ECOGRAPHY

Research

Influence of the geography of speciation on current patterns of coral reef fish biodiversity across the Indo-Pacific

Théo Gaboriau, Fabien Leprieur, David Mouillot and Nicolas Hubert

T. Gaboriau (http://orcid.org/0000-0001-7530-2204) (theo.gaboriau@umontpellier.fr), F. Leprieur and D. Mouillot, UMR MARBEC (CNRS, IRD, IFREMER, UM), Univ. de Montpellier, Montpellier, France. DM also at: Australian Research Council Centre of Excellence for Coral Reef Studies, James Cook Univ., Townsville, QLD, Australia. – N. Hubert, Inst. de Recherche pour le Développement (IRD), UMR 226 ISEM, Montpellier, France.

The role of speciation processes in shaping current biodiversity patterns represents a major scientific question for ecologists and biogeographers. Hence, numerous methods have been developed to determine the geography of speciation based on co-occurrence between sister-species. Most of these methods rely on the correlation between divergence time and several metrics based on the geographic ranges of sister-taxa (i.e. overlap, asymmetry). The relationship between divergence time and these metrics has scarcely been examined in a spatial context beyond regression curves. Mapping this relationship across spatial grids, however, may unravel how speciation processes have shaped current biodiversity patterns through space and time. This can be particularly relevant for coral reef fishes of the Indo-Pacific since the origin of the exceptional concentration of biodiversity in the Indo-Australian Archipelago (IAA) has been actively debated, with several alternative hypotheses involving species diversification and dispersal. We reconstructed the phylogenetic relationships between three species-rich families of coral reef fish (Chaetodontidae, Labridae, Pomacentridae) and calculated co-occurrence metrics between closely related lineages of those families. We demonstrated that repeated biogeographic processes can be identified in present-day species distribution by projecting co-occurrence metrics between related lineages in a geographical context. Our study also evidence that sister-species do not co-occur randomly across the Indo-Pacific, but tend to overlap their range within the IAA. We identified the imprint of two important biogeographic processes that caused this pattern in 48% of the sister-taxa considered: speciation events within the IAA and repeated divergence between the Indian and Pacific Ocean, with subsequent secondary contact in the IAA.

Introduction

Diversity is not evenly distributed on earth, but tends to accumulate in a few areas of exceptional species richness. In the marine realm, the Indo-Australian Archipelago (hereafter IAA, Fig. 1A), which encompasses Indonesia, Malaysia, the Philippines, Timor Leste, Papua New Guinea and northern Australia, is the richest part of the world with ca 2500 fish species and 600 coral species (Veron et al. 2009, Pellis-

(A) Location of the IAA

Western Indo-Pacific | Indo Australian Archipelago | Eastern Indo-Pacific

(B) Centre of Origin hypothesis

WIP | IAA | EIP

■ Divergence time
■ Asymmetry
□ Co-occurence
□ Overlap

(C) Centre of Overlap hypothesis

WIP | IAA | EIP

■ Divergence time
■ Asymmetry
□ Co-occurence
□ Overlap

(D) Centre of Accumulation hypothesis

WIP | IAA | EIP

■ Divergence time
■ Asymmetry
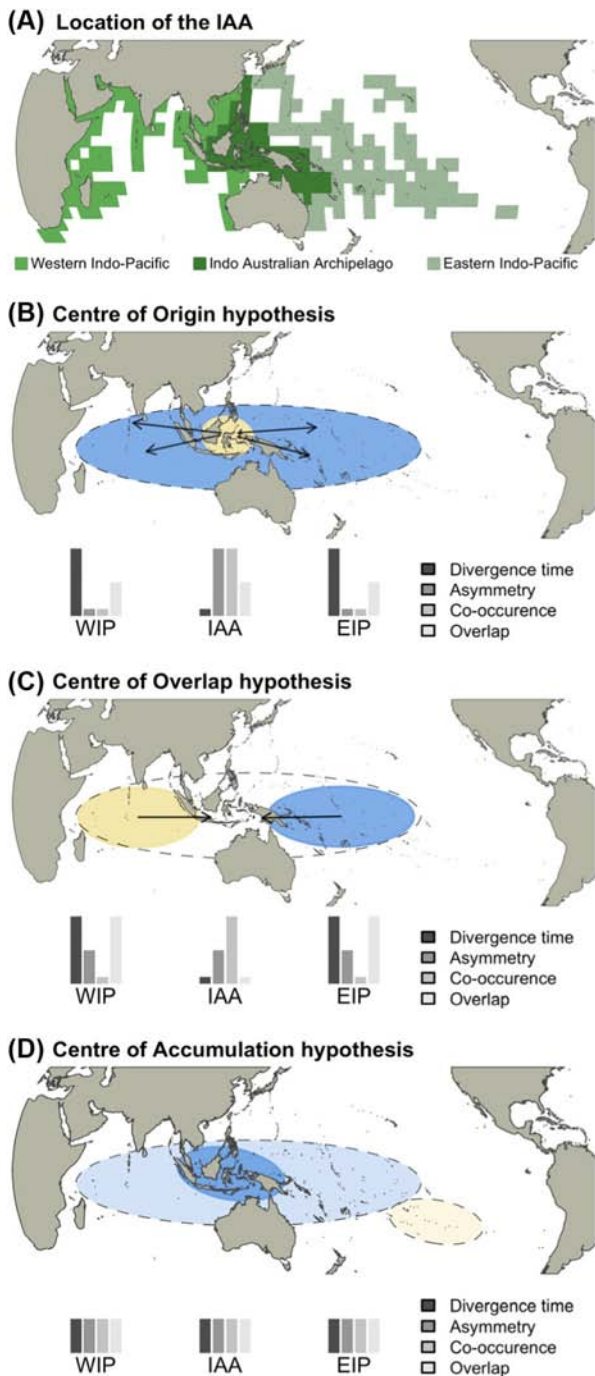□ Co-occurence
□ Overlap

Figure 1. Description of the three main hypotheses proposed to account for the origin of biodiversity patterns across the Indo-Pacific with their respective predictions regarding sister-species range overlap, range asymmetry, divergence time and co-occurrence. (A) Delineation of the Indo-Australian Archipelago (IAA) defined by Gaither and Rocha (2013) and the peripheral regions included in the analysis: western Indo-Pacific region (WIP) and eastern Indo-Pacific region (EIP). (B) Centre of origin hypothesis: species diverged within the boundaries of the IAA and dispersed towards peripheral areas leading to IAA-centred overlapping geographic ranges. The IAA is expected to display a higher number of co-occurrences between sister-taxa, a lower divergence time of co-occurring

sier et al. 2014). Consequently, the mechanisms at the origin of this exceptional concentration have been debated in a particularly rich literature, involving three main competing hypotheses (Cowman 2014). First, the 'centre of origin' hypothesis (Briggs 2000) posits that the IAA acted as a 'cradle of diversification' that hosted the earliest evolutionary history of many lineages, which dispersed towards peripheral areas (Fig. 1B). Second, the 'centre of overlap' hypothesis (Fig. 1C) proposes that biodiversity peaks in the IAA as a consequence of its geographic position at the junction between the Indian and Pacific Oceans, each hosting distinct reef fish faunas (Woodland 1983). Last, the 'centre of accumulation' hypothesis (Fig. 1D) posits that new species arose across the Indo-Pacific and further accumulated in the IAA due to a higher stability of coral reefs through times and lower extinction rates (Ladd 1960, Barber and Bellwood 2005, Pellissier et al. 2014).

Only recently, these hypotheses have been considered as not mutually exclusive and potentially all relevant (Cowman and Bellwood 2013, Cowman 2014). For instance, recent studies have emphasized that the IAA has been generating biodiversity since the Miocene (–23.5 my) (Cowman and Bellwood 2013) and acted as a major refugia during the Quaternary glaciations (Pellissier et al. 2014). From a phylogeographic perspective, there is also recent evidence that many species occurring in the IAA originated in peripheral regions through allopatric speciation across large-scale barriers (Hubert et al. 2011, 2012, Gaither and Rocha 2013). Thus, the challenge is no longer to determine which hypotheses are relevant, but to determine at which extent each mechanism left its imprint on present-day species distribution. This challenge resonates even more strongly today in front of the increasing threats, associated to climate change and overexploitation, that imperils the ability of coral reefs to act as refuges, particularly so in the IAA (Couce et al. 2013) and the Great Barrier Reef (Ainsworth et al. 2016).

