Does genetic distance between parental species influence outcomes of hybridization among coral reef butterflyfishes?

STEFANO R. MONTANARI,* JEAN-PAUL A. HOBBS,† MORGAN S. PRATCHETT,§ LINE K. BAY¶ and LYNNE VAN HERWERDEN**

*AIMS@JCU, Australian Institute of Marine Science, School of Marine and Tropical Biology, James Cook University, Townsville, Qld 4811, Australia, †Centre for Sustainable Tropical Fisheries and Aquaculture, James Cook University, Townsville, Qld 4811, Australia, ‡Department of Environment and Agriculture, Curtin University, Perth, WA 6845, Australia, §ARC Centre of Excellence for Coral Reef Studies, James Cook University, Townsville, Qld 4811, Australia, ¶Australian Institute of Marine Science, PMB 3, Townsville MC, Townsville, Qld 4810, Australia, **School of Marine and Tropical Biology, James Cook University, Townsville, QLD 4811, Australia

Abstract

Christmas Island is located at the overlap of the Indian and Pacific Ocean marine provinces and is a hot spot for marine hybridization. Here, we evaluate the ecological framework and genetic consequences of hybridization between butterflyfishes Chaetodon guttatissimus and Chaetodon punctatofasciatus. Further, we compare our current findings to those from a previous study of hybridization between Chaetodon trifasciatus and Chaetodon lunulatus. For both species groups, habitat and dietary overlap between parental species facilitate frequent heterospecific encounters. Low abundance of potential mates promotes heterospecific pair formation and the breakdown of assortative mating. Despite similarities in ecological frameworks, the population genetic signatures of hybridization differ between the species groups. Mitochondrial and nuclear data from C. guttatissimus × C. punctatofasciatus (1% divergence at cyt b) show bidirectional maternal contributions and relatively high levels of introgression, both inside and outside the Christmas Island hybrid zone. In contrast, C. trifasciatus × C. lunulatus (5% cyt b divergence) exhibit unidirectional mitochondrial inheritance and almost no introgression. Back-crossing of hybrid C. guttatissimus × C. punctatofasciatus and parental genotypes may eventually confound species-specific signals within the hybrid zone. In contrast, hybrids of C. trifasciatus and C. lunulatus may coexist with and remain genetically distinct from the parents. Our results, and comparisons with hybridization studies in other reef fish families, indicate that genetic distance between hybridizing species may be a factor influencing outcomes of hybridization in reef fish, which is consistent with predictions from terrestrially derived hybridization theory.

Keywords: Chaetodontidae, Christmas Island, hybrid ecology, introgression, NEWHYBRIDS, STRUCTURE

Introduction

Hybridization is often a significant evolutionary force that can erode genetic diversity in natural populations (Abbott et al. 2013), but can also contribute to creating and maintaining genotypic novelty (Seehausen 2004; Mallet 2007; Nolte & Tautz 2010; Abbott et al. 2013). Hybridization challenges the assumptions of the biological species concept: to provide a suitable framework for the interpretation of natural hybridization (Frankham et al. 2012), we define species as separate
‘genotypic clusters’ that remain stable in the face of gene flow (Mallet 2007). Hybridization can increase genotypic variation, which may be significant in enhancing adaptation to altered or novel environments (e.g. Darwin finches – Grant & Grant 2002). Further, hybridization can have significant effects on evolution through the formation of hybrid swarms (e.g. sticklebacks – Taylor et al. 2006) and stable hybrid lineages, which coexist in sympathy with parental species (e.g. sparrows – Hermansen et al. 2011). Albeit well understood in terrestrial and freshwater systems, the role of hybridization in shaping the evolution of marine organisms remains, with a few exceptions (e.g. corals – Willis et al. 2006), in need of thorough evaluation.

Several ecological and behavioural processes promote natural hybridization (Willis et al. 2013). Closely related species often share similar ecological niches (habitat, diet), and this can increase the frequency of heterospecific encounters (e.g. fire-bellied toads – MacCallum et al. 1998). Species in low abundance may choose to mate with close relatives when conspecific partners are not available; thus, rarity of one or both species within the contact zone might result in the formation of heterospecific social groups (e.g. Grant & Grant 2008), the breakdown of assortative mating (Arnold 1997) and hybridization. Through ecological observations, the abovementioned studies have identified conditions that favour hybridization in terrestrial systems, but quantitative ecological data are scarce in the marine hybridization literature (Montanari et al. 2012).

Studies have shown a negative correlation between frequency of hybridization and evolutionary divergence (Edmands 2002; Mallet 2005, 2007): genetic distance, with some exceptions (Edmands 2002), is considered a good predictor of reproductive isolation (Singhal & Moritz In press). Further, interspecific gene flow mediated by hybridization (introgression) can occur between species with varying levels of divergence, but appears to be strongest in more closely related species (Mallet 2005). The evolutionary proximity of the parental species facilitates hybridization because closely related species are more likely to be genetically compatible and therefore capable of producing viable hybrids (Mallet 2005). Conversely, if divergence is too extensive, successful hybridization might not be possible due to genetic incompatibility (Mallet 2005; Abbott et al. 2013). Geographical locations where hybridization is most prevalent are ideal to investigate the outcomes of hybridization in taxa with varying degrees of relatedness, because these narrow areas allow controlling for environmental variation that may influence patterns of hybridization (Avise 2000).

Suture zones are geographical locations where hybrid zones naturally cluster (Swenson & Howard 2004) and were defined by Remington (1968) as ‘[bands] of geographical overlap between major biotic assemblages, including some pairs of species or semi-species which hybridize in the zone’. In terrestrial suture zones, the extent of divergence and reproductive isolation between hybridizing species can vary greatly and influence the evolutionary consequences of hybridization (Moritz et al. 2009): here, we propose to test this terrestrially derived notion in marine species.

