Faunal breaks and species composition of Indo-Pacific corals: the role of plate tectonics, environment and habitat distribution

S. A. Keith1, A. H. Baird1, T. P. Hughes1, J. S. Madin3 and S. R. Connolly1,2

1Australian Research Council Centre of Excellence for Coral Reef Studies, and 2School of Marine and Tropical Biology, James Cook University, Townsville, Queensland 4811, Australia
3Department of Biological Sciences, Macquarie University, Australia, New South Wales 2109, Australia

Species richness gradients are ubiquitous in nature, but the mechanisms that generate and maintain these patterns at macroecological scales remain unresolved. We use a new approach that focuses on overlapping geographical ranges of species to reveal that Indo-Pacific corals are assembled within 11 distinct faunal provinces. Province limits are characterized by co-occurrence of multiple species range boundaries. Unexpectedly, these faunal breaks are poorly predicted by contemporary environmental conditions and the present-day distribution of habitat. Instead, faunal breaks show striking concordance with geological features (tectonic plates and mantle plume tracks). The depth range over which a species occurs, its larval development rate and genus age are important determinants of the likelihood that species will straddle faunal breaks. Our findings indicate that historical processes, habitat heterogeneity and species colonization ability account for more of the present-day biogeographical patterns of corals than explanations based on the contemporary distribution of reefs or environmental conditions.

1. Introduction

More than 100 hypotheses have been proposed to explain the systematic changes in species richness across latitude, longitude, altitude and depth, which generate gradients in biodiversity [1,2]. The need to understand the genesis and maintenance of such patterns is becoming ever more urgent as we seek to manage the effects of global environmental change on biodiversity [3]. However, despite the plethora of hypothesized explanations, the underlying mechanisms driving these patterns are elusive [4,5]. The species (and genus) richness of reef-building corals in the Indo-Pacific exhibits a peak in biodiversity in the Indo-Australian Archipelago (IAA), with a gradual decline in richness with increasing distance from the IAA hotspot, particularly in the highly depauperate eastern Pacific and at higher sub-tropical latitudes [6–11]. Multiple explanations have been proposed for this pattern, including hypotheses pertaining to ambient energy, habitat availability, dispersal limitation, geometric constraints and evolutionary time [7,8,12–14]. Some degree of empirical support has been offered for most of these hypotheses [8,14]. However, disentangling the relative importance of these different underlying processes remains a significant challenge.

The standard approach to evaluating competing hypotheses has been to investigate relationships between species richness and variables that serve as surrogates for the hypothesized mechanisms (e.g. temperature) using multiple regression and related techniques. However, this approach suffers from important limitations [1]. Foremost among these is that species richness is correlated with predictions of multiple hypotheses (i.e. multicollinearity). For example, in the Indo-Pacific, coral species richness is correlated with sea surface temperature (SST), habitat area, distance from the mid-domain [8] and with surrogate variables for historical
One approach that could improve the capacity to discriminate between competing hypotheses in biogeography is an expansion in focus from counts of species towards species composition more broadly (i.e. changes in species identities as well as species richness over geographical space) [1,7]. Marine biogeographical zones have been recognized in previous work as reflecting changes in species composition [17,18]. However, delineation of these zones has tended to be subjective, and few of these earlier studies have focused on the potential reasons for such large-scale changes across faunal boundaries. A more rigorous quantitative approach to the spatially explicit quantification of taxonomic composition has the potential to evaluate a richer set of predictions. Specifically, using species composition rather than species richness alone also allows investigation of how different traits of individual species influence where they reach their range boundaries and why range boundaries can co-occur to create provincial boundaries. For example, if dispersal limitation drives biogeographical patterns, then species traits that influence susceptibility to dispersal barriers should be good predictors of the locations of range limits [19]. Therefore, a focus on species composition and distribution through space can complement analyses focusing on species richness alone, and afford stronger tests of macroecological hypotheses [16].