Comparative methods have been developed to examine geographic modes of speciation by assessing the degree of

Figure 1. Continued
sister-taxa and a higher asymmetry between geographic ranges of co-occurring sister-taxa than peripheral areas. (C) Centre of overlap hypothesis: species diverged by isolation between the two sides of the IAA and dispersed towards the IAA leading to geographic ranges overlapping in the IAA but centred in peripheral areas. The IAA is expected to display a higher number of co-occurrences between sister-taxa, a lower divergence time of co-occurring sister-taxa and a lower overlap between geographic ranges of co-occurring sister-taxa than peripheral areas. (D) Centre of survival hypothesis: speciation events occurred anywhere in the Indo-Pacific, but species only survived in the IAA, which led to geographic ranges centred on the Indo Australian Archipelago with no prediction on sister-taxa geographic range structures. The ancestral geographic range of species is represented by the dotted lines. The ranges of sister-taxa are represented in blue and yellow. Pale blue and yellow represent local extinction, and the arrows represent the direction of post-speciation dispersion.

species range overlap and asymmetry across phylogenies (Fitzpatrick and Turelli 2006, Anacker and Strauss 2014 for mammals and plants, respectively, and Quenouille et al. 2011, Hodge et al. 2012, Hodge and Bellwood 2015 for coral reef fishes). For these authors, a correlation between divergence times and range overlap/asymmetry (age range correlation, hereafter ARC) indicates that the imprint of the geography of speciation is conserved in a significant proportion of sister-species ranges (Ribera et al. 2011). Sister-taxa that recently diverged in allopatry are likely to display no or limited range overlap (Barraclough and Vogler 2000), whereas a positive relationship between divergence time and range overlap is expected due to post-speciation range expansion (Losos and Glor 2003). By contrast, sister-taxa that have recently diverged in sympatry may exhibit a higher degree of over-lap between their geographic ranges, a trend that is expected to decrease over time as species ranges can extend or shrink. Similarly, a high degree of asymmetry between the range sizes may emerge among young sister-taxa if speciation occurred in peripheral areas as a consequence of low connectivity (i.e. peripatric speciation) or strong environmental gradients (i.e. parapatric speciation) (Barraclough and Vogler 2000). In the present study, we detailed the predictions made by the different mechanisms potentially underlying the distribution of fish diversity at the Indo-Pacific scale, and we compared them to the current patterns of geographic range overlap and asymmetry between sister-taxa.

The 'centre of origin' hypothesis predicts the presence of a high number of neo-endemics (i.e. species with a narrow geographic range due to their recent origin) within the IAA (Mora et al. 2003, Bellwood and Meyer 2009). Under this hypothesis, one might expect a high number of sister-taxa co-occurring in the IAA with a high geographic range asymmetry (Fig. 1B). On the other hand, the 'centre of overlap' hypothesis assumes that the IAA created a barrier between the Indian and Pacific Oceans that promoted allopatric speciation (Hubert et al. 2012, Gaither and Rocha 2013). Consequently, a high number of sister-taxa co-occurring in the IAA with a low overlap between their geographic ranges is expected under this hypothesis (Fig. 1C). However, these hypotheses predict that the imprint of past-speciation events on diversity patterns is still traceable. The 'centre of accumulation' hypothesis proposes an alternative explanation, based on post-speciation, that makes no predictions about the geography of speciation as the origin of the species has little impact on their survival (Fig. 1D). Finally, we may expect other patterns of co-occurrence between sister-taxa due to mechanisms that acted at smaller spatial scales, such as peripheral isolation (Hodge et al. 2012).

To explore the potential influence of the different mechanisms, we reconstructed the phylogenetic relationships between species of three speciose reef fish families (Pomacentridae, Chaetodontidae and Labridae) and estimated divergence times between lineages within those families. We also calculated several indices describing the co-occurrence between sister-taxa using a global database of tropical reef fish distribution (Parravicini et al. 2014, Pellissier et al. 2014). Then, we compared divergence times with the different metrics to test whether past speciation events have influenced present-day species distribution in the Indo-Pacific. Finally, we mapped all the estimated metrics to test the predictions derived from the main hypotheses in terms of species age, range overlap, range asymmetry and co-occurrence.

Methods

Phylogenetic reconstruction

DNA sequences available from previously published studies were collected from GenBank (Supplementary material Appendix 2), and loci were selected independently for each of the three families according to their levels of taxonomic coverage. Several DNA sequence alignments were built and further concatenated following a supermatrix approach that maximized the taxonomic overlap among loci (Thomson and Shaffer 2010). For the maximum likelihood reconstructions, we used outgroups selected for their phylogenetic relatedness from the latest fish super-phylogeny (Rabosky et al. 2013), and maximizing the overlap of the selected loci with the ingroup. For the Chaetodontidae, we included four mitochondrial loci (non-coding: 12S and 16S, coding: COI and Cytb) and four nuclear loci (ETS2, RAG2, tmo4c4 and S7 II) from 100 species that is a taxonomic coverage of 79% of the 127 nominal species of the family. The Chaetodontidae phylogenetic reconstructions were rooted using six outgroups belonging to the families Labridae and Kyphosidae. For the Pomacentridae, we collected five mitochondrial loci (non-coding: 12S and 16S, coding: COI, Cytb and ND3) and four nuclear loci (RAG1, RAG2, S7 II and ATP8/6) from 225 species that is 61% of the 371 nominal species of Pomacentridae. The phylogenetic reconstructions were rooted by using six outgroup taxa from the families Chaetodontidae, Labridae and Kyphosidae. For the Labridae, we assembled sequences of four mitochondrial loci (non-coding: 12S and 16S, coding: Cytb and COI) and three nuclear loci (RAG2, S7 II and tmo4c4) from 348 species that is 53% of the 653 nominal species of the family. We included 11 outgroup species from the family Cirrhitidae and Kyphosidae.

Sequences from each loci were independently aligned using MUSCLE (Edgar 2004) with standard settings, and the alignments were manually edited using MEGA 6 (Tamura et al. 2013). The individual alignments were further concatenated into a supermatrix for each family using SequenceMatrix (Vaidya et al. 2011) and phylogenetic reconstructions were performed using the maximum likelihood (ML) algorithm RAxML (Stamatakis 2014) as implemented in RAxMLGUI (Silvestro and Michalak 2011). Considering the size of the supermatrices, we used two parallel strategies of data partitioning. First, we detected the optimal

partitioning scheme and substitution models, without a priori consideration of the genomic position of the base pairs, using Partition Finder 2 (Lanfear et al. 2014, 2017). Second, we considered a partitioning scheme derived from the known genomic position of the base pairs and consisting of a single partition for each gene. For each gene partition, the most likely substitution model was selected among 88 substitution models according to the Bayesian information criterion (BIC) as implemented in Jmodeltest 2.1.6 (Guindon et al. 2010, Darriba et al. 2012). Once the most likely partitions and the substitution models were selected, we performed 10 parallel runs in RAxMLGUI using the GTR model to comprehensively explore the tree space, and we selected the best tree according to the likelihood scores.

Node age estimates were first computed using the penalized likelihood approach implemented in r8s 1.8 (Sanderson 2003) applied to the most likely trees obtained from the ML analyses. Outgroup species were deleted at this stage of the chronogram reconstructions. These chronograms were further used as starting trees for Bayesian inference of tree topologies and node ages through MCMC as implemented in Beast ver. 1.8.1 (Drummond et al. 2012). We ran six independent MCMC runs of 40×10^6 generations under the selected substitution models. We used the same fossil or biogeography-based calibrations as in Cowman and Bellwood (2011), with broader priors (Supplementary material Appendix 1). We concatenated each independent MCMC run, after removing a burn-in of 10 M steps by run, using LogCombiner ver. 1.8.1 (Drummond et al. 2012). Then, we calculated the maximum clade credibility tree for the combined tree sets using TreeAnnotator ver. 1.8.1 (Drummond et al. 2012). The median node ages and their 95% highest posterior density (HPD) intervals were plotted on the resulting chronograms.

