Seascape features, rather than dispersal traits, predict spatial genetic patterns in co-distributed reef fishes

Libby Liggins1,2,3*, Eric A. Treml4, Hugh P. Possingham3,5 and Cynthia Riginos3

ABSTRACT

Aim To determine which seascape features have shaped the spatial genetic patterns of coral reef fishes, and to identify common patterns among species related to dispersal traits [egg type and pelagic larval duration (PLD)].

Location Indian and Pacific Oceans, including the Indo-Australian Archipelago.

Methods We sampled coral reef fishes with differing dispersal traits (Pomacentrus coelestis, Dascyllus trimaculatus, Hailchoeres hortulanus and Acanthurus triostegus) and characterized spatial (mtDNA) genetic patterns using AMOVA-clustering and measures of genetic differentiation. Similarity in the spatial genetic patterns among species was assessed using the congruence among distance matrices method and the seascape features associated with the genetic differentiation of each species were identified using multiple regression of distance matrices (MRDM) and stepwise model selection.

Results Similar spatial genetic patterns were found for P. coelestis and H. hortulanus, despite their differing egg type (benthic versus pelagic). MRDM indicated that geographical distance was underlying their correlated genetic patterns. Species with pelagic eggs (A. triostegus and H. hortulanus) also had correlated patterns of genetic differentiation ($D_{est}$); however, a common underlying seascape feature could not be inferred. Additionally, the common influence of the Torres Strait and the Lydekker/Weber’s line was identified for the genetic patterns of differentiation for P. coelestis and A. triostegus, despite their differing dispersal traits, and the uncorrelated spatial genetic patterns of these species.

Main conclusions Our study demonstrates the value of a quantitative, hypothesis-testing framework in comparative phylogeography. We found that dispersal traits (egg type and PLD) did not predict which species had similar spatial genetic patterns or which seascape features were associated with these patterns. Furthermore, even in the absence of visually similar, or correlated spatial genetic patterns, our approach enabled us to identify seascape features that had a common influence on the spatial genetic patterns of co-distributed species.

Keywords biophysical model, comparative phylogeography, coral reef fish, early life history traits, Indo-Australasian Archipelago, Indo-Pacific Ocean, larval dispersal, pelagic larval duration, seascape genetics, statistical phylogeography

INTRODUCTION

Whether the spatial genetic patterns of co-distributed species are influenced by physical processes in similar, or predictable ways, is an area of great interest (Avise, 2000). In some cases, multiple species have concordant patterns of genetic structure that also align with certain physical features (phylogeographical concordance sensu Avise & Ball, 1990; Avise, 1992).
Conversely, when there is no evident genetic structure or overwhelming concordance among species, inferring what processes are responsible, and which processes have had a common influence on genetic patterns across species, is challenging. However, the influence of physical processes on genetic patterns is also mediated by the biological traits of species. Consequently, we would expect genetic patterns to vary predictably among species based on the interaction between their biological traits and the physical processes of interest.

For benthic marine organisms, meta-analytical and simulation approaches have highlighted the role of the early life-history dispersal traits (hereafter dispersal traits) in forming patterns of genetic structure and geographical scales of genetic differentiation [isolation-by-distance (IBD); Slatkin, 1993]. Specifically, fish species with benthic eggs tend to have greater genetic structure across their range than species with pelagic eggs (Riginos et al., 2011, 2014). Furthermore, the duration of a species’ pelagic larval stage (PLD) has been observed to have a weak positive relationship with the slope of their IBD relationship (Siegel et al., 2003; Selkoe & Toonen, 2011), but PLD has shown no relationship with genetic structure (excluding direct developers: Weersing & Toonen, 2009; Riginos et al., 2011). Nonetheless, simulations that incorporate dispersal traits clearly support an inverse relationship between PLD and genetic structure (Faurby & Barber, 2012) and PLD and evolutionarily significant levels of migration that would determine genetic structure (Treml et al., 2012).

Whereas meta-analytical studies are often spatially implicit and simulation studies have a contemporary focus, empirical genetic studies can quantify spatially explicit effects of historical and contemporary seascape features on genetic patterns. Such studies have demonstrated that co-distributed marine species sometimes have spatial genetic structure coincident with the same oceanographic features despite differences in dispersal traits (e.g. inter-island channels, Toonen et al., 2011). Other studies advocate a role of dispersal traits in determining whether historical processes have impacted spatial genetic patterns (e.g. Sherman et al., 2008), and suggest genetic patterns of high dispersal species are more associated with contemporary features than historic features (Pelc et al., 2009). These studies suggest that the simultaneous consideration of species’ dispersal traits, as well as historical and contemporary seascape features, are necessary for understanding the processes underlying the spatial genetic patterns of marine organisms.