An implicit assumption of many studies of diversity gradients is that species distributions are in equilibrium with contemporary abiotic conditions and the current geographical distribution of habitat. However, transient non-equilibrium dynamics can result in mismatches between present-day abiotic conditions and species distributions (e.g. lags in ongoing species range shifts in response to anthropogenic climate change [19,20]). Consequently, explanations for species distributions may require consideration of historical processes as well as contemporary environmental conditions. Within the Indo-Pacific, the influence of historical processes has received some attention in previous work: the region of highest species richness has shifted westwards over time, from the Mediterranean approximately 40 Ma to its present location in the IAA [21]. This suggests that the influence of long-term processes may be apparent in the structure of present-day Indo-Pacific biodiversity patterns. Here, we use biogeographical patterns of species composition across 290 sites throughout the Indo-Pacific to more comprehensively evaluate competing hypotheses that seek to explain the diversity gradients of reef-building scleractinian coral species. Specifically, we quantify spatial patterns in species composition (n = 719) and the extent to which range limits co-occur among species. We then examine those patterns to determine the influence of three potential causes that could generate and maintain biogeographical structure: contemporary environmental conditions, the present-day spatial distribution of reef habitat and geological history. Our approach identifies zones of abrupt change (breaks) in species composition, and assesses the extent to which the location of those breaks is consistent with geographical patterns in the potential explanatory variables. To complement this approach, we also assess whether species traits indicative of environmental tolerance, dispersal ability and evolutionary histories can provide further evidence for potential causes of abrupt change in species composition between faunal provinces.

2. Methods

(a) Coral range data

Our primary data consist of detailed maps of reefs and of the ranges of 732 scleractinian coral species. We compiled a database of digitized geographical range extents of Indo-Pacific coral species [22]. Individual reefs (23; n = 6919; electronic supplementary material, dataset S1) were aggregated within a radius of 100 km to produce 290 aggregated reef groups or sites (see the electronic supplementary material, dataset S2) to reduce false presences and to avoid exceeding the computational memory limits of ArcGIS during analysis. The species composition was determined for each site, and a sites-by-species incidence matrix was generated for analysis.

(b) Biogeographical structure

To identify how coral species composition changes through space within the Indo-Pacific, the spatial structure of scleractinian coral species distributions was quantified using elements of metacommunity structure (EMS) analysis [24,25]. EMS analysis discriminates between six idealized patterns for multiple species distributions (random, checkerboard, Clementsian, Gleasonian, nestedness, evenly spaced) according to three elements (coherence, species replacement, species range boundary clumping) of a sites-by-species matrix (see the electronic supplementary material, methods (a)). The estimate of the coherence element is biased when species with a range size of one site are included [26], so we removed all singular occurrences (n = 13; less than 2% of the total species), slightly reducing our dataset to 719 species. The estimate of species turnover within EMS analysis is prone to error [27], so we additionally calculated the more robust nested measure based on overlap and decreasing fills (NODF) metric [28]. Distinct faunal provinces and associated faunal breaks were identified and mapped on the basis of the EMS output with a distance-based method for cluster analysis (see the electronic supplementary material, methods (b), [29]).

(c) Geological, habitat and environmental clusters

We examined the co-location of faunal breaks with abrupt transitions in contemporary environmental conditions, present-day spatial distribution of reef habitat and broad-scale geological features. We identified geological clusters according to transitions in features indicative of different geological histories, which, if co-located with faunal breaks, would suggest that species distributions reflect long-term colonization, extinction or speciation processes. These geological clusters were defined by the extent of major tectonic plates [30] and mantle plume tracks [31–33], which share a similar geological history ([31,34,35]; electronic supplementary material, methods (e)). To account for the possibility that reefs on the same plate may have a different geological history due to the formation of more recent mantle plume tracks, we assigned an intermediate geological distance value to reflect similar ancient (plate), but dissimilar recent (track), geological history. We used a similar rule for analysing reefs along the Reunion mantle plume track that straddles the African and Indian plates i.e. reefs that share the track were assumed to have a similar geological history.

To investigate whether present-day spacing of reef habitat (potentially creating barriers to dispersal) is important, we
identified clusters of reefs separated by a given distance threshold. We repeated this analysis for a range of distance thresholds from 125 to 250 km, with increments of 25 km (see the electronic supplementary material, methods (d)). We chose the lower limit of this range to ensure the distance exceeded the aggregation distance used to generate reef groups comprising each of the 290 sites. Distances from 200 to 250 km generated identical clusters, therefore, we set our upper limit at 200 km.