Species geographic ranges

We gathered information on the presence and absence of reef fish species at 169 locations worldwide by examining almost 500 references and extracting information from published works, regional checklists, monographs on specific families or genera and reports (Parravicini et al. 2013). To account for the potential heterogeneity in the accuracy of the species records, species geographic ranges were modelled using convex hull polygons that were further used to score species occurrences across a grid of 5° × 5°, which corresponds to approximately 555 × 555 km at the Equator. This grid size proved to represent a good compromise between the desired resolution and the geographical density of the information (Parravicini et al. 2014, Pellissier et al. 2014). We excluded from our analyses all cells that were located offshore (i.e. below 75 m depth or without coral reefs) or that consisted of more than 90% of emerged lands. The bathymetric data were obtained from the SRTM30_PLUS bathymetry (Shuttle Radar Topography Mission, available at: < http://topex.ucsd.edu/WWW_html/srtm30_plus.html >), and information on the distribution of coral reefs was obtained from available information at the Millenium Coral Reef Mapping project (< http://data.unep-wcmc.org >).

Overlap and asymmetry between sister-taxa geographic ranges

We used several indices to explore the relationships between the geographic ranges of sister-taxa that diverged during the last 23 million yr. First, we estimated the degree of overlap (Eq. 1), in which $N_{A \cap B}$ is the number of cells where taxa A and B co-occur, and N_A and N_B are the total number of cells where taxa A and B occur, respectively. The degree of overlap varies from 0 (fully allopatric distribution) to 1 (fully sympatric distribution).

$$Ov = \frac{N_{A \cap B}}{\min(N_A, N_B)} \tag{1}$$

Second, we estimated the degree of asymmetry (Eq. 2), which varies from 0 (the geographic ranges of sister-taxa have the same size) to 1.

$$As = 1 - \frac{\min(N_A, N_B)}{\max(N_A, N_B)}. \tag{2}$$

These indexes were calculated across the phylogeny by approximating the geographic ranges of ancestral species as the union of the geographic ranges of all of their extant descendants (Barraclough and Vogler 2000). In the analysis, we used all pairs that included either an extant species and its ancestral sister-species or two extant species, but we excluded couples that contained two ancestral species. This approach offers the most comprehensive assessment of species co-occurrence patterns by providing geographic information for all the species available in the phylogeny. For the sake of clarity, we use the term sister-taxa to refer to all pairs including ancestral species or/and extant species, and the term sister-species to refer to pairs including only extant species.

For each cell in the grid, we calculated 1) the number of co-occurring sister-species (excluding the pairs having approximated ancestral ranges to allow comparison with a null model), 2) mean degree of overlap between co-occurring sister-taxa, 3) mean divergence time between co-occurring sister-taxa and 4) the mean range asymmetry index between co-occurring sister-taxa. We mapped the value of each index in every cell to visually explore the patterns of co-occurrence between sister-taxa. We tested if the values obtained for those indices are significantly different in the IAA than in the peripheral areas using a Wilcoxon–Mann–Whitney test. Finally, we tested if those indices are correlated with species richness at the Indo-Pacific scale using a Spearman rank sum test. Because a higher number of co-occurrences between sister-species is expected by chance in cells showing a higher species richness, we randomly re-assigned which species are considered as sister-species and we re-calculated the number of co-occurrence between those random pairs (thus maintaining the same species richness in each cell). Repeating this operation 1000 times allowed us to test if the observed

number of co-occurrence between sister-species is equal to the number of co-occurrence expected by chance.

Finally, we visually inspected the geographic range of each pair of sister-taxa and the distribution of their areas of overlap, to identify general geographic patterns. To explore the relative importance of each process through geological time, we also estimated the frequencies of the patterns identified according to the geological times associated with sister-taxa divergence (Pleistocene, Pliocene and Miocene).

Sensitivity analysis

Incomplete taxon sampling (ITS) is a known problem in historical inference that may lead to spurious spatial patterns and the misinterpretation of biogeographical processes (Cowman and Bellwood 2013). For instance, two species may be identified as sister-species simply by the absence of their closest relatives due to ITS, and their distribution may be interpreted as such. Similarly, missing taxa may induce an overestimation of species age if sister-species do not share a direct common ancestor (Hodge and Bellwood 2015). To estimate the potential impact of such biases, we repeated our analyses with truncated datasets that only included pairs of sister-taxa belonging to genera with a taxonomic coverage above 75%. The results were then compared with those derived from the entire datasets.

It has been repeatedly acknowledged that post-speciation dispersal and local extinction can blur the imprint of past speciation events on present day species distribution (Losos and Glor 2003, Fitzpatrick and Turelli 2006). This statement does not mean that those processes systematically erase the geographic signature of speciation modes (Fitzpatrick and Turelli 2006, Ribera et al. 2011). To test for the general conservation of an imprint of the speciation process on species distribution, we examined the correlation (Spearman's rank test) between divergence time and the degrees of overlap and asymmetry following the ARC method (Barraclough and Vogler 2000) in each family.

Data deposition

Data available from the Dryad Digital Repository: < http://dx.doi.org/10.5061/dryad.b0v7h > (Gaboriau et al. 2017).

Results

Phylogenetic inferences and node age estimates

The three supermatrices consisted of a total of 2594, 2908 and 3842 base pairs of mitochondrial loci and 2609, 1894 and 4556 base pairs of nuclear loci for Chaetodontidae, Pomacentridae and Labridae, respectively. Models that included 6 substitution categories (e.g. GTR, TrN, TIM2 and TIM3), the proportion of invariable sites (I) and four categories of site-specific substitution rates (Γ) were selected in most cases

(Supplementary material Appendix 1 Table A2). The three most likely trees were further used to infer node age through r8s with smoothing parameters S = 1 for Chaetodontidae, S = 1 for Pomacentridae and S = 100 for Labridae, as indicated by the cross-validation procedures. Those ML chronograms were further used as starting trees for the Bayesian inferences of the tree topologies and node ages using the selected substitution models in Beast ver. 1.8.1. In each family, the partitioning by gene provided estimated sample sizes (ESS) exceeding 200 for most parameters, indicating reasonably low standard errors. By contrast, the optimal partition schemes defined by Partition Finder 2 yielded poorly converging MCMC chains for Pomacentridae and Labridae.

Consequently, we kept the three chronograms based on the gene partition schemes (available in Dryad Digital Repository). Maximum credibility trees were statistically well supported considering that only 36 nodes of the 359 nodes of the three chronograms (i.e. 10% of the nodes) included in the spatial analyses were associated to low posterior probabilities (i.e. < 60%).

Geographic co-occurrence and overlap between sister-taxa

For each family, divergence time between sister-taxa that diverged during the Neogene and the Quaternary was correlated with their degree of overlap (Fig. 2, Spearman's rank test: Chaetodontidae, p = 0.001, ρ = 0.45; Pomacentridae, p < 0.001, ρ = 0.54; Labridae, p < 0.001, ρ = 0.35). Divergence time between sister-taxa that diverged during the Neogene and the Quaternary was only correlated with their degree of asymmetry in Pomacentridae (Spearman's rank test: Chaetodontidae, p = 0.639, ρ = 0.062; Pomacentridae, p = 0.047, ρ = –0.17; Labridae, p = 0.442, ρ = 0.059).

