these morphospecies complexes would likely represent a 'global syngameon,' with similar players in different geographic locations.

Here, we utilize multilocus genetic data in combination with the genotypic cluster species definition to explore the composition of these two species complexes throughout their Indo-Pacific distributions. Specifically, we first investigate 5–7 geographic locations for evidence of sympatric cryptic species within *A. cytherea* and *A. hyacinthus*. Then, we extend the concepts in the genotypic cluster species definition (Mallet 1995) to explore the concordance of cryptic species (i.e. genotypic clusters) among geographic locations. We do this by comparing the levels of genetic similarity among allopatric clusters to those seen among distinct sympatric clusters. Finally, we exploit the presence of sympatric species in multiple geographic locations to explore the consistency of introgression throughout species' sympatric distributions.

Our data suggest the presence of a broadly global syngameon. We find strong support for at least six

widespread cryptic species in these two complexes. Extensive sympatry is observed for species both within and between complexes, and patterns of introgression between species are consistent across geographic locations.

Methods

Sample collection

Coral tissue samples were collected from 12 reefs in seven geographic locations throughout the Indo-Pacific: Palmyra, American Samoa, Fiji, Pohnpei, Palau, northeastern Australia and Zanzibar (Table 1). Samples consisted of 2- to 5-cm-long coral fragments, which were preserved in 70–95% ethanol. *A. cytherea* was sampled from all seven locations; *A. hyacinthus* was sampled from five locations (excluding Fiji and Zanzibar). These seven locations are representative of the full-described range of the *A. cytherea* morphospecies (Wallace 1999; Veron 2000). The existence of *A. hyacinthus* in the Indian Ocean is uncertain; therefore, we restricted our analysis of this species to the Pacific (Veron 2000). Sample sizes, per morphospecies per location, ranged from 7 to 79 colonies (average/location: *A. cytherea* = 15, *A. hyacinthus* = 30) (Table 1). Within a location, samples were obtained from 1 to 3 distinct reefs, all located within 100 km. Samples were combined across reefs for analysis. The Australian samples are identical to those used in the study by Ladner (2012).

Data collection

DNA was extracted from all samples using a slightly modified version of the protocol developed by Wilson *et al.* (2002). Custom primers were used to PCR amplify and sequence 12 regions of the genome including a fragment of the mitochondrial control region (mtCR) and 11 nuclear exons (Table 2). Primers for the mtCR are as in the study by Vollmer & Palumbi (2002). Custom primers were used to amplify the 11 exons (Table S1, Supporting information). All loci were amplified with Fermentas Taq DNA polymerase using a touch-down PCR protocol consisting of an initial melting step of 94 °C for 5 min, 35–45 cycles of (i) 94 °C for 30 s, (ii) X °C for 30 s and (iii) 72 °C for 1 min (X begins at 62 °C and steps down to 48 °C at a rate of 1° per cycle).

PCR products were cleaned using AMPure beads (Agencourt) and sequenced either in-house on a 3100 Genetic Analyzer (ABI) using BigDye Terminator sequencing chemistry (ABI), or by Sequetech (Mountain View, CA, USA). Sequences were edited using Sequencher 4.8. Raw diploid sequences from nuclear regions were phased using a combination of the statistical package PHASE (Stephens & Donnelly 2003) and bacterial cloning. Specifically, individuals that could not be phased statistically with confidence ≥90% at all polymorphic sites were cloned for haplotype verification. At least three sequences were obtained from each clone library to control for PCR errors.

Final marker lengths were chosen to maximize sequence length while eliminating regions with common haplotypes that appear to have originated through recombination, identified as loops in haplotype networks created using TCS v1.21 (Clement *et al.* 2000). Given the large number of samples, it was not feasible to eliminate all regions exhibiting any recombinant haplotypes. Therefore, recombinant haplotypes that occurred only once in the entire data set were simply treated as missing data in the analyses of introgression, so as not to violate the nonrecombination assumption.

Table 1 GPS coordinates and sample sizes for each geographic location

Location	Reef	Latitude	Longitude	*Acropora cytherea*	*Acropora hyacinthus*
American Samoa	Ofu Island	14° 10′ 47″ S	169° 39′ 17″ W	0	37
	Vatia Bay	14° 14′ 57″ S	170° 40′ 25″ W	8	42
Australia	Nelly Bay, Magnetic Island	19° 09′ 29″ S	146° 51′ 34″ E	1	2
	North Orpheus, Palm Islands	18° 34′ 00″ S	146° 29′ 18″ E	15	15
	Pelorus, Palm Islands	18° 33′ 42″ S	146° 30′ 10″ E	0	29
Fiji	Bounty Island	17° 40′ 23″ S	177° 18′ 22″ E	7	0
Palau	Northern Reef	7° 54′ 29″ N	134° 37′ 29″ E	10	34
Palmyra	Sand Island	5° 52′ 33″ N	162° 6′ 25″ W	17	26
Pohnpei	Ahnt Atoll	6° 49′ 35″ N	157° 59′ 32″ E	15	14
	Manta Pass	7° 2′ 14″ N	158° 17′ 25″ E	22	14
Zanzibar	Fawatu Reef	6° 18′ 00″ S	39° 13′ 51″ E	3	0
	Tele Island	6° 02′ 45″ S	39° 06′ 49″ E	10	0
Total				108	213

Table 2 Genomic type, sequence lengths and the best known translated protein BLAST (BLASTx) hits for all loci

Locus	Genome	Length	Putative gene
Control region	mitochondrial	331	Mitochondrial DNA control region
Exon 783	nuclear	155	Rab3a, RAS oncogene family
Exon 2291	nuclear	307	—
Exon 2361	nuclear	249	—
Exon 2980	nuclear	135	—
Exon 3684	nuclear	249	—
Exon 3842	nuclear	284	Lariat debranching enzyme
Exon 4373	nuclear	216	Tubulin tyrosine ligase family member 2
Exon 4706	nuclear	397	Guanine nucleotide–binding protein alpha-12 subunit
Exon 5279	nuclear	135	—
Exon 5491	nuclear	337	Frizzled homolog 4
Exon PMCA	nuclear	545	Plasma membrane calcium ATPase

Dashes indicate the absence of a strong protein match (i.e. no known matches with bit scores >50).

Identification of cryptic species

Structure (Pritchard *et al.* 2000) was used to identify the number of genetic clusters within each location and to assign the individuals to these clusters. Details of our use of structure are included in Appendix S1. *Structure*-identified clusters within locations are considered to represent distinct species utilizing the 'genotypic cluster' species definition (Mallet 1995).

Identification of early-generation hybrids/backcrosses

NewHybrids (Anderson & Thompson 2002) was used to identify early-generation hybrid or backcrossed individuals between putative species within geographic locations. Individuals were considered to be potential early-generation hybrids/backcrosses if *structure* identified them as being ≥10% admixed. Seven ancestry classes were included in each *NewHybrids* analysis. Two correspond to pure individuals of species 1 and species 2, while the other five represent the most likely early-generation hybrid/backcross classes: first-generation hybrid (F1), first-generation backcrosses to species 1 and species 2 and second-generation backcrosses to species 1 and species 2. To increase the power to detect hybrids, all individuals with >90% estimated ancestry from a single putative species were stated to belong to the appropriate pure parental category using the 'z' option. Each run of *NewHybrids* is performed pairwise between two potential parental species. Therefore, analyses were not conducted for individuals that were significantly admixed between more than two species (i.e. ≥10% ancestry from ≥3 species); such individuals were excluded from further analyses owing to ambiguity in species assignments.