Boundaries in environmental conditions were established using an iterative self-organizing cluster procedure within ArcGIS ver. 10 [36] based on the spatial distribution of four environmental variables: SST mean, SST range, salinity and nitrate concentration [37]; electronic supplementary material, methods (c)). Mean SST, SST range and nitrate concentrations are common proxies for available energy, environmental stability and productivity [8,14,38]. The geographical location of transitions between environmental clusters indicated breaks in environmental conditions, and we used these boundaries to classify reefs into regions with similar environmental conditions.

Reef groups were assigned to the environmental, habitat and geological clusters with which they spatially co-occurred. The extent to which faunal provinces were correlated with clusters based on these three sets of potential explanatory variables was quantified using Mantel tests on group contrast matrices [39]. Mantel $R_{M}$ is consistently lower than the more familiar Pearson’s $r^{2}$ [40]. Specifically, a simulation study suggested that $r^{2} = 0.25$ was roughly equivalent to $R_{M} = 0.05$, and that a large coefficient of determination ($R^{2} \geq 0.80$) produced by trend surface analysis was equivalent to $R_{M} \geq 0.25$ [41]. However, this approach has been criticized because it assumes that locations are statistically independent of one another (i.e. it fails to account for spatial autocorrelation), leading to an increased type I error rate [42]. To address this limitation, we evaluated statistical significance using an area-weighted permutation procedure that randomly relocated breaks along the spatially ordered matrix generated using Mantel tests on group contrast matrices [39]. The best models were selected as those judged to be biologically feasible (see the electronic supplementary material, methods (g)). The null distribution of Mantel $R_{M}$ statistics $(\sqrt{R_{M}})$ generated by our approach takes better implicit account of spatial autocorrelation within the data, by preserving both the ordering of sites in the sites-by-species matrix (which arises from the spatial structure in species composition), and the size distribution of provinces. We adjusted for multiple tests using the formula $R_{M} = 1 - (1 - R_{M})^{1/n}$, where $\alpha$ is the corrected alpha level, $\alpha$ is the critical value (0.05) and $n$ is the number of comparisons.

(d) Species traits

We collated eight traits for each of 480 species for which information is available (67% of the species in the EMS analysis, electronic supplementary material, dataset S3), and tested whether any of these traits consistently conferred a higher or lower probability that a species would straddle faunal breaks. Coral species at faunal breaks may be selectively filtered out of the species pool by declines in the variety of available habitat types (i.e. habitat heterogeneity [43]), in which case we would expect generalist species to be more likely to traverse breaks. We evaluated this expectation using three species traits: depth range, wave exposure tolerance and turbidity tolerance. Depth range was measured in 10 m increments from 0–10 m to more than 50 m, after Carpenter et al. [44]. Wave and turbidity tolerance were collated from Diaz & Madin [45].

We also considered four reproductive traits that are associated with dispersal ability: reproductive mode (brooder versus spawner), the nutritional mode of larvae, indicated by the presence or absence of symbiotic zooxanthellae, egg size and sexuality (gonochore versus hermaphroditic). If dispersal limitation has acted individually on setting the range limits of coral distributions, we would expect the species that possess traits associated with greater dispersal abilities to be the most likely to traverse faunal breaks. For broadcast-spawning coral species, egg size is a proxy for development rate, because small eggs develop more rapidly [46]. Spawning species were placed in five categories of egg size ranging from less than 100 μm to more than 500 μm. Brooding species have an instantaneous development rate because larvae are ready to settle upon release [47] and, therefore, constituted an additional category that represented the fastest development rate.

The range sizes of coral species have been hypothesized to be larger for older taxa [10], but see [13]). If so, older species might be more likely to straddle faunal breaks. To test this hypothesis, we included the species trait of genus age, as determined from the fossil record (species level ages cannot be obtained reliably [48]). Age was partitioned into six categories (0–19, 20–39, 40–59, 60–79, 80–99 and more than or equal to 100 Ma) reflecting precision in the available estimates (Paleobiology Database, http://paleodb.org; M. Kosnik 2006, unpublished data).