The IAA displayed much greater co-occurrence between sister-taxa than the peripheral regions, a pattern that was consistently observed across the three families (Fig. 3B, Wilcoxon rank sum test: Chaetodontidae, p < 0.0001; Pomacentridae, p < 0.0001; Labridae, p < 0.0001). The number of co-occurrences showed a high correlation with species richness (Fig. 3A, Spearman's rank test: p < 0.0001, ρ = 0.99). However, the observed number of co-occurrences was significantly higher (p < 0.05) than expected by chance in most cells (Fig. 3C). Moreover, the mean overlap index was markedly lower for the pairs of sister-taxa that overlapped their range in the IAA compared to those that overlapped in the peripheral regions (Fig. 3D, Wilcoxon rank sum test: Chaetodontidae, p < 0.0001; Pomacentridae, p < 0.0001; Labridae, p < 0.0001). The mean overlap index was also negatively correlated to species richness (Spearman's rank test: p < 0.0001, ρ = –0.54), and most cells with low mean range overlap were located in the Sunda Shelf. Mean divergence time between co-occurring pairs of sister-taxa was negatively correlated with species richness (Spearman's rank test: p < 0.0001, ρ = –0.74) with significantly lower values in the IAA and its vicinity than in the peripheral areas (Fig. 3E, Wilcoxon rank sum test: Chaetodontidae, p < 0.0001; Pomacentridae,
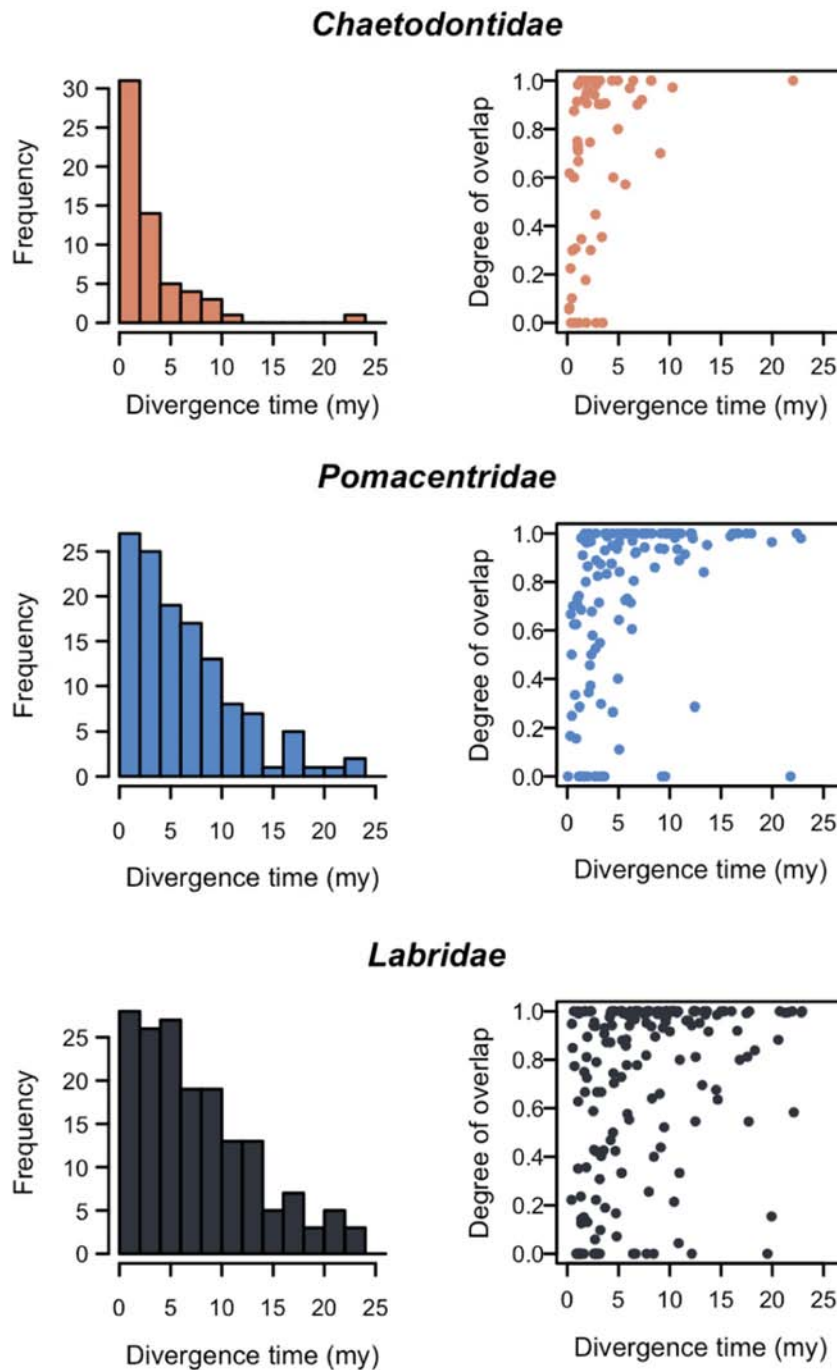
Figure 2. Age-range correlation of sister-taxa that originated during the Paleogene or the Quaternary, for the three families. For each family, we show the histogram of divergence time between sister-taxa at the top, and the relationship between the degree of overlap and divergence time between sister taxa at the bottom.

p < 0.0001; Labridae, p < 0.0001). The Red Sea, Hawaii and southwestern Australia exhibited the highest divergence times on average. Finally, the mean asymmetry indexes were significantly higher in the IAA than in the peripheral areas (Fig. 3E, Wilcoxon rank sum test: Chaetodontidae, p = 0.01; Pomacentridae, p < 0.0001; Labridae, p < 0.0001), with a few exceptions in Hawaii, southeast Australia, the Red Sea

and the Oman Sea that exhibited relatively high asymmetry index values. Deleting pairs of sister-taxa belonging to genera with phylogenetic sampling covering less than 75% of their species provided similar results because the distribution of the number of co-occurring pairs, the overlap index (excepting Pomacentridae) and the asymmetry index were similar (Supplementary material Appendix 1 Table A3).
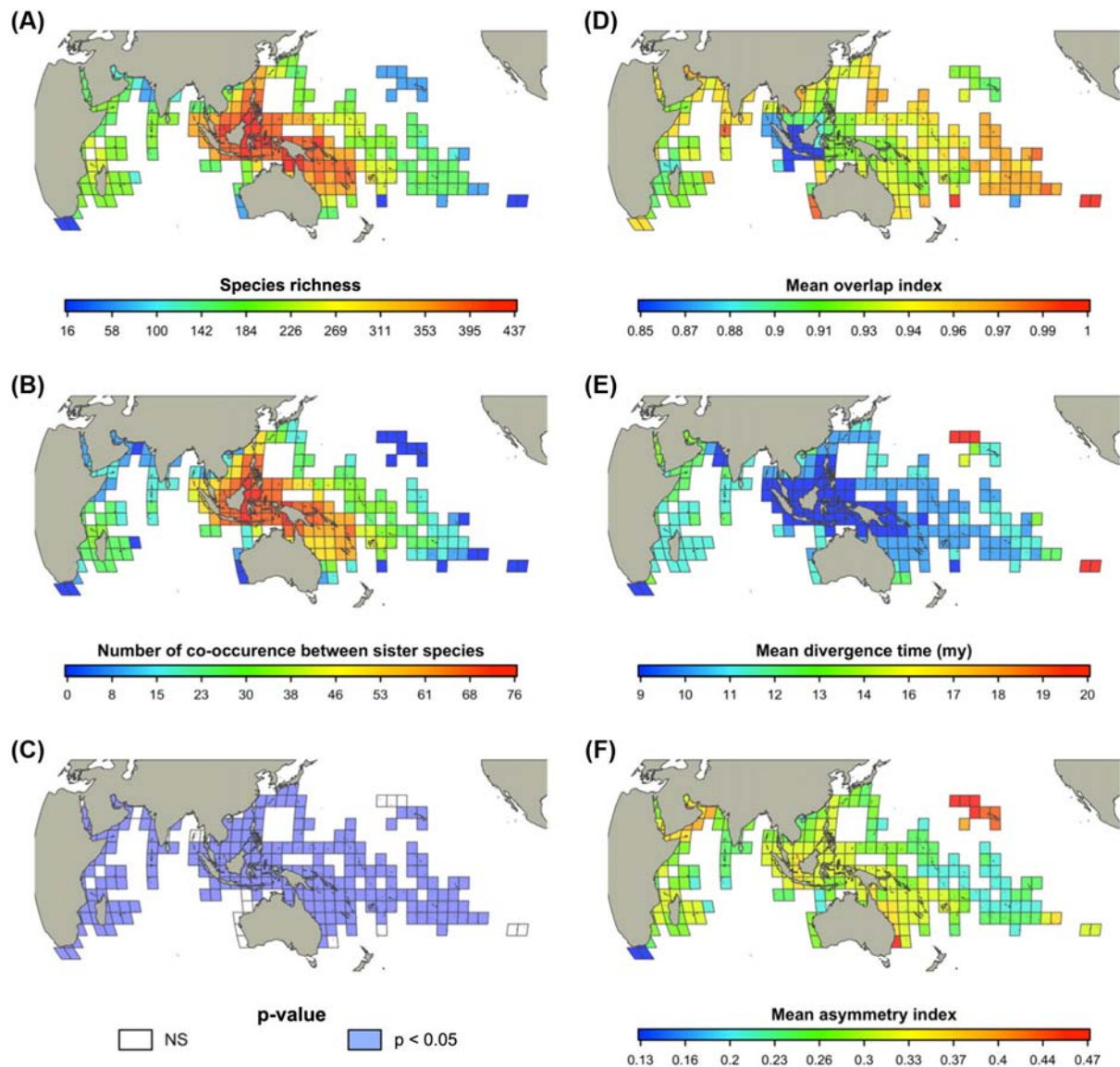
Figure 3. Spatial distribution of coral reef fish biodiversity for the three families (Labridae, Chaetodontidae and Pomacentridae) at the Indo-Pacific scale. (A) Species richness in each grid cell. (B) Number of co-occurrences between sister-species in each grid cell. (C) Significance of the differences between the observed numbers of co-occurring sister-species and those obtained under a null hypothesis where the number of co-occurring sister-species is due to species richness. (D) Mean degree of overlap in each grid cell. For a given cell, we calculated the mean degree of overlap between all sister-taxa co-occurring in that cell. (E) Mean divergence time in each grid cell. For a given cell, we calculated the mean divergence time between all sister-taxa co-occurring in that cell. (F) Mean asymmetry index in each grid cell. For a given cell, we calculated the mean range asymmetry index between all sister-taxa co-occurring in that cell.