Concordance of species across locations

To examine the potential concordance of cryptic species across locations, we expand upon the concept of 'distinguishable genotypic groups in sympatry' put forth in the genotypic cluster species definition (Mallet 1995). Specifically, we investigated the genetic similarity of allopatric *structure*-defined lineages using agglomerative hierarchical clustering (hclust, R v2.12). This method results in a population-level dendrogram where clusters/clades represent the groups of genetically similar populations. The presence of allopatric populations with greater genetic similarity to each other than to other sympatric groups is taken as evidence for these populations belonging to the same cryptic species. Additional species divisions may be present within such clades; however, we do not attempt to distinguish species without evidence from sympatric populations. Dissimilarity matrices were calculated using all *structure*-defined groups across the seven geographic locations using three different measures of genetic distance: (i) squared codominant genotypic distance calculated in *GENALEX* v6.41 (Peakall & Smouse 2006), (ii) Jost's D (Jost 2008) and (iii) G''_{ST} (Meirmans & Hedrick 2011), both calculated in *GenoDive* v2.0b21. Numerical haplotype codes were used (see Appendix S1, Supporting information), which do not take into account the degree of genetic differentiation among haplotypes. This method is preferable because of the large genetic distances often found between common haplotypes within *Acropora* species (Palumbi *et al.* 2011). This pattern is thought to be the result of incomplete lineage sorting and introgression, which lead to the maintenance of old allele groups within species for long periods of time. Therefore, allele identity is more informative than allele

similarity. Clustering was conducted using the average linkage method (UPGMA) (Sokal & Michener 1958).

Reliabilities for the clades in the hierarchical clustering dendrogram were calculated using *Phylip* v3.69 (Felsenstein 2005). Allele frequencies for each *structure*-identified population were calculated using the haplotype codes, and then UPGMA trees were formed (*neighbor*) using Cavalli-Sforza chord distances (*gendist*) for 1000 bootstrap data sets (*seqboot*). *Phyutility* v2.2 (Smith & Dunn 2008) was also used to calculate the leaf stability indices, which are used to examine the stability of individual populations within the tree.

Tests for reciprocal monophyly

After identifying cryptic species groups, we tested for reciprocal monophyly by checking for fixed sequence differences between cryptic species. Comparisons were made on a pairwise basis between all sympatric pairs of cryptic species (i.e. within geographic locations), as well as between each cryptic species after pooling individuals across all sampled locations. Phylogenetic trees were created with *Phylip* v3.69 (Felsenstein 2005) using parsimony criteria (dnapars). All default settings were used, except that 100 bootstrap data sets (seqboot) were run and the input order of individuals for each data set was randomly generated 10 times. For each locus, the resulting 'best' trees were then condensed into a single tree using extended majority rule (consense).

Tests for selection and linkage

After assigning individuals to cryptic species, we tested all polymorphic loci for signatures of selection within species in DNAsp v5 (Librado & Rozas 2009) using Tajima's D and Fu and Li's D* and F* statistics (Tajima 1989; Fu & Li 1993). Diploid loci were tested for Hardy–Weinberg equilibrium using Arlequin v3.11 (Excoffier *et al.* 2005). To ensure independence of all loci, tests for linkage disequilibrium between loci were conducted using Genepop 4.0 (Raymond & Rousset 1995).

Analysis of introgression

Rates of introgression were estimated between each pair of putative species within each location using *Isolation with Migration* (*IMa*) (Hey & Nielsen 2007). Appropriate priors (i.e. when the posterior probability approaches zero on both sides of the peak) were identified for each pair of species through iterative trial. However, when good peaks were unattainable (i.e. tails never approach zero) for the divergence time and ancestral population size parameters, maximum cut-offs were set at 40 and 115, respectively. We chose to conduct pairwise analyses of introgression for two reasons: (i) the difficulty of simultaneously estimating the large number of parameters contained in the models with more than two groups and (ii) lack of a reliable, rooted phylogenetic tree relating putative species within a location, which is required as input into *IMa2* (Hey 2010).

Once appropriate priors were identified, each M-mode analysis was repeated three times from different starting locations to assure convergence (details included in Appendix S1). Tree files from the three M-mode runs were then combined into a single L-mode run to calculate the overall marginal peak locations. Each posterior probability peak is output as a series of probabilities for 1000 evenly spaced bins. Migration parameters were considered to be significantly larger than zero if the 90% highest probability density interval does not include the smallest bin.

All 12 loci were used in the *IMa* analyses, and for most locations, the lengths utilized are identical to those used in the identification of cryptic diversity (Table 2). However, in Palmyra, exon 2980 was trimmed to 124 bp because of patterns of recombination that were particular to this location. Also, in order to utilize the maximum amount of information available, data for the Australian corals were augmented by that from Ladner (2012). This added information from exon 4843 as well as longer sequences for exons 3842 and 5491 (336 and 523 bp, respectively) for the Australian corals.

Results

Cryptic species

Structure analyses support the presence of two sympatric, cryptic species within *A. cytherea* in both Australia and Pohnpei (Fig. 1). By contrast, only a single genetic species of *A. cytherea* was found in the other five sampling locations. For *A. hyacinthus*, multiple sympatric, cryptic species occur in all five locations (Fig. 1). The minimum number of cryptic species in *A. hyacinthus* within a location is two (Palmyra and Pohnpei) and the maximum is four (Australia, Samoa). Species assignments for each individual can be found in Table S2 (Supporting information).

Putative hybrids

In total, 14 individuals, from three locations, were identified as potential early-generation hybrids/backcrosses in the *structure* analyses (six in Australia, four in Pohnpei and four in Samoa). *NewHybrids* clearly identified two of these individuals as early-generation hybrids/backcrosses: AOAH03 (99% probability with Jeffreys prior, 88% uniform) and MPAC01 (98% Jeffreys, 77%

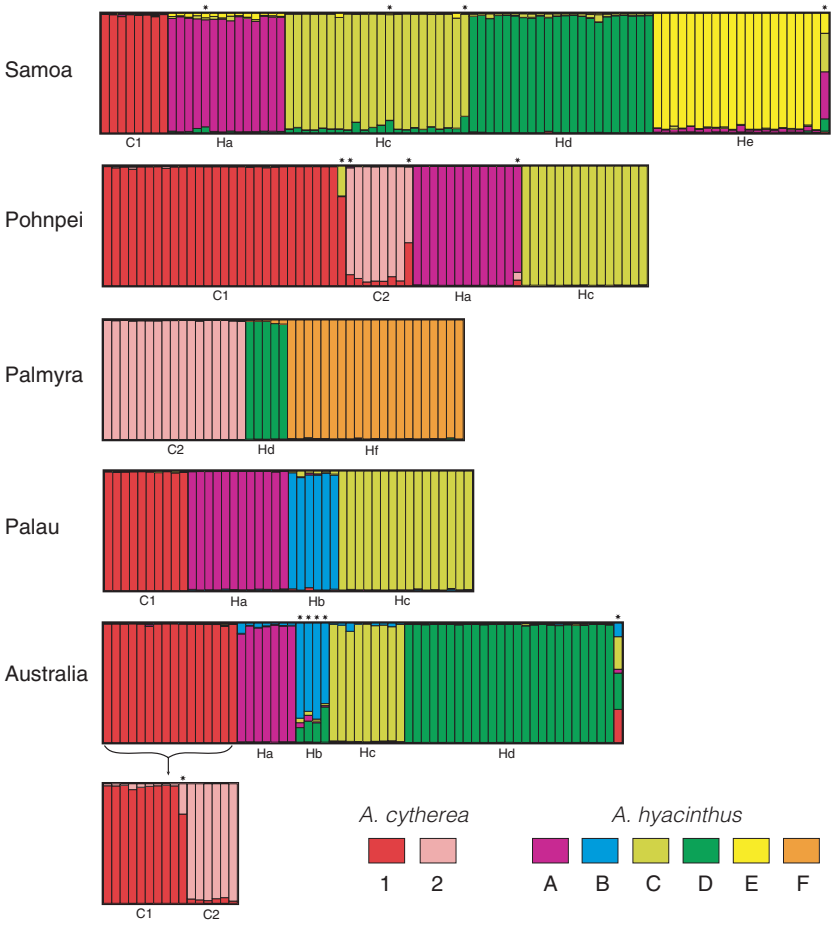

Fig. 1 *Structure* results (admixture model) for the five locations with >1 genetic species. Each vertical column represents one individual. In each case, results from the 10 replicate runs were combined into one figure using *clumpp* (Jakobsson & Rosenberg 2007) and *distruct* (Rosenberg 2004). Colours and labels indicate putative cryptic species assignments within the *Acropora cytherea* and *A. hyacinthus* species complexes, as illustrated in Fig. 3. Stars indicate individuals with ≥10% of their genome estimated as originating from ≥2 species.

uniform) (Fig. 2). Both individuals were sampled in Pohnpei. AOAH03 is most likely a first-generation backcross to *A. cytherea* 1 from a hybrid between *A. cytherea* 1 and *A. hyacinthus* C (76% Jeffreys, 43% uniform). MPAC01 is most likely an F1 hybrid between *A. cytherea* 1 and *A. cytherea* 2 (68% Jeffreys, 51% uniform). Because of their recent-generation hybrid/backcross status, these individuals were not included in the analyses of introgression.