The likelihood of species with particular traits consistently crossing faunal breaks was analysed with a generalized linear mixed model (GLMM; electronic supplementary material, methods (g)). The model included three random effects: faunal break, clade [49] and suborder [50]; the latter two were nested and included to account for any phylogenetic non-independence that was not explicitly captured within sexuality or genus age [51]. We investigated all combinations of main effects and interaction terms that were judged to be biologically feasible (see the electronic supplementary material, methods (g)). The best models were selected as those within three ΑICc (Akaike’s information criterion) units of the best model, and were model averaged to estimate effects of particular trait variables [51].

The significance of individual species traits within the best models was evaluated with 95 per cent CIs on model-averaged odds ratios. Odds ratios represent the likelihood that a species with a particular trait state will successfully traverse faunal breaks. This likelihood is expressed in comparison with the reference category for dichotomous variables, or as the change in likelihood with one unit difference for a continuous variable. An odds ratio of one indicates that all trait states are equally likely to cross faunal boundaries, less than one indicates the reference category or lowest value is more probable, and greater than one indicates that the non-reference category or highest value is more probable for species that successfully straddle faunal breaks compared with species that do not. Traits with confidence intervals that overlap one indicate that species with either trait state do not differ significantly in their likelihood of crossing faunal breaks.

3. Results

(a) Biogeographical structure

Our analysis of spatial patterns in coral species composition identified nine faunal provinces throughout the Indo-Pacific, within which species composition was distinct, separated by faunal breaks where multiple species reached their geographical range limits (figure 1a). The provinces were significantly nested (see the electronic supplementary material, table S1; EMS $p < 0.001; \text{NODF} = 74.99, p < 0.001$), which indicates that the loss of species through space occurs in a predictable sequence such that depauperate provinces comprise a subset of species present within adjacent more species-rich provinces [52]. In addition, species range limits co-occurred more frequently than expected by chance (Morisita’s index $\geq 15, p < 0.001$). The most species-rich province is centred on Indonesia and the Philippines (figure 1a). The additional eight provinces, from west to east, are: the Persian Gulf, Africa–India, Australia, Japan–Vietnam,
Figure 1. Indo-Pacific reef clusters derived from (a) species composition (i.e. faunal provinces), (b) geology, (c) distance and (d) environmental conditions. Black lines are faunal breaks (dashed for those separating Red Sea and Andaman–Nicobar Islands provinces). Results of the Mantel correlation of faunal provinces with clusters based on potential drivers are shown in the upper right corner of (b–d). The distance-based and environmental clusters depicted are for the best threshold (175 km) and variable combination (SST mean, nitrate, salinity), as defined from the number of clusters and the Mantel correlation with faunal provinces.
Table 1. Spatial co-occurrence of faunal provinces throughout the Indo-Pacific with clusters derived from geological features, distance and environmental conditions. Significance values were assessed based on a correction for multiple tests. The corrected 5% alpha value for distance-based clusters ($n = 4$) was 0.012, and for environmental clusters ($n = 11$) was 0.004, and significant correlations according to the corrected value are in bold.