Frequencies of geographic range patterns

Five main patterns of sister-taxa geographic distribution were identified. In the most frequent pattern ('IAA restricted distribution', Fig. 4A), the position of the geographic range midpoint of both sister-taxa is centred in the IAA, while at least one of the two sister-taxa is restricted to that region. It can be observed in 104 out 359 pairs, which represent 29% of pairs across the three families (Fig. 5). The second pattern ('WIP and EIP separated distribution', Fig. 4D) was identified by sister-taxa with allopatric

distributions on each side of the IAA. It was observed in 67 pairs (19%, Fig. 5). The third pattern ('EIP restricted distribution', Fig. 4C) was identified when at least one of the two sister-taxa was restricted to the eastern Indo-Pacific (EIP). The areas of range overlap or the boundaries between sister-taxa ranges were also restricted to the EIP. This pattern was observed in 66 pairs (18%, Fig. 5). The fourth pattern ('Widely overlapping distribution', Fig. 4B) follows a scheme where both sister-taxa are widely distributed over the Indo-Pacific realm. It can be observed in 65 pairs (18%, Fig. 5). Finally, the less common pattern, a
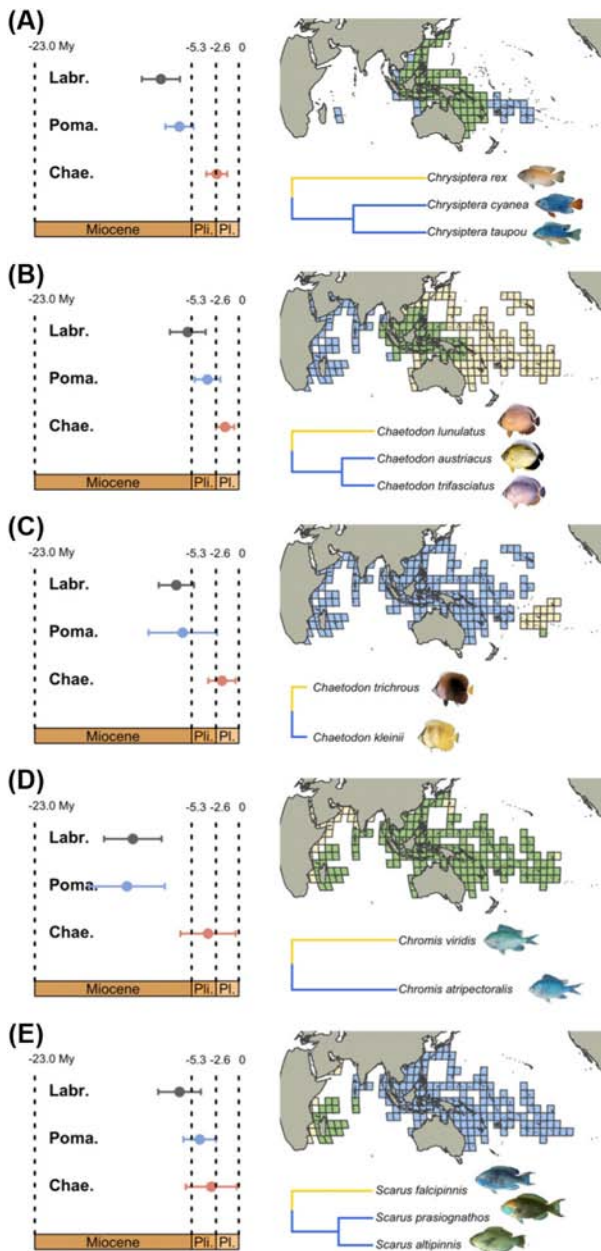
Figure 4. Illustrations of the five major patterns of geographic ranges involving pairs of sister-taxa from the families Pomacentridae, Chaetodontidae and Labridae in the Indo-Pacific Ocean. Each pattern is exemplified by a pair of sister-species from the present study their geographic range at the right. The blue and yellow cells correspond to areas of occurrence for the highlighted lineages while the green cells correspond to areas of co-occurrence between the two lineages. On the left, we presented the mean divergence time between sister-taxa pairs that match each pattern along with confidence intervals. (A) IAA restricted range; (B) WIP and EIP separated range; (C) EIP restricted range; (D) widely overlapping range; (E) WIP restricted range. All the pictures were obtained from Randall (1997).

'WIP restricted distribution' (similar to the 'EIP restricted distribution' on the western side of the Indo-Pacific region, Fig. 4E) was observed in 57 pairs (16%, Fig. 5).
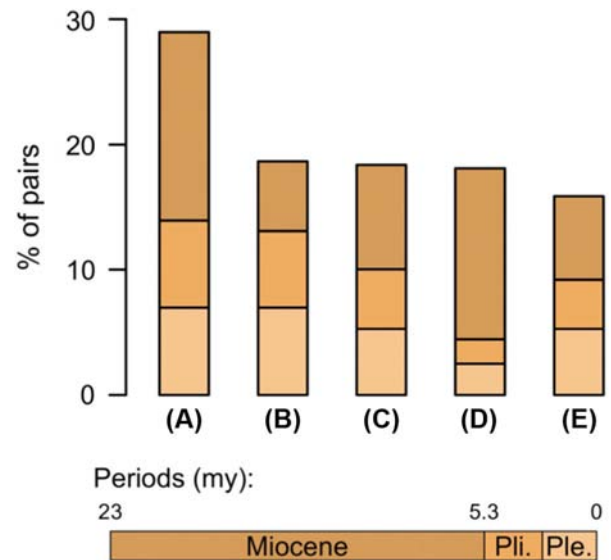

Figure 5. Frequency of each pattern identified between sister-taxa pairs from the three families along a gradient of time. Shades of orange represent the time scale divided in three periods (Miocene, Pliocene and Pleistocene). (A) IAA restricted range; (B) WIP and EIP separated range; (C) EIP restricted range; (D) widely overlapping range; (E) WIP restricted range.

Overall, the frequencies of each pattern were evenly distributed, but their relative frequencies varied across geological time (Fig. 5). The 'EIP restricted distribution' (22%) and the 'WIP and EIP separated distribution' (27%) were the most frequent patterns associated with speciation events during the Pleistocene (–1.8 my to present, n = 67). Among the pairs that diverged during the Pliocene (–5.3 to –1.8 my, n = 115), the 'IAA restricted distribution' (31%) and the 'WIP and EIP separated distribution' (25%) were the most frequently observed. Conversely, the 'Widely overlapping distribution' was frequently observed (25%) among pairs that diverged during the Miocene (–23.5 to –5.3 my, n = 165), but the 'IAA restricted distribution' was the most frequent (32%). Overall, divergence time between sister-taxa that present a 'Widely overlapping distribution' was higher than divergence time between all other pairs, and divergence time between pairs that showed a 'IAA restricted distribution' was significantly older than 'WIP and EIP separated' pairs (Wilcoxon rank sum test p = 0.0004).