In this study, we focus on spatial genetic patterns across the tropical Indian and Pacific Oceans, and in particular the juncture of the oceans around the Indo-Australasian Archipelago (IAA). The IAA is a tropical marine biodiversity hotspot (Roberts et al., 2002) and many of the hypotheses explaining this phenomenon also predict high phylogeographical structure (Bowen et al., 2013). Several suture zones delineate species boundaries according to historical continental affiliations (including Lydekker’s and Weber’s lines, Fig. 1; Lydekker, 1896; Weber, 1902), but these suture zones also affect intraspecific lineages (suggested by Avise, 1992; and found by DeBoer et al., 2014). Most population genetic structuring in this region, however, has been attributed to the climate oscillations of the Pleistocene (c. 2.5 Ma–12 ka, reviewed in Carpenter et al., 2011), when pockets of shallow marine habitat became isolated and connectivity among the oceans was restricted to a narrow passage of water between the Sahul and Sunda shelves (Fig. 1). Since c. 7 ka, there has also been a point of contact between the Indian and Pacific oceans via the Torres Strait (Voris, 2000; Reeves et al., 2008). Biophysical models of dispersal suggest high population connectivity

Figure 1 Map of the study area including sampled locations and geographical features (cylindrical equal-area projection). Dark gray represents present-day landmasses; light gray represents greatest landmass extent during the Last Glacial Maximum including the Sunda Shelf and the Sahul Shelf. Points represent sampled locations; coloured points denote locations for which all four study species were co-sampled. Full description of sampled locations can be found in Appendix S1 in the Supporting Information.
between the Pacific and Indian Oceans via the Torres Strait, the Indonesian Through Flow that runs between the continental shelves (Fig. 1) and several smaller currents within the IAA (Kool et al., 2011; Treml et al., 2012). Thus, contemporary dispersal may reconnect populations that were isolated during other periods in history.

Comparing Indian and Pacific Ocean populations, some taxa show strong genetic differentiation across the IAA (e.g. Lavery et al., 1996; Gaither et al., 2010), whereas others do not (e.g. Lessios et al., 2003; Gaither et al., 2010). Moreover, the position of genetic breaks within the IAA often varies across species (reviewed in Carpenter et al., 2011). However, these empirical results are difficult to compare quantitatively. Each study has its own sampling design, generally focuses on only one static spatial predictor of the genetic patterns (such as the entire IAA, e.g. Gaither et al., 2010; or the Torres Strait, e.g. Mirams et al., 2011), or describes only distinct genetic breaks (e.g. Barber et al., 2006). Yet, the impacts of historical and contemporary processes may manifest as patterns of genetic differentiation, rather than discrete genetic breaks, particularly in high dispersal marine species.

Here, we assess congruence in the spatial genetic patterns of four reef fishes and address which seascape features have had a common influence on their patterns. We expect that egg type will correlate with geographical structure established over time (Riginos et al., 2011, 2014) and that contemporary patterns of gene flow among populations will vary in predictable ways across species (related to PLD, Faurby & Barber, 2012; Treml et al., 2012). These expectations underlie four null hypotheses based on the egg type and PLD of species (hereafter conferring low dispersal or high dispersal, Fig. 2). We predict: greater genetic structure in low dispersal species than in high dispersal species (H01); species with similar dispersal traits will have similar spatial genetic patterns (H02); similar spatial genetic patterns will reflect the common influence of seascape features (H03); and for low dispersal species, we anticipate patterns of genetic differentiation will reflect the impact of historical seascape features predominantly, whereas the impact of contemporaneous seascape features will be more evident in the patterns of genetic differentiation for high dispersal species (H04).

MATERIALS AND METHODS

Study design

We examined four reef fishes [Pomacentrus coelestis (Jordan & Starkes, 1901), Dascyllus trimaculatus (Rüppell, 1829), Halichoeres hortulanus (Lacépède, 1801) and Acanthurus triostegus (Linnaeus, 1758)] that vary in egg type and PLD (Fig. 2). Fishes were collected from 19 locations (Fig. 1, see Appendix S1 in the Supporting Information). Species were only collected if they were present in high abundance; at seven locations all four species were collected (coloured points, Fig. 1). We attempted to avoid two caveats common to comparative phylogeographical studies: we focus our comparative analyses on locations where all species have been co-sampled (hereafter co-sampled range) to ensure our sampling design is balanced (Dawson, 2014); and for each species we examined genetic patterns over the widest possible range (hereafter broad-range) thus providing geographical context for the focal genetic patterns (Rocha et al., 2007; Bowen et al., 2014).

