CHAPTER 4

Biogeography of Butterflyfishes: A Global Model for Reef Fishes?

Is this correct?

Michel Kulbicki,[1] Laurent Vigliola,[2] Laurent Wantiez,[3] Nicolas Hubert,[4] Sergio R. Floeter[5] and Robert F. Myers[6]

Corresponding author?

INTRODUCTION

The biogeography of coral reef fishes is a complex issue still poorly understood largely because i) the knowledge of geographic ranges of many species is still far from complete; and ii) identification remains problematic for several taxa, especially within some families such as the Muraenidae, Ophichtidae, Apogonidae, Gobiidae and Trypterygiidae. There have been a number of attempts to understand the regional distribution of coral reef

---

[1] Centre IRD (Laboratoire Arago), BP 44-66651, Banyuls, France.
E-mail: michel.kulbicki@ird.fr
[2] Centre IRD (Nouméa), 101 Promenade Roger Laroque, Anse Vata, BP A5-98848 Noumea, New Caledonia.
E-mail: laurent.vigliola@ird.fr
[3] Université de la Nouvelle-Calédonie, BP R4-98851 Nouméa cedex, New Caledonia.
E-mail: laurent.wantiez@univ-nc.nc
[4] Université Montpellier II, Case courrier 63, Institut des Sciences de l'Evolution, Dpt Conservation et Domestication, 34095 Montpellier.
E-mail: nicolas.hubert@ird.fr
[5] Departamento de Ecologia e Zoologia, Universidade Federal de Santa Catarina, Florianópolis, Santa Catarina 88010-970, Brazil.
E-mail: floeter@ccb.ufsc.br
[6] Seaclicks/CoralGraphics, 9273 Silent Oak Circle, Wellington FL 33411-6634 United States.
E-mail: robmyers1423@gmail.com

fishes, in the Atlantic (e.g., Floeter et al., 2008), in the Indo-west Pacific (e.g., Springer, 1982; Bellwood and Hughes, 2001; Santini and Winterbottom, 2002; Allen, 2000, 2007) and in the Tropical Eastern Pacific (e.g., Robertson and Cramer, 2009), however there have been very few attempts to understand such patterns on a worldwide scale. A notable exception is Roberts et al. (2002), which investigated the distribution of 1700 species of coral reef fishes worldwide to compare the overlap between diversity hotspots and species geographic range. At present there is little general understanding of the distribution of coral reef fishes, except for the major patterns of diversity and endemism, and on the factors that may lead to this distribution at a global scale.

Coral reef fishes (abbreviated hereafter as "reef fish") are of particular interest for a worldwide biogeography analysis because they are distributed according to a number of large-scale gradients that can be related to several major hypotheses in ecology (see Whittaker et al., 2001 for a review of hypotheses and Ricklefs, 2004 for a comprehensive framework). Among the numerous gradients along which reef fish diversity varies, five are of particular interest: latitude, distance to biodiversity centres, island size, connectivity (specifically, the degree of exchange between populations) and large-scale coral diversity. In addition reef fishes are diverse and display a wide range of life-history traits. Such properties make these fishes ideal candidates for biogeographical and macro-ecological studies (see reviews by Kent, 2005, 2007). Coral reef fishes are by definition closely associated with coral reefs, which are in turn restricted to warm (the lower temperature limit for scleractinian corals is 17°C) oligotrophic waters. Water temperature varies substantially/markedly with latitude and depth or next to specific features such as cold currents or upwellings. Latitudinal gradients in species diversity are well established for both terrestrial and aquatic systems (Willig et al., 2003; Vázquez and Stevens, 2004), with these patterns being related to variation in temperature, all other factors being equal (Gaston, 2000). Therefore fish diversity should be highest where water temperature is high, which is usually associated with low latitudes and shallow waters. The second major gradient, at least in the Indo-Pacific, is a decrease in diversity with distance from the centre of biodiversity, represented by the Indo-Australian Archipelago (Bellwood and Hughes, 2001; Carpenter and Springer, 2005; Allen, 2007) and commonly referred to as the "Coral Triangle". This diversity centre represents an area of species accumulation more than speciation (Santini and Winterbottom, 2002; but see Mora et al., 2003 for another view).

Island size and connectivity also have important implications for reef fish biogeography. Island biogeography theory (MacArthur and Wilson, 1967) predicts that the number of habitats is a function of island size and as a consequence the number of species will increase with island size. Similarly, isolated islands will shelter less species than relatively connected ones (e.g., Moilanen and Nieminen, 2002) due to the lower rates of exchange amongst populations. Corals are the foundation species (or ecosystem engineers) of coral reefs and many reef fishes are closely associated with corals for food and/or shelter. Therefore the diversity of reef fishes should be highest where the diversity and/or cover of corals is high. In support of this Bellwood and Hughes (2001) demonstrated a correlation between reef fish diversity and coral diversity across the Indo-Pacific, and Roberts et al. (2002) reported similar relationships across a global scale. This raises the question of whether this congruence is the result of a co-evolution of the two groups or is just the result of a stochastic process (Gaston, 2000).

Butterflyfishes (family Chaetodontidae) represent an almost ideal model for the analysis of the major biogeographical trends of reef fishes: there are few identification problems with this family, their diversity is sufficient in most regions to enable statistical comparisons, the geographic range of most species is well known and there is information on the biology and phylogeny for a large proportion of species. At present there is no other coral reef fish family possessing all these advantages. Additionally, there is a good correlation between *Chaetodon* diversity and reef fish diversity in the Pacific (Kulbicki et al., 2005). Nonetheless, the Chaetodontidae are imperfect as a model group because the family is not represented by all possible reef fish life-history traits. In particular they have a limited range of maximum sizes, their diets are not very diversified compared to the other reef fishes (e.g., Labridae: Bellwood et al., 2006), and their range of social behaviours and levels of endemism are usually low. In the present chapter we will describe/quantify the worldwide distribution of the Chaetodontidae, and evaluate some of the major hypotheses underlying the distribution of these fishes and, by extension, that of coral reef fishes in general. We will first define a restricted number of biogeographical regions based on the species composition of Chaetodontidae in a large number of localities across the world. Once these regions are defined we will test the effects of large-scale factors (i.e., latitude, island size, isolation, distance to biodiversity centres, and large scale coral diversity) and life-history traits (i.e., diet, body size and behaviour) on the diversity (alpha and beta) of the Chaetodontidae. Here, we define alpha diversity as the number of species found within a single 'site'. Therefore the area covered by this diversity may vary from a small island to large continental expanses. Beta diversity is a measure of the diversity gradient between alpha diversities from various sites, in other words by how much does alpha diversity vary from one location to the next within a region. In particular, alpha diversity will be analysed as both "absolute" alpha diversity (total number of Chaetodontidae species per site), and "relative" alpha diversity (proportion of Chaetodontidae amongst reef fish species per site). Two types of beta diversities will also be considered

??

(Magguran, 2004), one independent of species composition (Whittaker, 1972) and another one taking into account the variations in species composition amongst regions (Legendre et al., 2005). Finally, the available information on the phylogeny of butterflyfishes will be used to try to understand some of the patterns found in the previous analyses.

BIOGEOGRAPHICAL REGIONS AND DIVERSITY PATTERNS

Worldwide diversity patterns of butterflyfishes were studied from species presence-absence datasets gathered from both the literature and our own field surveys at 175 sites spanning all of the world's oceans (Appendix 1, Fig. 4.1). Sites for which we could gather data had relatively even spatial distribution with a butterflyfish species checklist available at every ~1000 km or less, subsequently more sites were censused in the larger Pacific Ocean (108 species lists) than in the smaller Indian and Atlantic Oceans (35 lists each). Although their distribution was relatively even in space, our sampling units (i.e., sites) were of very different size, ranging from large regions (e.g., North-West Australia) to individual reefs (e.g., Tulear reef in Madagascar).

We defined a biogeographic region as a geographically continuous area inhabited by a relatively homogeneous fauna. From this definition, it results that the search for the boundaries of a biogeographic region simply consists of clustering ecologically similar sites that are part of the same continuous geographic zone. Cluster analysis requires two algorithms, a first one to measure the ecological similarity (or distance) between sites and a second one to decide when sites are similar enough to become members of the same cluster. The Raup and Crick (1979) probabilistic similarity coefficient (S_{27} in Legendre and Legendre, 1998) was selected to estimate distances (specifically, Raup and Crick's probabilistic coefficient of similarity in species composition) and Ward's (1963) minimum variance method was chosen to group sites. To ensure that ecologically similar sites were part of the same geographically continuous region and to minimise the effects of different sampling unit size and spatial overlap in our dataset, Ward's algorithm was only used as a basis for a step-by-step manual clustering procedure. First, the Raup and Crick's distance matrix among all 175 sites was calculated and a consensus tree build based on 999 bootstraps trees (+ the original tree) and the Ward's algorithm. Second, sites that were most ecologically similar (first rank clusters), geographically close (continuous region), and with high consensus indices (strong clusters) were merged and a new presence-absence dataset with fewer but larger sites was built. From this new presence-absence dataset a new consensus tree was generated, sites were merged into another new presence-absence dataset, and the procedure was repeated until all sites formed a single cluster. Once the

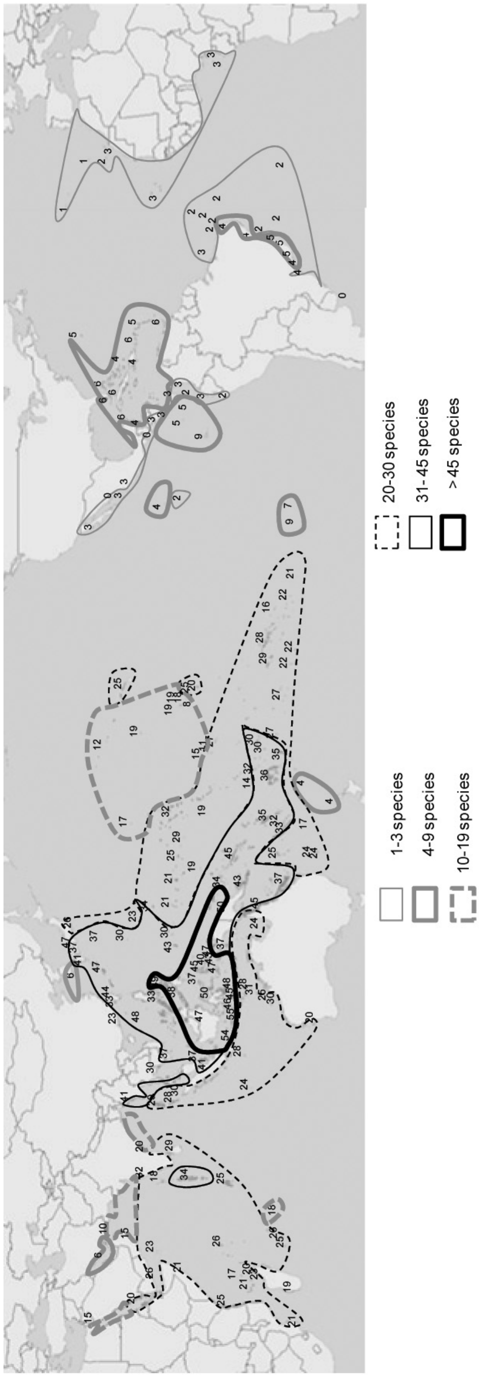

Fig. 4.1 World wide map of the alpha diversity distribution of butterflyfishes.

Colour image of this figure appears in the colour plate section at the end of the book.

biogeographic regions were defined, these were represented in a Cartesian coordinate system by multidimensional scaling (MDS), to represent distance relationships (Raup and Crick's distance) among the biogeographic regions (Legendre and Legendre, 1998).