Discussion

Robustness and consistency of phylogenies

The inferred phylogenetic relationships are consistent with previously published phylogenetic hypotheses. Among the 100 species of Chaetodontidae analysed (Supplementary material Appendix 1 Fig. A1), the two main clades (Bannerfishes and Butterflyfishes) are observed, and the

genus *Chaetodon* is separated into three main clades as previously described (Cowman and Bellwood 2011). Overall the tree topology and the node ages are concordant with previous reconstructions as the divergence between bannerfishes and butterflyfishes is inferred to happen during the early Miocene and most extant species appeared during the last 5my (Fig. 2).

Among the 348 species of Labridae analysed, our phylogenetic inferences present a better taxonomic coverage compared to previous reconstructions (Cowman and Bellwood 2011, Choat et al. 2012, Hodge et al. 2012), particularly for the most recently derived clades, and a similar one compared to the latest reconstruction (Baliga and Law 2016). The main difference is the placement of *Thalassoma* at the root of the labrines. The estimated age of the most recent common ancestor (MRCA) of Labridae is consistent (62 my) with previous reconstructions (Baliga and Law 2016). The high Smoothing value obtained from the r8s analyses in Labridae suggests that substitution rates largely vary among major clades, a result that may explain the difficulties to obtain concordant node ages in previous phylogenetic studies. Overall, a large portion of extant species was estimated to arise during the last 5 my whereas the overall repartition of species age is much more heterogeneous than in other families (Fig. 2).

Among the 225 species of Pomacentridae analysed, the tree topology and node age estimates are concordant with previous phylogenetic reconstructions (Quenouille et al. 2004, Cooper et al. 2009, Cowman and Bellwood 2011, Frédérich et al. 2013), and the subfamilies identified by James Cooper (2009, Stegastinae, Lepidozyginae, Chrominae, Abudefdufinae and Pomacentrinae) are monophyletic in our reconstructions.

The imprint of biogeographic events

The present study highlights that broad-scale patterns of co-occurrence between sister-taxa follow several major trends despite post-speciation dispersal and local extinction. In the literature, the effect of these two processes has often been used to criticize comparative analysis of sister-lineage's co-occurrence patterns (Losos and Glor 2003, Fitzpatrick and Turelli 2006). Over time, species range movements and local extinction may blur the signature of large-scale biogeographic processes, which may compromise the reliability of biogeographic inferences from present-day species distribution. Here, we showed that a significant part of the information might be conserved in present-day patterns of co-occurrence between sister-taxa for Indo-Pacific coral-reef fishes. The positive correlation between divergence time and the degree of overlap between sister-taxa (Fig. 2) indicates that the co-occurrence patterns have not been completely randomized over time (Quenouille et al. 2011, Ribera et al. 2011). Second, our results show marked patterns of co-occurrence between sister-taxa in the Indo-Pacific, significantly different from random expectations, which may indicate that a fair proportion of species ranges are sufficiently conserved to keep the record of important

biogeographic processes. Third, even though the imprint of speciation has been masked by dispersal and local extinction for several sister-taxa, as illustrated by the important number of old sister-taxa displaying pattern B (Fig. 4B), we can still observe repeated patterns of co-occurrence that are informative about the historical roles played by particular regions on present-day species distribution (e.g. pattern D, Fig. 4D). Here, replacing the co-occurrence pattern in a spatial framework to confront it to historical knowledge, allowed us to go further the limitations of traditional ARC analysis by identifying trends in co-occurrence, overlap and asymmetry between sister-taxa, at the Indo-Pacific scale.

Neo-endemism in the IAA

The most frequent pattern of co-occurrence (Fig. 4A, IAA restricted sister-taxa) coupled with the high values of asymmetry between geographic ranges of sister-taxa co-occurring in the IAA (Fig. 3F), and their low divergence time (Fig. 3E), confirms the presence of neo-endemics as previously suggested in Labridae (Barber and Bellwood 2005, Puckridge et al. 2015), Pomacentridae (Timm et al. 2008) and other families (Tornabene et al. 2014). This pattern of endemism can either be due to a higher rate of speciation in the IAA or to a lower rate of extinction in this area than in peripheral regions. These two hypotheses are consistent with the geological and climatic history of the region since the Neogene, a period during which most species from the families Chaetodontidae, Pomacentridae and Labridae originated (Cowman 2014; see also Supplementary material Appendix 1). It is marked by the collision of the Australian and Eurasian plates at approximately 23 my, a process that promoted the emergence of patchily distributed islands and carbonate platforms in the Indo-Australian region (Lohman et al. 2011), which then led to isolation and further evolutionary divergence between populations (Williams and Duda 2008, Carpenter et al. 2011). It has been recently showed that refuges formed by stable coral habitats during the Quaternary glacial maxima, acted as survival areas and influenced the present-day distribution of species richness (Pellissier et al. 2014). In addition, Quaternary refuges have been separated by several deep seas that promoted divergence between populations within the IAA (e.g. South China Sea, Banda Sea and Celebes Sea; Borsa et al. 2012, Tornabene et al. 2014, Puckridge et al. 2015). However, the high number of co-occurring sister-taxa (significantly higher than expected by chance, Fig. 3B and C) suggests that speciation and subsequent secondary contacts in the IAA have been predominant sources of diversity in the area.

Isolation between EIP and WIP

This high number of co-occurrence along with the low values of overlap between co-occurring sister-taxa observed in the Java sea (Fig. 3D) are also consistent with the idea of an area of secondary contact between sister-taxa in this sea

(Hubert et al. 2011, 2012, Gaither and Rocha 2013). The accumulation of sediments in the IAA and sea level variations caused by the glacial cycles of the Pliocene and Pleistocene (Naish et al. 2009), also contributed to repeated separations of the Indian and Pacific Oceans. However, The high proportion of sister-taxa overlapping only in the IAA (Fig. 4D), observed among pairs that diverged during the Pliocene and Pleistocene (28 and 24%), illustrates the imprint left by the barrier formed by the IAA during the sea level low stands of the Quaternary. Similar patterns have been previously identified (*Naso brevirostris* and *Naso unicornis,* Horne et al. 2008) and marked genetic structures between the Indian and the Pacific populations of numerous widespread species have been described (Horne et al. 2008, Leray et al. 2010, Gaither et al. 2011, Hubert et al. 2012), which highlights that this pattern is likely not the result of only dispersal and extinction processes. On the other hand, a few cases of fast post-speciation range expansions from pattern D (Fig. 4D) to pattern B (Fig. 4B) have been described (Quenouille et al. 2011, Dibattista et al. 2012), highlighting that the influence of the IAA as a barrier to dispersal might have been underestimated in our results.

Speciation and extinction in peripheral regions

In WIP and EIP, the observed patterns are much more heterogeneous. The mean degree of overlap and divergence time between sister-taxa are high across large portions of the Indian and Central Pacific Oceans whereas a visual exploration suggests that the proportion of speciation events observed in the WIP and in the EIP has increased in recent geological times (Miocene: 34%, Pliocene: 37% and Pleistocene: 44%, Fig. 5). This could imply that secondary contact between recently diverged lineages hasn't been reached (due to long post-speciation re-colonization or remaining barriers to dispersal) or that high local extinction hide recently diverged species. Furthermore, sister-taxa that co-occur in those areas show little asymmetry in the size of their geographic ranges, which suggests that mostly widespread sister-taxa co-occur near the boundaries of the Indo-Pacific region, as predicted by 'centre of overlap' and 'centre of origin' hypotheses. On the other hand we observe the imprint of other biogeographic patterns such as peripheral isolation in Hawaii, the Red Sea and the Gulf of Aden, regions that are highly isolated (Hawaii) or have been repeatedly isolated over time (Bosworth et al. 2005, Hall 2009). The common explanation for the Hawaii archipelago is that rare long-distance dispersal events allowed species from the Western Pacific Ocean to colonize and diverge because of reduced gene flow between populations (Craig et al. 2010, Hodge et al. 2014). Concerning the Red Sea, the continental shelf may have isolated peripheral populations during sea level fluctuations that led to evolutionary divergence (Hodge et al. 2014). Old endemics also match with the period when the Red Sea became isolated from the Mediterranean Sea (15 Mya, Hodge et al. 2014). Additionally, we identified patterns corresponding to peripheral isolation (Hodge et al. 2012) in several peripheral locations such as southern Australia (eastern and western), Pitcairn and Easter Islands and southern east Africa. Finally, the region of Madagascar shows co-occurrence and species richness patterns similar to the IAA, which suggests that the same processes may have acted here at a regional scale. Because phylogenetic sampling appears to be lower in those peripheral locations for the families investigated (Cowman 2014), we may also have underestimated the importance of peripheral speciation on reef fish diversification across the Indo-Pacific region. It is also likely that barriers that have not been identified before influenced divergence between species in those peripheral regions (Hodge and Bellwood 2016). For example, we identified 8 pairs in which the sister-taxa are separated by the limits between the Indo-Australian Plate and the Pacific Plate, following a similar phylogeographic separation than corals (Keith et al. 2013).