The best-known tropical marine suture zone is located at the Indo-Pacific biogeographical border, in the eastern Indian Ocean (Hobbs et al. 2009). Here, the fish fauna is characterized by an admixture of Indian and Pacific Ocean taxa (Hobbs & Salmond 2008). Typically allopatric sister species make secondary contact at this border, where they form the highest number of reef fish hybrids reported from any marine location (Hobbs et al. 2009). Christmas Island, Australia, is an oceanic seamount located on the Indo-Pacific biogeographical border (Allen et al. 2007), and its reefs provide a unique location to apply terrestrially derived theory to test ecological frameworks and evolutionary consequences of hybridization in a tropical marine suture zone.

Butterflyfishes inhabit coral reefs worldwide, are dependent on live coral for food (Cole et al. 2008) and shelter (Wilson et al. 2013), readily respond to changes in reef environments (Pratchett et al. 2008) and thus are ideal candidates to examine effects of hybridization on adaptation (Grant & Grant 2002). Butterflyfishes are well known for their propensity to hybridize, with more than 50% of species in the family involved in heterospecific pairing and/or interbreeding (Hobbs et al. 2013). Hybrids occur mostly along zones where major biogeographical provinces overlap (Hobbs et al. 2013), including at least eight butterflyfish species that form hybrids at Christmas Island (Hobbs et al. 2009, 2013). Chaetodon butterflyfishes are an ideal system to investigate reef fish hybridization because many species are monogamous (Yabuta 1997; Pratchett et al. 2006a). Further, even though there may be instances where hybrids go undetected (Hobbs et al. 2013), butterflyfish hybrids are generally easy to recognize through intermediate coloration (McMillan et al. 1999; Montanari et al. 2012; Hobbs et al. 2013).

In a previous study of hybridization between Chaetodon trifasciatus and Chaetodon lunulatus at Christmas Island (Montanari et al. 2012), we hypothesized that the magnitude of divergence between hybridizing parents might influence patterns of introgression in reef fishes based on comparisons of our results to those from the literature (incorporating several geographical locations and reef fish families). By examining hybridization between Chaetodon guttatissimus Bennett, 1832 and Chaetodon punctatofasciatus Cuvier, 1831 at the Indo-Pacific
marine suture zone, the present study allows us to control for taxon- and location-specific factors that may influence patterns of introgression in reef fishes. Therefore, the aims of this article are to (i) determine the ecological and behavioural context of hybridization between *C. guttatissimus* and *C. punctatofasciatus* by assessing abundance, spatial and dietary overlap, and breeding pair formation in parental species and hybrids; (ii) investigate the genetic mechanisms and evolutionary consequences of hybridization between these species through analyses of mitochondrial (mt) and nuclear microsatellite DNA; (iii) discuss similarities and differences in ecology, genetics and potential evolutionary trajectories of *C. guttatissimus* × *C. punctatofasciatus* and *C. trifasciatus* × *C. lunalatus* (Montanari et al. 2012) at the Indo-Pacific suture zone. Specifically, we evaluate whether genetic distance between hybridizing species influences maternal inheritance and introgression in tropical marine fish.

**Material and methods**

**Study location and species**

This study was conducted in October–November 2010 at Christmas Island, Australia, in the northeastern Indian Ocean (10°25′–10°34′S, 105°32′–105°42′E) (Fig. 1, inset). The peppered butterflyfish, *Chaetodon guttatissimus* (Fig. 2A), is wide-ranging in the Indian Ocean, occurring from the East coast of Africa to the Indo-Pacific biogeographical border at Christmas and Cocos (Keeling) Islands (Allen et al. 1998) (Fig. 1). The spot-band butterflyfish, *C. punctatofasciatus* (Fig. 2A), is wide-ranging in the Indian Ocean, occurring from the East coast of Africa to the Indo-Pacific biogeographical border at Christmas and Cocos (Keeling) Islands (Allen et al. 1998) (Fig. 1). The spot-band butterflyfish, 

---

Fig. 1 Map showing the distribution of *Chaetodon guttatissimus* (solid line) and *Chaetodon punctatofasciatus* (dashed line), in the Indian and Pacific Oceans, respectively. Asterisks represent sampling locations outside the Christmas Island hybrid zone (detailed sample sizes are given in Material and Methods). The star symbol identifies the position of Christmas Island within the area of overlap (darker shade of grey) between the two species. Inset shows details of the Christmas Island study sites used for the distribution surveys (black circles) and north coast area covered during the Global Positioning System-assisted surveys (thicker grey line).

Fig. 2 (A) *Chaetodon guttatissimus* (CG) and *Chaetodon punctatofasciatus* (CP) observed in a heterospecific pair at Christmas Island. (B) A hybrid (GPHYB) of this species complex, paired with *C. guttatissimus* (CG) at Christmas Island: the circle highlights the distinguishing maze-like dorsal pattern (cf the clear, straight lines of *C. punctatofasciatus* in photograph A). Maze-like patterns, such as these, have been shown to be characteristic of natural fish hybrids (Miyazawa et al. 2010).
Chaetodon punctatofasciatus (Fig. 2A), is distributed throughout the Western Pacific Ocean, from Indonesia to the Line Islands and from the Ryukyu Islands to the Great Barrier Reef (Allen et al. 1998) (Fig. 1). Christmas Island is the edge of the respective distributions of these butterflyfishes (Allen et al. 1998; Hobbs & Salmond 2008) (Fig. 1), which form heterospecific pairs at this location (Hobbs et al. 2009) (Fig. 2A). Importantly, putative hybrids with coloration intermediate to C. guttatissimus and C. punctatofasciatus (Fig. 2B) are seen at Christmas Island.