A total of 134 species belonging to 11 genera were recorded in our dataset (Appendix II). Three undescribed species were not included in the analysis as both their status and geographical distribution are uncertain. Taxonomic richness varied across localities from a low of one genus and one species at Madeira and the Azores in the eastern Atlantic to a maximum of eight genera and 55 species at Bali, Indonesia (Fig. 4.1). Both species and genus richness were highest in the Indo-Australian Archipelago, with 6–8 genera and over 40 species at each of the Banda, Celebes, Java, Sulu, Philippines, China, Arafura, Timor, Bismark, Solomon, and Coral Sea sites. The existence of a biodiversity centre for coral reef organisms (including fishes) centred around the Indo-Australian archipelago has been well described (see Carpenter and Springer, 2005; Bellwood et al., 2005; Allen, 2007), however there is some debate over the precise boundaries of this centre. Taxonomic richness decreased both eastward and westward of this Indo-Pacific biodiversity centre, as has been described previously for many reef organisms, in particular corals and fishes (e.g., Bellwood and Hughes, 2001; Bellwood et al., 2005). Diversity was much lower in the eastern Pacific and the western Atlantic, with only 2–3 genera and less than 5–6 species at each site, and lowest in the eastern Atlantic with only 2–3 species belonging to 1–2 genera.

The cluster analysis revealed 19 biogeographic regions (Fig. 4.2) and 39 sub-regions (Appendix I). In general there is little overlap between the boundaries of these regions and the diversity isopleths in Fig. 4.1. There were nine regions in the Pacific, four in the Indian, five in the Atlantic, and one across the Indian and Pacific Oceans. Regions in the Indo-Pacific were generally inhabited by more homogeneous butterflyfish faunas than those in the Atlantic, as indicated by differences in the within-region Ward's minimum variance distance. As will be discussed later, this can be in part related to the very low levels of endemism across most of the Indo-Pacific, whereas in the Atlantic endemism is much higher. Nevertheless, heterogeneous regions also existed in the Indo-Pacific, notably in the eastern Pacific, remote Hawaii and sub-tropical New Zealand regions, where endemism levels may also be high.

see p. 184

In the Pacific, the three most speciose regions were those around Indonesia (Indonesia, Philippines, Indochina, Peninsula), China-Japan (China, Korea, Japan), and Melanesia (Papua New Guinea, Great Barrier Reef, Melanesia). The Indonesian and Melanesian regions were the most homogeneous as shown by a very small ($\sim 10^{-13}$) within region Ward's minimum variance distance (Fig. 4.2). In comparison, the China-Japan

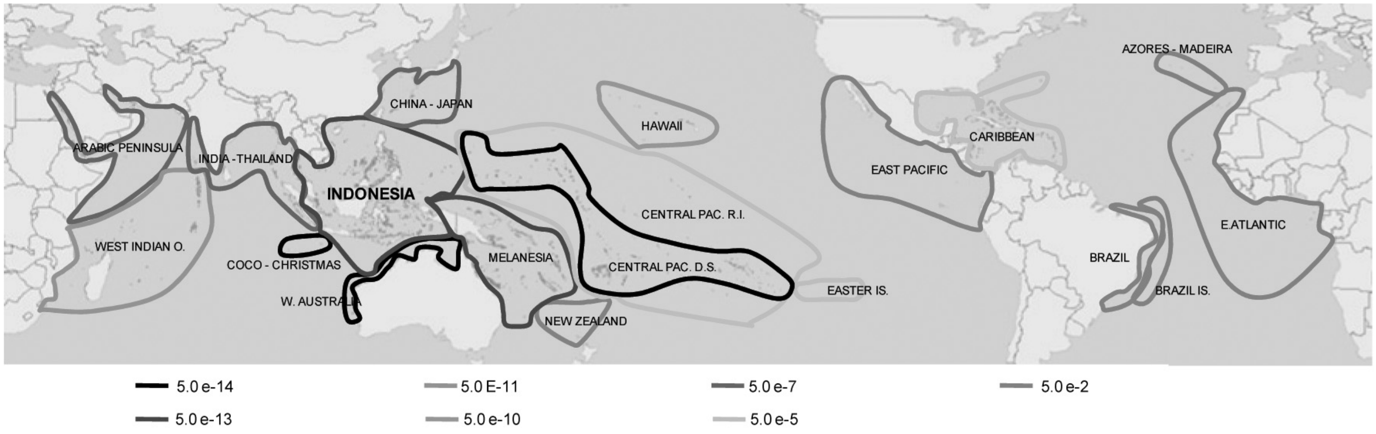

Fig. 4.2 Map of the 19 biogeographic regions determined by stepwise clustering of 175 butterflyfish species lists around the tropics. The lines indicate Ward's minimum variance distance levels.

Colour image of this figure appears in the colour plate section at the end of the book.

region was relatively heterogeneous with a clustering distance of ~10^{-7}. This region also presents the highest beta diversity of all regions, which indicates a strong variation in species richness and composition from one site to another. East of these three speciose regions lies a very large and relatively heterogeneous (~10^{-5}) region that encompasses most of the central Pacific. However, stepwise cluster analysis revealed that this large central Pacific region was formed by a very homogeneous central stripe (~10^{-14}) surrounded by a rather heterogeneous belt (~10^{-5}). The stripe was named "Central Pacific Dorsal Spine" and follows a corridor formed by the NW-SE alignment of the Mariana, central Caroline, Marshall, Gilbert, Ellice (Tuvalu), Wallis, Samoa, Fiji, Tonga, Society, Tuamotu, and Gambier Islands. Most of these islands are atolls (except the Society Islands and some of the Cook Islands), are of similar size, and have a similar coral diversity and seawater temperature regime. The belt was named "Central Pacific Remote Islands" as it was formed by a number of very scattered and geographically isolated islands on both sides of the stripe. The heterogeneity found in this "belt" may be related to the very low connectivity of these islands and also to their physical characteristics, e.g., the presence of high islands, such as the Marquesas or Rapa Islands, and atolls, such as the Line Islands. Three distinct and relatively isolated regions lie at the border of the central Pacific region, the Hawaiian region at the extreme north, the New Zealand region at the extreme south, and the Easter Island region at the extreme east. These regions are characterised by several endemic species and are at the limit of the habitat range of butterflyfishes. Separated from other Pacific regions by over 3000 km of deep waters and complex current systems, the East Pacific biogeographic region includes all sites of the western American coasts and islands. In particular, there is very little species overlap with the central and west Pacific; more than 50% of the species are endemic to the East Pacific region.

The Indian Ocean included four regions. A small and very homogeneous (~10^{-14}) "Coco-Christmas" region was formed by the Cocos-Keeling and Christmas Islands on the western border of the Indonesian Pacific region. The fauna of this region is intermediate between typical Indian Ocean fauna and Pacific fauna. Similarly, the fauna of the western Australian region contained a higher proportion of species centred on the Pacific than on the Indian Ocean. In this respect, islands off the western Australian coast (Rowley Shoals, Ashmore Reef, and Scott Reef) had a fauna that included them in the Indonesian region rather than into the western Australian region despite a closer distance to the latter. Westwards, India, the Laccadive Sea, the Bay of Bengal, the Andaman Sea, and the west coasts of Malaysia and Sumatra formed a relatively homogeneous (~10^{-7}) "India-Thailand" region. The Lakshadweep and western Indian coast share many similarities with the Maldives, but their butterflyfish fauna is less diverse. Despite the natural barrier formed by the freshwater inputs in the Bay of Bengal the similarity on both sides of this barrier remains high. This is in opposition with the other freshwater barrier formed by the Indus Delta, which creates an eastern boundary for the region grouping the Red Sea, the Persian Gulf and the western Arabian Sea. This "Arabian Peninsula" region ($\sim 10^{-7}$) is rather heterogeneous due to numerous endemic species and to the presence of an upwelling site in Oman that is believed to separate the Persian Gulf reef fish fauna from the Red Sea and Yemen (Randall, 1995; Kemp, 1998). In contrast, the southeast African coast, Madagascar, and the central Indian islands (up to the Maldives) formed a very homogeneous ($\sim 10^{-11}$) "Western Indian Ocean" region. This region could be a "sink", or accumulation centre, as species may either originate from the Arabian Peninsula or the India-Thailand regions. This is in agreement with Connolly et al. (2003) who indicated that reef fish diversity along the eastern African coast was above predicted levels. It should be noted that *Chaetodonmarleyi*, which is endemic to the "Western Indian Ocean", is related to the eastern Atlantic species *Chaetodonrobustus*, thus establishing a link between these oceans.

The most speciose and homogeneous ($\sim 10^{-5}$) region of the Atlantic, the "Caribbean" region, included the Caribbean Sea and the Gulf of Mexico. There were two other regions along the Atlantic coast of the Americas. The "Brazil" region included all sites along the coast of Brazil whereas the "Brazilian Islands" included remote islands offshore of South America. On the east side of the Atlantic, a large "East Atlantic" region included the western African coast and its offshore islands whereas a smaller "Azores-Madeira" region included the Azores and Madeira Islands. However for the latter two, only vagrant fish have been observed and no permanent population is known.

The consensus tree of the 19 biogeographic regions indicated highly dissimilar butterflyfish fauna in the Indo-Pacific and the Atlantic Oceans (Fig. 4.3). Within the Atlantic, the Caribbean and Brazil clustered first, quickly joined by the Brazilian islands, then by Azores-Madeira, and finally by the East Atlantic region that appeared as a relatively distinct region. Floeter et al. (2008) performed a biogeographical analysis of the entire Atlantic reef fish fauna. The regions they defined are broadly similar to those found using only butterflyfish, but with some differences. For example, the Brazilian coast and the islands off Brazil were clustered into a unique region and not two regions as in the present analysis. This could arise in our analysis due to the presence of an endemic species (*Prognathodesobliquus*) and lower diversity in the oceanic islands. Another difference is the closer distance of Azores-Madeira to the Caribbean and Brazilian regions than to the other eastern Atlantic sites in the present study. This is presumably because of the presence of a single vagrant species from the western Atlantic in the Azores-Madeira region that may have been transported by

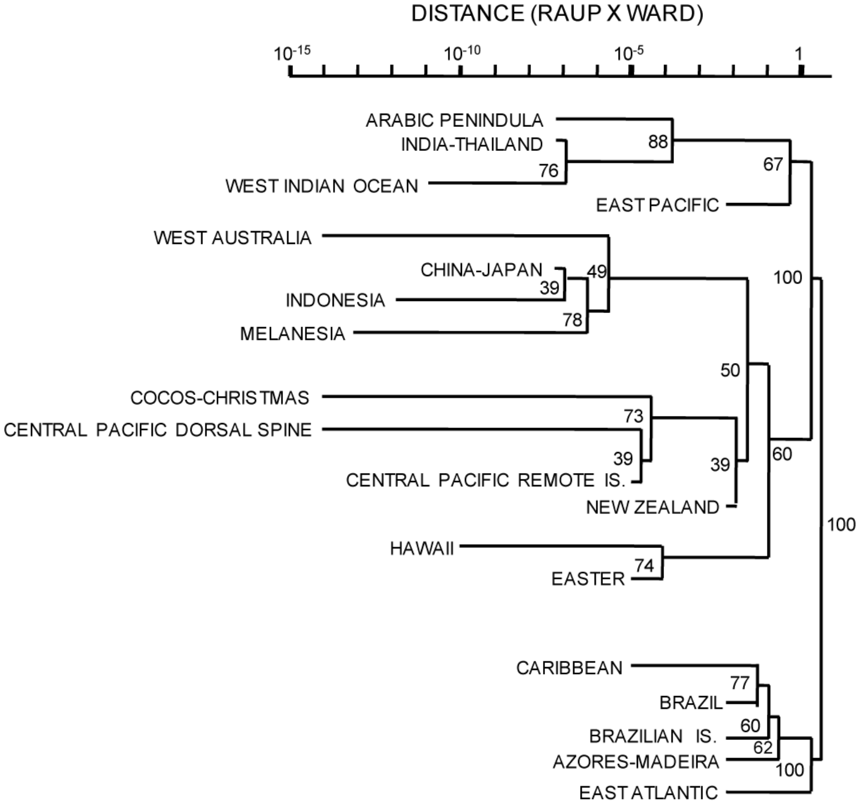

Fig. 4.3 Consensus tree of the 19 biogeographic regions. Values at nodes represent bootstrap support (% of 1000 replicates).

the Gulf Stream. Conversely, Floeter et al. (2008) separated the Ascension and St Helena islands in the tropical eastern Atlantic could be separated from the eastern Atlantic, but our analysis did not distinguish these regions due to the presence of 2 endemic species. The clustering of these islands with the remaining eastern Atlantic is due to one of these species, *Chaetodonsanctaehelenae*, which has been reported from the Canaries (Brito and Falcon, 1996).