Potential limitations and perspectives

In addition to range shifts due to dispersion and local extinction, our approach suffers from the same limitations as traditional ARC analysis. First, we approximated the geographic range of ancestral species as the union of the ranges of its descendants. However, the range of a lineage composed by several species is not expected to shift more than the range of a single species of the same age. Hence, the limitations that apply to species range are the same as the ones that apply to ancestral ranges. Second, incomplete phylogenetic sampling and uncertainties in phylogenetic reconstructions could have led us to consider pairs of sister-taxa that do not share a direct common ancestor. The sensitivity analysis revealed that the only index affected by taxon pruning is the divergence time between sister-taxa. This is not unexpected because missing taxa may lead to the overestimation of divergence times (Hodge and Bellwood 2015). For instance, most pairs that we inferred to have diverged during the Miocene belong to genera that are still poorly sampled (phylogenetic sampling < 75%), which could arguably affect our interpretations. The high frequency of widely overlapping distributions observed among pairs that diverged during the Miocene, for example, could either be due to a long history of post-speciation dispersion that led to fully overlapping distributions or to the low phylogenetic sampling of peripheral species, whereas widespread species are well sampled (Cowman 2014). Hence, pairs of sister-taxa from under-sampled genera often include two widespread species. However, remotely distributed sister-taxa may not be sampled yet. That would lead us to underestimate the importance of peripheral isolation.

Finally, this approach does not detect barriers that are still isolating species or adaptive radiations. Yet, such mechanisms have previously been identified in coral reef fishes. For example, the role of the Mid-Indian Ocean Barrier has been recently identified by Hodge and Bellwood (2016) using the distribution of non-overlapping sister-taxa. The identification of recent barriers using this method can be complementary to the use of overlapping ranges between sister-taxa. Likewise,

adaptive radiation has already been identified as an alternative cause of speciation in coral reef fishes. In *Amphiprion*, for instance, adaptive shifts may have been caused by mutualism in sea anemones with parallel processes being observed in the Indian and Pacific Oceans (Litsios et al. 2014). In *Dascyllus* spp., Leray et al. (2010) have highlighted that *Dascyllus albisella* recruits on branching corals due to the paucity of anemones in the Hawaii archipelago while a similar trend is also observed in the easternmost lineages within the *Dascyllus trimaculatus* species complex. Finally, in *Scarus* and *Chlorurus*, the potential causes of a rapid diversification of those genera are still debated, but sexual selection and ecological specialization have been previously proposed as potential mechanisms that promoted adaptive divergence and ecological speciation (Choat et al. 2012).

Conclusions

We showed that geographic concordance of speciation events can be identified using large-scale patterns of co-occurrence between sister-taxa, despite geographic range movements. A detailed analysis of sister-taxa geographic ranges in three families of coral reef fish allowed us to identify the Indo-Australian Archipelago as the main area of secondary contact between sister-taxa in the Indo-Pacific realm. Apparently, this pattern is the result of fragmented habitats that have characterized the IAA since the early Miocene and that promoted divergence between populations of the different seas within the Indo-Australian Archipelago and between populations of the Indian and Pacific Oceans. This analysis, however, highlights that peripheral regions also played an important role in producing and maintaining biodiversity in the Indo-Pacific. More precisely, according to the pattern presented in Fig. 5, we suggest that the three main competing hypotheses explaining current biodiversity patterns across the Indo-Pacific may be nested through time and that complementary historical mechanisms sequentially occurred over different geological periods.

References

Ainsworth, T. D. et al. 2016. Climate change disables coral bleaching protection on the Great Barrier Reef. – Science 352: 338–342.

Anacker, B. L. and Strauss, S. Y. 2014. The geography and ecology of plant speciation: range overlap and niche divergence in sister species. – Proc. R. Soc. B 281: 20132980.

Baliga, V. B. and Law, C. J. 2016. Cleaners among wrasses: phylogenetics and evolutionary patterns of cleaning behavior within Labridae. – Mol. Phylogenet. Evol. 94: 424–435.

Barber, P. H. and Bellwood, D. R. 2005. Biodiversity hotspots: evolutionary origins of biodiversity in wrasses (Halichoeres: Labridae) in the Indo-Pacific and new world tropics. – Mol. Phylogenet. Evol. 35: 235–253.

Barraclough, T. G. and Vogler, A. P. 2000. Detecting the geographical pattern of speciation from species-level phylogenies. – Am. Nat. 155: 419–434.

Bellwood, D. R. and Meyer, C. P. 2009. Searching for heat in a marine biodiversity hotspot. – J. Biogeogr. 36: 569–576.

Borsa, P. et al. 2012. Population genetic structure of blue-spotted maskray *Neotrygon kuhlii* and two other Indo-West Pacific stingray species (Myliobatiformes: Dasyatidae), inferred from size-polymorphic intron markers. – J. Exp. Mar. Biol. Ecol. 438: 32–40.

Bosworth, W. et al. 2005. The Red Sea and Gulf of Aden Basins. – J. Afr. Earth Sci. 43: 334–378.

Briggs, J. C. 2000. Centrifugal speciation and centres of origin. – J. Biogeogr. 27: 1183–1188.

Carpenter, K. E. et al. 2011. Comparative phylogeography of The Coral Triangle and implications for marine management. – J. Mar. Biol. 2011: 1–14.

Choat, J. H. et al. 2012. Patterns and processes in the evolutionary history of parrotfishes (Family Labridae). – Biol. J. Linn. Soc. 107: 529–557.

Cooper, J. W. et al. 2009. Exploring the radiation of a diverse reef fish family: phylogenetics of the damselfishes (Pomacentridae), with new classifications based on molecular analyses of all genera. – Mol. Phylogenet. Evol. 52: 1–16.

Couce, E. et al. 2013. Future habitat suitability for coral reef ecosystems under global warming and ocean acidification. – Global Change Biol. 19: 3592–3606.

Cowman, P. F. 2014. Historical factors that have shaped the evolution of tropical reef fishes: a review of phylogenies, biogeography, and remaining questions. – Front. Genet. 5: 1–15.

Cowman, P. F. and Bellwood, D. R. 2011. Coral reefs as drivers of cladogenesis: expanding coral reefs, cryptic extinction events, and the development of biodiversity hotspots. – J. Evol. Biol. 24: 2543–2562.

Cowman, P. F. and Bellwood, D. R. 2013. The historical biogeography of coral reef fishes: global patterns of origination and dispersal. – J. Biogeogr. 40: 209–224.

Craig, M. T. et al. 2010. Origins, ages and population histories: comparative phylogeography of endemic Hawaiian butterflyfishes (genus *Chaetodon*). – J. Biogeogr. 37: 2125–2136.

Darriba, D. et al. 2012. jModelTest 2: more models, new heuristics and parallel computing. – Nat. Methods 9: 772.

Dibattista, J. D. et al. 2012. Phylogeography of two closely related Indo-Pacific butterflyfishes reveals divergent evolutionary histories and discordant results from mtDNA and microsatellites. – J. Hered. 103: 617–629.