Hybrid zone ecology

Abundance, depth distribution, and diet surveys. To assess the abundance of all taxa, underwater visual censuses (UVCs) were conducted at nine sites along the accessible coasts (Fig. 1). In face of the relative rarity of the focal species, transect size was increased (Thompson 2004) during additional abundance surveys along the north coast (Fig. 1). Transect length varied (ranging from 162.1 to 580.5 m), but all data were standardized, with densities presented as the number of fishes per 3000 m². Surveyors swam unidirectionally along depth contours while towing a body board fitted with a Global Positioning System (GPS) receiver. The total area sampled for each of 14 replicate transects was calculated based on independent measures of each GPS track (see Data S1, Supporting Information); t-tests were used to assess significant differences in abundance between parental species and hybrids.

To assess depth distribution of the parent species and hybrids, the depth at which individual fishes were first sighted was recorded during UVCs (n = 30 individuals for all taxa). Depth data were examined using a one-way analysis of variance (ANOVA), comparing the mean depth occupied by parent species and hybrids.

In situ 3-minute feeding observations (following Pratchett 2005) were conducted for all individuals recorded during the depth distribution UVCs. To examine dietary overlap between parent species and hybrids, we recorded the number of bites taken from different benthic prey or substrates. Prey items included predominantly scleractinian corals that were categorized based on genus and growth form sensu Montanari et al. (2012). Dietary composition was analysed using a multivariate analysis of variance (MANOVA), comparing the proportion of bites taken from each prey category by the parents and hybrids. Feeding rates (number of bites over 3 min) were compared between parents and hybrids using a one-way ANOVA, to further identify differences (if any) in feeding behaviour.

Pairing behaviour surveys. During UVCs along the north coast, pair composition was recorded to determine the frequency of assortative pairing behaviour in the C. guttatissimus group. Pairings were noted for all focal fishes encountered, regardless of whether both partners were within the transect area, and therefore included in the abundance counts. Unpaired fishes were small (<70 mm TL) and most likely juveniles. For each parent species and hybrids, expected pairing frequencies were calculated by multiplying the proportional observed abundances by the number of paired individuals, and observational data were analysed for departures from expectations using a chi-square-test.

Hybrid zone genetics

Sampling and DNA extraction. Samples of C. guttatissimus (n = 25), C. punctatofasciatus (n = 18) and C. guttatissimus × C. punctatofasciatus hybrids (n = 16) were collected within the Christmas Island hybrid zone. Chaetodon guttatissimus samples from outside the hybrid zone were collected at Cocos (Keeling) Islands (n = 18) and Zanzibar (n = 1). Similarly, putative purebred C. punctatofasciatus were collected from the Marshall Islands (n = 7) and Guam (n = 1) in the Pacific Ocean. Individual fish were speared whilst SCUBA diving and fin clips were preserved in 80% ethanol for later genetic analysis. Purebred parental species from locations as far as 7500 km away from the hybrid zone were useful in phylogenetic analyses, to tease apart species-specific genetic signals from the signal obtained from the hybrid zone. Chaetodon citrinellus from Lizard Island were used to root all phylogenetic analyses described below (Fessler & Westneat 2007). DNA was extracted from fin clips using 5% Chelex-100 (Walsh et al. 1991).

Mitochondrial and microsatellite genotypes. Mitochondrial cytochrome (cyt) b primers (McMillan & Palumbi 1995), previously utilized in hybridization studies of Chaetodon butterflyfishes (Montanari et al. 2012), were used to amplify 566 bp of the cyt b gene in all samples. Sequences from Montanari et al. (2012) were also used to redraw the relevant haplotype network. Polymerase chain reactions (PCR), PCR evaluation, product purification, sequencing, alignment and manual editing were conducted as described in Montanari et al. (2012). Twenty microsatellite markers developed for C. punctatofasciatus (Montanari et al. 2013) were used to further examine hybridization in the C. guttatissimus × C. punctatofasciatus group. PCR and genotyping were performed as described in Montanari et al. (2013).

Phylo- and population genetic analyses. To identify species-specific and hybridization signals, phylogenetic relationships were inferred using four approaches:
neighbour-joining (NJ), maximum parsimony (MP), Bayesian inference (BI) and maximum likelihood (ML). All phylogenetic model parametrizations and cyt b haplotype network constructions were carried out following Montanari et al. (2012) to allow direct comparisons and minimize model-related variation (see Data S1, Supporting Information). Population genetic analyses followed the same protocols as described in Montanari et al. (2012) and did not include Zanzibar C. guttatissimus ($n = 1$) or Guam C. punctatofasciatus ($n = 1$), due to small sample sizes (see Data S1, Supporting Information).

Microsatellite genotypes were partitioned into clusters assuming an admixture model with independent allele frequencies between populations, using STRUCTURE v2.3.4 (Pritchard et al. 2000). Each value of $k$ (set from 1 to 10) was independently evaluated 20 times, with 1 500 000 iterations following a 100 000-long burn-in (Gilbert et al. 2012). The best fit model was chosen with the Evanno method (Evanno et al. 2005) implemented in STRUCTURE HARVESTER v0.6.93 (Earl & vonHoldt 2012), and values of $\Delta k$ plotted and presented as Supporting Information. Admixture coefficients ($Q$), averaged over the 20 independent runs, were visualized by means of a barplot with credibility regions for $k = 2$ (corresponding to the parental species irrespective of geographical origin). Posterior probabilities, based on microsatellite genotypes, of individuals belonging to six classes (pure parental species, F1 or F2 hybrids and backcrosses in either direction) were calculated using NEWHYBRIDS (Anderson & Thompson 2002). Populations outside the hybrid zone were designated as pure parental species as prior information, and the chain was run for 1 500 000 iterations, after 150 000 burn-ins. Probabilities were subsequently averaged at population level. A discriminant analysis of principal components (DAPC) (Jombart et al. 2010) was run on all loci to investigate the relationship between the sampled populations.