<table>
<thead>
<tr>
<th></th>
<th>$R_p^2$</th>
<th>$R_m^2$</th>
<th>$r_M$</th>
<th>$P$</th>
<th>clusters</th>
</tr>
</thead>
<tbody>
<tr>
<td>geology</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>tectonic plates</td>
<td>0.551</td>
<td>0.304</td>
<td>&lt;0.001</td>
<td>13</td>
<td></td>
</tr>
<tr>
<td>and mantle plume</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>track</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>distance-based</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>125 km</td>
<td>0.389</td>
<td>0.151</td>
<td>&lt;0.001</td>
<td>23</td>
<td></td>
</tr>
<tr>
<td>150 km</td>
<td>0.164</td>
<td>0.027</td>
<td>0.005</td>
<td>14</td>
<td></td>
</tr>
<tr>
<td>175 km</td>
<td>0.322</td>
<td>0.104</td>
<td>&lt;0.001</td>
<td>12</td>
<td></td>
</tr>
<tr>
<td>200 km</td>
<td>0.297</td>
<td>0.088</td>
<td>&lt;0.001</td>
<td>8</td>
<td></td>
</tr>
<tr>
<td>environment</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>SST mean, SST</td>
<td>0.167</td>
<td>0.028</td>
<td>&lt;0.001</td>
<td>12</td>
<td></td>
</tr>
<tr>
<td>range, salinity</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>nitrate</td>
<td>0.122</td>
<td>0.015</td>
<td>0.052</td>
<td>9</td>
<td></td>
</tr>
<tr>
<td>SST mean, SST</td>
<td>0.228</td>
<td>0.085</td>
<td>&lt;0.001</td>
<td>14</td>
<td></td>
</tr>
<tr>
<td>range, nitrate</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>salinity</td>
<td>0.122</td>
<td>0.015</td>
<td>0.052</td>
<td>9</td>
<td></td>
</tr>
<tr>
<td>SST mean, SST</td>
<td>0.228</td>
<td>0.085</td>
<td>&lt;0.001</td>
<td>14</td>
<td></td>
</tr>
<tr>
<td>range, salinity</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>SST mean, nitrogen</td>
<td>0.208</td>
<td>0.043</td>
<td>0.011</td>
<td>5</td>
<td></td>
</tr>
<tr>
<td>SST mean, nitrate</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>SST mean, salinity</td>
<td>0.216</td>
<td>0.046</td>
<td>0.004</td>
<td>8</td>
<td></td>
</tr>
<tr>
<td>SST range, salinity</td>
<td>0.254</td>
<td>0.065</td>
<td>&lt;0.001</td>
<td>7</td>
<td></td>
</tr>
<tr>
<td>SST range, nitrate</td>
<td>0.020</td>
<td>&lt;0.001</td>
<td>0.137</td>
<td>3</td>
<td></td>
</tr>
<tr>
<td>Salinity, nitrate</td>
<td>0.208</td>
<td>0.043</td>
<td>0.011</td>
<td>5</td>
<td></td>
</tr>
</tbody>
</table>

Fiji–Caroline Islands, Tonga–Samoa, Polynesia and Hawaii–Line Islands.

Of the 290 sites, 59 were of sufficient similarity to any other to be placed within a province (figure 1a). The reef areas encompassed by both the Red Sea and the Andaman and Nicobar Islands were very large, comparable to several of the provinces identified elsewhere (see the electronic supplementary material, figure S1). These two locations were not identified as ‘provinces’ simply because the reefs within these regions were so close together that they clustered as a single site. Consequently, we designated the Red Sea and the Andaman–Nicobar Islands as two additional provinces (figure 1a).

(b) Geological, habitat and environmental clusters

Geological clusters, based on tectonic plate boundaries and mantle plume tracks, were strikingly coincident with the faunal provinces (figure 1b; Mantel test, $R_p^2 = 0.301; p < 0.05$, table 1; and electronic supplementary material, figure S2), which suggests species distributions reflect long-term colonization, extinction or speciation processes. Moreover, the number of geological clusters ($n = 13$) was very similar to the number of faunal provinces ($n = 11$). The additional geological clusters arose because the Africa–India faunal province in the western Indo-Pacific was split into three geological clusters: the African plate, the Indian plate and the Reunion Track. Furthermore, in the east, the Fijian plate isolates Fiji from the Caroline Islands, splitting another faunal province. Conversely, two geological clusters combine faunal provinces: the Eurasian plate contains the Indonesian and Andaman–Nicobar Islands provinces, while the Pacific plate cluster links the Tonga–Samoa province with the Caroline Islands.

The spatial distribution of reef habitat is a substantially weaker predictor of provincial structure than geological features, indicating that species are unlikely to be distributed according to dispersal processes, because faunal breaks do not coincide with large expanses of ocean lacking reef habitat. A moderate distance-threshold (175 km) produced clusters that were the most similar to faunal provinces, based on correlation and on the overall number of clusters, but the correlation was approximately one-third of that found for geological features ($R_m^2 = 0.104, p < 0.001; n = 12$; figure 1c and table 1). Larger distance thresholds generated fewer clusters and had a lower correlation with faunal provinces ($n = 8, R_m^2 = 0.088$). Conversely, smaller distance thresholds generated clusters that far exceeded the number of provinces ($n = 23$ for 125 km), or had a much lower correlation ($R_m^2 = 0.027$ for 150 km).

