THE RAFFLES BULLETIN OF ZOOLOGY 2007 Supplement No. **14**: 77-86
Date of Publication: 31 Jan.2007
© Academia Sinica and National University of Singapore

MOLECULAR PHYLOGENY OF *CHAETODON* (TELEOSTEI: CHAETODONTIDAE) IN THE INDO-WEST PACIFIC: EVOLUTION IN GEMINATE SPECIES PAIRS AND SPECIES GROUPS

Kui-Ching Hsu
Department of Life Sciences, Cheng Kung University, Tainan, Taiwan 701

Jeng-Ping Chen
National Museum of Marine Biology and Aquarium, Taiwan 944

Kwang-Tsao Shao
Research Centre for Biodiversity, Academia Sinica, Nankang, Taipei, Taiwan 115
Fax: 886 227883463;
Email: zoskt@gate.sinica.edu.tw
(Corresponding Author)

***ABSTRACT.* –** We present a molecular phylogeny of *Chaetodon* derived from the analysis of sequences from mitochondrial cytochrome (cyt) *b* and 12S rRNA genes for 41 ingroup *Chaetodon* species. One species of *Chelmon* and one species of *Prognathodes* were used as outgroups to root the molecular trees. Two different computer software packages were utilized, namely, MEGA 2 and PHYLIP version 3.6a2. Multiple methods such as neighbor-joining (NJ), minimum evolution (ME), maximum likelihood (ML) and maximum parsimony (MP) were used. Cytochrome *b* sequence data provided greater resolution than 12S rRNA. Of the nine *Chaetodon* subgenera sampled, all were in the monophyletic group, but their phylogenetic relationships differed significantly from those inferred from morphological characters. *Lepidochaetodon* was the basal subgenus followed by *Exornator* and others. Molecular data support that *Corallochaetodon* and *Citharoedus* are sister groups, as suggested by morphological analysis. We found no geographical pattern of speciation events in the Indo-West Pacific and speciation occurred throughout the Indo-West Pacific as a result of diverse mechanisms. The evolutionary patterns of species pairs or species groups were more complex than previous envisioned as evidence from estimated divergences between species pairs suggests that the speciation events have occurred many times.

KEY WORDS. – Indo-West Pacific, speciation, biodiversity, *Chaetodon*.

INTRODUCTION

The vast Indo-West Pacific (IWP) region, stretching from the Red Sea and shores of East Africa to the islands of Polynesia, is the world's richest marine biological province. Over three-quarters of fish species in the world inhabit this region. The modern distribution of species richness in tropical oceans is the product of interactions between origination, extinction and migration of species and many hypotheses have been put forth to emphasize one or a combination of these processes (Bellwood & Wainwright, 2002; Connolly et al., 2003). There are two major questions related to speciation in the marine environment. Firstly, what are the prevailing modes and the geographical patterns of speciation? Many studies have supported the role of geographical isolation involving large-scale barriers (Colborn et al., 2001; Lessios et al., 2001), but divergence might also occur in sympatry (Vacquier, 1998), or require only transient allopatry (Hellberg, 1998). Allopatry may be achieved by vicariant division or by founder dispersal (Paulay & Meyer, 2002). Speciation events may be concentrated at the periphery of the region (as suggested by center-of-accumulation models of the diversity focus; Jokiel & Martinelli, 1992) or in the species rich center (the center-of-origin model; Briggs, 1999), or scattered across the region (Bellwood & Wainwright, 2002). Secondly, how old are the speciation events? Modern species may be the product of tectonic events during the Miocene and Pliocene, or alternatively of sea level fluctuations during the glacial cycles of the past three million years (Bellwood & Wainwright, 2002).

The genus *Chaetodon* from the family Chaetodontidae, or butterflyfishes, are widely distributed over major regions like the Western Pacific and the Indian Ocean. There have been

many reports of what appears to be closely related species pairs of the genus *Chaetodon* (Allen, 1979). The species pairs reported seem to have evolved relatively recently judging by their similarity in morphology, especially in color pattern. Randall (1998) lists seven pairs of geminate species of *Chaetodon* in his article of zoogeographical pattern of Indo-Pacific shorefishes. Here, we generate a molecular phylogeny inferred from the analysis of mitochondrial DNA (mtDNA) cytochrome *b* and 12S rRNA genes for *Chaetodon*. Our aim is to examine the mode, geography and timing of speciation in *Chaetodon*. In this study, we present a molecular phylogeny that includes all the 41 recognized taxonomic species in the genus. We combine the results with detailed distribution data to examine the speciation mechanism in the IWP region.