The STRUCTURE, NEWHYBRIDS and DAPC analyses described above were also run on genotypes from Montanari et al. (2012) and added as Supporting Information. By choosing a number of PCs equal to the number of individuals divided by three and a number of DA eigenvectors corresponding to the number of populations minus one in both analyses, the genotypic variability retained in DAPC was similar between the two hybridizing butterflyfish groups. This allowed direct comparisons, thus highlighting key differences in the evolutionary consequences of hybridization in the two Chaetodon species groups at the Christmas Island suture zone.

**Results**

**Hybrid zone ecology**

*Abundance, depth distribution and diet.* Hybrid Chaetodon guttatissimus × Chaetodon punctatofasciatus were relatively common at Christmas Island ($2 \pm 0.47$ SE individuals per 3000 m$^2$) and at least as abundant as the least common parental species, C. punctatofasciatus ($2 \pm 0.47$ SE) ($t_{(26)} = 0.42, P = 0.68$) (Table 1). Chaetodon guttatissimus was significantly more abundant ($40 \pm 4.5$ SE) than C. punctatofasciatus ($t_{(13)} = 8.32, P < 0.0001$) (Table 1).

The ecology (specifically habitat use and dietary composition) of C. guttatissimus, C. punctatofasciatus and their hybrids was very similar. There was no significant difference in depth distribution between C. guttatissimus (average depth $16.6 \pm 0.49$ SE) and C. punctatofasciatus ($15.5 \pm 0.49$ SE) ($F_{1,89} = 3.14, P = 0.08$) (Table 1).

<table>
<thead>
<tr>
<th>Chaetodon guttatissimus group</th>
<th>C. trifasciatus group</th>
</tr>
</thead>
<tbody>
<tr>
<td>Parental species abundance</td>
<td>One parent rare (2 individuals per 3000 m$^2$)</td>
</tr>
<tr>
<td>Hybrid abundance</td>
<td>As abundant as rare parent</td>
</tr>
<tr>
<td>Parental depth distribution</td>
<td>Range: 13–17 m; largely overlapping (&gt;93%)</td>
</tr>
<tr>
<td>Hybrid depth distribution</td>
<td>Overlapping (&gt;99%) with parents</td>
</tr>
<tr>
<td>Parental species diet</td>
<td>Generalist corallivores; largely overlapping (&gt;73%)</td>
</tr>
<tr>
<td>Hybrid diet</td>
<td>Generalist corallivore; overlapping (&gt;76%) with parents</td>
</tr>
<tr>
<td>Parental species pairing behaviour</td>
<td>Nonassortative</td>
</tr>
<tr>
<td>Hybrid pairing behaviour</td>
<td>Pairing with both parents; nonassortative</td>
</tr>
</tbody>
</table>

© 2014 John Wiley & Sons Ltd
The parental species occupied relatively narrow, largely overlapping, depth ranges (Table 1). The depth distribution of the hybrids (16.2 m ± 0.50 SE) was not statistically different from that of either parent species ($F_{2.29} = 1.47, P = 0.235$). Similarly, dietary composition was not significantly different between parent species ($Pillai’s \; Trace_{(34)} = 0.51, P = 0.085$) (Table 1). Both parental species most frequently fed on encrusting Montipora and massive Porites, which are among the most common coral genera at Christmas Island. The hybrids fed largely on the same prey as their parental species ($Pillai’s \; Trace_{(34)} = 0.51, P = 0.085$) (Table 1). The feeding rates (number of bites per 3 min observation) of parent species and hybrids were not significantly different ($F_{2.29} = 2.03, P = 0.14$).

Pairing behaviour. The relative number of individuals that paired with conspecifics, heterospecifics or hybrids was generally proportional to the abundance of these individuals (Fig. 3A, Table 1). The pairing behaviour of *C. punctatofasciatus* did not significantly deviate from the frequencies expected based on abundances ($\chi^2_{(2, n=30)} = 2.89, P > 0.24$), indicating that this species is pairing nonassortatively under these conditions (Fig. 3A, Table 1). *Chaetodon guttatissimus* appeared to actively choose to pair heterospecifically (disassortative mating) ($\chi^2_{(2, n=264)} = 14.91, P < 0.001$), but this may be a statistical artefact of the large sample size for this species (Fig. 3A). Hybrids were never observed paired together and formed pairs with the parental species nonassortatively ($\chi^2_{(2, n=28)} = 3.25, P > 0.19$) (Fig. 3A). This indicates that hybrids are likely choosing partners based on their prevalence rather than phenotype.

**Hybrid zone genetics**

Five hundred and sixty-six base pairs (bp) of the mitochondrial cyt $b$ region were resolved for a total of 86 individuals in the *C. guttatissimus* group. The alignment contained 92 parsimony informative sites and identified 49 discrete haplotypes (Fig. 3B). Twenty microsatellite loci reliably amplified and were scored in 83 individuals: one *C. punctatofasciatus* from Christmas Island was excluded due to >20% missing data. Population level tests showed significant departures from HWE in 26 of 100 tests after sequential Bonferroni correction ($\alpha = 0.01$) (Table S1, Supporting information). Eighteen (69%) of these HWE departures were concentrated at five loci (Cpun3, 4, 7, 9 and 13) (Table S1, Supporting information). Null alleles contributed to departures from HWE in all abovementioned loci. *Chaetodon punctatofasciatus* from Christmas Island had the most private alleles (17) compared with all other taxa in this group (Table S1, Supporting information).

**Phylogenetic relationships**

Congruent phylogenetic relationships were inferred with four methodologies (NJ, MP, BI and ML), and a clear separation between the two parental clades was strongly supported by all analyses (Fig. 3B). Six fixed nucleotide changes (1% divergence at cyt $b$) separated the two parental species, despite evidence of some interspecific mtDNA exchange (Fig. 3B). All *C. punctatofasciatus* individuals and three of 44 individuals (7%) identified in the field as *C. guttatissimus* based on coloration were contained in a single clade (Fig. 3B). Two of these individuals were from cocos (Keeling) Islands, outside the hybrid zone of Christmas Island (Fig. 3B).