Laboratory methods

For all species, we targeted the mtDNA control region (CR) to infer genetic patterns using DNA extraction and amplification protocols detailed in Mirams et al. (2011). A suitable length of CR proved difficult to amplify for A. triostegus, so we used ATPase subunit 6 and 8 (ATP6-8) amplified with ATP6.2 and CO3.2 (Lessios & Robertson, 2006). Amplicons were purified, sequenced, and sequences were manually checked and aligned as per Liggins et al. (2015). Datasets were augmented with sequences previously published by the authors available on GenBank (see Data Accessibility).

Genetic patterns

To observe the genealogical relationships among locations, we constructed minimum spanning networks in PopART
Patterns of genetic structure were estimated using two analysis of molecular variance (AMOVA, Excoffier et al., 1992) – clustering methods: based on haplotype identities (Genodive 2.0B23, Meirmans & van Tienderen, 2004); and considering pairwise distance among haplotypes and the geographical location of populations (Samova 1.0, Duplanloup et al., 2002). These methods determined how many significant groups (k) were found across the species’ broad- and co-sampled ranges, so we could address whether there was greater genetic structure in low dispersal species than in high dispersal species (H_0). Using Genodive, 2 to N-2 groups were selected using the Calinski & Harabasz (1974) pseudo-F-statistic with 100,000 steps and 20 repeats. Using Samova, analyses were run using 2 to N-2 groups, in turn (10,000 simulated annealing processes each), and 10,000 permutations of populations among groups were conducted to infer significance.

Pairwise relationships among populations were described using Φ_ST (based on the Tamura–Nei distance as selected by the Bayesian information criterion in jModeltest 2.1.4, Darriba et al., 2012; 10,000 permutations), F_ST (based on haplotype identities; using ARL_EQUN 3.5, Excoffier & Lischer, 2010) and D_out (Jost, 2008; using Genodive).

### Congruence among species in genetic patterns

The similarity in the patterns of genetic structure and differentiation among species was tested using the congruence among distance matrices (CADM, Legendre & Lapointe, 2004) method in ape 3.0-8 (Paradis et al., 2004; executed in R 2.15.3, R Core Team). The overall coefficient of concordance reported is Kendall’s W, a test statistic that varies from 0, indicating no congruence to 1, indicating high congruence (Kendall & Smith, 1939; using the function CADM_global). Pairwise similarity among matrices within the comparison is then assessed a posteriori (r_mst, using the CADM_post function). The analysis was repeated for distance matrices based on each of the genetic measures (k-clusters, F_ST, Φ_ST and D_out) for the co-sampled range of all the species. Distance matrices for the k-clusters were binary matrices based on cluster membership, where populations of the same cluster were assigned an inter-population distance of 0, and populations of different clusters were assigned a distance of 1. The Holm (1979) method was used to correct P-values following multiple testing (recommended by Legendre & Lapointe, 2004; 10,000 permutations). We had no prior expectation for the overall level of congruence in genetic patterns among species; only that species pairs that had the highest correlation also had similar dispersal traits (H_0).

### Seascape predictors of genetic differentiation

We used multiple regression of distance matrices (MRDM; Legendre et al., 1994) to examine the association of seascape features with patterns of genetic differentiation across the sampled ranges of the fish species. The MRDM method has been identified as one of the best for predicting genetic differentiation (Balkenhol et al., 2009); it appropriately facilitates the simultaneous analysis of distances and barriers as predictors (Paquette & Lapointe, 2009) and has been demonstrated to be effective in removing the influence of historical covariates so that contemporary influences may also be deduced (Dyer et al., 2010).

Predictor matrices in the MRDM included: (1) phylogeographical structure, (2) Lydekker/Weber’s line, (3) Torres Strait, (4) oceanographic distances and (5) larval dispersal distances. Phylogeographical structure (1) was only included in the model if haplotype networks revealed a distinct lineage that was restricted to geographically proximal locations. The rationale for incorporating phylogeographical structure was not to test its effect, but to remove its effect, so that predictor matrices relevant to the residual variance could be identified.

Phylogeographical structure and the geographical boundaries [Lydekker/Weber’s line (2) and the Torres Strait (3), Fig. 1] were represented as binary matrices in the analysis (as described above). The last two predictor matrices were weighted distances. First, contemporary least-cost oceanographic distances (4) were estimated using MARMAP 0.5 in R (Pante & Simon-Bouhet, 2013) by restricting the measurement of distance between locations to the habitable ocean less than 35° latitude. Second, we used a biophysical model of larval dispersal (5) to quantify the relative dispersal strength among the co-sampled locations for each species independently. These models include coral reef habitat (Spalding et al., 2001), oceanographic data describing sea surface currents for 3 years (HYCOM, Chassignet et al., 2009) and several biological parameters describing the known or expected dispersal characteristics of each species (Fig. 2; for details, see Treml et al., 2012 and Appendix S2).