MATERIALS AND METHODS

Live fish for the experiments were obtained from field collections around Taiwan and aquariums. A total of 142 specimens, representing 41 species and 11 subgenera of the genus *Chaetodon* were included in this study. There were at least 2 (range = 2 - 8) individuals representing each species of the genus *Chaetodon.* We used 2 other butterflyfishes, *Chelmon rostratus* and *Prognathodes aculeatus* as outgroups to root our phylogenetic analysis. All of the specimens were used for partial mtDNA cytochrome (cyt) *b* and 12S rRNA sequencing. The species used and their sample sizes are listed in Table 1.

Approximately 0.5 g of muscle tissue was obtained from each specimen and was homogenized in 2.0 ml of cold grinding buffer (0.2 M NaCl, 0.05M EDTA, pH 8.0). Tissue previously stored in 95% ethanol was soaked in 5.0 ml of grinding buffer on ice for approximately 1 hour before grinding. High molecular weight DNA was isolated by proteinase K/phenol extraction procedures (Duda & Palumbi, 1999). The cyt *b* gene was amplified with primers cyt *F* (5'-ATGGCCAGCT TACGTAAAAC-3') and cyt *R* (5'-AAGAACGTATCCTACAAAG-3'). The 12S rRNA gene was amplified with primers 12SF (5'-TAAACATTGACAGTAAATAA-3') and 12SR (5'-TGCTTACTGCTAAATCCTC-3'). Polymerase chain reaction (PCR) amplification was carried out in an MJ PTC-100 Thermal Cycler (MJ Research, Inc.). Each 100 μL reaction mixture contained 10 ng of template DNA, 10 μL of 10_ reaction buffer, 10 μL $MgCl_2$ (25 mM), 10 μL dNTP mix (10 mM, 10 pmol of each primer) and 4 units of *Taq* polymerase (Promega, Madison, USA). PCR amplification conditions were: 35 cycles of denaturation at 94°C for 45 seconds, annealing at 53°C for 30 seconds and extension at 72°C for 45 seconds, followed by an extension at 72°C for 10 minutes. Storage was at 4°C. PCR products were separated and eluted using agarose gel purification (Boehringer Mannheim). Products were then sequenced directly with fluorescently-labeled dye terminators (Applied Biosystems, Inc.).

Nucleotide sequences were aligned with the program CLUSTALV (Higgins et al., 1992). Using Molecular Evolutionary Genetics Analysis program (MEGA) version 2.1 (Kumar et al., 2001), we calculated the base compositional frequencies and pairwise transition / transversion (TS / TV) ratios for all taxa. The potential for saturation of transitions is evaluated by comparing TS / TV ratios with percentage divergences for each pair of taxa. We examined the phylogenetic relationship between species based on the cyt *b* and 12S rRNA sequence by using neighbor-joining (NJ), minimum evolution (ME), maximum likelihood (ML) and maximum parsimony (MP) methods. Parsimony analyses were done by considering the TSs and TVs equally and using an *a priori* weighting scheme in which a TS / TV weighting of 1:2 was employed. NJ and ME trees were generated according to pairwise distance obtained by the Kimura's two-parameter model (Kimura, 1980). The ML tree was constructed according to Tamura & Nei's substitution model (Tamura & Nei, 1993), which recovers slightly biased branch lengths of lineages (Håstad & Björklund, 1998). The analyses were performed in MEGA2 (Kumar et al., 2001) and Phylogeny Inference Package (PHYLIP) version 3.6a2 (Felsenstein, 1993) programme packages. We tested the confidence of reconstructed clades by bootstrapping (Felsenstein, 1985) with 1,000 replicates using unweighted characters. The nodes with bootstrap values greater than 70, as a rule of thumb, were significantly supported with 95 % probability (Hillis & Bull, 1993).