Hybrids in the *C. guttatissimus* group shared haplotypes with both parental clades, indicating a bidirectional maternal contribution to hybridization (Fig. 3B). This contrasts with the *Chaetodon trifasciatus* group, where redrawn haplotype relationships from Montanari et al. (2012) show 5% divergence between the parent species at cyt $b$, and all hybrids occur in only one of the two parental clades (unidirectional maternal contribution – Fig. S1B, Supporting information).

**Population genetic structure**

Cytochrome $b$ haplotype ($h$) and nucleotide ($\pi$) diversities, as well as gene diversity based on microsatellites (1-Q inter) within the Christmas Island hybrid zone, were high for all taxa in the *C. guttatissimus* group (Table S2, Supporting information). The AMOVA fixation index for mtDNA cyt $b$ was $\Phi_{st} = 0.48 \; P < 0.0001$. Microsatellites indicated a clear separation between parental species and hybrids and had raw $F_{st} = 0.038, P < 0.0001, D_{est} = 0.115$ and ENA-corrected values that were comparable to raw values, indicating low confounding effects from null alleles (Table S4, Supporting information). Nearly all pairwise $F_{st}$ tests were significant for mitochondrial and nuclear markers, and this was further confirmed with $D_{est}$ (Table S3, Supporting information). Genetic structure was evident between parental species irrespective of geographical location (Tables S3, Supporting information). Analyses of cyt $b$ did not detect significant intraspecific structure between populations of either *C. guttatissimus* or *C. punctatofasciatus* (Table S3A, Supporting information). Microsatellites indicated weak intraspecific structure between *C. guttatissimus* populations, but not between *C. punctatofasciatus* populations (Table S3B, Supporting information), possibly due to small sample size of the Marshall Island population. The hybrid population significantly differed from all other populations (Table S3, Supporting information). Mitochondrial data indicated that hybrids are less differentiated from the parental species.
Fig. 3 (A) Pairing frequencies of *Chaetodon guttatissimus* (yellow), *C. guttatissimus × Chaetodon punctatofasciatus* hybrids (orange) and *C. punctatofasciatus* (red). All three taxa are colour-coded according to the legend below. Bars represent observed pairings from Christmas Island, and dots represent expected pairing frequencies based on observed taxon abundance. Observed pairing does not statistically deviate from expectations, indicating that taxa are pairing nonassortatively. (B) MST showing haplotype relationships in the *C. guttatissimus* group. Each circle represents one individual and is colour-coded for taxon and geographical origin. Each black dot on connecting branches represents one substitution (bp). Bootstrap support values for phylogenetic relationships inferred by NJ, MP, maximum likelihood and posterior probabilities from BI are shown for the partition between the two major clades in the species group. (C) Scatterplot of DAPC (Jombart et al. 2010) performed on 20 microsatellite loci for five populations of the *C. guttatissimus* group. Populations are shown by colours and 95% inertia ellipses, squares represent individual genotypes. Axes show the first two discriminant functions, and eigenvalues the genetic information retained by discriminant functions. (D) Barplot of STRUCTURE admixture coefficients based on 20 microsatellite loci in five populations of the *C. guttatissimus* group. Bars represent individuals, black lines are 90% credibility regions, and subdivisions show the genotypic admixture between clusters (k = 2, representing the parent species). Colour coding as well as taxon and geographical location abbreviations are valid throughout all panels: CG, *C. guttatissimus*; CP, *C. punctatofasciatus*; GPHYB, *C. guttatissimus × C. punctatofasciatus*; CK, Cocos (Keeling) Islands; GUA, Guam; RMI, Republic of Marshall Islands; XMAS, Christmas Island; ZAN, Zanzibar.
within the hybrid zone (Table S3A, Supporting information). However, microsatellite data showed that the hybrids are less differentiated from *C. punctatofasciatus* within the hybrid zone, but more from *C. guttatissimus* (Table S3B, Supporting information).

STRUCTURE identified two clusters (Fig. S2A, Supporting information): some admixture was detected between parental species (Fig. 3D), both outside and within the hybrid zone – consistent with previous mtDNA and microsatellite analyses. Interestingly, the *C. guttatissimus* population from Cocos (Keeling) showed a slightly higher level of admixture with *C. punctatofasciatus* than the Christmas Island population (Fig. 3D). Most notably, the *C. punctatofasciatus* population in the contact zone showed greater levels of admixture than *C. guttatissimus* (Fig. 3D), but with high levels of variability in the estimates. The hybrids’ intermediacy was evident compared to both parental species, particularly *C. guttatissimus* (Fig. 3D). Two clusters were also identified in the *C. trifasciatus* group (Montanari et al. 2012) (Fig. S2B, Supporting information): this data set shows lower levels of parental admixture (particularly in the *Chaetodon lunulatus* populations) and hybrid intermediacy is clear in this group (Fig. S1D, Supporting information). In both butterflyfish groups, however, STRUCTURE lacks the resolution to reliably detect backcrossing and hybrid classes (possibly as a result of the small sample sizes and limited number of molecular markers).