Within the Indo-Pacific, regions of the Indian Ocean clustered together, with the exception of the Cocos-Christmas region that clustered with the central Pacific regions. Similarly, regions of the Pacific Ocean clustered together, with the exception of the eastern Pacific region, which joined the Indian Ocean cluster. Many analyses have investigated the definition of biogeographic regions of reef fishes or have mapped the distribution of reef fish diversity in the Indo-Pacific (e.g., Springer, 1982; Kemp, 1998; Mora et al., 2003; Santini and Winterbottom, 2002; Allen, 2007; Kulbicki, 2007).

Springer (1982) was the first to suggest the importance of the tectonic plate margins in the distribution of reef fishes in the Indo-Pacific, in particular the very sharp drop in diversity at the western margin of the Pacific Plate. Subsequently, the first definition of major biogeographic regions for reef fishes was proposed that matched the limits of these plates. Following on from this work Thresher (1991) attempted a worldwide classification of 21 reef fish sites based on species composition. Indo-Pacific sites did not cluster according to their geographic affinities but more according to their physical environments, with island size, island type, and connectivity seemingly important factors. Santini and Winterbottom (2002) arbitrarily defined 18 regions in the Indo-Pacific, mainly based on geological features, but their study did not encompass the central Pacific (i.e., Polynesia, Hawaii, and the Line Islands). More recently, Kulbicki (2007) examined the reef fish fauna of 65 sites in the west and central Pacific and identified nine biogeographical regions that displayed a remarkable overlap with geological features such as tectonic plate margins.

In the present study, we found strong congruence with the previous classifications of biogeographic regions for both corals and reef fishes. There was, however, some variation in the relatedness of the various regions. For example, in the Indian Ocean, India-Thailand and the West Indian Ocean regions were most ecologically similar and were later joined by the Arabian Peninsula region. In contrast, Santini and Winterbottom (2002) found that the Arabian Peninsula first clustered with eastern Africa, then the Mascarene and Seychelles islands, and only after that clustered with India. Similarly, Kemp (1998) found on a smaller scale that the Socotra fauna (in the Gulf of Aden) was linked first to the Red Sea, then Somalia, and then to western India. Veron (2000) found that corals of the Arabian Peninsula clustered with the Mascarene and the eastern Indian Ocean regions on a similar level, which is similar to the present study on butterflyfish.

Multi-dimensional scaling of the 19 biogeographic regions further outlined existing gradients in butterflyfish biodiversity across the world (Fig. 4.4). The analysis revealed four independent provinces, the Atlantic, East Atlantic, Indo-Pacific, and East Pacific. The presence of two centres of diversity, one in Indonesia and the other in the Caribbean, raises the question of the importance of these centres in the distribution of species among regions, and the relationships between regions (Briggs, 2005). These centres appear to neither be centres of endemism (Hughes et al., 2002; Allen, 2007; but different opinions expressed by Roberts et al., 2002, Mora et al., 2003) nor centres of origin (Santini and Winterbottom, 2002) and the factors contributing to these peaks in diversity are complex as will be discussed below.

In the Atlantic the total number of species within the Chaetodontidae is low (15 species) and it is therefore difficult to draw strong conclusions.

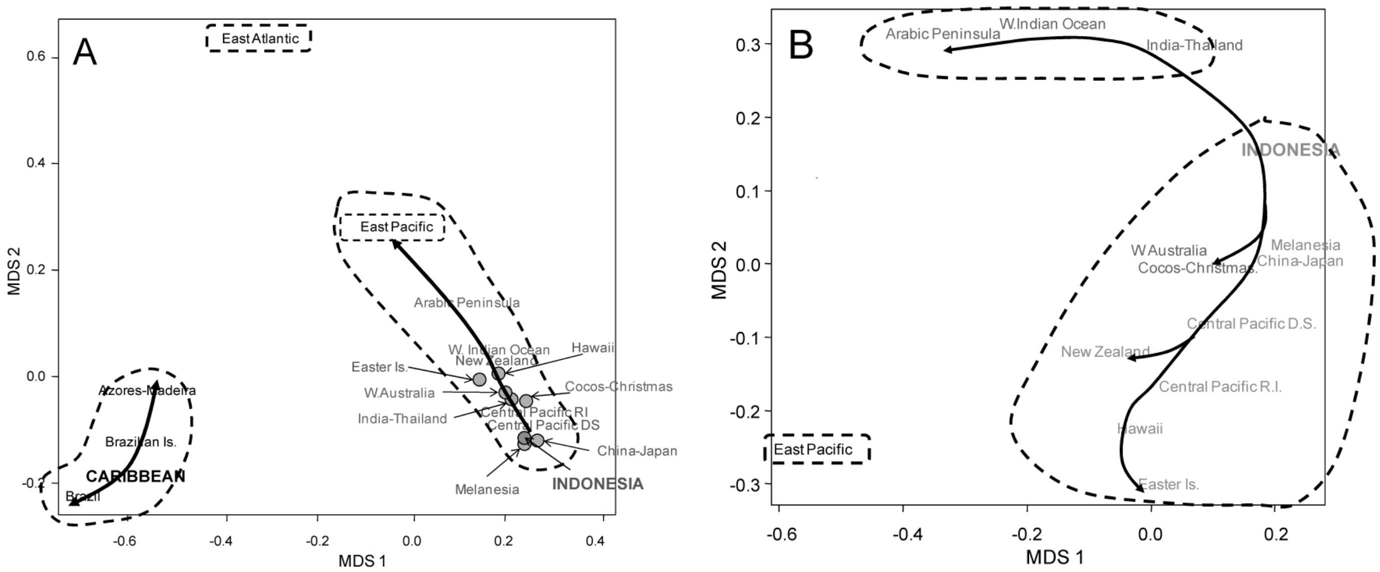

Fig. 4.4 Multidimensional scaling **(A)** all 19 biogeographic regions, **(B)** only Indo-Pacific regions. Regions from the Indian Ocean are in red; regions from the west and central Pacific are in blue. MDS1 and MDS2 are axes of the multidimensional scaling. The arrows indicate the shortest pathways from one region to another.

Colour image of this figure appears in the colour plate section at the end of the book.

However the eastern Atlantic has a unique fauna, its five species being not found anywhere else in the Atlantic, which explains its isolated position on Fig. 4.4A. Brazil has the most species in common with the Caribbean, and while there is a continental continuum between these two regions the Amazon and Orinoco Rivers' fresh water flows represent barriers to species dispersal. Recently however, it has been proposed that some species are able to use deep-water reefs beneath these river plumes as stepping-stones between the two regions (Rocha, 2003). More surprising is the paucity of butterflyfish species on the islands off Brazil. These islands are small and isolated which could in part explain this paucity but other factors such as current patterns and coral diversity may have also contributed to the observed patterns.

In the Indo-Pacific (including East Pacific) two features are remarkable (Fig. 4.4). First, the Indian Ocean and west Pacific faunas show a high degree of similarity with Indonesia and this similarity decreases regularly in both oceans with increasing distance from Indonesia. Second, the eastern Pacific shares very little in common with the rest of the Indo-Pacific, reflecting its isolated position (Fig. 4.4B). In the Indo-Pacific, ecological similarity to Indonesia may decrease with geographic distance for different reasons in the two oceans. In the Pacific, the lowest similarity is found with small and isolated island systems (Easter Island, Hawaii) whereas in the Indian Ocean the Arabian Peninsula, which presents the lowest similarity with Indonesia, has a continental continuum with Indonesia (Fig. 4.4B). This continuum is however disturbed by three physical barriers (two river systems and one upwelling) that have generated differences in similarity higher than those created by isolation or island size in that ocean.

The peculiar situation of the East Pacific region as an independent region located at the extreme east of the eastward Pacific gradient implies that this relatively independent region has recently started to link with the rest of the Pacific. Detailed data analysis indicates that the stepping stone islands are probably the Galapagos as all seven butterflyfish species found in the East Pacific are present in the Galapagos whereas there are two species in the Galapagos that are present in the Indo-Pacific but absent in the East-Pacific.

Alpha Diversity

"Absolute" alpha (α) diversity

The number of butterflyfish species at a site varies from one (*Chaetodonsedentarius,* a vagrant species in the islands of Azores in the Atlantic) to 55 species (Bali–Indonesia) with strong differences between and within regions (Fig. 4.1). In particular there are far less species in the Atlantic and East Pacific than anywhere in the Indo-Pacific. A number of studies on reef fishes indicate very similar geographical patterns with the distribution of butterflyfishes (e.g., Randall, 1998; Roberts et al., 2002; Bellwood et al., 2005; Allen, 2000, 2007). One notices also the very abrupt gradient at the limits of the geographic range of butterflyfishes in many areas of the Indo-Pacific; in some cases the number of species drops from over 20 species to 0 with little or no transition. Within the Indo-Pacific diversity decreases as the distance from the Indo-Australian archipelago increases, with these decreases coinciding with increasing latitude, decreasing island size, and increase disolation. Similar patterns have been shown for scleractinian corals (Veron, 2000) and given the high dependence of many butterflyfishes on corals for food and/or habitat (See Cole and Pratchett, Chapter 5) links may exist between butterflyfish and coral diversity on a large scale. Here the joint effect of five large-scale factors (i.e., latitude, distance to biodiversity centres, island size, connectivity, and regional coral diversity) on the distribution of α diversity of butterflyfishes was analysed on a worldwide scale using GLMs. Biodiversity centres were defined for each ocean as the areas where the coral reef fish diversity is the highest: the Coral Triangle for the Indo-Pacific (Carpenter and Springer, 2005; Allen, 2007), the southern Caribbean for the Atlantic (Briggs, 2005; Floeter et al., 2008), the Panama isthmus for the Eastern Pacific (Mora and Robertson, 2005; Robertson and Cramer, 2009). Connectivity was defined on a semi-quantitative scale: low (islands more than 400 km from the nearest island), medium (200–400 km), high (100–200 km), very high (< 100 km). Regional scale coral diversity was extracted from the maps published by Veron (2000).

Regional coral diversity was the factor that best explained the distribution of butterflyfishes (Table 4.1). With the exception of the East Pacific, the number of Chaetodontidae species increased significantly as the regional diversity of coral increased (Fig. 4.5). Bellwood and Hughes (2001) found a significant correlation between coral diversity and reef fish diversity (represented by 13 families) in the Indo-Pacific (including the East Pacific; $r = 0.65$, $N = 113$ sites, $p < 0.001$). This correlation proved even stronger using our worldwide butterflyfish dataset (non-parametric Spearman rank order correlation $r = 0.87$, $N = 178$ sites, $p < 0.001$). However to our knowledge, no one has demonstrated if this result is a result of similar environmental requirements for both butterflyfishes (or reef fishes in general) and corals, or if there is co-evolution of the two groups. If these congruent patterns were a result of co-evolution we may expect the proportion of obligate coral-feeding butterflyfishes to be highest where coral diversity is highest (assuming there is a correlation between coral diversity and coral abundance). However, the proportion of both coral feeding species and coral specialists tended to be constant (slope of regression not significantly different from 0) as coral

Table 4.1 Correlation between butterflyfish diversity and large-scale factors. N: number of sites; NS: not significant; *: $p < 0.05$; **: $p < 0.01$; ***: $p < 0.001$. The numbers represent the correlation coefficient of a single regression, except in the last column. In the last column the numbers represent the total r^2 for GLMs involving all the previous factors. Factors are symbolised by C: connectivity; D: distance to biodiversity centres; L: latitude. Factors are ranked by increasing importance.