A well-documented evolutionary rate is needed for estimating divergence between species. The rate was estimated to be 2% per million years for the mtDNA cyt *b* gene (Johns & Avise, 1998). In this study, using this evolutionary rate, the divergence times of *Chaetodon* species were estimated based on the average of the number of nucleotide substitutions (K) among lineages by removing sequences that deviated from a molecular clock model. Estimates of divergence times between species were then calculated using penalized likelihood in PHYLIP, as implemented in the r8s program. A 95% confidence interval was calculated for two relevant nodes using an algorithm included in r8s with a cut-off value of 4 (Sanderson, 2002). At least one dated reference point was needed to calibrate the output from r8s program into actual age estimates. The calibration point should have a date assigned, which is based on other data such as fossil information. However, we could not identify fossils for *Chaetodon* and the penalized likelihood analysis was therefore performed on a dataset representing divergence between lineages. In our analysis, we used the other two butterflyfishes (*Chelmon rostratus* and *Prognathodes aculeatus*) to root the tree and used the divergence time of the overall species estimated above based on a molecular clock, as a calibration point. The maximum likelihood-based PHYLIP version 3.6a2 (Felsenstein, 1993) was used to estimate branch lengths with the general time reversible model. The best tree recovered was input to r8s. An unconstrained penalized likelihood analysis (Sanderson, 2002) was conducted with the Powell algorithm (Chen et al., 1998).

Table 1. The subgenera list of the genus *Chaetodon* and the species number of each subgenus from Blum (1988), sample size and collection source used in this study. The values in parentheses represent the number of each subgenera used in this study.

Taxon	Species Number	Sample Size	Source
Order Perciformes			
Family Chaetodontidae			
Genus *Chaetodon*	84 (42)		
Subgenus *Chaetodon*	7 (0)		
Subgenus *Rabdophorus*	26 (15)		
C. mesoleucos		3	Aquarium
C. rafflesi		3	Aquarium
C. collare		3	Aquarium
C. flavirostris		3	Aquarium
C. auriga		6	North Taiwan
C. decussatus		3	Aquarium
C. vagabundus		8	Taiwan
C. ephippium		5	North Taiwan
C. ulietensis		2	Taiwan
C. lineolatus		5	Taiwan
C. oxycephalus		3	Aquarium
C. lunulatus		3	Aquarium
C. melannotus		5	Taiwan
C. ocellicaudus		3	Aquarium
C. semilarvatus		3	Aquarium
Subgenus *Roaops*	6 (1)		
C. tinkeri		3	Aquarium
Subgenus *Exornator*	22 (9)		
C. citrinellus		3	Aquarium
C. quadrimaculatus		3	Aquarium
C. argentatus		5	Taiwan
C. xanthurus		5	Taiwan
C. madagascariensis		3	Aquarium
C. paucifasciatus		5	Aquarium
C. punctatofasciatus		5	Aquarium
C. guttatissimus		3	Aquarium
C. pelewensis		5	Aquarium
Subgenus *Lepidochaetodon*	3 (3)		
C. unimaculatus		3	Aquarium
C. kleinii		3	Taiwan
C. trichrous		3	Aquarium
Subgenus *Megaprotodon*	2 (1)		
C. trifascialis		3	Aquarium
Subgenus *Gonochaetodon*	3 (2)		
C. baronessa		3	Taiwan
C. larvatus		3	Aquarium
Subgenus *Tetrachaetodon*	4 (2)		
C. plebeius		3	Taiwan
C. speculum		3	Taiwan
Subgenus *Discochaetodon*	4 (3)		
C. aureofasciatus		3	Aquarium
C. octofasciatus		5	Taiwan
C. rainfordi		3	Aquarium
Subgenus *Corallochaetodon*	4 (3)		
C. austriacus		3	Aquarium
C. trifasciatus		3	Aquarium
C. lunulatus		3	Aquarium
Subgenus *Citharoedus*	3 (3)		
C. meyeri		3	Aquarium
C. ornatissimus		3	Aquarium
C. reticluatus		3	Aquarium