MRDM models were first constructed across the broad-range for each species. This was done separately for each of the pairwise genetic differentiation measures (F_ST, Φ_ST, D_out). Second, models were constructed for the co-sampled range for each genetic differentiation measure, so that models could be compared across species. Analyses proceeded by fitting a model via the stepwise approach (Legendre et al., 1994). Terms were only added to the model if their significance did not exceed a Bonferroni-corrected P-value. We anticipated that for species identified as having congruent spatial genetic patterns (in the previous CADM analysis), their final MRDM models would attribute their patterns of genetic differentiation to the same seascape features (H_3). We hypothesized that the effect of historic seascape features [Lydekker/Weber’s line and the Torres Strait: (2, 3)] would be most evident in the patterns of genetic differentiation for the low dispersal species (P. coelestis and D. trimaculatus); whereas, we expected that the effect of more contemporary seascape features [oceanographic and dispersal distances: (4, 5)] would be most evident in the patterns of genetic differentiation of the higher dispersal species (H. hortulanus and A. triostegus; H_4).
RESULTS

Genetic patterns

The final datasets included sequences from: 237 individuals of *Pomacentrus coelestis* (336 bp, 13 locations); 165 *D. trimaculatus* (375 bp, 12 locations); 147 *H. hortulanus* (359 bp, 10 locations); and 190 *A. triostegus* (830 bp, 10 locations; see Appendix S1). All species had high levels of genetic diversity that was shared across many locations (Fig. 3). Phylogeographical structure was evident in *A. triostegus*, corresponding to the suggested *A. triostegus sandvicensis* subspecies of Hawaii (HAW; Randall, 1956). *Dascyllus trimaculatus* also had geographical and genetically distinct lineages in both the Chagos archipelago (CHA) and the Cook Islands (COK). The distinct lineage of the Cook Islands could be due to hybridization with *D. auripinnis*, where the two species are suggested to not be *bona fide* (Bernardi et al., 2003). Matrices representing the phylogeographical structure (1) of *A. triostegus* and *D. trimaculatus* were included in subsequent MRDM analyses across their broad-range (Table 1a).

The genetic variance attributed to clusters and the number of clusters found in each species across the co-sampled locations did not consistently decrease according to increasing dispersal potential, but varied depending on the clustering method (Fig. 4). Nonetheless, *P. coelestis* did have the greatest variance described by the clustering (*Genodive* $F_{CT}$ = 0.040, *Samova* $F_{CT}$ = 0.232) and *D. trimaculatus* had the highest number of significant clusters ($k = 5$).

Congruence among species in genetic patterns

Analyses of CADM based on spatial clusters for all four species had low levels of concordance (*Genodive* $W = 0.24$, $\chi^2 = 19.48$, adjusted $P_{perm} = 0.5963$; *Samova* $W = 0.36$, $\chi^2 = 28.48$, adjusted $P_{perm} = 0.10$). However, there was complete concordance in the spatial clusters of *P. coelestis* and *H. hortulanus* (*Samova* $r_m = 1.0$, adjusted $P_{perm} = 0.0438$). Based on measures of genetic differentiation ($F_{ST}$, $\Phi_{ST}$ and $D_{est}$), the coefficient of concordance across the co-sampled range of all species was also low and insignificant ($F_{ST}$: $W = 0.15$, $\chi^2 = 11.86$, adjusted $P_{perm} = 0.7830$; $\Phi_{ST}$: $W = 0.22$, $\chi^2 = 17.82$, adjusted $P_{perm} = 0.6070$; $D_{est}$: $W = 0.23$, $\chi^2 = 18.59$, adjusted $P_{perm} = 0.5488$). Again, the highest correlation was found for the spatial patterns of *P. coelestis* and *H. hortulanus* ($\Phi_{ST}$: $r_m = 0.85$, adjusted $P_{perm} = 0.0018$; $F_{ST}$: $r_m = 0.85$, adjusted $P_{perm} = 0.0018$).

![Minimum spanning haplotype networks for the study species.](image-url)
Table 1 (a) Summary of the final models for the multiple regression of distance matrices analyses for each measure of genetic differentiation ($F_{ST}$, $\Phi_{ST}$, $D_{est}$) for each species. The top panel presents the models for the seven locations that were co-sampled for all species; the bottom panel is for all sampled locations within each species (referred to as the broad-range in text). Models were chosen via the stepwise procedure, with a significance threshold of a Bonferroni-corrected $P$-value. Beta coefficients are presented on a standardized scale. Predictors not included in the final models are denoted by a dash (i.e. –). NA represents where a predictor is not applicable to the dataset and was excluded. (b) Summary of each measure of genetic differentiation ($F_{ST}$, $\Phi_{ST}$, $D_{est}$) regressed against the biophysically derived dispersal distance the co-sampled populations for each species.