Drummond, A. J. et al. 2012. Bayesian phylogenetics with BEAUti and the BEAST 1.7. – Mol. Biol. Evol. 29: 1969–1973.

Edgar, R. C. 2004. MUSCLE: multiple sequence alignment with high accuracy and high throughput. – Nucleic Acids Res. 32: 1792–7.

Fitzpatrick, B. M. and Turelli, M. 2006. The geography of mammalian speciation: mixed signals from phylogenies and range maps. – Evolution 60: 601–615.

Frédérich, B. et al. 2013. Iterative ecological radiation and convergence during the evolutionary history of damselfishes (Pomacentridae). – Am. Nat. 181: 94–113.

Gaboriau, T. et al. 2017. Data from: Influence of the geography of speciation on current patterns of coral reef fish biodiversity across the Indo-Pacific. – Dryad Digital Repository, < http://dx.doi.org/10.5061/dryad.b0v7h >.

Gaither, M. R. and Rocha, L. A. 2013. Origins of species richness in the Indo-Malay-Philippine biodiversity hotspot: evidence for the centre of overlap hypothesis. – J. Biogeogr. 40: 1638–1648.

Gaither, M. R. et al. 2011. Phylogeography of the reef fish *Cephalopholis argus* (Epinephelidae) indicates Pleistocene isolation across the Indo-Pacific Barrier with contemporary overlap in The Coral Triangle. – BMC Evol. Biol. 11: 189.

Guindon, S. et al. 2010. New algorithms and methods to estimate maximum-likelihood phylogenies: assessing the performance of PhyML 3.0. – Syst. Biol. 59: 307–321.

Hall, R. 2009. Southeast Asia's changing palaeogeography. – Blumea J. Plant Taxon. Plant Geogr. 54: 148–161.

Hodge, J. R. and Bellwood, D. R. 2015. On the relationship between species age and geographical range in reef fishes: are widespread species older than they seem? – Global Ecol. Biogeogr. 24: 495–505.

Hodge, J. R. and Bellwood, D. R. 2016. The geography of speciation in coral reef fishes: the relative importance of biogeographical barriers in separating sister-species. – J. Biogeogr. 43: 1324–1335.

Hodge, J. R. et al. 2012. The role of peripheral endemism in species diversification: evidence from the coral reef fish genus *Anampses* (Family: Labridae). – Mol. Phylogenet. Evol. 62: 653–663.

Hodge, J. R. et al. 2014. Temporal evolution of coral reef fishes: global patterns and disparity in isolated locations. – J. Biogeogr. 41: 2115–2127.

Horne, J. B. et al. 2008. High population connectivity across the Indo-Pacific: congruent lack of phylogeographic structure in three reef fish congeners. – Mol. Phylogenet. Evol. 49: 629–638.

Hubert, N. et al. 2011. Community assembly and diversification in Indo-Pacific coral reef fishes. – Ecol. Evol. 1: 229–277.

Hubert, N. et al. 2012. Cryptic diversity in Indo-Pacific coral-reef fishes revealed by DNA-barcoding provides new support to the centre-of-overlap hypothesis. – PLoS One 7: e28987.

Keith, S. et al. 2013. Faunal breaks and species composition of Indo-Pacific corals: the role of plate tectonics, environment and habitat distribution. – Proc. R. Soc. B 280: 20130818.

Ladd, H. S. 1960. Origin of the Pacific Island molluscan fauna. – Am. J. Sci. 258: 137–150.

Lanfear, R. et al. 2014. Selecting optimal partitioning schemes for phylogenomic datasets. – BMC Evol. Biol. 14: 82.

Lanfear, R. et al. 2017. PartitionFinder 2: new methods for selecting partitioned models of evolution for molecular and morphological phylogenetic analyses. – Mol. Biol. Evol. 34: 772.

Leray, M. et al. 2010. Allopatric divergence and speciation in coral reef fish: the three-spot dascyllus, *Dascyllus trimaculatus*, species complex. – Evolution 64: 1218–1230.

Litsios, G. et al. 2014. The radiation of the clownfishes has two geographical replicates. – J. Biogeogr. 41: 2140–2149.

Lohman, D. J. et al. 2011. Biogeography of the Indo-Australian Archipelago. – Annu. Rev. Ecol. Evol. Syst. 42: 205–226.

Losos, J. B. and Glor, R. E. 2003. Phylogenetic comparative methods and the geography of speciation. – Trends Ecol. Evol. 18: 220–227.

Mora, C. et al. 2003. Patterns and processes in reef fish diversity. – Nature 421: 933–936.

Naish, T. et al. 2009. Obliquity-paced Pliocene west Antarctic ice sheet oscillations. – Nature 458: 322–328.

Parravicini, V. et al. 2013. Global patterns and predictors of tropical reef fish species richness. – Ecography 36: 1254–1262.

Parravicini, V. et al. 2014. Global mismatch between species richness and vulnerability of reef fish assemblages. – Ecol. Lett. 17: 1101–1110.

Pellissier, L. et al. 2014. Quaternary coral reef refugia preserved fish diversity. – Science 344: 1016–1019.

Puckridge, M. et al. 2015. The role of peripheral endemism and habitat associations in the evolution of the Indo-West Pacific tuskfishes (Labridae: Choerodon). – Mol. Phylogenet. Evol. 84: 64–72.

Quenouille, B. et al. 2004. Molecular systematics of the damselfishes (Teleostei: Pomacentridae): Bayesian phylogenetic analyses of mitochondrial and nuclear DNA sequences. – Mol. Phylogenet. Evol. 31: 66–88.

Quenouille, B. et al. 2011. Speciation in tropical seas: allopatry followed by range change. – Mol. Phylogenet. Evol. 58: 546–552.

Rabosky, D. L. et al. 2013. Rates of speciation and morphological evolution are correlated across the largest vertebrate radiation. – Nat. Commun. 4: 1958.

Randall, J. E. 1997. Randall's tank photos. Collection of 10000 large format photos (slides) of dead fishes. – < www.fishbase.org/Refences/FBRefSummary.php?ID=286187 >.

Ribera, I. et al. 2011. The geography of speciation in narrow-range endemics of the "Haenydra" lineage (Coleoptera, Hydraenidae, Hydraena). – J. Biogeogr. 38: 502–516.

Sanderson, M. J. 2003. Evolution and divergence times in the absence of a molecular clock. – Bioinformatics 19: 301–302.

Silvestro, D. and Michalak, I. 2011. raxmlGUI: a graphical front-end for RAxML. – Org. Divers. Evol. 12: 335–337.

Stamatakis, A. 2014. RAxML version 8: a tool for phylogenetic analysis and post-analysis of large phylogenies. – Bioinformatics 30: 1312–3.

Tamura, K. et al. 2013. MEGA6: Molecular evolutionary genetics analysis version 6.0. – Mol. Biol. Evol. 30: 2725–2729.

Thomson, R. C. and Shaffer, H. B. 2010. Sparse supermatrices for phylogenetic inference: taxonomy, alignment, rogue taxa, and the phylogeny of living turtles. – Syst. Biol. 59: 42–58.

Timm, J. et al. 2008. Contrasting patterns in species boundaries and evolution of anemonefishes (Amphiprioninae, Pomacentridae) in the centre of marine biodiversity. – Mol. Phylogenet. Evol. 49: 268–276.

Tornabene, L. et al. 2014. Support for a "Center of Origin" in the Coral Triangle: cryptic diversity, recent speciation, and local endemism in a diverse lineage of reef fishes (Gobiidae: Eviota). – Mol. Phylogenet. Evol. 82: 200–210.

Vaidya, G. et al. 2011. SeqenceMatrix: cladistics multi-gene datasets with character set and codon information. – Cladistics 27: 171–180.

Veron, J. E. N. et al. 2009. Delineating the Coral Triangle. – Galaxea J. Coral Reef Stud. 11: 91–100.

Williams, S. T. and Duda, T. F. 2008. Did tectonic activity stimulate Oligo-Miocene speciation in the Indo-West Pacific? – Evolution 62: 1618–1634.

Woodland, D. J. 1983. Zoogeography of the Siganidae (Pisces): an interpretation of distribution and richness patterns. – Bull. Mar. Sci. 33: 713–717.