	Island Size	**Latitude**	**Connectivity**	**Distance to Biodiversity Centres**	**Regional Coral Diversity**	**Multiple regression R^2; factors**
Atlantic (N=34)	0.19 NS	0.17 NS	0.57***	–0.68***	0.70***	0.62; D
East Pacific (N=15)	–0.36 NS	–0.21 NS	–0.61*	–0.13 NS	0.22 NS	0.33; C
Indian (N=31)	–0.21 NS	–0.53**	0.01 NS	–0.61***	0.58***	0.41; D, L
Pacific (N= 91)	0.34**	–0.27*	0.69***	–0.68***	0.78***	0.52; C, D

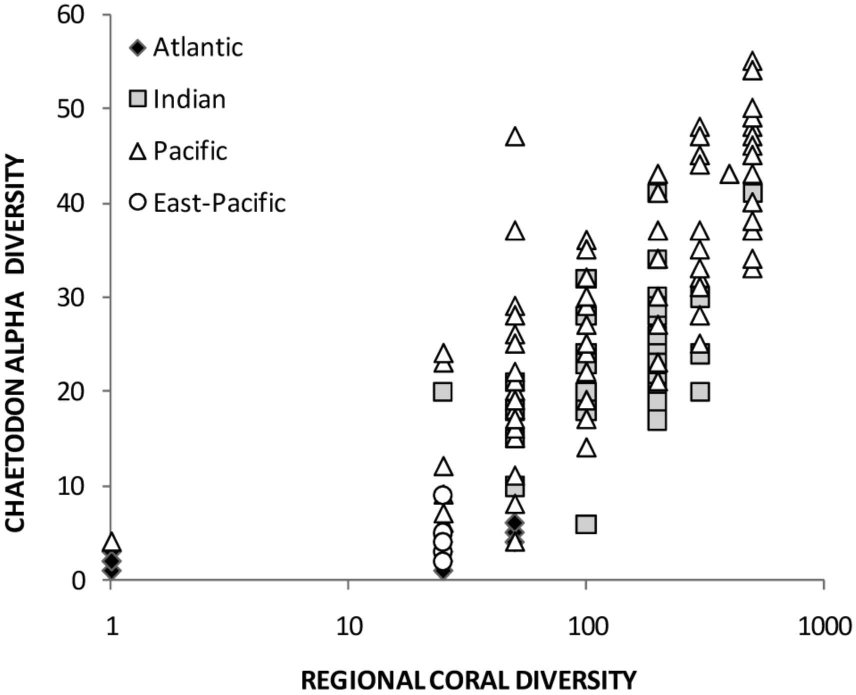

Fig. 4.5 Relationship between the alpha diversity of Chaetodontidae and regional coral diversity.

diversity increased (Fig. 4.6). While not conclusive, this suggests that these congruent patterns are likely to reflect similar environmental requirements as opposed to co-evolution.

The four other large-scale factors tested had varying effects on butterflyfish diversity among the different ocean basins (Table 4.1). The most consistent factor was the distance to the biodiversity centre with a significant decrease in diversity as this distance increases. This occurred in all oceans except the East Pacific. Such an effect has been demonstrated for both reef fish and corals in the Indo-Pacific, including the East Pacific (Bellwood and Hughes, 2001). Connolly et al. (2003) investigated if such patterns could be attributed to the "mid-domain" effect (MDE), which predicts a peak or a plateau in species richness near the centre of ranges within a bounded domain (reviewed by Zapata et al., 2003; Colwell et al., 2004; Pimm and Brown, 2004). They found that the diversity of reef fish and corals was higher along the eastern coast of Africa and in the Indo-Australian archipelago and lower in the Pacific, (especially the East Pacific) than predicted by the MDE. Their major explanation was a link with oceanic circulation. Bellwood et al. (2005) extended this to investigate the relative importance of MDE,

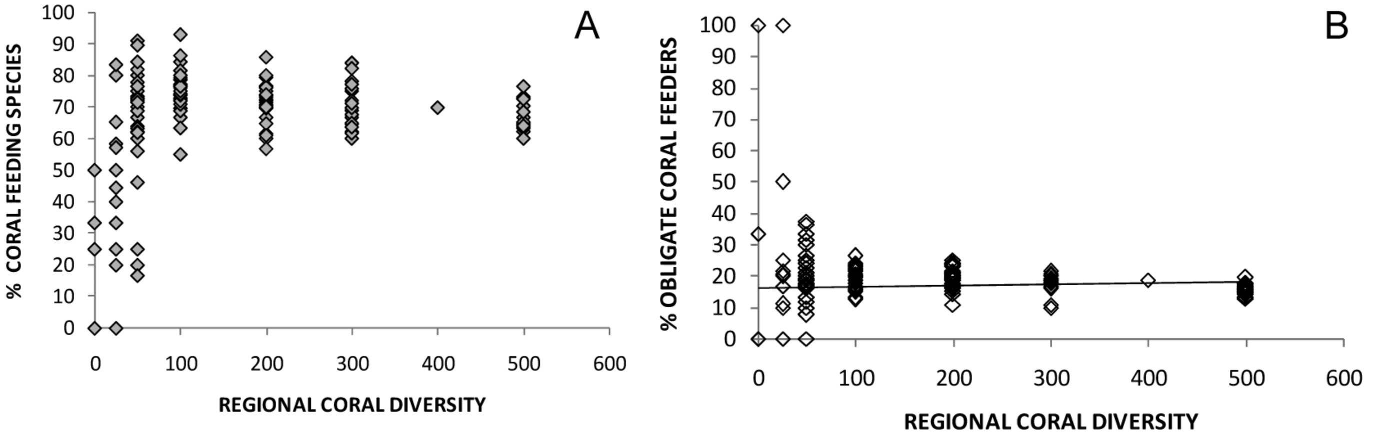

Fig. 4.6 Proportion of **(A)** coral feeding butterflyfish species and **(B)** coral specialist (obligate corallivore) butterflyfish species as regional coral diversity increases.

energy and habitat area and found that reef area and MDE best explained the distribution of diversity, and that energy (sea surface temperature and solar irradiance) had a relatively minor effect. Our analysis of butterflyfishes may provide further insights. In particular, connectivity, latitude, and island size had different effects in the Indian Ocean compared to the Pacific Ocean. In the latter, diversity was highly correlated to connectivity and island size. The diversity of butterflyfishes decreased as islands became smaller and more isolated, whereas in the Indian Ocean these factors had no significant effect—butterflyfish diversity decreased only with latitude. There are striking differences in the distribution of continents and islands in these oceans and there are different types of physical barriers to the geographical distribution of coral reef fishes, which may explain the observed distributions. In the Indian Ocean there are few islands but a continuous coastline from Indonesia to South Africa. Three natural barriers interrupt the progression of diversity along this coastline: the delta system in the Bay of Bengal, the delta system in Pakistan, and an upwelling area in Oman. Diversity drops westwards of these barriers suggesting that they are effective dispersal barriers for these fishes. However, diversity increases again some distance from these barriers suggesting that these barriers may not have been effective all the time or that other mechanisms are at play. The island system formed by the Lakshadweep, Maldives, and Chagos groups may have represented a stepping-stone towards the Mascarene group and the Seychelles, and then towards Madagascar and South Africa. If this view were correct, one would expect an accumulation in the southern end of eastern Africa as species could come from two directions, north from the Arabian Peninsula and east via the aforementioned islands. In such a case neither connectivity nor island size would be significant factors, as observed. In contrast, in the western Pacific there was a very strong overlap between the gradient of island size, island isolation, and decrease in diversity with increasing distance from the biodiversity centre (as indicated by the exclusion of island size as a significant factor in the multiple regression—Table 4.1). In such a case it is clear that diversity could be explained by the dispersal and colonising capacities of the various species. In particular, the lowest diversities found in the most remote areas such as the Hawaiian Archipelago or Easter Island, are supported by this hypothesis. Latitude was also shown to play a minor but significant role, suggesting that the temperature tolerance of these fishes may be important, especially at the southern and northern extremes of their distributions, such as in northern New Zealand or Korea.

In the Atlantic, the effects of large-scale factors were somewhat similar to what was observed in the west Pacific, a decrease in diversity with distance to the biodiversity centre (Caribbean Sea–Briggs, 2005; Floeter et al., 2008) and with isolation (remote islands supported less diversity). Such a pattern has already been described indirectly by Floeter et al. (2008) for reef fishes in general in the Atlantic; in particular, these authors showed the role of barriers such as the Amazon, the Benguela current, and the width of the Atlantic.

In the East Pacific the distribution of butterflyfish diversity was counter-intuitive, with diversity being negatively correlated with connectivity. The highest diversity is indeed found in the Galapagos with an accumulation of both species from the central Pacific and East Pacific endemics. This suggests that this archipelago is a sink for these two sources. For this reason, there is no real biodiversity centre for butterflyfishes in that region and therefore, logically, no correlation with such a factor. There was very little gradient in island size in this region, with either small islands or a continent and the higher diversity observed on islands was not significantly different (t-tests, $p > 0.05$) from the diversity on the continent. Mora and Robertson (2005) investigated the factors behind the latitudinal variations in shore fish diversity in the East Pacific. They found that diversity varied according to the geographic range of species, the diversity of species with wide ranges being mainly explained by MDE and the diversity of restricted range species being explained by habitat variables. Energy was not a significant variable as there was a confounding effect with MDE. No latitudinal variations for butterflyfishes were found in that region, which suggests that they do not follow the general distribution of shores fishes in this region. One potential reason is that these butterflyfish have specific dispersal and colonising capacities that are different for species coming from the Indo-Pacific and those endemic to the East Pacific.

"Relative" α diversity

Butterflyfishes are only one of many families of reef fishes. The relative importance of this family amongst reef fishes may fluctuate according to identifiable factors. An increase in the proportion of butterflyfish according to one factor would suggest that these fish are better adapted than the 'average' reef fish to the conditions imposed by this factor. For instance, if their relative α diversity increases as islands become smaller and more isolated, then this would be an indication of an above average dispersal and colonising capacity. Bellwood and Hughes (2001) analysed the proportion of 13 families of coral and reef fishes in the Indo-Pacific (including the East Pacific). One of their major findings was that the proportion of these families tended towards a constant as total diversity increased. We adopted a similar approach to analyse the butterflyfishes on a worldwide scale, each ocean being analysed separately as there may be specific effects in each ocean as observed for "absolute" α diversity. One of the problems in defining the proportion of butterflyfishes is to properly assess the total reef fish fauna (Mora et al., 2008). Many families, in particular small and/or cryptic species such as Muraenidae, Scorpaenidae, Apogonidae, Blenniidae, or Gobiidae, are very difficult to census completely. Therefore it was decided to use only families that are easily sampled (as Bellwood and Hughes, 2001; Roberts et al., 2002) to represent the total reef fish fauna (see Appendix I for the list of selected families). The proportion of butterflyfishes was clearly not constant within an ocean (Fig. 4.7). Within each ocean there was a decreasing trend of this proportion as total reef fish diversity increased; sites with low % of butterflyfishes showed a large variance in total species richness. The next focus of this chapter, then, is to understand the factors that generated this variance.