Three putatively hybrid individuals (OIAC09, AOAC01 and AOAC07) had strong probability (≥82%) of assignment to one of the pure parental species with both priors (Table S3, Supporting information). Three additional samples (AHVO01, AHVO03 and OfuH18) were clearly identified as pure parentals with the uniform θ prior (≥96% probability), but were estimated to be some type of early-generation hybrid/backcross with approximately 43–84% probability with Jeffreys prior. In this case, because of the discrepancy between the two priors, we chose to accept the null hypothesis that these individuals are not early-generation hybrids/backcrosses. In all cases, individuals were assigned to the species that *structure* identified as the major genetic contributor.

All four individuals assigned to *A. hyacinthus* B in Australia (see below) were identified as potentially admixed by *structure*. *NewHybrids* is inconclusive about the ancestry of these four individuals; however, this is not surprising given the limited sample size for this species. Furthermore, the lack of any members without potential admixture, which could be utilized by the program to assist in the estimation of parental species allele frequencies, suggests caution about these assignments. Owing to these limitations, we continue to analyse these four as pure individuals of *A. hyacinthus* B, but future work is needed, with larger sample sizes, to better explore the relationship between this species and the others within *A. hyacinthus*. The final two individuals (MIAH02 and OfuH71) were significantly admixed between ≥3 species in the *structure* analysis and were therefore discarded from further analyses.

Geographic concordance of cryptic species

In the hierarchical clustering analysis, several distinct clades were formed, each with members from multiple geographic locations (Fig. 3). Although the exact order of joining differed, the three measures of genetic

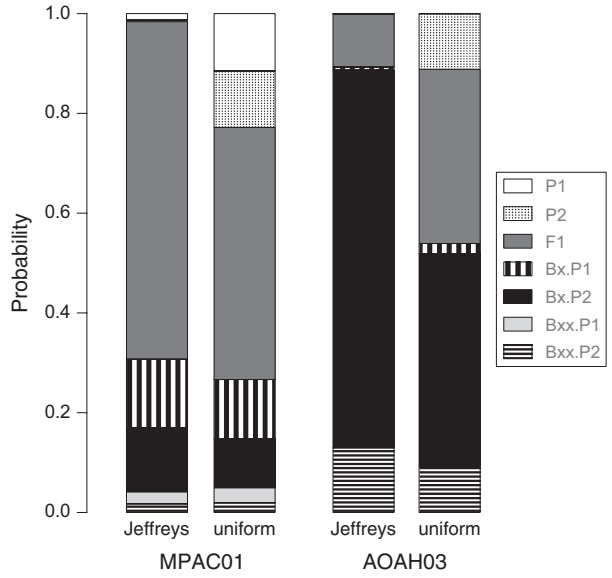

Fig. 2 Results of the *NewHybrids* analysis for the two colonies from Pohnpei found to be early-generation hybrids. Bars represent posterior probabilities of assignment to the seven ancestry categories: P = pure parental species, F1 = first-generation hybrid, Bx = first-generation backcross and Bxx = second-generation backcross. Results are shown using both Jeffreys and uniform priors for the allele frequency parameter. For MPAC01, P1 = *A. cytherea* 2 and P2 = *A. cytherea* 1. For AOAH03, P1 = *A. hyacinthus* C and P2 = *A. cytherea* 1.

distance resulted in identical clades (not shown). Utilizing information from the analyses of cryptic species within locations, the dendrogram can be divided into at least six distinct clades, each with ≥2 locations represented (Fig. 3). Most of these clades have good bootstrap support, especially after the removal of two outlier lineages (see below). The exception to this is *A. hyacinthus* D (bootstrap = 48). This low confidence is driven entirely by the population from Palmyra (excluding Palmyra, the Australian and Samoan corals form a clade with 84% bootstrap support).

These six clades represent putative species under our geographic extension of the genotypic cluster species definition, and they correspond perfectly to the six putative species previously described in Australia (Ladner 2012). We continue to use the naming scheme previously employed for these cryptic species in Australia, that is, numbers for cryptic species of *A. cytherea* (1–2) and letters for *A. hyacinthus* (A–D).

Two cryptic populations of *A. hyacinthus* (one in Samoa and one in Palmyra) cannot be clearly assigned to any of the four cryptic species of *A. hyacinthus*. These outlier lineages may represent additional species that have only been sampled in a single location, and therefore, they have been given unique identifiers (E and F, respectively). These lineages are also two of the three most unstable lineages in the tree (leaf stability: $E = 0.747$, $F = 0.714$), and removal of these two lineages greatly increases the overall stability of the tree (not shown) as well as bootstrap confidence levels for several of the cryptic species clades (Fig. 3).

Tests for reciprocal monophyly

Despite the clear distinctions between cryptic species in the *structure* analyses (Fig. 1), few loci exhibit fixed sequence differences between species (Table S4, Fig. S1, Supporting information). Of 480 pairwise species comparisons within locations (40 comparisons/12 loci), only 27 (5.6%) exhibited fixed sequence differences between species (across seven loci), and 25 (92.6%) of these were between morphospecies complexes (i.e. an *A. cytherea* species vs. an *A. hyacinthus* species). Reciprocal monophyly was even more rare when sequences from each putative cryptic species were pooled across the sampled geographic locations, with only two loci showing any evidence for reciprocal monophyly (pairwise) and no loci exhibiting monophyly of any one species when compared to all other species. *A. cytherea* 1 has one fixed difference at exon 5279 as compared to *A. hyacinthus* B, C and E (bp 108) and it has four fixed bases at the same locus as compared to *A. hyacinthus* D (bp 85, 94, 108, 129). *A. hyacinthus* A and F have one fixed difference (bp 23) in the mitochondrial control region. Overall, these results illustrate a general lack of reciprocal monophyly between cryptic species, especially within a morphospecies complex.

Tests for selection and linkage

None of the 66 locus pairs displayed significant ($P < 0.05$) evidence of linkage disequilibrium (Table S5, Supporting information). A total of 264 tests of Hardy–Weinberg equilibrium were carried out, and no significant deviations were detected ($P < 0.05$) after Bonferroni correction at the level of geographic location. Eight tests (3%) were significant before correction (Table S6, Supporting information). For each of the three tests of neutrality, 253 analyses were conducted (Table S6, Supporting information). Without correcting for multiple tests, 10 (4%), 13 (5.1%) and 11 (4.3%) tests were significant for Tajima's D and Fu and Li's D* and F*, respectively. After correcting for multiple tests, only six analyses are potentially significant (i.e. uncorrected $P < 0.02$, no lower bound specified in *DNAsp*). This includes two D* and four F* analyses that are spread across three cryptic species, three geographic locations and four loci. Therefore, given the overall low number of potentially significant tests, we decided to utilize the full data set in the analyses of introgression.