### Co-sampled locations

<table>
<thead>
<tr>
<th>Factors</th>
<th>Pomacentrus coelestis</th>
<th>Dascyllus trimaculatus</th>
<th>Halichoeres hortulanus</th>
<th>Acanthurus triostegus</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>$F_{ST}$</td>
<td>$\Phi_{ST}$</td>
<td>$D_{est}$</td>
<td>$F_{ST}$</td>
</tr>
<tr>
<td>(2) Lydekker/Weber’s line</td>
<td>–</td>
<td>0.270*</td>
<td>–</td>
<td>–</td>
</tr>
<tr>
<td>(3) Torres Strait</td>
<td>–</td>
<td>0.753***</td>
<td>–</td>
<td>–</td>
</tr>
<tr>
<td>(4) Oceanographic distance</td>
<td>–</td>
<td>0.946***</td>
<td>–</td>
<td>–</td>
</tr>
<tr>
<td>(5) Dispersal distance</td>
<td>–</td>
<td>0.426**</td>
<td>0.411**</td>
<td>–</td>
</tr>
<tr>
<td>$F_{ST}$</td>
<td>160.40</td>
<td>42.19</td>
<td>–</td>
<td>–</td>
</tr>
<tr>
<td>$\Phi_{ST}$</td>
<td>0.894</td>
<td>0.824</td>
<td>–</td>
<td>–</td>
</tr>
<tr>
<td>$D_{est}$</td>
<td>0.946***</td>
<td>0.800***</td>
<td>–</td>
<td>–</td>
</tr>
<tr>
<td>$r^2/R^2$</td>
<td>–</td>
<td>0.011</td>
<td>0.040</td>
<td>–</td>
</tr>
</tbody>
</table>

### All sampled locations

<table>
<thead>
<tr>
<th>Factors</th>
<th>Pomacentrus coelestis</th>
<th>Dascyllus trimaculatus</th>
<th>Halichoeres hortulanus</th>
<th>Acanthurus triostegus</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>$F_{ST}$</td>
<td>$\Phi_{ST}$</td>
<td>$D_{est}$</td>
<td>$F_{ST}$</td>
</tr>
<tr>
<td>(1) Phylogeographical</td>
<td>NA</td>
<td>NA</td>
<td>NA</td>
<td>NA</td>
</tr>
<tr>
<td>structure</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>(2) Lydekker/Weber’s line</td>
<td>–</td>
<td>0.187**</td>
<td>–</td>
<td>–</td>
</tr>
<tr>
<td>(3) Torres Strait</td>
<td>–</td>
<td>0.762***</td>
<td>0.778***</td>
<td>–</td>
</tr>
<tr>
<td>(4) Oceanographic distance</td>
<td>–</td>
<td>0.157***</td>
<td>–</td>
<td>–</td>
</tr>
<tr>
<td>$F_{ST}$</td>
<td>105.20</td>
<td>92.61</td>
<td>45.56</td>
<td>407.00</td>
</tr>
<tr>
<td>$\Phi_{ST}$</td>
<td>0.581</td>
<td>0.712</td>
<td>0.416</td>
<td>0.928</td>
</tr>
<tr>
<td>$D_{est}$</td>
<td>0.011</td>
<td>0.040</td>
<td>0.020</td>
<td>0.008</td>
</tr>
<tr>
<td>$r^2/R^2$</td>
<td>0.02</td>
<td>0.01</td>
<td>0.000</td>
<td>0.00</td>
</tr>
</tbody>
</table>

* $P < 0.05$, ** $P < 0.01$, *** $P < 0.001$. 
Many of the final MRDM models retained only one predictor, and in some cases none of the predictors were found to be associated with the patterns of genetic differentiation (Table 1). All correlations among predictor matrices were of an acceptable level ($r < 0.72$). The variance inflation factors (VIF) were inspected to ensure that the predictor variables did not exhibit multicollinearity that would affect model selection (max. VIF observed was 1.78, below the recommended guideline of 5; Rogerson, 2001). Significant final models are presented below, by species.

For A. triostegus, the effect of the Torres Strait (3) on patterns of genetic differentiation defined by $F_{ST}$ was significant across both the broad- and co-sampled range, and for $D_{est}$ across the co-sampled range (Table 1a). Disjunction associated with the Lydekker/Weber’s line (2) also had a significant effect on $\Phi_{ST}$ and $D_{est}$ over both the broad- and co-sampled ranges.