RESULTS

Among a total of 670 bp of aligned nucleotide sequences (324 bp of cyt *b*, 346 bp of 12S rRNA), 233 sites were variable, with 219 being phylogenetically informative (121 bp of cyt *b*, 98 bp of 12S rRNA). For cyt *b*, the average frequencies of nucleotides for all taxa are as follows: A = 23.8%; C = 28.1%; G = 17.3%; T = 30.8%. This partial cyt *b* gene of the genus *Chaetodon* is A + T biased (54.6%) and the average TS / TV ratio across all pairwise sequence comparisons in the data set was 3.2. This level of transition bias of cyt *b* is within the range of biases previously reported for other vertebrates and serves as a basis for the transition / transversion weighting ratios used in phylogenetic reconstruction. For 12S rRNA, the average frequencies of nucleotides for all taxa are as follows: A = 30.2%; C = 25.9%; G = 22.2%; T = 21.7%. This partial 12S rRNA gene of the genus *Chaetodon* is A + T biased (51.9%) and the average TS / TV ratio across all pairwise sequence comparisons in the data set was 2.3. The average TS / TV ratio suggests that variations at most sites for cyt *b* and 12S rRNA have not reached saturation. Third codon positions of cyt *b* and 12S rRNA sequences were retained in all analyses.

The phylogenetic trees were reconstructed by four methods (MP, ME, NJ and ML) for cyt *b*, 12S rRNA and combined cyt *b* + 12S rRNA genes. The MP, ME and NJ trees have similar topology. The data fully supports the monophyletic grouping of all the individuals representing the same species, except for the *C. punctatofasciatus* species group (Fig. 1). Hybridization between closely-related *Chaetodon* species has long been suspected. For example, McMillan et al. (1999) proposed overwhelming evidence from behavioural, genetic and phenotypic between *C. pelewensis* and *C. punctatofasciatus*, making it impossible to justify their continued species-level status.

Figure 1 represents reconstructions of the phylogeny of the genus *Chaetodon* inferred from 12S rRNA (Fig. 2A), cyt *b* (Fig. 2B) and combined trees (Fig. 2C) by the NJ method

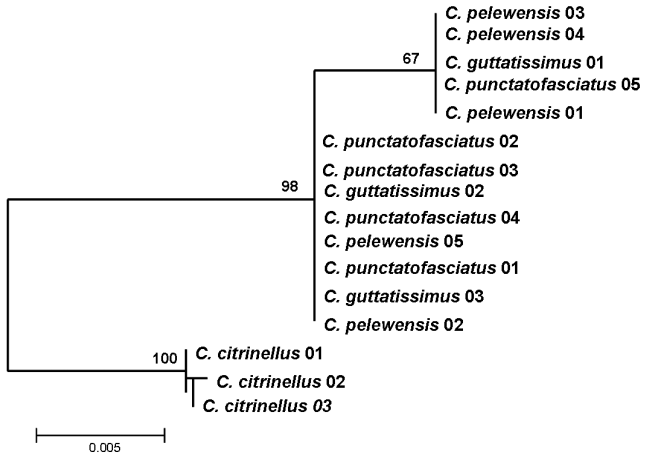

Fig. 1. Neighbor-joining phylogenetic tree based on the partial *cyt* b gene for the haplotypes of *C. punctatofasciatus* species group with *C. citrinellus* as outgroup taxa. Numbers above branches indicate bootstrap values based on 1,000 replications.

(MP and ME trees are not shown). However, trees obtained from three runs of the 12S rRNA and cyt *b* data were not identical and showed large differences in topology in the form of poorly supported branches. According to the phylogenetic trees inferred from 12S rRNA, the subgenera *Rabdophorus* and *Corallochaetodon* are polyphyletic groups while the others are poorly supported. Nevertheless, the trees obtained by analyses of cyt *b* sequences and combined cyt *b* + 12S rRNA were very similar. The subclade, including *Lepidochaetodon*, *Corallochaetodon* and *Citharoedus,* was the most basal subgenus. Based on cyt *b* and combined trees, all the nine *Chaetodon* subgenera sampled belonged to a monophyletic group. However, their phylogenetic relationships differed significantly from those inferred from morphological characters. Based on the ML phylogenetic tree (Fig. 3), *Lepidochaetodon* is the basal subgenus followed by *Exornator* and others. Molecular data support that *Corallochaetodon* and *Citharoedus* were sister groups. However, this data set could not resolve the relationships among subgenera within the genus *Chaetodon*, as the relationships among the subgenera are not well supported by bootstrapping.