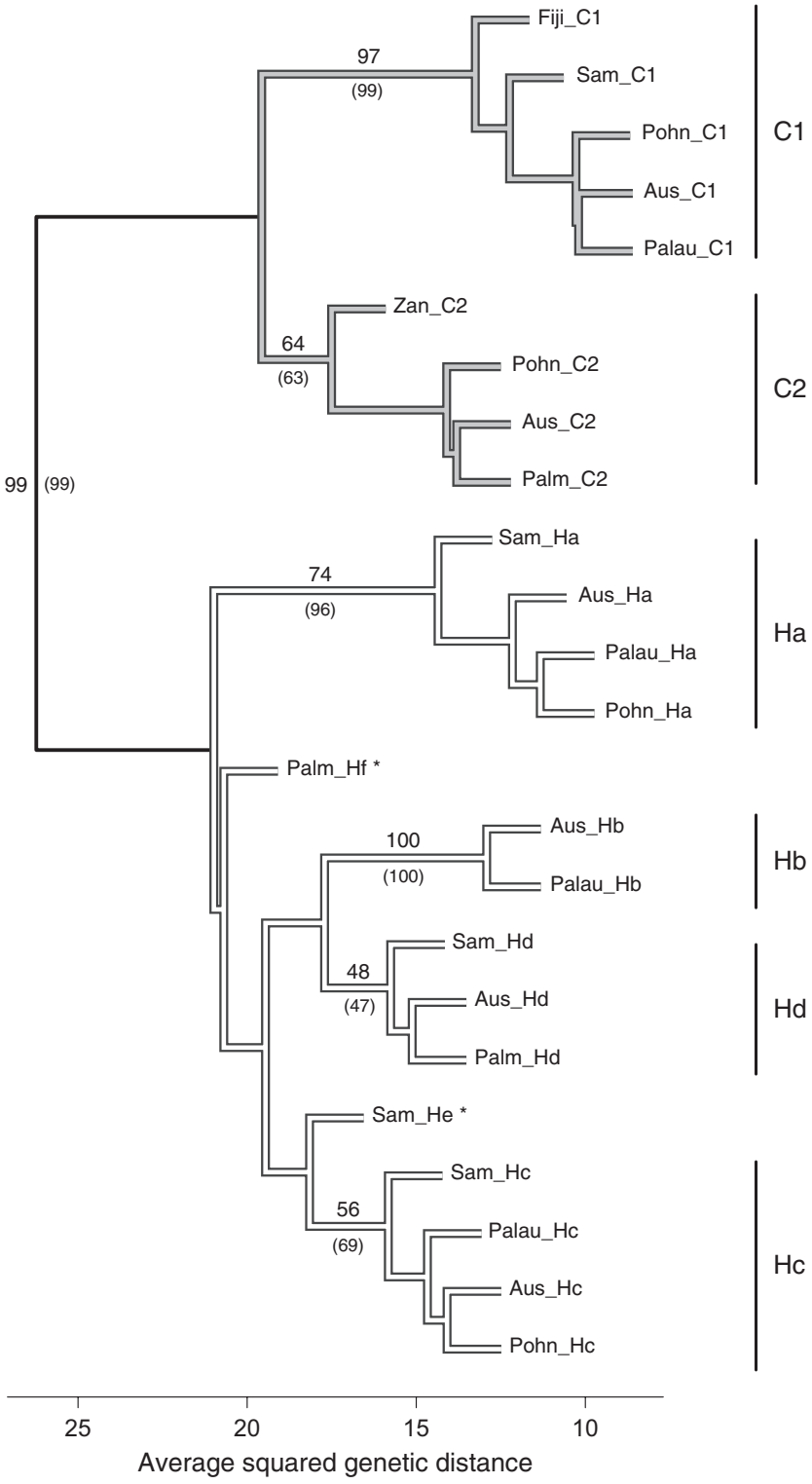

Fig. 3 Dendrogram from the hierarchical clustering of all *structure*-identified groups across geographic locations. The average linkage method was used with squared codominant genotypic distances. Grey branches = *Acropora cytherea*, white = *A. hyacinthus*. Branch labels indicate geographic location of each population. Stars indicate outlier lineages that do not fall clearly within any identified species group. Numbers above the branches indicate bootstrap confidences for each clade using the full data set; numbers under the branches in parentheses indicate bootstraps after removing the two outlier lineages.

Patterns of introgression

In total, 40 pairwise analyses of introgression were conducted across the five geographic locations with ≥2 genetic species sampled (i.e. excluding Fiji and Zanzibar). The effective sample size for L[P] was >10 500 for all *IMa* runs (>20 000 for approximately 96% of runs), and results were consistent across the three replicates for each pairwise comparison.

Genetic introgression has been common, in all locations, between species both within and between the *A. cytherea* and *A. hyacinthus* complexes (Fig. 4; Tables S7

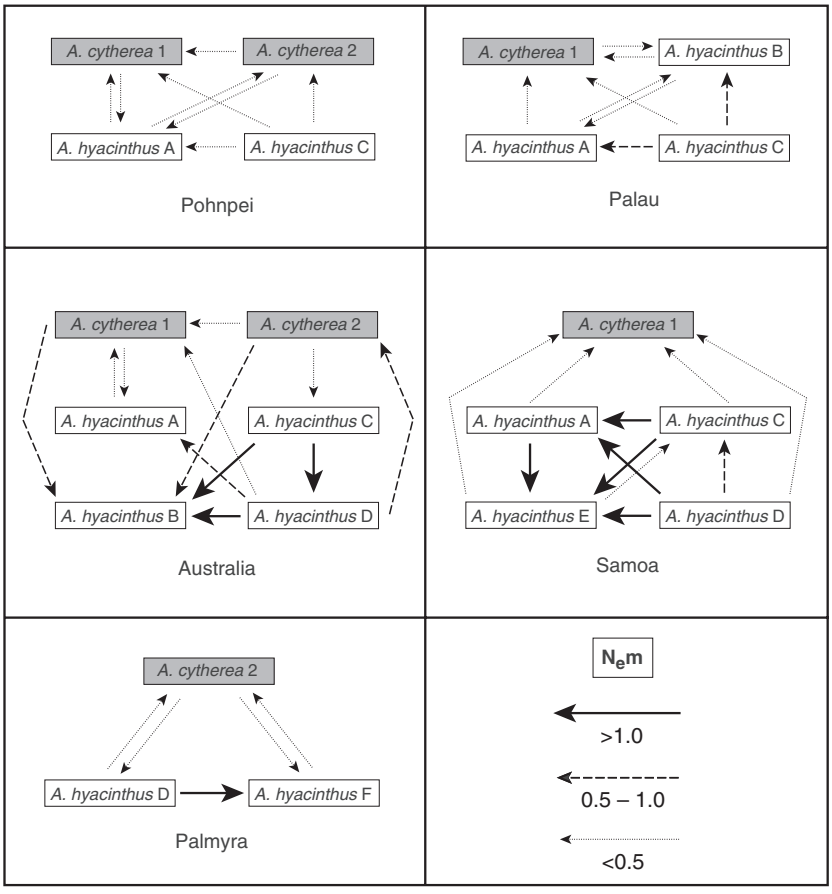

Fig. 4 Significant pathways of introgression detected in each geographic site containing multiple sympatric genetic species. The thickness and dash type of the arrows indicate the frequency of introgression measured as the effective number of migrants ($2N_em$). Grey boxes = *Acropora cytherea* species, white boxes = *A. hyacinthus* species.

and S8, Supporting information). In fact, significant introgression was detected in at least one direction between all pairs of species in Palau, Palmyra, Pohnpei and Samoa and in all but four pairwise comparisons in Australia (11 of 15; Fig. 4). Eight analyses (including seven different species pairs across all five sites) detected bidirectional introgression, 36 analyses (19 pairs, five sites) support a history of unidirectional introgression, and four analyses (four pairs, one site) found no significant evidence for introgression.

In order to compare the patterns of introgression across geography, it is important for pairs of cryptic species to occur together in multiple locations. There are theoretically 28 possible pairwise comparisons that can be made among the eight identified groups (six cryptic species and two outlier lineages) (Fig. 3). Of these, our data include 21 pairs. Three pairs were found in four different geographic locations, 10 were found in two locations, and eight were seen in only a single location.

In general, patterns of introgression between species were consistent across locations (Fig. 4). Each species pair fell into one of the three categories based on the similarity of introgression in different locations: (i) identical patterns (i.e. exactly the same significant introgression pathways in all locations), (ii) broadly consistent patterns (i.e. patterns were not conflicting, but not all inferred pathways for introgression were significant at all sites) and (iii) conflicting patterns (i.e. opposite unidirectional introgression in different locations). Of 13 species pairs sampled in ≥2 locations, four exhibited identical patterns (30.8%), seven were broadly consistent (53.8%) and two were conflicting (15.4%).