NEWHYBRIDS assigned over 95% of *C. guttatissimus* individuals to their pure species, in both populations of origin (Fig. 4). As also suggested in STRUCTURE, the Cocos (Keeling) population had a somewhat greater probability of introgression than the Christmas Island population (Fig. 4). The hybrids were clearly intermediate and were mostly either assigned to *C. guttatissimus* or designated as F2 hybrids (Fig. 4). Likewise, a similarly high probability of being F2 hybrids (almost 30%) was assigned to the Christmas Island population of *C. punctatofasciatus*, consistent with the suggested pattern of introgression observed in the STRUCTURE analysis (Figs 3D and 4). This contrasts with the *C. trifasciatus* group (Montanari et al. 2012), in which both parental species were assigned to their respective pure clusters with >92% probability irrespective of geographical location (Fig. S3, Supporting information). The hybrids in this group had a range of probabilities of being assigned to either parental cluster, F1, F2 or either backcross (10–25%) (Fig. S3, Supporting information). Moreover, approximately 60% of assignments were to interpupal crosses, the remainder being to pure parental clusters (Fig. S3, Supporting information). In both NEWHYBRIDS analyses, the standard deviation around the mean posterior probabilities was negligible for all taxa, except for the hybrids, underlining the uncertainty associated with assigning these intermediate individuals.

DAPC examined the relationship between clusters, predefined as combinations of taxon and geographical location (Fig. 3C). The hybrid population was distinct from all others, and hybrid genotypes were intermediate between parental species’ genotypes (Fig. 3C). Little partitioning was evident between populations of the same species (Fig. 3C), consistent with other analyses. *Chaetodon guttatissimus × C. punctatofasciatus* hybrids occupied a broad parameter space close to their parental clusters, and confidence ellipses were shared in seven of 16 individuals (Fig. 3C). In contrast, microsatellite data from Montanari et al. (2012), presented in a reparametrized DAPC (see Material and Methods), show almost no overlap of hybrid and parental genotypes in the *C. trifasciatus* group (Fig. S1C, Supporting information).

### Fig. 4 Posterior probabilities, based on microsatellite data, of individuals of the *Chaetodon guttatissimus* group belonging to six classes: pure parental species, F1 or F2 hybrids and backcrosses (Bx) in either direction. Individual data were averaged within population of origin. Colour codes for the six classes are given in the legend. Each bar represents one population and is designated by species and geographical location (for sample sizes refer to Material and Methods). CG, *C. guttatissimus*; CP, *Chaetodon punctatofasciatus*; GP, *G. punctatofasciatus* × *C. punctatofasciatus*; CK, Cocos (Keeling) Islands; RMI, Republic of Marshall Islands; XMAS, Christmas Island.
Discussion

Hybrid zone ecology

*Chaetodon guttatissimus* and *Chaetodon punctatofasciatus* as well as *Chaetodon lunulatus* and *Chaetodon trifasciatus* (Montanari et al. 2012) have come into secondary contact at the tropical marine suture zone of Christmas Island. Our results highlight several ecological factors that are likely to contribute to the propensity of these species to hybridize. Some degree of habitat overlap is a necessary precursor to hybridization in sexual vertebrates (e.g. Bombina toads – Vines et al. 2003). The distributions of *C. guttatissimus* and their hybrids largely overlap at Christmas Island: all taxa occupy sites with similar exposure (north coast) and have relatively narrow and consistent depth ranges. Habitat overlap was also reported between *C. lunulatus* and *C. trifasciatus* (Montanari et al. 2012) and has been documented for a large number of hybridizing marine fishes (Nichols 1918; Norman 1934; Schultz & Smith 1936; Gosline 1948; Randall 1956; Feddern 1968; Hettler 1968; Fischer 1980; Rao & Lakshmi 1993; Frisch & van Herwerden 2006; Yaakub et al. 2006, 2007; Marie et al. 2007; Hobbs et al. 2013). Such overlap increases the chance of heterospecific encounters between hybridizing butterflyfishes at the Indo-Pacific suture zone.

*Chaetodon guttatissimus*, *C. punctatofasciatus*, *C. trifasciatus* and *C. lunulatus* are relatively specialized obligate coralivores (Cole et al. 2008; Pratchett 2013), and their feeding mode has been confirmed through both gut content analyses (Harmelin-Vivien 1989; Sano 1989) and direct observations (Pratchett 2005). This study and data from Montanari et al. (2012) indicated that, in each hybridizing group, the two parental species and their respective hybrids fed on the same suite of coral prey. Further, gut content analyses and direct feeding observations in tropical marine fishes belonging to the Acanthuridae (Randall 1956), Pomacanthidae (Feddern 1968) and Serranidae (Fischer 1980) showed, in all cases, that the diets of hybridizing parents and hybrids were essentially the same. In synergy with overlap in habitat use, dietary overlap further increases encounter probability between hybridizing butterflyfishes at Christmas Island.

Rarity of conspecific mates is considered a promoting factor in hybridization among terrestrial organisms (e.g. Darwin finches – Grant & Grant 2002) and reef fishes (Randall et al. 1977; Pyle & Randall 1994; van Herwerden et al. 2002; Maruska & Peyton 2007; Hobbs et al. 2009). Although *C. guttatissimus* is relatively common at Christmas Island, its sister taxon, *C. punctatofasciatus*, is rare. At Christmas Island, *C. punctatofasciatus* occurs in densities 40–100 times lower than those found at locations near the centre of its distribution range (e.g. Indonesia and Palau – Findley & Findley 2001). The local rarity of *C. punctatofasciatus* may explain why many of these individuals are found in heterospecific pairs. *Chaetodon trifasciatus* and *C. lunulatus* are both rare at Christmas Island (Montanari et al. 2012), and their abundances are one to three orders of magnitude lower compared with any other location for which abundance data are available (Adrim & Hutomo 1989; Findley & Findley 2001; Pratchett et al. 2004, 2006b; Pereira & Videira 2005).