The correlation of faunal provinces with clusters of sites generated from combinations of environmental variables was statistically significant ($p < 0.004$) for five variable combinations, all of which included either the mean or range of SST. The highest correlation was with environmental clusters based on SST mean, nitrate and salinity ($R_m^2 \leq 0.100, p < 0.05$; figure 1d) which generated 11 clusters—equal to the number of faunal provinces. However, these environmental clusters of sites were geographically widespread, and many of the reefs within the same cluster were far apart (e.g. the dark blue cluster in figure 1d includes reefs around Madagascar, Fiji and Japan). Furthermore, the Mantel correlation was three times lower than the correlation with geological clusters.

Overall, these results suggest that faunal breaks coincide predominantly with geological features, and to a lesser (but statistically significant degree) with contemporary environmental conditions and the spatial distribution of habitat. In all tests, our modified permutation test produced null distributions that were substantially more conservative than the traditional approach that assumes no spatial structure in the data (see the electronic supplementary material, figure S2).
Within suborder, accounted for less than 1 per cent of the breaks. Phylogenetic effects in the likelihood of species ranges spanning the individual among-break variation in intercept estimates from the model were explained by baseline differences (specifically, among-break variation in intercept estimates from the model) in the likelihood of species ranges spanning the individual breaks. Phylogenetic effects, represented by clade nested within suborder, accounted for less than 1 per cent of the variation within the model. However, sexuality, in addition to providing information about colonization potential, is highly conserved within clades and, therefore, may implicitly incorporate a phylogenetic signal. This interpretation is supported by the fact that phylogenetic random effects become more important if sexuality was dropped from the analysis.

(c) Species traits and faunal breaks
Among the species traits we examined, depth range was a strong predictor of the likelihood that a species occurred on both sides of a faunal break (see figure 2 and electronic supplementary material, table S2). Specifically, for each 10 m increase in depth range, species were 27 per cent more likely to straddle faunal breaks. In other words, as provinces decrease in richness, the species within them become skewed towards a greater proportion of depth generalists. Three other species traits (tolerance to wave exposure and turbidity, propagule nutritional mode and reproductive mode) were included within the set of best predictive models for whether species crossed faunal breaks, but do not appear to strongly influence biogeographical patterns (estimates of their effects were small in magnitude and 95% CIs for odds ratios overlapped 1; figure 2). In addition, the influence of larval development rate and genus age on biogeographical patterns differed among gonochores and hermaphrodites. Gonochores with fast larval development rates were more likely to occur on both sides of faunal breaks (28% more likely for each decrease in development rate class), but this trait did not affect the likelihood of hermaphrodite species spanning faunal breaks. However, the occurrence of hermaphrodites across faunal breaks was well predicted by the final trait we examined, which was genus age: species were 27 per cent more likely to straddle breaks with each 20 Ma increase in genus age. Genus age was not an important predictor for gonochores.

At the two most peripheral faunal breaks we considered (between Africa–India and the Persian Gulf; and between Fiji–Caroline Islands and the Hawaii–Line Islands), species were less likely to occur on both sides than at the more centrally located boundaries (see the electronic supplementary material, figure S3). This suggests that these faunal boundaries are more distinctive. Overall, 42 per cent of the variance within the model was explained by baseline differences (specifically, among-break variation in intercept estimates from the model) in the likelihood of species ranges spanning the individual breaks. Phylogenetic effects, represented by clade nested within suborder, accounted for less than 1 per cent of the variation within the model. However, sexuality, in addition to providing information about colonization potential, is highly conserved within clades and, therefore, may implicitly incorporate a phylogenetic signal. This interpretation is supported by the fact that phylogenetic random effects become more important if sexuality was dropped from the analysis.

4. Discussion
Indo-Pacific corals are spatially assembled at large scales, generating distinct faunal provinces that are delineated by sharp faunal breaks at which multiple species reach their geographical range limits. These provinces and breaks have not been identified in previous analyses of corals, which have focused principally on species richness [14], or on species turnover at smaller scales [54,55]. The provinces we have identified share some similarities with the global marine provinces proposed by Spalding et al. [17]. However, their zoning scheme identified more than 30 Indo-Pacific provinces, most of which are subsets of the Indonesian and Australian provinces, which our analyses indicate have a relatively homogenous coral fauna. Moreover, Spalding et al.’s provinces were principally based on qualitative assessment of the global distribution of endemic species, whereas endemics are relatively uncommon among scleractinian coral faunas in most parts of the Indo-Pacific [56]. Other previous schemes differ more substantially from our own, classifying most of the Indo-Pacific as a single province [18,57]. These disparities point to the need for a more rigorous data-driven, quantitative approach to identifying large-scale marine zones or biogeographical provinces, such as that presented here.