Discussion

Species complexes across space

Previous work in Australia found that two common and widespread corals represent complexes of morphologically cryptic species (Ladner 2012). Our geographically expanded data set clearly demonstrates that the presence of multiple, sympatric cryptic species within *Acropora cytherea* and *A. hyacinthus* is not peculiar to Australia, especially for *A. hyacinthus*, in which multiple species were found at every geographic location. Sympatric species were also present in *A. cytherea*, although less commonly, with multiple cryptic species found only in Australia and Pohnpei. While sample sizes are clearly sufficient to detect multiple cryptic species in most locations, we do not presume that they

are large enough to guarantee that all species within these complexes have been described. Instead, we chose to focus our efforts on the relatively common species in each location.

Despite a general lack of reciprocal monophyly (Table S4, Fig. S1, Supporting information), *structure* analyses using multiple loci were able to clearly differentiate cryptic species within locations (Fig. 1). However, these *structure* analyses alone are insufficient for determining the relationships among these groups across locations. Therefore, we augmented these analyses by clustering all *structure*-identified groups based on allele frequencies, thus allowing us to examine the relationships among both sympatric and allopatric populations.

There are two potential explanations for the widespread occurrence of cryptic diversity within these morphospecies: (i) repeated instances of locally restricted or endemic speciation or (ii) a handful of cryptic species, each with widespread distributions. Many supposedly widespread species have turned out to be artefacts of overly conservative systematics (Klautau *et al.* 1999). This is especially true for morphologically simple organisms (Klautau *et al.* 1999; Boissin *et al.* 2008; Barroso *et al.* 2010), but could likely play a role in *Acropora* as well. Geographic variation in morphology is well known throughout the genus (Veron 2000). It is possible that this geographic variation could actually be representative of species divisions. Under the scenario of narrow-range/endemic species, sympatric species would most likely represent local episodes of speciation similar to the radiation events that have been described for *Anolis* lizards in the Caribbean (Losos *et al.* 1998) and cichlid fish in East Africa (Seehausen 2006). In this case, *structure*-defined groups should cluster by geographic location in the hierarchical clustering analysis. Alternatively, this pattern could result from the presence of just a few widespread species within each morphospecies complex. In this case, the same genetic species would be present in multiple locations, and therefore, cryptic species should not cluster based on geography.

The data strongly support the second hypothesis of a handful of wide-ranging cryptic species. The dendrogram from the hierarchical clustering analysis (Fig. 3) exhibits six distinct clades, and each includes only a single population from any geographic location. Specifically, the data support the presence of at least two cryptic species of *A. cytherea* and at least four cryptic species of *A. hyacinthus*, with each species present in 2–4 locations (Fig. 5). The geographically widespread nature of these cryptic species helps to reinforce the divisions among cryptic groups. Sympatric distinctions provide the strongest evidence for species divisions because they demonstrate a lack of interbreeding in the absence of geographic barriers (Mallet 1995). Therefore, the presence of two closely related species within the same location provides strong support for a real biological distinction. For instance, *A. hyacinthus* B from Palau would likely have been included in the *A. hyacinthus* D cluster had it not been for the sympatric occurrence of *A. hyacinthus* B and D in Australia.

It is possible that there are additional cryptic species present within the six clades. For example, the population from Zanzibar is by far the most genetically divergent within the *A. cytherea* 2 clade (Fig. 3). Corals in the Indian Ocean may have been isolated from those in the Pacific long enough that they have evolved species-level differences. Similarly, large genetic breaks have also been described in several reef-dwelling organisms, which are thought to have resulted from repeated events of isolation between the Indian and Pacific Ocean basins because of sea level drops during periods of glaciation (Benzie 1999). Typically, allopatric species are identified using somewhat arbitrary cut-offs on the degree of genetic or morphological divergence (Klautau *et al.* 1999; Veron 2000). This technique is difficult to apply in these *Acropora* complexes because of the gen-

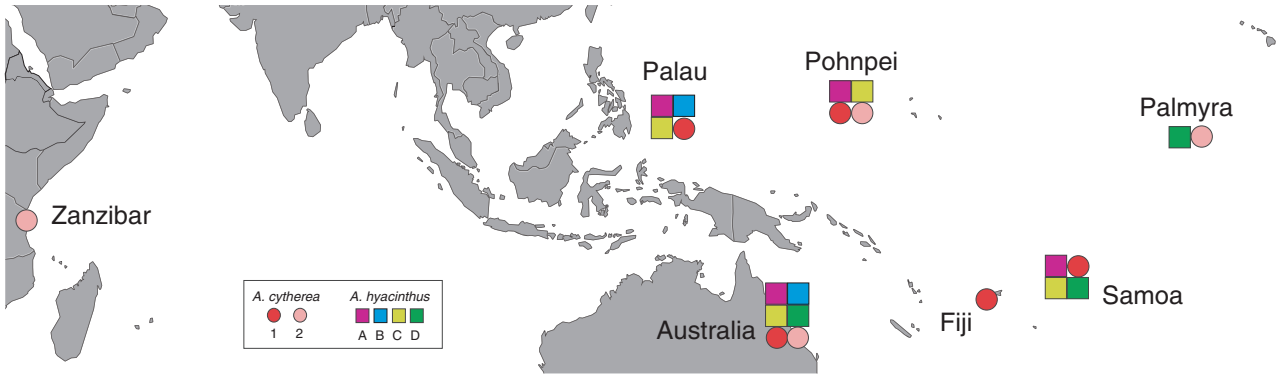

Fig. 5 Geographic distributions of the six cryptic species that were sampled in ≥2 locations. *Acropora cytherea* species = circles, *A. hyacinthus* species = squares.

eral absence of fixed genetic differences between species and the lack of fine-scale morphological characterizations in most locations. Therefore, we have chosen to follow the advice of Darwin (1859) and Mallet (1995), thereby adopting the null hypothesis of 'naturalists having sound judgment' and calling populations conspecific unless there is evidence of distinction from sympatric overlap.

Outlier lineages

Two outlier lineages are also present within the *A. hyacinthus* species complex, currently referred to as species E-F. These two lineages do not clearly associate with any of the clades in the clustering analysis. They likely represent additional cryptic species, which have only been sampled from a single geographic location, perhaps indicating rather restricted distributions. It is also possible that these species do occur in some of the other focal locations, but were not sampled owing to rarity. However, both outlier lineages were relatively common within at least one reef from their respective locations (*A. hyacinthus* E is the most common species in the back-reef pools on Ofu Island, American Samoa, and *A. hyacinthus* F is the most commonly sampled species of *A. hyacinthus* in Palmyra). Another possibility is that these species are more strongly morphologically divergent at other geographic locations and were therefore not considered part of the *A. hyacinthus* morphospecies. In hindsight, it has become clear that *A. hyacinthus* E is the most morphologically distinct species in the *A. hyacinthus* complex within Samoa. Following genetic identification, this species has successfully been identified in the field by the presence of slightly thicker, longer and more distinct branches (Fig. S2, Supporting information). This finding illustrates the enhanced utility of both genetic and morphological data when used in combination.

These two outliers are two of the least stable 'leaves' in the dendrogram from the clustering analysis (Fig. 3). In fact, uncertainty about the placement of these populations results in substantial destabilization of several cryptic species clades. This is especially true for *A. hyacinthus* E (Samoa). Removal of this single population increases the average overall stability of the dendrogram by approximately 2.4% (0.836–0.856). This cumulative effect is larger than that of either of the tree's less stable 'leaves' (*A. hyacinthus* F and *A. hyacinthus* D from Palmyra). *A. hyacinthus* A and C are the two clades that are most strongly stabilized by the removal of *A. hyacinthus* E (bootstraps with/without sp. E: A = 74/95, C = 56/64). This effect is likely the result of high levels of introgression between *A. hyacinthus* species in Samoa. Rates of introgression from *A. hyacinthus* A and C into species E are the two highest we have measured across all species pairs and geographic locations ($2N_em$: 4.14 and 2.84, respectively).