In this study, we examined 11 possible species pairs of the genus *Chaetodon* (Table 2), as proposed by previous literature (Burgess, 1978; Steene, 1978; Allen, 1979; Blum, 1989; Kuiter, 1995; Allen et al., 1998). Among these species pairs, four pairs are sympatrically distributed, three pairs are parapatrically distributed and the remaining four pairs are allopatrically distributed (Table 2). The geographical distributions of *Chaetodon* species are shown together within their phylogenetic relationships in Figure 5. The pairwise sequence divergences of the sympatric (average = 5.3%) and parapatric (average = 4.3%) distribution species pair are higher than the allopatric distribution (average = 1.5%). We calculated the separation times for these grouped species pairs. Separation times have been implied by using a divergence rate of 2% per million years (Johns & Avise, 1998) for this data set. The separation time between sympatric pairs will be approximately 1.5 to 4.0 million years ago (mya) and similar to the time between parapatric species pairs. Separation times between allopatrically distributed species pairs ranged from 0.5 to 1.0 mya (Table 2). Consequently, based on the speciation time of the species pairs of allopatric distribution being more recent than that of the sympatrically- and parapatrically distributed species pairs, the phylogeography of coral reef fishes in the IWP may have experienced vicariance. In this study, the relative rate tests via pairwise comparisons of mtDNA sequences (using *Chelmon rostratus* as an outgroup) revealed that most lineages were not consistent with the molecular clock.

DISCUSSION

Before reconstructing the phylogenetic relationships, it is necessary to choose an applicable molecular tool. In this study, we use cyt *b*, 12S rRNA and combined trees as phylogenetic trees to confer the evolution within the genus *Chaetodon*. The topologies of the cyt *b* and combined trees

Table 2. The pairwise divergence values of possible species pairs examined in this study. Genetic distances calculations were made by the Kimura's two-parameter model. Standard errors were also indicated.

Species Pair	Distribution Pattern	Reference	Pairwise Divergence
C. oxycephalus / C. lineolatus	Sympatric	Burgess, 1978, Allen, 1979	0.03 ± 0.01
C. vagabundus / C. decussatus	Sympatric	Burgess, 1978, Allen, 1979	0.05 ± 0.01
C. melannotus / C. ocellicaudus	Sympatric	Burgess, 1978, Allen, 1979	0.05 ± 0.01
C. xanthurus / C. argentatus	Sympatric	Blum, 1988	0.08 ± 0.01
C. meyeri / C. ornatissimus	Parapatric	Allen et al., 1998	0.05 ± 0.01
C. pelewensis / C. punctatofasciatus	Parapatric	Steen, 1978	0.01 ± 0.01
C. lunulatus / C. trifasciatus	Parapatric	Kuiter, 1995	0.07 ± 0.02
C. xanthurus / C. madagascariensis	Allopatric	Allen et al., 1998	0.02 ± 0.01
C. kleinii / C. trichrous	Allopatric	Allen et al., 1998	0.01 ± 0.01
C. guttatissimus / C. punctatofasciatus	Allopatric	Blum, 1988	0.02 ± 0.01
C. paucifasciatus / C. madagascariensis	Allopatric	Blum, 1988	0.01 ± 0.01

were very similar but the 12S rRNA topology differed significantly. The 12S rRNA tree showed that some subgenera of *Chaetodon* were polyphyletic groups and it also showed that the genus *Chaetodon* was not a monophyletic group. 12S rRNA provided considerably less resolution with poor nodal supports, relatively shorter internal branch lengths and it is unlikely that the gene will provide a cost-effective marker in further resolving the *Chaetodon* phylogeny. Furthermore, the secondary structure of RNA is strongly conserved over long time periods and natural selection is acting to maintain a structure that is essential for the function of these molecules (Nicholas et al., 2001). We conclude that the 12S rRNA phylogenetic tree of *Chaetodon* was not applicable to investigate the evolution of *Chaetodon*. Some studies (e.g. Rosenberg, 2002; Dettman et al., 2003; Slatkin & Pollack, 2006) on the concordance of gene trees and species trees have indicated that the correct species tree may be inferred if several unlinked genes are studied. Therefore, we used the combined phylogenetic tree (Fig. 3) to investigate the evolution in geminate species pairs and species groups of *Chaetodon*.