The distribution of the sites in the Indo-Pacific may be divided into sites with low total reef fish diversity (< 200 species) and those with higher diversity (following Bellwood and Hughes, 2001). A correlation analysis between the proportion of butterflyfishes (within the reef fish species pool) and the distance to the biodiversity centre, island size, connectivity, and latitude showed a striking difference between the Indian and Pacific oceans. In the Indian Ocean the proportion of butterflyfish species was significantly and negatively correlated (r= –0.45; p=0.0012) to the distance

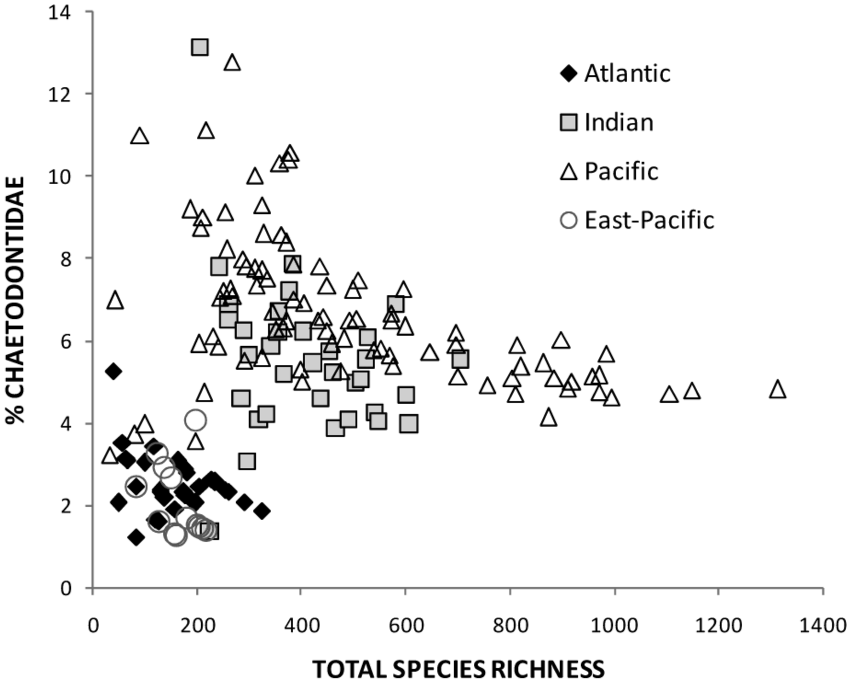

Fig. 4.7 Contribution of butterflyfishes to total reef fish diversity. Each point is one of 174 sites.

to the biodiversity centre and to island size ($r=-0.37$, $p=0.02$). Latitude and connectivity were also significant factors, but in a GLM ($r^2=0.42$, $N=33$) the component of variance they explained was no longer significant once distance to the biodiversity centre was taken into account. In contrast, in the western Pacific none of the previous factors had a significant effect (GLM $r^2<10^{-5}$, $N=81$). In addition there was no relationship between the proportion of butterflyfish species and the average values of their life-history traits in the western Pacific whereas in the Indian Ocean several traits were significant (see Appendix III for the distribution of life-history traits). In particular, on sites where the proportion of butterflyfish species was high, these species tended to be larger, more species fed on non-coral invertebrates, and they tended not to school. In the Atlantic there was a negative correlation (multiple regression $r^2=0.17$, $N=34$) with connectivity and latitude ($p=0.02$ for both factors) and no significant relationship with life-history traits. In the East Pacific connectivity was a significant factor ($r^2=0.84$, $N=15$, $p< 10^{-5}$) but there was a very strong correlation with life-history traits. In particular when the proportion of butterflyfish species was high, most species were small and tended to feed on coral or/and plankton, but did not feed on algae. These results suggest that different mechanisms are at play in each ocean but at present it is difficult to identify these mechanisms and their causes.

Beta Diversity

There are many ways of defining and estimating beta diversity (Koleff et al., 2003). Here we used two indices. The first, β_I, is based only on the gradient in butterflyfish species numbers within a region and was estimated as $\beta_I = \gamma/\alpha$ (Whittaker, 1972), where γ is the total number of species within a region (or sub-region) and α the average number of species across sites within the region (or sub-region). The second index, β_{II}, takes into account the butterflyfish species composition and was estimated as the variance of the presence/absence matrix (Legendre et al., 2005).

Beta diversity is difficult to analyse in the context of this chapter because the number of sites per region or sub-region varied considerably and was usually low. Subsequently only sub-regions with at least 3 sites were retained for the analysis. In general, there was little correlation between beta diversity and most factors (Table 4.2). β_I diversity was negatively correlated with coral diversity and the proportion of coral feeding species. This suggests that the presence of a high coral diversity will favour a rather homogeneous distribution of species, whereas low coral diversity will generate a stronger gradient in species diversity. In the Pacific Ocean body size increased with β_I diversity, which suggests that species heterogeneity is highest where species are the largest, which coincides with a lower

Table 4.2 Correlation between beta diversity (estimated at the sub-region level) and life-history traits and large scale factors. Only Spearman correlation coefficients were estimated as the data set was not sufficient for more complex analyses, such as GLMsNS: not significant; *: $p < 0.05$; **: $p < 0.01$; ***: $p < 0.001$. β_I: beta diversity estimated from Whittaker (1972); β_{II}: beta diversity from Legendre et al. (2005). "World": all sub-regions (>2 sites/sub-region). "Pacific": sub-regions in the Pacific Ocean (> 2 sites/sub-region), the Easter Island sub-region being excluded. N: number of sub-regions. Note that the factors, '% algae feeding species', 'number of diet items', 'schooling behaviour', and 'latitude' are not shown because they are never significantly correlated to beta diversity.

	γ Diversity	**% Coral feeding Species**	**% Mobile. Invert. feeding species**	**Species Size**	**Average Species Geographic Range**	**Geological Age**	**Island Size**	**Connectivity**	**Distance to Biodiversity Centres**	**Coral Diversity**
β_I World N= 27	NS	NS	NS	NS	NS	NS	NS	NS	NS	–0.40*
β_I Pacific N=17	NS	–0.51*	NS	0.67**	NS	NS	NS	NS	NS	–0.37*
β_{II} World N= 27	0.77***	0.56**	–0.58**	NS	0.43*	–0.50*	NS	NS	NS	NS
β_{II} Pacific N= 17	0.51*	NS	NS	NS	–0.61**	NS	0.67**	0.61**	–0.62**	NS

proportion of coral feeding species (coral feeders being on average smaller than other species). The level of correlation was much higher for β_{II} diversity that takes into account the difference in species composition amongst sites. High β_{II} diversity, reflecting high spatial heterogeneity in species composition and variance in species numbers, was found in the Pacific on large and/or connected islands (or continents) and was associated with islands near the biodiversity centre where average species geographical areas are small. These relationships suggest that high spatial heterogeneity in the species composition of butterflyfishes is found in areas where α and ɤ diversities are high. This raises the question of species "saturation" as regional diversity increases (Loreau, 2000; Hillebrand, 2005). In the case of butterflyfishes there is little evidence to suggest saturation of species diversity at such scales, even at the highest regional diversities there was very little variation in the rate of increase of the local diversity (Fig. 4.8). While the asymptotic relationships between α and ɤ cannot be considered as an indication of local species saturation (Loreau, 2000), the lack of an asymptote can be considered as an absence of local species saturation.

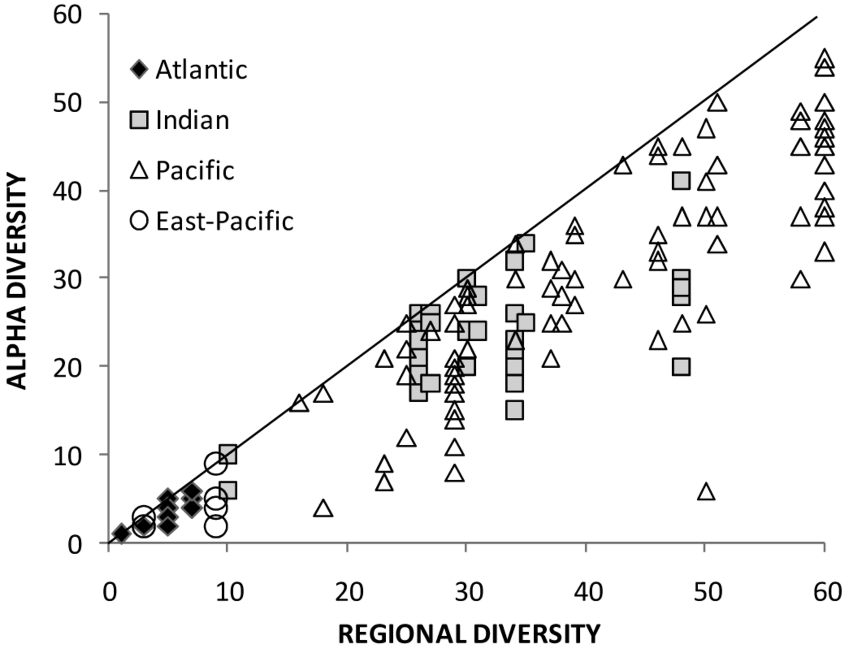

Fig. 4.8 Relationship between local (alpha) and regional diversity. Each point represents a different site. The diagonal line represents the one for one relationship (local = regional diversity).

Geographic Range

In biogeography there is often a special interest oriented towards endemic species, as they may be important to understand the mechanisms behind colonisation. Conversely, there is often little attention directed towards widespread species that may have life-history traits that are the most efficient for dispersal and colonisation. Endemism, by definition, is scale-dependent. As the size of the defined area increases, the rate of endemism will increase (Fig. 4.9). For this chapter, geographic ranges were estimated based on the number of sites where each species is recorded (rather than projections of the total geographic range) to avoid any erroneous extrapolations based on patchy records of occurrence. The number of endemics were then assess based on whether species are recorded at only one site, within just one sub-region, or a single ocean basin. There are only 6 species of Chaetodontidae that are endemic at the scale of sites, versus 34 endemic species at the level of sub-regions. A high proportion of the butterflyfish species (58% of all species) are endemic to one ocean, and all the Atlantic species are endemic at that scale. Within the Atlantic, the highest endemism is found in the East Atlantic sub-region (5 species, 100% endemic) followed by the Caribbean (42%). In the East Pacific, endemism is high in the islands (22%) but nil along the coast. The Red Sea has by far the highest endemism (23.5%) within the Indian Ocean and the largest number of endemic species (8) across all of the world's oceans. Only 3 other sub-regions in the Indian Ocean have endemic species, Western Australia having the highest proportion (6.7%). The rate of endemism is the lowest in the Pacific; only two sub-regions have an endemism rate above 10% (the Easter Island group: 17.4%; and Hawaii:

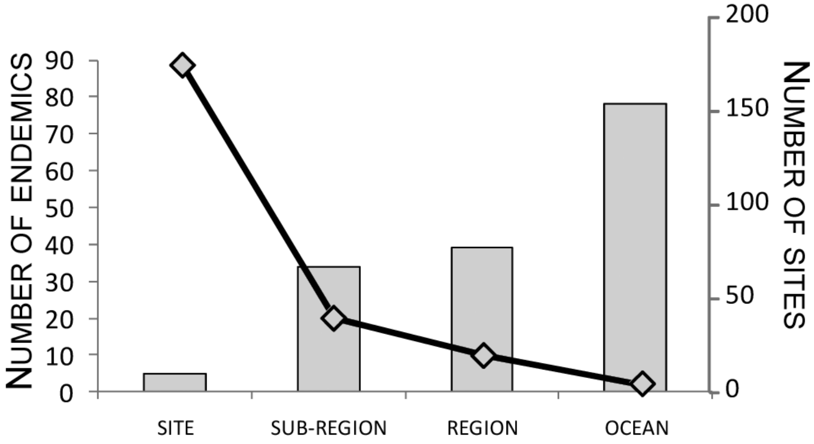

Fig. 4.9 Number of endemic species (bars) according to the area level considered (squares; linked by a line for clarity).