Unexpectedly, the provincial boundaries we identified do not coincide with breaks in contemporary environmental conditions or present-day habitat distribution. Instead, faunal boundaries show striking concordance with geological features (i.e. tectonic plates and mantle plume tracks), which suggests a role for long-term historical processes in the generation of biogeographical structure of Indo-Pacific corals [58].

Figure 2. Association of species traits with the likelihood of a species crossing Indo-Pacific faunal breaks. Values plotted are model-averaged odds ratios (points) with 95% CIs (solid horizontal lines). Where confidence intervals do not encompass 1 (vertical dotted line), the trait state (label within box) was significantly more likely among species that crossed breaks. Genus age and development rate interacted significantly with sex, so gonochores and hermaphrodites are shown separately for these traits.
The influence of history on the distribution of marine organisms has long been recognized [11,34,35,58]. However, attempts to identify the effect of history on present-day coral biogeography have been hindered by the use of species richness rather than composition as the principal response variable [14]. Our analyses of species traits also revealed that the depth range over which a species occurs, the rate of larval development in gonochorists, and the genus age of hermaphrodites are important determinants of the likelihood that species will straddle faunal breaks.

(a) Historical mechanisms
Understanding the mechanisms that link biogeographical structure and geological features requires consideration of ecological and evolutionary processes that operate over long temporal scales. At convergent plate boundaries, one plate is subsumed under another forming a subduction zone, which is characterized topographically by a trench on the subducting side, and by island arcs (linear archipelagos) on the consuming side (e.g. west of Sumatra [31]). These arcs, which run parallel to plate boundaries, can act as stepping stones for dispersal and colonization within a tectonic plate over long geological timescales [35]. For example, the Izu–Bonin–Volcanic Island arc connects otherwise isolated islands parallel to the Filipino plate boundary within the Japan–Vietnam province. Mantle plume tracks, which form volcanic island chains as tectonic plates move over a stationary plume [35], could also link reefs within the Africa–India province, Hawaii–Line Islands province, and the Polynesian province. Stepping stone colonization has received empirical support. On some islands, taxa are older than the geological age of those islands, indicating that these taxa originated elsewhere and later colonized the younger islands. For example, two islands along the track within the Africa–India province are 2–3 million years old, but host an endemic plant family that molecular dating has estimated as 52 million years old [59].

During the process of migration along stepping stones, species with strong dispersal and/or establishment abilities may be favoured, observable as selective colonization. This mechanism is often invoked to explain nested diversity structures [60] and therefore fits well with our findings. For example, Tonga and Samoa are geologically young (approx. 4–5 Ma), so may have experienced selective colonization of species from the geologically older islands of Fiji (approx. 35 Ma [35]), potentially explaining the faunal break between these islands. Moreover, when an island is created through geological separation from the mainland, resultant vicariant speciation events can be followed by selective colonization of a new island species back to the mainland. Molecular evidence suggests that such ‘reverse colonization’ has occurred for some species of Indian Ocean chameleons, which originated in Madagascar after the break-up of Gondwana, and subsequently colonized the mainland of Africa and India, and the Seychelles [61].

A logical prediction of selective colonization as a mechanism for provincial differences is that species with efficient colonization abilities and broad environmental tolerance would be disproportionately present in low richness provinces. Our species trait analysis is generally consistent with this prediction. Generalist species with broader depth ranges were more likely to straddle faunal breaks, which suggests that species are selectively filtered according to their environmental tolerance. In contrast, gonochoristic species with slower development rates (and therefore, potentially longer pelagic larval durations), indicated by a large egg size, were less likely to cross faunal breaks. While this result is inconsistent with the conventional view that long pelagic larval durations should lead to more long-distance dispersal, it is consistent with more recent work suggesting that species with shorter pelagic durations are more successful at consolidating range expansions because the offspring of colonizers are more likely to recruit back to the new population [19,46], as well as recent work showing weak pelagic larval duration–dispersal links in fishes [62]. Hermaphroditic coral species within younger genera are also more likely to reach range limits at faunal breaks compared with hermaphroditic species within older genera, which may be explained by the increased time available for species of an older age to have colonized peripheral regions [10].