There are many studies about the *Chaetodon* phylogeny based on morphology and molecular data (e.g., Blum, 1988; Smith et al., 2003; Littlewood et al., 2004). In the absence of members of the subgenera *Roaops* and *Chaetodon*, the subgenus *Lepidochaetodon* was the most basal clade and confirmed the close relationship between *C. kleinii* and *C. trichrous*, as discussed by Blum (1988). The close relationship between *Lepidochaetodon* and *Exornator* is also well supported by morphological analyses (Blum, 1988; Smith et al., 2003). The sister group relationship between the monophyletic subgenera *Corallochaetodon* and *Citharoedus* was confirmed with molecular data in our study and Littlewood et al.'s hypothesis (2004). The resolution of *Tetrachaetodon* and *Discochaetodon* as sister taxa were also supported by morphological data (Littlewood et al., 2004). According to the previous studies described above, different analysis with different samples yielded different topologies. Therefore, wider taxonomic sampling remains preferable. Consequently, we would like to perform a re-analysis of *Chaetodon* phylogeny in the future.

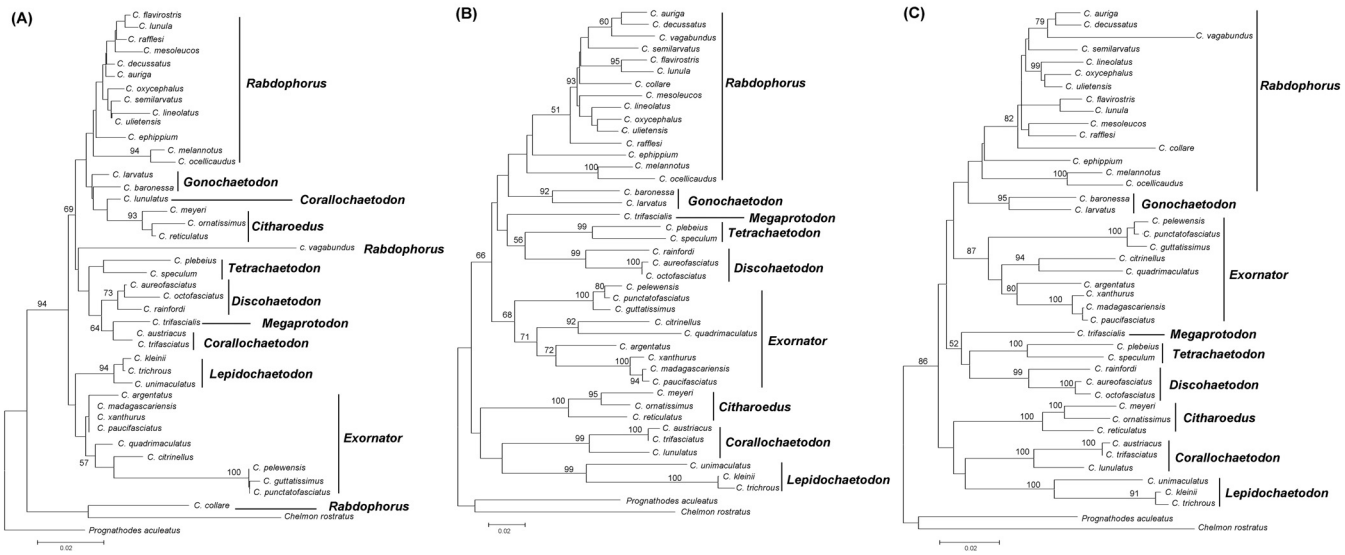

Fig. 2. Neighbor-joining tree of the genus *Chaetodon* inferred from the DNA sequence variations of A) 12S rRNA, B) cyt *b*, C) combined 12S rRNA and cyt *b*. Numbers at nodes indicate the bootstrap values estimated from 1,000 replications.