The frequency of heterospecific pair formation was proportional to the abundance of parent species in both the *C. guttatissimus* and *C. trifasciatus* (Montanari et al. 2012) hybrid groups at Christmas Island. This supports the hypothesis that rare species (and hybrids) are forming heterospecific pairs based on encounter rates and that a rare species (or hybrid) will choose a partner based on availability rather than the phenotypic identity of the individual. A breakdown in assortative mate choice has been reported for other pair-forming *Chaetodon* butterflyfishes that are known to hybridize (McMillan et al. 1999; Hobbs et al. 2013). The parent species and hybrids examined in this study and in Montanari et al. (2012) belong to subgenera thought to be exclusively monogamous (Pratchett et al. 2006a; Craig et al. 2010), and indeed, examination of the gonads of heterospecific pairs at Christmas Island revealed that these pairs always comprised a mature male and a mature female (Hobbs unpublished data). Therefore, the observed heterospecific breeding pairs are likely producing the hybrids seen at Christmas Island. Overall, our observations indicate that the ecological and behavioural processes that set the scene for hybridization are similar across *Chaetodon* butterflyfish hybrid groups at Christmas Island and probably explain the onset of hybridization in pair-forming butterflyfishes elsewhere (Hobbs et al. 2013).

Hybrid zone genetics

Mitochondrial and nuclear DNA analyses confirmed hybridization in both the *C. guttatissimus* and *C. trifasciatus* (Montanari et al. 2012) groups. However, despite similarities in the ecological context of hybridization in the two complexes, the genetic mechanisms are clearly different. In *C. guttatissimus* – *C. punctatofasciatus*, which are 1% divergent at cyt b as measured in this study, hybrids shared mtDNA with both parental clades, indicating bidirectional maternal contribution to hybridization, a mode previously reported in reef fishes (McMillan et al. 1999; van Herwerden and Doherty 2006). This is consistent with field observations of heterospecific pairs in which females were identified as
either *C. guttatissimus* or *C. punctatofasciatus*, but in contrast with unidirectional mitochondrial inheritance in *C. trifasciatus* × *C. lunulatus*, 5% divergent at cyt *b* (Montanari et al. 2012). In previous studies of reef fish hybridization, most or all hybrids reportedly shared haplotypes with the more abundant parental species, suggesting sneak mating by males of the rare species with females of the common species, likely due to rarity suggesting sneak mating by males of the rare species haplotypes with the more abundant parental species, supporting introgression in ongoing and that hybrids are still contributing to interspecific gene flow.

An alternative interpretation of our detection of mtDNA and nDNA introgression between *C. guttatissimus* and *C. punctatofasciatus* is incomplete lineage sorting. Recent and robust phylogenies of the Chaetodontidae based on two mtDNA and rRNA markers unequivocally partition the two sister species, suggesting that the lineages have sorted completely (Littlewood et al. 2004; Hsu et al. 2007). Moreover, our phylogenetic analyses have shown that *C. guttatissimus* and *C. punctatofasciatus* populations sampled from locations most distant from the hybrid zone have distinct, species-specific mtDNA haplotypes. However, detection of introgressed individuals outside the hybrid zone points to possible incomplete lineage sorting, because allopatric populations of these species show some degree of admixture, irrespective of the geographical distance between them. To discriminate between this scenario and introgressive hybridization, further studies should include more samples across the distribution ranges of these species and apply genotyping-by-sequencing techniques to increase resolution.

**Consequences of hybridization**

Contrary to what has been observed in the Solomon Islands-Papua New Guinea hybrid zone involving *C. punctatofasciatus* and *Chaetodon pelewensis* (McMillan et al. 1999) and in another hybridizing fish, *Acanthurus leucosternon*, at Christmas Island (Marie et al. 2007), the introgressive hybridization between *C. guttatissimus* and *C. punctatofasciatus* is not strong enough to swamp species-specific signals. Although this pattern could be the result of chance given the small sample size, our data suggest that divergence between *C. guttatissimus* and *C. punctatofasciatus* is decreasing within the hybrid zone and gene flow mediated by the hybrids appears to be ongoing. Persistence of hybrids and introgressed individuals at Christmas and Cocos (Keeling) Islands may eventually confound species signals in the *C. guttatissimus* hybrid group, resulting in a hybrid swarm (sensu Taylor et al. 2006). Alternatively, the presence of novel genotypes (and the high genetic diversity) in the hybrid population at Christmas Island may one day enable hybrids to exploit niches not occupied by parent species. This process was documented in terrestrial *Geospiza* Darwin finches – Grant & Grant 2002) and freshwater environments (cichlids – Seehausen 2004)
and can lead to the formation of new species (Seehausen 2004). Long-term monitoring of the reef fish suture zone at Christmas Island (Hobbs et al. 2009; Arnold & Martin 2010), through regular assessment of hybrid prevalence and genotypic make up across a wide range of taxa, could further elucidate the ecological and evolutionary relevance of hybridization in reef fishes.

The scenario emerging from the *C. trifasciatus* hybrid group (Montanari et al. 2012) appears different to that of the *C. guttatissimus* group. Lack of introgression, evident both in mtDNA and in microsatellites, and unidirectional mtDNA inheritance in the *C. trifasciatus* group indicate that interspecific gene flow mediated by hybrids is minimal at Christmas Island. Even though failure to detect significant levels of introgression in this group could be due to sample size, the sample sizes in the two groups were similar, leading us to expect similar power of detection in both hybridizing groups. Interestingly, in the *C. trifasciatus* group, a Zanzibar individual identified in the field as *C. trifasciatus* showed almost 2% divergence at cyt b from its putative species clade (Montanari et al. 2012): this could be a rare backcross with a hybrid formed between *C. trifasciatus* and other members of Corallochaetodon that occur in that area (e.g. *C. melapterus* – B. Bowen, pers. comm.). This needs further work to be confirmed, but if found to be true could indicate that barriers to gene flow are permeable in Corallochaetodon, despite the apparent lack of backcrossing at Christmas Island (Montanari et al. 2012).