(b) Contemporary environmental conditions
The relatively low concordance between faunal breaks and transitions in environmental conditions suggests that geographical variation in contemporary environment is unlikely to play a dominant role in determining the changes in species composition that occur along Indo-Pacific biodiversity gradients, despite the fact that some previous work has found SST to be a strong statistical predictor of coral species richness [8,14]. Our finding illustrates the insight that can be gained by expanding the focus of analysis beyond species richness alone, to encompass other aspects of species composition. The Mantel correlation coefficients between species composition and environmental conditions in our study are similar in magnitude to those found between oceanographic conditions (temperature, pH and productivity) and bivalve occurrence on a global scale [55], although our modification of the Mantel permutation test to account for spatial autocorrelation produces a much lower estimate of statistical significance than the standard test. Moreover, our analysis found that, for corals, species composition exhibits a much stronger relationship with geological clusters than with present-day environmental conditions.

The relatively weak association between faunal breaks and transitions in environmental conditions does not necessarily imply that environmental conditions have no effect on faunal breaks in the Indo-Pacific at all. For instance, ecologically, localities with high habitat heterogeneity tend to be associated with the presence of both generalist and specialist species because a greater variety of habitats are available to accommodate the specialists, and this contributes to a high diversity of species [43,63]. The Indonesian and Philippine region holds a diverse array of different reef habitats, including large areas of shallow shelves and areas where reefs flourish in turbid waters. Moving away from this region, habitat heterogeneity decreases. For example, within the central and eastern Pacific, fringing reefs and atolls are the dominant reef type [31,35]. Thus, as the variety of available habitats progressively decreases in these more peripheral regions, specialist species would be expected to be filtered out of the regional species pool [43]. This expectation is consistent with our results, which show that specialist species are indeed disproportionately filtered out at faunal breaks.

(c) Dispersal limitation
The present-day spatial distribution of reef habitat appears to be of relatively low importance for structuring broad-scale
coral composition. While distance-based clusters of habitat were significantly correlated with faunal provinces for all distance thresholds, either the number of clusters was far in excess of the faunal provinces or the Mantel correlation was low. However, low correlations may reflect the limitations of using distance as a proxy for dispersal. For example, a more complex dispersal model, that included physical oceanic processes, predicted that larval connectivity is disrupted between northern and southern Papua New Guinea [64]. This predicted disruption was not evident in our distance-based clusters, but its location aligns with the faunal break we identified between the Indonesian and Australian provinces (figure 1a,c). The far western and eastern faunal breaks, between the Persian Gulf and Africa–India, and the Fiji–Caroline Islands and the Hawaii–Line Islands, were consistently aligned with distance-based breaks, regardless of the specific distance thresholds we used. Interestingly, these were the same two breaks that had markedly different intercepts within the GLMM, suggesting that the relationship between species traits and the baseline probability of crossing these breaks differs from the rest of the Indo-Pacific.

5. Conclusion

Our results identify a striking, unexpected concordance between geological features and coral biogeographical structure, suggesting a substantial role for long-term processes in the generation and maintenance of coral biodiversity on Indo-Pacific reefs. The weaker, but still statistically significant, concordance with contemporary environmental conditions and spatial distribution of habitat suggests that these factors also play a role. Selective colonization as a hypothesized mechanism to explain the importance of geological features is supported by our biological traits analyses, which shows that depth generalist species cross faunal breaks more readily than specialists. More broadly, our study shows how the ability to distinguish between competing hypotheses for species richness patterns can be enhanced by expanding the focus of macroecological analysis from species richness patterns to species composition, with explicit consideration of species range limits. In particular, a focus of research effort on faunal breaks, transitions in species composition, and the species traits associated with those transitions are fruitful areas for future research into mechanisms that structure species diversity gradients.

6. Data accessibility

— Digitized reef polygons uploaded as electronic supplementary material;
— coral species range boundaries: Hughes et al. [22];
— centroids of merged reef groups uploaded as electronic supplementary material; and
— species traits data uploaded as electronic supplementary material.

We thank the Australian Research Council for financial support. We also acknowledge the support from the Queensland and Australian government for the high-performance computing facility at James Cook University.

References