Extensive sympatry

In comparison with reports in other taxa (e.g. Klautau *et al.* 1999; Brown *et al.* 2007; Caputi *et al.* 2007; Elmer *et al.* 2007; Ross *et al.* 2010), these cryptic species of *Acropora* appear to be unusual in their widespread distributions and extensive sympatry. Even in *A. cytherea*, although the two cryptic species were only sampled sympatrically in two of seven locations, both species exhibit extensive distributions (e.g. *A. cytherea* 2 occurs from the eastern central Pacific to the Western Indian Ocean) (Fig. 5). Therefore, these species are likely to have additional locations of overlap. In fact, *A. cytherea* is often relatively rare within locations (Ladner and Palumbi, personal observations), and therefore, it is possible that one of the *A. cytherea* species could simply have been overlooked in some of the locations, especially those with a limited number of *A. cytherea* samples (e.g. Fiji = 7, Samoa = 8). Alternatively, differences in habitat preference and/or variation in the intensity of competition could explain the rarity of sympatrically sampled cryptic species of *A. cytherea*. Even within Pohnpei, there is a strong frequency difference between *A. cytherea* 1 and 2 in the two sampled reefs, with the collection from Ahnt Atoll consisting of even numbers of species 1 and 2, while only species 1 was found in Manta Pass, ~50 km away. Multiple sympatric species of *A. hyacinthus* were sampled from each geographic location, but the number and composition of species present was variable. Two of the cryptic species exhibit distributions throughout the central Pacific (A and C), whereas the other two seem to have more restricted ranges. *A. hyacnithus* B was only found in two locations, both on the Western portion of the geographic area sampled. *A. hyacinthus* D was found in three different locations, which correspond to the southeastern portion of the sampled range.

Impacts on estimates of dispersal

Cryptic diversity, if uncharacterized, can have important implications for the estimation of dispersal in natural populations. Inaccurate estimates could result from unknowingly pooling multiple distinct species, which may each have different dispersal capabilities, and will likely be present in different proportions in different locations. One striking example within our data set is for *A. cytherea* in Pohnpei, which was sampled from two reefs ~45 km apart. Species 1 is present in both locations, while species 2 is only present in one location.

Without knowledge of the cryptic species divisions, one would infer surprisingly low levels of dispersal between these two nearby reefs with five of 12 loci exhibiting significant differentiation between locations (significant F_{ST} values range from 0.036 to 0.47). However, there are no significant F_{ST}s when only *A. cytherea* 1 individuals are included in the comparison. Similar situations would arise when comparing distant locations, which may contain different subsets of cryptic species. For instance, a comparison of *A. hyacinthus* in Palmyra (species D and F) with that in Pohnpei (A and C) is likely to greatly underestimate the long-distance dispersal capabilities of the individual species. An in-depth analysis of dispersal is beyond the scope of this study, but our results provide a framework for future exploration of dispersal in these complexes.

Letting genetics define the boundaries of cryptic species

Many discoveries of cryptic species, including foundational work in corals (Knowlton *et al.* 1992), have been predicated on prior observations of behaviourally or morphologically distinctive groups within recognized species (Mendelson & Shaw 2002). In many cases, genetics are then used to test the evolutionary distinctiveness of these predefined groups (e.g. Mendelson & Shaw 2002; Fukami *et al.* 2004; Stefani *et al.* 2007). One of the risks with this type of analysis is that additional cryptic groups may be overlooked if they exhibit no diagnostic characters according to the recognized phenotypic criteria. It is also possible for the recognized phenotypic criteria to be imperfectly correlated with species boundaries, such as when morphological diversity is shared between species. For example, imagine two morphologically identical corals, each containing two colour morphs: pink and blue. If one species contains 90% blue colonies while the other contains 90% pink, a genetic comparison of groupings based on colour is likely to yield a significant result; however, species assignments based on colour will misidentify one of every 10 individuals. A different approach is to allow the genetic data to directly determine the number of cryptic species (i.e. distinct genotypic clusters; Mallet 1995) and to assign individuals to these cryptic groups (e.g. Caputi *et al.* 2007; Piggott *et al.* 2011). Genetically driven species identification/assignment should eliminate potential biases associated with predefined phenotypic groupings. Confidence in these genetic groupings can then be bolstered by investigations into other potential distinctions between groups, including morphological, ecological, ethological and reproductive traits. The genetic method is strongest when many individuals have been sampled per species (≥5 individuals seem to be necessary for clear distinctions among our corals) and with data from many independent, highly polymorphic markers (the number of necessary markers will depend on the level of divergence between species).

Patterns of introgression

We see significant evidence for widespread introgression both within and between the two morphological species complexes at all geographic locations (Fig. 4; Tables S7 and S8, Supporting information). The highest rates of introgression were measured in Samoa and involve gene flow from an *A. hyacinthus* species (D, C and A, respectively) into *A. hyacinthus* E ($2N_em$ = 2.3, 2.8 and 4.1). In the two most extreme cases, the hierarchical clustering results imply that introgression may be affecting the distinctness of species boundaries. The majority of introgression, however, has been occurring at a frequency that is likely to slow the evolution of reciprocal monophyly (Wright 1931; Anderson 1949; Slatkin 1985) and allow the exchange of adaptive diversity (Anderson 1949), but is unlikely to result in the fusion of species (for 93% of significant migration parameters, $2N_em$ = 0.04–2) (Haldane 1930). We also see an abundance of unidirectional gene flow between species pairs, with 28 of 36 pairwise comparisons of introgressing species (~78%) exhibiting unidirectional exchange.

The combination of geographically widespread species, extensive sympatry and introgression provides strong evidence that *A. cytherea* and *A. hyacinthus* may comprise a wide-ranging, or 'global,' syngameon with hybridization and introgression occurring at multiple geographic locations. Measured patterns of historical introgression are largely consistent with this hypothesis (i.e. patterns of introgression among the same species are similar in different locations) (Fig. 4). The consistency of gene flow across locations is particularly striking for the three species pairs that were each sampled from four geographic locations. Unidirectional gene flow from *A. hyacinthus* C to both *A. hyacinthus* A and *A. cytherea* 1 was significant in three of four locations, while gene flow in the opposite direction was never seen. Similarly, gene flow from *A. hyacinthus* A to *A. cytherea* 1 was significant in all four locations; significant gene flow in the opposite direction was also detected in two of these locations. These results are consistent with introgression occurring in the same directions throughout the large sympatric distributions of these species. This result is in striking contrast to the majority of introgressing taxa in terrestrial systems, which are typically characterized by a narrow zone of hybridization (e.g. Barton & Hewitt 1985; Martinsen *et al.* 2001; Stein & Uy 2006; Hird & Sullivan 2009). However, it is also possible that introgression has only occurred at a subset of locations and that gene flow

within species and/or a range expansion has geographically distributed the genetic diversity that was inherited through introgression. To date, there is no evidence of this location-dependent introgression, but further work investigating contemporary patterns of introgression, as well as the biogeography of these cryptic species, should help to distinguish these hypotheses.

Contemporary exchange

The *IMa* analyses clearly demonstrate that introgression has occurred between many of these corals at some point in time following their initial divergence; however, this does not necessarily mean that introgression is ongoing. The best way to investigate the potential for ongoing introgression is to look for genetic signatures of recent-generation hybrids/backcrosses. The multilocus genetic data clearly support the presence of at least two early-generation hybrids/backcrosses (Fig. 2), thus suggesting that introgression is likely ongoing between, at least, a subset of the investigated species. Furthermore, one individual is most likely a first-generation backcross, and the inferred direction of introgression (*A. hyacinthus* C to *A. cytherea* 1) is consistent with the direction of historical introgression detected with *IMa*. This suggests a continuation of the same pattern of introgressive hybridization that may have been shaping these corals for thousands of years. The overall rarity of hybrid individuals (two of 321) is consistent with the low number of introgression events per species per generation (Table S7, Supporting information).