The rarity of both parent species and hybrids in the *C. trifasciatus* group may prevent detection of introgression and bidirectional maternal contribution at Christmas Island (Montanari et al. 2012). In this group, the measured 5% divergence at cyt b (Montanari et al. 2012) appears to be large enough to generate genotypic novelty in the form of a persistent sympatric hybrid taxon, albeit small enough to warrant successful hybridization (Mallet 2005). Nuclear microsatellite DNA data were particularly informative for this group, confirming the hybrids’ status as hybrids rather than aberrant colorations of *C. lunulatus*, a possibility not ruled out by previous mtDNA analyses (Montanari et al. 2012). Microsatellites further showed that hybrid genotypes are intermediate and different to those of the parent species, even within the hybrid zone, thus maintaining their genotypic identity despite extensive ecological, behavioural and reproductive contact with parental species. Hybrid genotypes or hybrid species sometimes colonize environments distinct to those of their parents, as observed for example in cichlids and sculpins (Seehausen 2004; Nolte et al. 2006). However, sympatric hybrid coexistence with parental forms does occur (sparrows – Hermansen et al. 2011; swallowtail butterflies – Kunte et al. 2011), and this could be the case for *C. trifasciatus* × *C. lunulatus* hybrids at Christmas Island.

The apparent negative interaction between extent of divergence and introgression highlighted in this study finds further validation when data from other hybridizing reef fishes are examined. As noted in Montanari et al. (2012), for example, in the Solomon Islands, hybridization between *C. punctatofasciatus* and *C. pleuro-ensis* (McMillan et al. 1999), divergent by 0.7% at cyt b (McMillan & Palumbi 1995), results in extensive bidirectional introgression (McMillan et al. 1999). This interaction holds true even in families other than the Chaetodontidae. In the Labridae, bidirectional introgression was detected in hybridizing *Thalassoma jansenii* and *T. quinquevittatum* (Yaakub et al. 2006), divergent by <2% at cyt b (Bernardi et al. 2004). Conversely, in *Halichoeres garnoti* and *H. bifasciatus*, divergent by >5.5% based on three mtDNA markers (Barber & Bellwood 2005), hybridization did not result in introgression (Yaakub et al. 2007). In the Acanthuridae, hybridization between *A. leucosternon* and *Acanthurus nigricans*, 1% divergent at mtDNA COI, was introgressive and bidirectional (Marie et al. 2007). In hybridizing Serranids *Plectropomus leopardus* and *Plectropomus maculatus*, 1% divergent based on two nuclear and two mtDNA markers (Craig & Hastings 2007), hybridization was highly introgressive, but the maternal contribution was unidirectional (van Herwerden et al. 2006).

Further enquiry should be aimed at evaluating the relative importance of divergence levels in shaping the evolutionary outcomes of reef fish hybridization, and to test whether reef fish have a threshold of divergence beyond which their ability to hybridize is lost, as suggested for terrestrial species (Mallet 2005). Given their position at the Indo-Pacific marine suture zone, Christmas and Cocos (Keeling) Islands could provide an ideal location for these future studies. Further, application of genomic tools may identify adaptive genes that are differentiated between hybridizing reef fish species, which will provide insights into adaptation and selection for hybrid genotypes in environments that are novel compared with those inhabited by the parental species outside the hybrid zone.

**Acknowledgements**

This study was financially supported by AIMS@JCU (SRM) and the ARC CoE for Coral Reef Studies (MSP). We thank Giamcomo Bernardi, Shane Blowes, Sean Connolly and five anonymous reviewers for comments that improved this manuscript; Jeremiah Plass-Johnson and Ulrich Frye for providing Zanzibar collections; all staff at MEEL Townsville, AGRF Brisbane, GGF Athens and MACROGEN Seoul for assistance with laboratory work, sequencing and genotyping; Parks Australia and Wet N’ Dry Adventures for logistic support at Christmas Island.
References


Hettler WF (1968) Artificial fertilization among yellowfin and gulf menhaden (Brevoortia) and their hybrid. Transactions of the American Fisheries Society, 97, 119–123.


Hsu KC, Chen JP, Shao KT (2007) Molecular phylogeny of Chaetodon (Teleostei: Chaetodontidae) in the Indo-West

© 2014 John Wiley & Sons Ltd


Norman JR (1934) A Systematic Monograph of the Flatfishes (Heterostomate). Natural History Section of the British Museum, London, UK.


Supporting information

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

Fig. S1 Pairing behaviour (A), mtDNA haplotype network (B), scatterplot of DAPC (C) and STRUCTURE admixture plot (D) for the Ch. trifasciatus group (redrawn from Montanari et al. 2012).

Fig. S2 Δk plots (Evanno et al. 2005) from STRUCTURE HARVESTER (Earl & vonHoldt 2012) showing a sharp decline in the rate of change of the log probability of data for values of k > 2.

Fig. S3 Posterior probabilities, based on microsatellite data from Montanari et al. (2012), of individuals of the Ch. trifasciatus group belonging to six classes: pure parental species, F1 or F2 hybrids and backcrosses (Bc) in either direction.

Data S1 Material and Methods.

Table S1 Summary statistics for 20 microsatellite loci (Montanari et al. 2013) used to genotype the Ch. guttatissimus group: for each adequately sampled population we report sample size (n), number of alleles (Na), private alleles (Pa), average inbreeding coefficient (Fis), observed (He) and expected (Ht) heterozygosities, and probability of departure from HWE (p).

Table S2 Sample sizes, cyt b number of haplotypes (nh), haplotype (h) and nucleotide (n) diversities and intra-population gene diversity based on 20 microsatellite loci (1-Q Intert).

Table S3 Pairwise population comparisons in the Ch. guttatissimus group.

Table S4 Raw and ENA-corrected (Chapuis & Estoup 2007) population differentiation and estimator of actual differentiation (Jost 2008) (Dint) presented locus-by-locus and as a mean over all loci.