For P. coelestis, the effects of the Lydekker/Weber’s line (2) and the Torres Strait (3) were apparent across both geographical scales when patterns of genetic differentiation were characterized using $D_{est}$. When focusing on $\Phi_{ST}$, only the

---

**Figure 4** Map of sampled locations for each study species and significant genetic clusters ($k$) identified using Genodive (a) and Samova (b). Significant clusters across all the sampled locations for each species are represented by the grey lines, and significant clusters based only on the co-sampled locations are represented by the black dashed lines. Significance of the clustering is denoted by asterisks: *$P < 0.05$, **$P < 0.01$, ***$P < 0.001$. 

---

$r_M = 0.70$, adjusted $P_{perm} = 0.0562$). However, for measures based on $D_{est}$, the spatial patterns in H. hortulanus were significantly correlated with those of A. triostegus ($r_M = 0.54$, adjusted $P_{perm} = 0.0398$).

**Seascape predictors of genetic differentiation**

Many of the final MRDM models retained only one predictor, and in some cases none of the predictors were found to be associated with the patterns of genetic differentiation (Table 1). All correlations among predictor matrices were of an acceptable level ($r < 0.72$). The variance inflation factors (VIF) were inspected to ensure that the predictor variables did not exhibit multicollinearity that would affect model selection (max. VIF observed was 1.78, below the recommended guideline of 5; Rogerson, 2001). Significant final models are presented below, by species.

For A. triostegus, the effect of the Torres Strait (3) on patterns of genetic differentiation defined by $F_{ST}$ was significant across both the broad- and co-sampled range, and for $D_{est}$ across the co-sampled range (Table 1a). Disjunction associated with the Lydekker/Weber’s line (2) also had a significant effect on $\Phi_{ST}$ and $D_{est}$ over both the broad- and co-sampled ranges.

For P. coelestis, the effects of the Lydekker/Weber’s line (2) and the Torres Strait (3) were apparent across both geographical scales when patterns of genetic differentiation were characterized using $D_{est}$. When focusing on $\Phi_{ST}$, only the
Torres Strait (3) was significant over the broad-range, and only dispersal distance (5) over the co-sampled range. For _D. trimaculatus_, none of the predictor matrices were significantly associated with the patterns of genetic differentiation across the co-sampled range. Across the broad-range, oceanographic distance (4) was found to be a predictor of differentiation based on $\Phi_{ST}$.

For _H. hortulanus_, oceanographic distance (4) was the best predictor in final models across the co-sampled ($\Phi_{ST}$ only) and broad-range ($F_{ST}$, $\Phi_{ST}$, and $D_{ST}$). Over the broad-range of _H. hortulanus_ the Torres Strait (3) was also associated with the patterns of genetic differentiation defined by $\Phi_{ST}$.

**DISCUSSION**

The high dispersal ability and large population sizes of many marine organisms often contribute to weak spatial genetic structure (Waples, 1998). Consequently, inferring common spatial genetic patterns and underlying processes can be challenging. Few studies have attempted to quantitatively examine the underlying causes of genetic differentiation for co-distributed marine species (but see Selkoe et al., 2010; Dawson et al., 2014). Such an approach requires a common sampling design (Dawson, 2014) and robust statistical analysis. Our study focused on four reef fishes sampled across the IAA. We aimed to infer whether species had shared spatial genetic patterns and which seascape features were associated with these patterns. We yielded two significant results that were contrary to our hypotheses based on conventional wisdom. First, we found that species with shared spatial genetic patterns across the seascape did not necessarily have similar dispersal traits (e.g. _P. coelestis_ and _H. hortulanus_ based on egg type and PLD). Second, we demonstrated that the common influence of seascape features can be inferred, even when species appear to have different patterns of spatial genetic structure (i.e. the association of Lydekker/Weber's line, Torres Strait; _H. hortulanus_ and _A. triostegus_ despite their disparate patterns of genetic clustering; Table 1a, Fig. 4).

**Unpredictable and congruent genetic patterns**

We found mixed support for the hypothesis that species of low dispersal potential have greater population genetic structure (H_0). The species likely to have the lowest dispersal potential (based on egg type and PLD), _P. coelestis_, had the greatest amount of genetic variance attributed to clusters (Genodive $F_{CT} = 0.04$, Samova $F_{CT} = 0.23$), and the second lowest disperser, _D. trimaculatus_, had the greatest number of significant clusters ($k = 5$) across the co-sampled range. However, simultaneously considering all species, there was no consistent association between the dispersal traits and the amount of genetic variance explained, or number of significant clusters found. Furthermore, the spatial genetic patterns of clustering differed among all species (using the CADM, based on clusters inferred using Genodive $W = 0.24$, adjusted $P_{perm} > 0.05$; Samova $W = 0.36$, adjusted $P_{perm} > 0.05$). The only species pair found to have similar spatial structure was _H. hortulanus_ and _P. coelestis_, that differ substantially in dispersal traits (Samova $r_M = 1.0$, adjusted $P_{perm} < 0.05$).