Conclusion

Multilocus genetic data indicate that the *Acropora cytherea* and *A. hyacinthus* complexes are composed primarily of a handful of geographically widespread, morphologically cryptic species, which occur in extensive sympatry. Genetic introgression has been common between most pairs of species in these two complexes, and the patterns of exchange are relatively consistent across geographic locations. Furthermore, the presence of multiple early-generation hybrids/backcrosses strongly suggests that introgression is likely ongoing. The combination of large distributions, extensive sympatry and consistent patterns of introgression suggests that these two species complexes compose a widespread syngameon with introgression potentially occurring at many geographic locations.

Acknowledgements

We thank the U.S. National Park of American Samoa, the Micronesia Conservation Trust, the Palau International Coral Reef Center and the Nature Conservancy for access to research sites. Funding came from Conservation International, the Woods Institute for the Environment at Stanford University, Stanford's Center for African Studies, the Schmidt Ocean Institute, the Australian Institute of Marine Science and the Gordon and Betty Moore Foundation. Special thanks to the University of Dar es Salaam's Institute of Marine Sciences, especially Dr. Narriman Jiddawi, for assistance in the field and to the Palumbi laboratory for discussions of the data.

References

Anderson E (1949) *Introgressive Hybridization*. J. Wiley, New York, New York.

Anderson EC, Thompson EA (2002) A model-based method for identifying species hybrids using multilocus genetic data. *Genetics*, **160**, 1217–1229.

Barroso R, Klautau M, Solé-Cava AM, Paiva PC (2010) Eurythoe complanata (Polychaeta: Amphinomidae), the 'cosmopolitan' fireworm, consists of at least three cryptic species. *Marine Biology*, **157**, 69–80.

Barton N, Hewitt G (1985) Analysis of hybrid zones. *Annual Review of Ecology and Systematics*, **16**, 113–148.

Benzie JAH (1999) Genetic structure of coral reef organisms: ghosts of dispersal past. *American Zoologist*, **39**, 131–145.

Bickford D, Lohman DJ, Sodhi NS *et al.* (2007) Cryptic species as a window on diversity and conservation. *Trends in Ecology & Evolution*, **22**, 148–155.

Boissin E, Féral JP, Chenuil A (2008) Defining reproductively isolated units in a cryptic and syntopic species complex using mitochondrial and nuclear markers: the brooding brittle star, *Amphipholis squamata* (Ophiuroidea). *Molecular Ecology*, **17**, 1732–1744.

Brown DM, Brenneman RA, Koepfli K-P *et al.* (2007) Extensive population genetic structure in the giraffe. *BMC Biology*, **5**, 57.

Caputi L, Andreakis N, Mastrototaro F *et al.* (2007) Cryptic speciation in a model invertebrate chordate. *Proceedings of the National Academy of Sciences, USA*, **104**, 9364–9369.

Clement M, Posada D, Crandall K (2000) TCS: a computer program to estimate gene genealogies. *Molecular Ecology*, **9**, 1657–1659.

Collins FH, Paskewitz SM (1996) A review of the use of ribosomal DNA (rDNA) to differentiate among cryptic Anopheles species. *Insect Molecular Biology*, **5**, 1–9.

Darwin C (1859) *On the Origin of Species by Means of Natural Selection, or the Preservation of Favoured Races in the Struggle for Life*. J. Murray, London, UK.

DeBach P (1969) Uniparental, sibling, and semi-species in relation to taxonomy and biological control. *Israel Journal of Entomology*, **4**, 11–28.

Elmer KR, Dávila JA, Lougheed SC (2007) Cryptic diversity and deep divergence in an upper Amazonian leaflitter frog, Eleutherodactylus ockendeni. *BMC Evolutionary Biology*, **7**, 247.

Excoffier L, Laval G, Schneider S (2005) Arlequin (version 3.0): an integrated software package for population genetics data analysis. *Evolutionary Bioinformatics Online*, **1**, 47–50.

Felsenstein J (2005) PHYLIP (Phylogeny Inference Package) version 3.6. *Distributed by the Author*. Department of Genome Sciences, University of Washington, Seattle, Washington.

Forsman ZH, Barshis DJ, Hunter CL, Toonen RJ (2009) Shape-shifting corals: molecular markers show morphology is

evolutionarily plastic in Porites. *BMC Evolutionary Biology*, **9**, 45.

Fu Y, Li W (1993) Statistical tests of neutrality of mutations. *Genetics*, **133**, 693–709.

Fukami H, Budd A, Levitan D *et al.* (2004) Geographic differences in species boundaries among members of the Montastraea annularis complex based on molecular and morphological markers. *Evolution*, **58**, 324–337.

Goodman SM, Maminirina CP, Weyeneth N *et al.* (2009) The use of molecular and morphological characters to resolve the taxonomic identity of cryptic species: the case of Miniopterus manavi (Chiroptera, Miniopteridae). *Zoologica Scripta*, **38**, 339–363.

Grundt HH, Kjølner S, Borgen L, Rieseberg LH, Brochmann C (2006) High biological species diversity in the arctic flora. *Proceedings of the National Academy of Sciences, USA*, **103**, 972.

Hafez M, Doutt RL (1954) Biological evidence of sibling species in *Aphytis maculicornis* (Masi). (Hymenoptera, Aphelinidae). *The Canadian Entomologist*, **86**, 90–96.

Haldane J (1930) A mathematical theory of natural and artificial selection. (Part VI, Isolation.). *Mathematical Proceedings of the Cambridge Philosophical Society*, **26**, 220–230.

Hebert PD, Penton EH, Burns JM, Janzen DH, Hallwachs W (2004) Ten species in one: DNA barcoding reveals cryptic species in the neotropical skipper butterfly Astraptes fulgerator. *Proceedings of the National Academy of Sciences, USA*, **101**, 14812–14817.

Heraty JM, Woolley JB, Hopper KR *et al.* (2007) Molecular phylogenetics and reproductive incompatibility in a complex of cryptic species of aphid parasitoids. *Molecular Phylogenetics and Evolution*, **45**, 480–493.

Hey J (2010) Isolation with migration models for more than two populations. *Molecular Biology and Evolution*, **27**, 905–920.

Hey J, Nielsen R (2007) Integration within the Felsenstein equation for improved Markov chain Monte Carlo methods in population genetics. *Proceedings of the National Academy of Sciences, USA*, **104**, 2785.

Hird S, Sullivan J (2009) Assessment of gene flow across a hybrid zone in red-tailed chipmunks (*Tamias ruficaudus*). *Molecular Ecology*, **18**, 3097–3109.

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

Jost L (2008) GST and its relatives do not measure differentiation. *Molecular Ecology*, **17**, 4015–4026.

Klautau M, Russo CAM, Lazoski C *et al.* (1999) Does cosmopolitanism result from overconservative systematics? A case study using the marine sponge Chondrilla nucula. *Evolution*, **53**, 1414–1422.

Knowlton N (1993) Sibling species in the Sea. *Annual Review of Ecology and Systematics*, **24**, 189–216.

Knowlton N, Weil E, Weigt LA, Guzman HM (1992) Sibling species in *Montastraea annularis*, coral bleaching, and the coral climate record. *Science*, **255**, 330.

Ladner JT (2012) 'Hidden diversity in corals and their endosymbionts', Ph.D. dissertation, Stanford University.

Librado P, Rozas J (2009) DnaSP v5: a software for comprehensive analysis of DNA polymorphism data. *Bioinformatics*, **25**, 1451–1452.

Losos JB, Jackman TR, Larson A, de Queiroz K, Rodriguez-Schettino L (1998) Contingency and determinism in replicated adaptive radiations of Island lizards. *Science*, **279**, 2115–2118.

Mallet J (1995) A species definition for the modern synthesis. *Trends in Ecology & Evolution*, **10**, 294–299.

Martinsen G, Whitham T, Turek R, Keim P (2001) Hybrid populations selectively filter gene introgression between species. *Evolution*, **55**, 1325–1335.

Mayr E (1949) *Systematics and the Origin of Species from the Viewpoint of a Zoologist*. Columbia University Press, New York, New York.

Mayr E (1970) *Populations, Species, and Evolution: An Abridgment of Animal Species and Evolution*. Belknap Press of Harvard University Press, Cambridge, Massachusetts.