16%) and 16 out of 22 sub-regions have no endemic butterflyfish species. There is, therefore, no relationship between the endemism and diversity of butterflyfishes. This distribution confirms previous findings that there is no relationship between centres of diversity and endemism for reef fishes (Hughes et al., 2002; Allen, 2007; but see Roberts et al., 2002 and Mora et al., 2003). It also shows that in the Indo-Pacific the areas of high endemism are at the limit of the butterflyfish distribution; these areas are characterised by a high level of isolation (remote islands in the case of Easter Island and Hawaii; isolated sea in the case of the Red Sea).

A geographic range index (GRI), based on the average geographic range of the species within a site, was estimated to try to understand the relationship between diversity and geographic range. A high GRI is indicative of a large proportion of wide ranging species and, on the other hand, a low GRI may be associated with large proportions of species with restricted ranges and possibly high endemism. Butterflyfish species with a restricted range are the most numerous (Fig. 4.10). This is the most common type of distribution (Gaston, 2003) and has been documented for several types of reef organisms (Roberts et al., 2002). Butterflyfish diversity was significantly correlated to GRI (r=0.60; N=171) on a global scale, but this relationship was not significant in most oceans except the Pacific (r= –0.35; N=91). The latter negative correlation can be expected as butterflyfish diversity decreases as connectivity decreases in the Pacific, thus favouring species with high dispersal or colonising capacities. The absence of

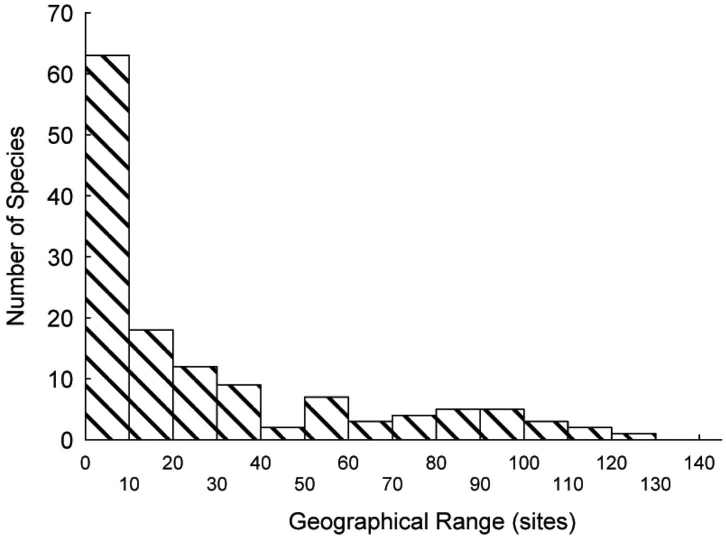

Fig. 4.10 Number of species according to geographical range (measured by the number of sites where the species has been recorded).

correlation in the other oceans is probably linked to the potential dispersal along large landmasses, which does not favour species with high dispersal capacities. This is confirmed by the analysis of the relationship between GRI and several large-scale factors (Table 4.3). In both the Indian and West Pacific Oceans, island size, latitude, and connectivity were significant negatively related to GRI. In other words, the geographic range of species tended to be greater on smaller, isolated islands at high latitudes. Such relationships may be associated with higher dispersal and colonising abilities. The low level of correlation with the distance to the biodiversity centres is in fact due to a dome shaped curve (Fig. 4.11), indicating that low GRI values in the West Pacific are found either near the biodiversity centre or far from it.

The next question is to determine if specific life-history traits may be related to geographic range (Robertson, 2001). Despite a number of significant correlations (Table 4.4), the only consistent relationship across the four oceans was a negative correlation between geographic range and the proportion of schooling species. In other words, species with the highest schooling capacities have either a low dispersal capacity or a low colonising success. Restricting the analysis to the Indo-Pacific where many species (52) are found in both the Indian and Pacific Oceans reveals a much stronger structure in the Pacific, where almost all life-history traits were very significantly correlated; in the Indian Ocean, only body size reached a high level of correlation. In areas of the Pacific where GRI is low there is a lower proportion of butterflyfish species feeding on coral and algae, a lower proportion of coral specialists, a smaller number of diet items, and a higher proportion of schooling and plankton feeding species. This means that in peripheral regions butterflyfish species will tend to have small geographic ranges and be characterised by smaller body sizes, lower dependency on coral in their diets, and a stronger tendency to school. This may be expected since these regions have low coral diversity. This also suggests that whenever coral diversity is sufficient, butterflyfish species will colonise since most corallivorous species have a wide geographic range and demonstrate a high adaptation to coral.

Table 4.3 Spearman correlation coefficients between the geographic range index (GRI) and large-scale factors, within each of the major geographic locations. N: number of sites; NS: not significant; *: $p < 0.05$; **: $p < 0.01$; ***: $p < 0.001$.

	Island Size	**Latitude**	**Connectivity**	**Distance to Biodiversity Centres**	**Regional Coral Diversity**
Atlantic (N=34)	–0.04 NS	–0.06 NS	0.25 NS	–0.46*	–0.01 NS
East Pacific (N=15)	0.17 NS	0.07 NS	–0.35 NS	0.21 NS	0.25 NS
Indian (N=31)	–0.50**	–0.36*	–0.47**	–0.38*	0.10 NS
Pacific (N= 91)	–0.29**	–0.29**	–0.40***	0.16 NS	–0.30**

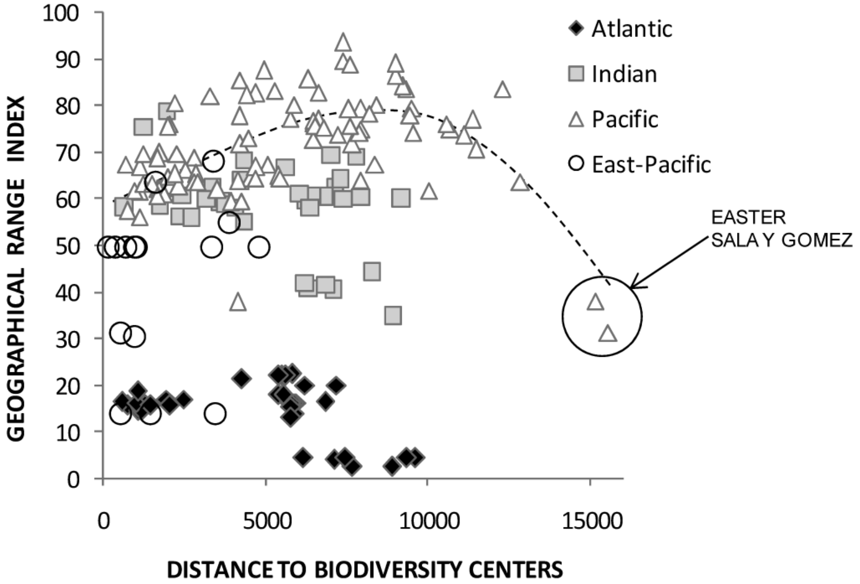

Fig. 4.11 Relationship between the geographical range index (GRI) and the distance (km) to the biodiversity centres for the four oceans. A 2° polynomial regression is significant at $p < 10^{-5}$ for the Pacific Ocean.

Geological Age and Phylogeography

To identify if butterflyfish diversity in an area is linked to the geological history, a first step is to determine if the distribution of the phylogenetic diversity supports the definition of the biogeographical regions based on the species distribution. The phylogeny of the butterflyfishes has been established recently for 71 species among the 134 species of the family (Fessler and Westneat, 2007). This well-resolved and well-supported topology was used to assess the phylogenetic structure of the 19 biogeographic regions defined above (Fig. 4.2). The sequences for 71 chaetodontids from the study of Fessler and Westneat (2007) were downloaded (GenBank accession numbers EF616820 to EF617232), aligned for each data set with Clustal X (Thompson et al., 1993) and concatenated into a single file for subsequent analyses. A total of 3223 nucleotides were aligned and a neighbour-joining tree was built following the bioNJ algorithm (Gascuel, 1997) using the JC model for distance computations as implemented in PAUP* 4.0b10 (Swofford, 2002). The tree and the presence/absence data matrix were used for subsequent analyses of the phylogenetic structure of the communities using Phylocom (Webb et al., 2008).

Table 4.4 Significance level of Spearman correlation coefficients between geographic range and geological age versus dietary composition (e.g., major dietary items), body size, schooling behaviour and depth range. N: number of sites; NS: not significant; *: $p < 0.05$; **: $p < 0.01$; ***: $p < 0.001$.

	Algae	Coral	Other Sessile Invertebrates	Mobile. Invert	Plankton	Coral specialists	Body Size	Solitary	Pairing	School	Depth range
Geographic range											
Atlantic (N=34)	NS	***	***	***	NS	**	–*	***	NS	–***	–*
East Pacific (N=11)	–***	NS	–***	NS	NS	NS	NS	***	*	–**	NS
Indian (N=35)	NS	NS	–**	***	NS	**	***	NS	NS	–***	NS
Pacific (N= 91)	***	***	**	–**	–**	***	***	NS	***	–**	NS
TOTAL (N=175)	***	***	–***	–***	–*	**	**	NS	NS	–*	–*
Geological Age											
Atlantic (N=34)	NS	NS	–**	NS	–**	NS	**	NS	NS	**	NS
East Pacific (N=11)	***	–***	*	*	–*	–*	**	–***	NS	***	*
Indian (N=35)	NS	*	NS	*	–**	NS	**	NS	**	–*	NS
Pacific (N= 91)	NS	–*	–***	*	**	–**	NS	NS	–**	**	NS
TOTAL (N=175)	**	–***	**	***	–*	–***	***	–**	–***	**	*

The distribution of the phylogenetic distances between species from the same biogeographic region or belonging to distinct biogeographic regions provided contrasted patterns (Fig. 4.12). In each case, a bimodal distribution was found, the first mode corresponding to comparisons among species within genera and the second mode among species from distinct genera. The mode at 0.08 within regions and the mode starting at 0.09 amongst regions involved comparisons between *Chaetodon-Prognathodes* and the other genera (*Coradion, Chelmonops, Chelmon, Forcipiger, Hemitaurichthys, Johnrandallia*, and *Heniochus*). Strikingly, phylogenetic distances among species of the same genera (i.e., the first mode) were lower within- than among-regions, indicating a closer phylogenetic relatedness among congeneric species from the same region. This result suggests that the speciation events occurring at the region's boundaries tend to be older on average since phylogenetic distances are larger among region than within region.

The ordination of the biogeographic provinces according to the average phylogenetic distances among species, revealed a poorly resolved cluster (Fig. 4.13). Only some of the Atlantic regions clustered together (Brazil, Azores-Madeira, Brazilian Islands, and the Caribbean), indicating that species from that area are more closely related among them than with those from the Indian and Pacific Oceans. The two most basal *Chaetodon* species (*C. hoefleri* and *C. robustus*) are found in the eastern Atlantic, thus promoting

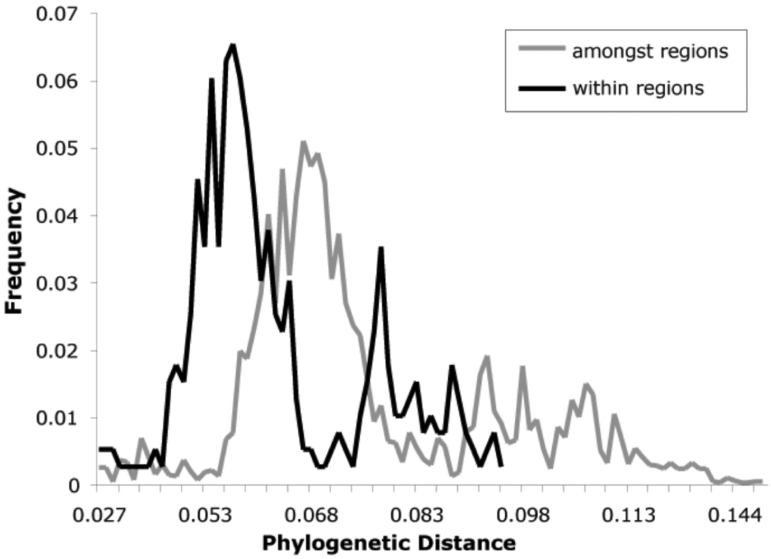

Fig. 4.12 Distribution of phylogenetic distances between species belonging to the same biogeographic region (within regions) and species belonging to distinct biogeographic region (amongst regions). Frequencies are expressed as percent of the total number of distances computed, namely 399 within regions and 7182 amongst regions.