Shifting our focus from patterns of genetic structure to genetic differentiation (based on $F_{ST}$, $\Phi_{ST}$ and $D_{ST}$), we found instances of concordance across species both with, and without, similar dispersal traits (_H. hortulanus_). Patterns of spatial differentiation were significantly correlated (adjusted $P_{perm} < 0.05$; based on CADM) for _H. hortulanus_ and _A. triostegus_, our two high dispersal species (for $D_{ST}$), but also _H. hortulanus_ and _P. coelestis_ (for $\Phi_{ST}$), that have different dispersal traits. For both _H. hortulanus_ and _P. coelestis_, patterns in $\Phi_{ST}$ were highly associated with weighted distances (oceanographic and dispersal distance, respectively, according to MRDM analyses; Table 1) providing support for our hypothesis that similar patterns of genetic structure and differentiation reflect the common influence of certain seascape features ($H_0$). In contrast, for _A. triostegus_ and _H. hortulanus_, we were unable to identify any seascape features responsible for the correlation in their patterns of genetic differentiation based on $D_{ST}$.

Patterns of genetic differentiation for our high dispersal species were not consistently associated with contemporary seascape features (i.e. oceanographic and dispersal distance); nor were the patterns of genetic differentiation of our low dispersal species consistently associated with historical features (i.e. the Lydekker/Weber’s line, Torres Strait; _H_4). Although the best fit MRDM models for _H. hortulanus_, our second highest disperser, consistently contained oceanographic distance, models for _A. triostegus_, our species of greatest dispersal potential, were most affected by the Torres Strait and the Lydekker/Weber’s line (Table 1). These same historic seascape features were predicted to be most important in driving patterns of genetic differentiation in _P. coelestis_ (based on $D_{ST}$), one of our lowest dispersers. Thus, common seascape features were found to be associated with the genetic patterns of _A. triostegus_ and _P. coelestis_, despite their different dispersal potential and their apparently disparate spatial genetic patterns (based on spatial genetic structure, Fig. 4, and CADM).

The use of multiple independent loci per species, rather than a single locus, would be preferred to understand the historical and contemporary connectivity among populations for each of our study species. However, in comparative phylogeography, there is often a trade-off in replication across species versus number of loci used within species (Avise, 2000). In these cases, the use of mtDNA has been advocated (Bowen et al., 2014). Despite our best efforts to use the same gene region across species, it is conceivable that the differing mutation rates of CR and ATP6-8 (used only in _A. triostegus_) have influenced our results. Nonetheless, both CR and ATP6-8 are on the mitochondrial locus and therefore have the same mode of inheritance.
The role of the ocean in driving genetic differentiation

The existence of boundaries to dispersal in a fluid environment is non-intuitive, yet our study joins the growing number of marine focused studies that demonstrate genetic differentiation associated with straits, inundated continental shelves and oceanographic distance (reviewed in Riginos & Liggins, 2013). Contrary to results from other recent studies that have included biophysical model predictions (reviewed in Liggins et al., 2013), species-specific larval dispersal distances were not the most significant predictor of mtDNA genetic differentiation patterns for our study species (Table 1b; also found in Crandall et al., 2014). It is likely that the tumultuous evolutionary history of the lineages across the IAA has pre-empted the influence of contemporary seascape features, and in particular the combined effect of present-day ocean currents and species dispersal ability.

In our study, the ability to directly compare species was possible by including co-sampled locations (Dawson, 2014); however, the relevance of these inferred patterns for each species was better understood within the context of the wider geographical range (as suggested by Rocha et al., 2007; Bowen et al., 2014). A change in the importance of seascape features over geographical scale may be indicative of different patterns of genetic differentiation across a species range (Slatkin, 1993). For P. coelestis, where the most significant predictor of genetic differentiation based on \( \Phi_{ST} \) changed from dispersal distance over the co-sampled range, to the Torres Strait over the broad-range, a lack of an IBD relationship in the southern range periphery of this species is probably the cause (Liggins et al., 2015). In contrast, whereas none of our spatial predictor matrices were associated with the patterns of genetic differentiation across the co-sampled range of D. trimaculatus, oceanographic distance was associated with the patterns defined by \( \Phi_{ST} \) over the broad-range.