Meirmans PG, Hedrick PW (2011) Assessing population structure: FST and related measures. *Molecular Ecology Resources*, **11**, 5–18.

Mendelson TC, Shaw KL (2002) Genetic and behavioral components of the cryptic species boundary between *Laupala cerasina* and *L. kohalensis* (Orthoptera: Gryllidae). *Genetica*, **116**, 301–310.

Miles MA, Feliciangeli MD, de Arias AR (2003) American trypanosomiasis (Chagas' disease) and the role of molecular epidemiology in guiding control strategies. *BMJ*, **326**, 1444–1448.

van Oppen M, McDonald B, Willis B, Miller D (2001) The evolutionary history of the coral genus *Acropora* (Scleractinia, Cnidaria) based on a mitochondrial and a nuclear marker: reticulation, incomplete lineage sorting, or morphological convergence? *Molecular Biology and Evolution*, **18**, 1315–1329.

van Oppen M, Willis B, van Rheede T, Miller D (2002) Spawning times, reproductive compatibilities and genetic structuring in the *Acropora aspera* group: evidence for natural hybridization and semi-permeable species boundaries in corals. *Molecular Ecology*, **11**, 1363–1376.

Palumbi SR, Vollmer S, Romano S, Oliver T, Ladner J (2011) The role of genes in understanding the evolutionary ecology of reef building corals. *Evolutionary Ecology*, **26**, 317–335.

Paris CA, Wagner FS, Wagner WH (1989) Cryptic species, species delimitation, and taxonomic practice in the homosporous ferns. *American Fern Journal*, **79**, 46–54.

Peakall ROD, Smouse PE (2006) genalex 6: genetic analysis in excel. Population genetic software for teaching and research. *Molecular Ecology Notes*, **6**, 288–295.

Pfenninger M, Schwenk K (2007) Cryptic animal species are homogeneously distributed among taxa and biogeographical regions. *BMC Evolutionary Biology*, **7**, 121.

Piggott MP, Chao NL, Beheregaray LB (2011) Three fishes in one: cryptic species in an Amazonian floodplain forest specialist. *Biological Journal of the Linnean Society*, **102**, 391–403.

Pinzon JH, LaJeunesse TC (2011) Species delimitation of common reef corals in the genus *Pocillopora* using nucleotide sequence phylogenies, population genetics and symbiosis ecology. *Molecular Ecology*, **20**, 311–325.

Pritchard J, Stephens M, Donnelly P (2000) Inference of population structure using multilocus genotype data. *Genetics*, **155**, 945–959.

Raymond M, Rousset F (1995) GENEPOP (version 1.2): population genetics software for exact tests and ecumenicism. *Journal of Heredity*, **86**, 248–249.

Rosen D (1986) The role of taxonomy in effective biological control programs. *Agriculture, Ecosystems & Environment*, **15**, 121–129.
Rosenberg NA (2004) DISTRUCT: a program for the graphical display of population structure. *Molecular Ecology Notes*, **4**, 137–138.
Ross KG, Gotzek D, Ascunce MS, Shoemaker DD (2010) Species delimitation: a case study in a problematic ant taxon. *Systematic Biology*, **59**, 162–184.
Sáez AG, Probert I, Geisen M *et al.* (2003) Pseudo-cryptic speciation in coccolithophores. *Proceedings of the National Academy of Sciences, USA*, **100**, 7163.
Seehausen O (2006) African cichlid fish: a model system in adaptive radiation research. *Proceedings of the Royal Society B: Biological Sciences*, **273**, 1987.
Slatkin M (1985) Rare alleles as indicators of gene flow. *Evolution*, **39**, 53–65.
Smith SA, Dunn CW (2008) Phyutility: a phyloinformatics tool for trees, alignments and molecular data. *Bioinformatics*, **24**, 715–716.
Sokal R, Michener C (1958) A statistical method for evaluating systematic relationships. *University of Kansas Science Bulletin*, **38**, 1409–1438.
Souter P (2010) Hidden genetic diversity in a key model species of coral. *Marine Biology*, **157**, 875–885.
Stefani F, Benzoni F, Pichon M, Cancelliere C, Galli P (2007) A multidisciplinary approach to the definition of species boundaries in branching species of the coral genus *Psammocora* (Cnidaria, Scleractinia). *Zoologica Scripta*, **37**, 71–91.
Stein A, Uy J (2006) Unidirectional introgression of a sexually selected trait across an avian hybrid zone: a role for female choice? *Evolution*, **60**, 1476–1485.
Stephens M, Donnelly P (2003) A comparison of Bayesian methods for haplotype reconstruction from population genotype data. *The American Journal of Human Genetics*, **73**, 1162–1169.
Stuart BL, Inger RF, Voris HK (2006) High level of cryptic species diversity revealed by sympatric lineages of Southeast Asian forest frogs. *Biology Letters*, **2**, 470.
Tajima F (1989) Statistical method for testing the neutral mutation hypothesis by DNA polymorphism. *Genetics*, **123**, 585–595.
Trontelj P, Fišer C (2009) Cryptic species diversity should not be trivialised. *Systematics and Biodiversity*, **7**, 1–3.
Veron JEN (2000) *Corals of the World*. Australian Institute of Marine Science, Townsville MC, Queensland.
Vollmer SV, Palumbi S (2002) Hybridization and the evolution of reef coral diversity. *Science*, **296**, 2023–2025.
Wallace CC (1999) *Staghorn Corals of the World: A Revision of the Coral Genus Acropora (Scleractinia; Astrocoeniina; Acroporidae) Worldwide, with Emphasis on Morphology, Phylogeny and Biogeography*. CSIRO, Collingwood, Victoria.
Wilson K, Li Y, Whan V *et al.* (2002) Genetic mapping of the black tiger shrimp *Penaeus monodon* with amplified fragment length polymorphism. *Aquaculture*, **204**, 297–309.
Winker K (2005) Sibling species were first recognized by William Derham (1718). *Auk*, **122**, 706.
Wright S (1931) Evolution in Mendelian populations. *Genetics*, **16**, 97–159.
Yoder AD, Burns MM, Génin F (2002) Molecular evidence of reproductive isolation in sympatric sibling species of mouse lemurs. *International Journal of Primatology*, **23**, 1335–1343.

This work represents a portion of J.L.'s PhD dissertation. In general, J.L. is interested in the use of genetic and genomic data to understand the evolutionary histories of populations. He is currently a post-doctoral fellow at the U.S. Army Medical Research Institute of Infectious Diseases. S.P. is a marine population biologist with a particular interest in genetics and conservation in the oceans.

Data accessibility

DNA sequences: Dryad entry doi:10.5061/dryad.1bm52p41 (all loci) and GenBank accessions JQ640583–JQ645963 (only the nine loci ≥200 base pairs).

Table with GenBank accessions listed by individual: Dryad entry doi:10.5061/dryad.1bm52p41.

Structure and IMa input files: Dryad entry doi:10.5061/dryad.1bm52p41.

Supporting information

Additional supporting information may be found in the online version of this article.

Table S1 Primer sequences (5′–3′) used to amplify the exon sequences used in this study, along with their amplicon lengths (in base pairs).

Table S2 Individual sample IDs listed by location of origin and cryptic species assignment.

Table S3 Results from the *NewHybrids* analysis. The posterior probability of assignment to each of the seven ancestry categories is reported.

Table S4 Fixed sequence differences between cryptic species compared pairwise within geographic locations.

Table S5 Results from tests of linkage disequilibrium between pairs of loci across all species.

Table S6 Tests for selective neutrality and Hardy-Weinberg (H-W) equilibrium.

Table S7 Rates of introgression measured from *IMa* analyses.

Table S8 Effective population size parameters measured from *IMa* analyses ($\theta = 4N_e\mu$).

Fig. S1 Maximum parsimony trees for each of the 12 loci.

Fig. S2 Colony photos of *Acropora hyacinthus* from Ofu Island, American Samoa.

Appendix S1 Methods.

Please note: Wiley-Blackwell are not responsible for the content or functionality of any supporting information supplied by the authors. Any queries (other than missing material) should be directed to the corresponding author for the article.