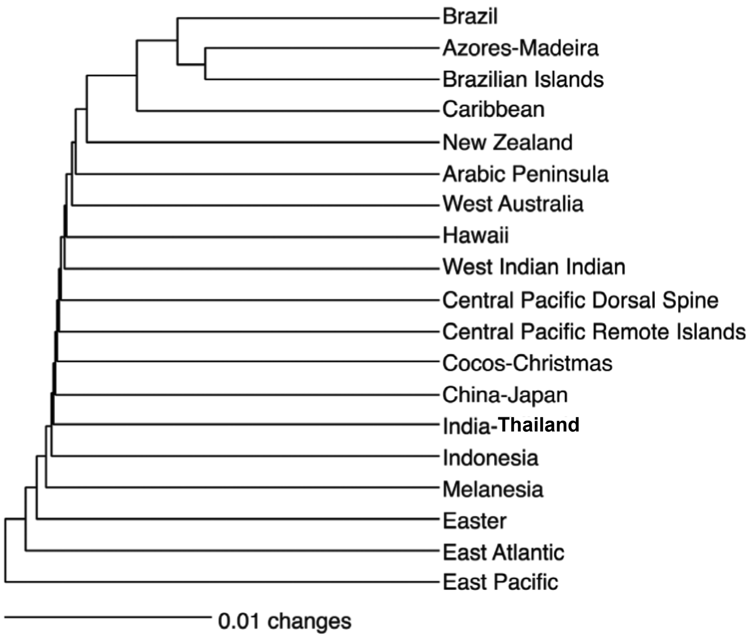

Fig. 4.13 UPGMA dendrogram of the 19 regions computed based on the mean pair wise phylogenetic distance between species communities. UPGMA: Unweighted Pair Group Method with Arithmetic Mean is a clustering algorithm that assesses the dissimilarity between the new cluster and the existing cluster by the distance between the means.

distinctiveness from the western Atlantic. By contrast, phylogenetic communities from the other regions diverged from one another on average by the same phylogenetic distance. This suggests that dispersal has been important enough to promote the spatial expansion of older lineages throughout the biogeographical provinces aforementioned, and hide the signal in the communities' phylogenetic structure of the chronology of differentiation of these regions (Fig. 4.12). Despite the apparent homogeneity of the phylogenetic structure of the biogeographical regions, we further explored the phylogenetic composition of the 19 regions by using several metrics of communities' phylogenetic structure (Fig. 4.14).

We first computed Faith's (1992) index of phylogenetic diversity (PD) for each region. This index is the proportion of total branch length in the phylogeny associated with the taxa in a particular region. The Net Relatedness Index (NRI) and the Nearest Taxon Index (NTI) according to Webb (2000) and Webb et al. (2002) were then computed. These indices describe the differences between the average phylogenetic distances (NRI) and the nearest neighbour distance (NTI) in observed and null communities

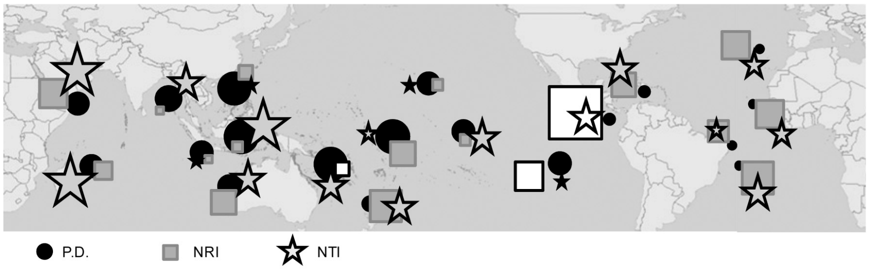

Fig. 4.14 Regional values of PD, NRI and NTI for Chaetodontidae across the world. PD: phylogenetic distance, radius of circle is proportional to value of PD. NRI: Net Relatedness Index, size of square is proportional to NRI value, grey represents >0 and white <0 values. NTI: Nearest Taxon Index, size of star is proportional to NTI value, black represents values <–0.3; +0.3>, grey represents >0.3 and white <–0.3 values.

see below

Colour image of this figure appears in the colour plate section at the end of the book.

(no phylogenetic structure), standardized by the standard deviation of phylogenetic distances in the null communities. High values of NRI indicate that species are more phylogenetically related than expected by a null model, and negative values are found when species are more phylogenetically divergent than expected by chance. High values of NTI indicate that the closest related species are closer than expected and negative values suggest species further than expected.

Phylogenetic diversity (PD) was found to be largely heterogeneous amongst the 19 regions (Fig. 4.14). The highest PD values were observed at the centre of biodiversity (Indonesia) and nearby regions (Melanesia, China-Japan) with very low values in the Atlantic and East Pacific. This means that regions with the highest species diversities had also the highest phylogenetic diversities. This was confirmed by the NRI where values were generally higher in the Atlantic indicating that species were more phylogenetically related than expected under a null community model. This suggests that within the Atlantic regions most species had a common history. The East Pacific and Easter Island regions were characterised by high phylogenetic divergence. For the East Pacific this may be related to the mix of endemic species with species originating from the West Pacific. For the Easter Island region the high phylogenetic divergence could be due to the grouping of Pitcairn-Dulcie (21 species) with species-poor Easter and Sala-Gomez islands (9 and 7 species respectively).

The Net Taxon Index (NTI) was higher in the Indian Ocean compared to random values. In other words, there tends to be more phylogenetic homogeneity in that ocean than elsewhere. Associated with the higher-than-expected species diversity (Connolly et al., 2003) in the western part of the

Indian Ocean, this could be the reflection of a common dispersal history. However, the highest NTI value was found in Indonesia, indicating that species in that region harbour a very low divergence to their closest relative compared to other regions. This is in contrast with the surrounding regions where NTI is low. This may be an indication that speciation may have also occurred in the area. Together, these phylogenetic structure results indicate that the highest phylogenetic diversity was found in the biodiversity centre (Indonesia) but also that frequently, species from this region are closely related. This apparent discrepancy suggests that this biodiversity centre hosts very divergent lineages that increase the phylogenetic diversity values, but also hosts very speciose lineages with a high level of closely related species. This result converges with the findings of Barber and Bellwood (2005) that suggested *Halichoeres* (Labridae) communities in the coral triangle resulted from a mixed influence of old speciation events from the middle Miocene at 12 million years ago (Ma) and recent speciation events during the Pliocene around 3 Ma.

The distribution of species divergence suggests that speciation events are older amongst regions than within regions (Fig. 4.12). The pattern of relationships amongst regions based on average phylogenetic divergence, however, remains unclear (Fig. 4.13) suggesting that dispersal frequently occurred during the diversification of the Chaetodontidae. Thus, ecological factors that are likely to affect dispersal and establishment of species in new areas may be related to historical parameters. The recent phylogenetic study of the family (Fessler and Westneat, 2007) provided a very detailed timeframe for the diversification of this group of fish. This set covers a majority of species (71 species of ~134 species), but in some areas the number of species covered may be low compared to the total number of species. We computed the average species age within islands according to the species age estimates from Fessler and Westneat (2007) as an indicator of the relative contribution of recent and older speciation events to extant communities. Generally, the average species age was related to several life-history traits (Table 4.4). In particular sites with a high average species age supported species that tended not to feed on coral, were above average body size, and tended to school. Average species age was also negatively related to regional coral diversity and distance from biodiversity centres, and positively related to island size. If average species age is ranked according to longitude and ocean, the homogeneity of this age for the Indo-Pacific is evident as well as the superior ages of the Atlantic and East Pacific and the peculiar position of the East Atlantic (Fig. 4.15).

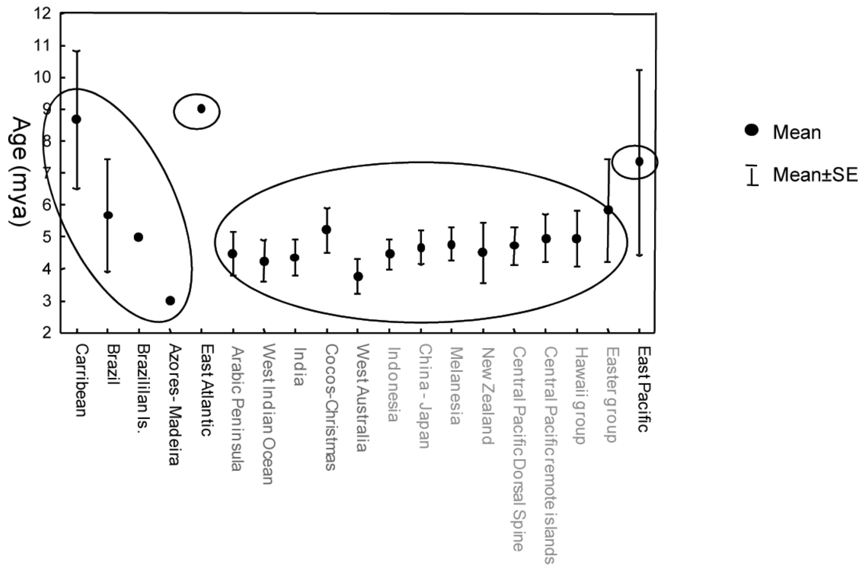

Fig. 4.15 Average age (mya) of species in each of the 19 biogeographic regions. Regions are ordered according to longitude. Indian Ocean regions are in red, west Pacific regions are in blue.

Colour image of this figure appears in the colour plate section at the end of the book.

CONCLUSION

Butterflyfishes are a conspicuous component on coral reefs around the world. Even though they do not necessarily represent a large proportion of the total reef fish diversity or abundance, their link to coral makes them an ideal model to understand the relationships between coral reef fishes and coral on a global basis. Based on checklists from 175 sites across the world, the geographical distribution of 134 species of butterflyfishes and their major life-history traits were analysed, with specific attention upon the potential effects of large-scale factors such as island size and connectivity, distance to centres of biodiversity, latitude, and regional coral diversity. The strength of butterflyfishes as a model for macro-ecological analyses of coral reef fishes is confirmed by the general convergence of the findings for butterflyfishes with previous findings from analyses based on larger sets of species. This strongly suggests that a number of findings based only on butterflyfishes presented in this chapter may be representative of coral reef fish in general, but of course this will need to be validated as our knowledge on the spatial distribution, biology, ecology, and phylogeny of other coral reef fish progresses.