In many cases, the final models changed according to which genetic differentiation measure was used, and in doing so, provided a nuanced view of the processes driving the genetic patterns. For example, the effect of oceanographic distance was not observed when patterns of genetic differentiation were based on \( D_{st} \) or \( F_{ST} \) (except for H. hortulanus, Table 1). In some cases, the turnover of haplotypes among populations may be too high for \( F_{ST} \) or \( D_{st} \) to be informative over vast oceanographic distances. In contrast, \( \Phi_{ST} \) can be informative, even when no haplotypes are shared among populations, and thus may be especially useful for wide-ranging species such as Indo-Pacific reef fishes. Patterns as described by \( \Phi_{ST} \) more closely resemble changes in nucleotide diversity and thus reflect the processes of mutation accumulation and drift. In instances of contemporary migration, however, \( F_{ST} \) can capture variance in haplotype frequencies without the ‘noise’ introduced by considering the genealogical relationships among haplotypes (Bird et al., 2011). Based on these expectations, the final models for A. triostegus suggest that differentiation associated with the Torres Strait is indicative of more contemporary migration limitation (indicated by \( F_{ST} \)) than differentiation across the Lyddekker/Weber’s line (evident in models based on \( \Phi_{ST} \)).

Other predictors of spatial genetic patterns in the sea

Ultimately testing the role of early life-history dispersal traits in determining which seascape features are important in shaping spatial genetic patterns requires the analysis of more than the four species studied here. With the addition of more species, suites of responses to oceanographic features or spatial patterns across species might emerge (as seen in Pelc et al., 2009; and Selkoe et al., 2014; respectively). However, a focus solely on egg type and PLD is also overly simplistic. The dispersal of pelagic larvae is known to be influenced by species traits such as their sense of smell, homing behaviour, swimming ability (Leis et al., 2007) and reproductive timing (Carson et al., 2010; Treml et al., 2015). Alternatively, spatial genetic patterns may be unrelated to life-history traits (Bird et al., 2007; Galarza et al., 2009).

Comparative spatial genetic patterns of benthic marine organisms have also been studied in relation to other species-specific traits such as habitat specificity, diet and ecology. Once again, whereas some studies find support for a role of distinct species-related traits (diet specialization: Lawton et al., 2011; species ecology or habitat specialization: Rocha et al., 2002; Ayre et al., 2009) others do not (species ecology, Crandall et al., 2008). Thus, it is likely that the primarily stochastic processes of colonization, population size fluctuation and extinction have a large role in determining spatial genetic patterns in benthic marine organisms, and cannot be easily predicted across species (Selkoe et al., 2014; but see Hart & Marko, 2010; Dawson et al., 2014 for analytical suggestions).

CONCLUSIONS

Our study aimed to identify seascape features that have a common influence across species, rather than focusing on concordant spatial genetic patterns per se. Certainly, the aim of many multi-species genetic studies is to identify spatial patterns, not the underlying processes; the spatial delineation of regions of high genetic turnover and high genetic diversity is important for the prioritization of regions for conservation purposes (Moritz & Faith, 1998; Rocha et al., 2007). However, an understanding of the seascape features that have led to, and maintain, such patterns is also important. Despite our study being focused on only four species, our analytical approach revealed some unexpected commonalities in the seascape features underlying spatial genetic patterns. Our study suggests that inference based on the observation of spatial genetic patterns alone may overlook the common role of some seascape features in shaping the spatial genetic patterns of co-distributed species.
ACKNOWLEDGEMENTS

We are grateful for the logistic, permitting and research support provided by several individuals, research institutions, governmental and non-governmental organizations. Full acknowledgment of these groups and permit information is provided in Appendix S3 in the Supporting Information. Funding for this work was provided by the Australian Research Council (DP0878306, to C.R. and H.P.P.) and an Explorer’s Club Exploration Fund (to L.L.). Many of the ideas discussed here grew out of work funded by the Sea World Research & Rescue Foundation (SWR/1/2012, to C.R. and L.L.), a Paddy Pallin Foundation, The Foundation for National Parks & Wildlife Science Grant, an Ecological Society of Australia Student Research Grant, the Lerner Gray Memorial Fund of the American Museum of Natural History and a Great Barrier Reef Marine Park Authority’s Science for Management Award (to L.L.). We are grateful for comments on an earlier version of this manuscript provided by H.B. Harrison and two anonymous referees.

REFERENCES


Statistical comparative phylogeography of coral reef fishes


**SUPPORTING INFORMATION**

Additional Supporting Information may be found in the online version of this article:

**Appendix S1** Sampled locations and genetic diversity analyses.

**Appendix S2** Biophysical models of ‘dispersal distance’.

**Appendix S3** Full acknowledgements.

**DATA ACCESSIBILITY**


**BIOSKETCH**

Libby Liggins is a lecturer in marine ecology at Massey University, Auckland, New Zealand. Her research uses genetic information to help understand population demography and biogeography, with a particular interest in edge-of-range populations and transitional communities.

Author contributions: All authors contributed to the idea conception; L.L., E.A.T. and C.R. collected the data; L.L. analysed the data; L.L. led the writing and all other authors contributed comments on the manuscript.

Editor: David Bellwood
Seascape features, rather than dispersal traits, predict spatial genetic patterns in co-distributed reef fishes

Liggins L

2015