Based upon butterflyfish species composition, 19 regions and 39 subregions were defined. This is the first attempt to define biogeographical regions for coral reef fishes on a worldwide basis, and revealed congruence with most previous findings across smaller geographical scales. Butterflyfishes therefore, could represent an interesting model for large-scale reef fish biogeography, except in the Atlantic and eastern Pacific where the diversity of these fishes is low. The highest diversities of butterflyfishes were found in the Indonesian-Philippine region for the Indo-Pacific Ocean and in the Caribbean in the Atlantic Ocean. The rate of endemism of these fishes was highest in the Atlantic, but there were several spots of high endemism in the Indo-Pacific, in particular in the Red Sea, Hawaii, and Easter Island. The distribution of these endemic species confirms the absence of overlap between biodiversity centres and diversity hotspots for reef fishes. The diversity of butterflyfishes was related to several large-scale factors. Regional coral diversity was the best factor explaining the geographical distribution of these fishes, with the distance to biodiversity centres being the second best factor, except for the East Pacific Ocean where this factor played no role. Island size, connectivity, and latitude had similar influences on butterflyfish diversity in both the Atlantic and Pacific Oceans, but only latitude was significant in the Indian Ocean.

Most butterflyfish species tend to have a rather restricted geographic range. There is little correlation between the average range of species on a site and the life-history traits of the species, except in the Pacific where diet, size, and schooling behaviour change significantly with geographic range (Pratchett, Chapter 9).

Phylogenetic analyses of the 71 chaetodontids from the study of Fessler and Westneat (2007) suggested that Chaetodontidae might have dispersed at a fast rate in the Indo-Pacific with little local speciation, whereas in the Atlantic there was a clear separation between East and West. On the other hand, the highest phylogenetic divergence was found in the East Pacific. Phylogeny pointed also at a probable species accumulation rather than speciation in the Indonesian-Philippine biodiversity centre, which could be related to a mid-domain effect. In the Indian Ocean, phylogeny suggested that dispersal has been rather homogeneous compared to the Pacific. Geological ages are higher and more heterogeneous in the Atlantic and East Pacific than in the Indo-Pacific. In addition there was a negative correlation between average species age and coral feeding or coral dependence, which suggests that the dependence of butterflyfishes to coral is a relatively recent event (see also Bellwood and Pratchett, Chapter 1).

REFERENCES

Allen, G.R. 2000. Indo-Pacific coral-reef fishes as indicators of conservation hotspots. *Proceedings of the 9th International Coral Reef Symposium, Bali*, pp. 921–926.

Allen, G.R. 2007. Conservation hotspots of biodiversity and endemism for Indo-Pacific coral reef fishes. *Aquatic Conservation: Marine and Freshwater Ecosystems* 18: 541–556.

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

Bellwood, D.R. and T.P. Hughes. 2001. Regional-Scale Assembly Rules and Biodiversity of Coral Reefs. *Science* 292: 1532–1535.

Bellwood, D.R., T.P. Hughes, S.R. Connolly and J. Tanner. 2005. Environmental and geometric constraints on Indo-Pacific coral reef biodiversity. *Ecology Letters* 8: 643–651.

Bellwood, D.R., P.C. Wainwright, C.J. Fulton and A.S. Hoey. 2006. Functional versatility supports coral reef biodiversity. *Proceedings of the Royal Society B* 273: 101–107.

Briggs, J.C. 2005. Coral reefs: conserving the evolutionary sources. *Biological Conservation* 126: 297–305.

Brito, A. and J.M. Falcon. 1996. Capture of the St. Helena butterflyfish, *Chaetodonsanctaehelenae* (Chaetodontidae) in the Canary Islands. *Cybium* 20: 99–100.

Carpenter, K.E. and V.G. Springer. 2005. The center of the center of marine shore fish biodiversity: the Philippine Islands. *Environmental Biology of Fishes* 72: 467–480.

Colwell, R.K., C. Rahbek and N.J. Gotelli. 2004. The Mid-Domain Effect and Species Richness Patterns: What Have We Learned So Far? *American Naturalist* 163: E1–E23.

Connolly, S.R., D.R. Bellwood and T.P. Hughes. 2003. Indo-Pacific biodiversity of coral reefs: deviations from a mid-domain model. *Ecology* 84: 2178–2190.

Faith, D.P. 1992.Conservation evaluation and phylogenetic diversity. *Biological Conservation* 61: 1–10.

Fessler, J.L. and M.W. Westneat. 2007. Molecular phylogenetics of the butterflyfishes (Chaetodontidae): Taxonomy and biogeography of a global coral reef fish family. *Molecular Phylogenetics and Evolution* 45: 50–68.

Floeter, S., L.A. Rocha, D. Robertson, J. Joyeux, W.F. Smith-Vaniz, P. Wirtz, A. Edwards, J. Barreiros, C. Ferreira and J. Gasparini. 2008. Atlantic reef fish biogeography and evolution. *Journal of Biogeography* 35: 22–47.

Gascuel, O. 1997. BIONJ: an improved version of the NJ algorithm based on a simple model of sequence data. *Molecular Biology and Evolution* 14: 685–695.

Gaston, K.J. 2000. Global patterns in biodiversity. *Nature* 405: 220–227.

Gaston, K.J. 2003. *The structure and dynamics of geographic ranges*, Oxford University Press, Oxford, UK.

Hillebrand, H. 2005. Regressions of local on regional diversity do not reflect the importance of local interactions or saturation of local diversity. *Oikos* 110: 195–198.

Hughes, T.P., D.R. Bellwood and S.R. Connolly. 2002. Biodiversity hotspots, centres of endemicity, and the conservation of coral reefs. *Ecology Letters* 5: 775–784.

Kemp, J. 1998. Zoogeography of the coral reef fishes of the Socotra Archipelago. *Journal of Biogeography* 25: 919–933.

Kent, M. 2005. Biogeography and macroecology. *Progress in Physical Geography* 29: 256–264.

Kent, M. 2007. Biogeography and macroecology: now a significant component of physical geography. *Progress in Physical Geography* 31: 643–657.

Koleff, P., K.J. Gaston and J.J. Lennon. 2003. Measuring beta diversity for presence-absence data. *Journal of Animal Ecology* 72: 367–382.

Kulbicki, M. 2007. Biogeography of reef fishes of the French territories in the South Pacific. *Cybium* 31: 275–288.

Kulbicki, M., Y. Bozec and A. Green. 2005. Implications of biogeography in the use of butterflyfishes (Chaetodontidae) as indicators for Western and Central Pacific areas. *Aquatic Conservation: Marine and Freshwater Ecosystems* 15: S109–S126.

Legendre, P. and L. Legendre. 1998. *Numerical ecology*. 2nd ed., Elsevier, Amsterdam, The Netherlands.

Legendre, P., D. Borcard and P.R. Peres-Neto. 2005. Analyzing beta diversity: partitioning the spatial variation of community composition data. *Ecological Monographs* 75: 435–450.

Loreau, M. 2000. Are communities saturated? On the relationship between α, ß and ɤ diversity. *Ecology Letters* 3: 73–76.

MacArthur, R.H. and E.O. Wilson. 1967. *The theory of island biogeography*, Princeton University Press, Princeton, USA.

Magurran, A.E. 2004. *Measuring biological diversity*, Wiley-Blackwell, Malden, USA.

Moilanen, A. and M. Nieminen. 2002. Simple connectivity measures in spatial ecology. *Ecology* 83: 1131–1145.

Mora, C. and D.R. Robertson. 2005. Causes of latitudinal gradients in species richness: a test with fishes of the Tropical Eastern Pacific. *Ecology* 86: 1771–1782.

Mora, C., P.M. Chittaro, P.F. Sale, J.P. Kritzer and S.A. Ludsin. 2003. Patterns and processes in reef fish diversity. *Nature* 421: 933–936.

Mora, C., D.P. Tittensor and R.A. Myers. 2008. The completeness of taxonomic inventories for describing the global diversity and distribution of marine fishes. *Proceedings of the Royal Society B: Biological Sciences* 275: 149–155.

Pimm, S.L. and J.H. Brown. 2004. Domains of diversity. *Science* 304: 831–833.

Randall, J.E. 1995. *Coastal fishes of Oman*, Crawford House Publishing, Bathurst, Australia.

Randall, J.E. 1998. Zoogeography of Shore Fishes of the Indo-Pacific Region. *Zoological Studies* 37: 227–268.

Raup, D.M. and R.E. Crick. 1979. Measurement of faunal similarity in paleontology. *Journal of Paleontology* 53: 1213–1227.

Ricklefs, R.E. 2004. A comprehensive framework for global patterns in biodiversity. *Ecology Letters* 7: 1–15.

Roberts, C.M., C.J. McClean, J.E.N. Veron, J.P. Hawkins, G.R. Allen, D.E. McAllister, C.G. Mittermeier, F.W. Schueler, M. Spalding and F. Wells. 2002. Marine biodiversity hotspots and conservation priorities for tropical reefs. *Science* 295: 1280–1284.

Robertson, D.R. 2001. Population maintenance among tropical reef fishes: inferences from small-island endemics. *Proceedings of the National Academy of Sciences* 98: 5667–5670.

Robertson, D.R. and K.L. Cramer. 2009. Shore fishes and biogeographic subdivisions of the Tropical Eastern Pacific. *Marine Ecology Progress Series* 380: 1–17.

Rocha, L.A. 2003. Patterns of distribution and processes of speciation in Brazilian reef fishes. *Journal of Biogeography* 30: 1161–1171.

Santini, F. and R. Winterbottom. 2002. Historical biogeography of Indo-western Pacific coral reef biota: is the Indonesian region a centre of origin? *Journal of Biogeography* 29: 189–205.

Springer, V.G. 1982. Pacific plate biogeography, with special reference to shore fishes. *Smithsonian Contributions to Zoology* 367: 1–182.

Springer, V.G. and J.T. Williams. 1990. Widely distributed Pacific plate endemics and lowered sea-level. *Bulletin of Marine Science* 47: 631–640.

Swofford, D.L. 2002. *Phylogenetic analysis using parsimony (* and other methods)*, Sinauer Associates, Sunderland, USA.

Thompson, J.D., D.G. Higgins and T.J. Gibson. 1994. CLUSTAL W: improving the sensitivity of progressive multiple sequence alignment through sequence weighting, position-specific gap penalties and weight matrix choice. *Nucleic Acids Research* 22: 4673–4680.

Thresher, R.E. 1991. Geographic variability in the ecology of coral reef fishes: evidence, evolution and possible implications. In *The Ecology of Fishes on Coral Reefs*, P.F. Sale (ed.). Academic Press, Inc., San Diego, USA, pp. 401–435.

Vàzquez, D.P. and R.D. Stevens. 2004. The latitudinal gradient in niche breadth: concepts and evidence. *American Naturalist* 164: E1–E19.

Veron, J.E.N. 2000. *Corals of the World*, Australian Institute of Marine Science, Townsville, Australia.

Ward Jr., J.H. 1963. Hierarchical grouping to optimize an objective function. *Journal of the American Statistical Association* 58: 236–244.

Webb, C.O. 2000. Exploring the phylogenetic structure of ecological communities: an example for rain forest trees. *American Naturalist* 156: 145–155.

Webb, C.O., D.D. Ackerly, M.A. McPeek and M.J. Donoghue. 2002. Phylogenies and community ecology. *Annual Review of Ecology Evolution and Systematics* 33: 475–505.

Webb, C.O., D.D. Ackerly and S.W. Kembel. 2008. Phylocom: software for the analysis of phylogenetic community structure and trait evolution. *Bioinformatics* 24: 2098–2100.

Whittaker, R.H. 1972. Evolution and measurement of species diversity. *Taxon* 21: 213–251.

Whittaker, R.J., K.J. Willis and R. Field. 2001. Scale and species richness: towards a general, hierarchical theory of species diversity. *Journal of Biogeography* 28: 453–470.

Willig, M.R., D.M. Kaufman and R.D. Stevens. 2003. Latitudinal gradients of biodiversity: pattern, process, scale, and synthesis. *Annual Review of Ecology Evolution and Systematics* 34: 273–309.

Zapata, F., K. Gaston and S. Chown. 2003. Mid-domain models of species richness gradients: assumptions, methods and evidence. *Journal of Animal Ecology* 72: 677–690.