The Indo-Pacific region is a place where speciation may have occurred because of a barrier to the East-West dispersal of marine fishes. This is a result of the lowering of sea-levels during glacial periods. Within the Indo-West Pacific, vicariance has been invoked to explain pairs of sister species in the Indian and Pacific Oceans (McMillan & Palumbi, 1995; Williams & Benzie, 1998; Benzie, 1999). However, the number, location and timing of these vicariance events in the Indian and Pacific Oceans were unknown.

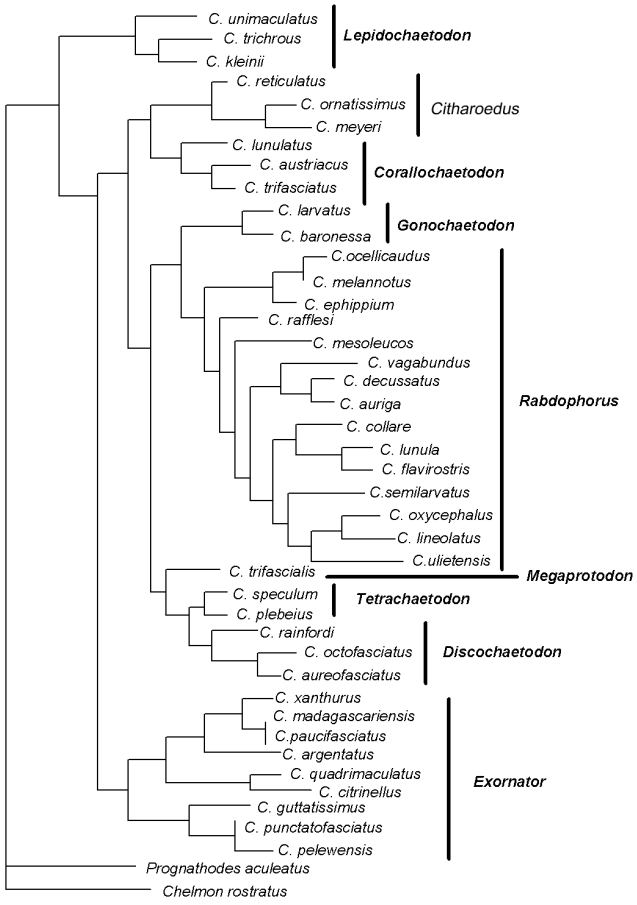

Fig. 3. A combined molecular phylogenetic analysis of *Chaetodon* from a maximum likelihood (ML) analysis of 12S rRNA and cyt *b* sequences.

Blum (1989) infers 12 barriers based on the distributions of *Chaetodon* species complexes and suggests that these barriers may be sufficient to isolate the species (Fig. 4). These barriers might influence the dispersal and speciation mode of *Chaetodon*. In this study, most geminate species pairs were distributed on each side of these barriers. We suggest that these barriers are vicariance events that disturbed the dispersal of the *Chaetodon* species. We suggest that the fall in sea levels associated with glacial maxima resulted in a massive loss of shallow inner reefs and lagoons. It may have exposed large areas of continental shelf and reduced reef habitats to narrow fringes, or, may even have increased it in some areas by exposing submerged islands. Thus, habitat loss and increased/reduced circulation have all contributed to the isolation between the adjacent regions. There is evidence that climatic oscillations during the late Pliocene and Pleistocene have been important causes of marine speciation, especially in the Indo-West Pacific (Palumbi, 1997; Benzie, 1999). Molecular phylogenetic studies of small, closely related clades also show that sister species diverged in the last one to five million years (McCartney et al., 2000; Williams, 2000; Lessios et al., 2001; Harrison, 2004). In *Chaetodon*, divergence times between species pairs were mostly 0.5 to 5 mya (Table 2). However, we could not estimate the evolutionary rates for *Chaetodon* by molecular tools as the Tajima's relative rates test (Tajima, 1989) rejected clock-like behaviour.

The previous study on the rate of control-region evolution in Pacific butterflyfishes also supports that the evolutionary rate

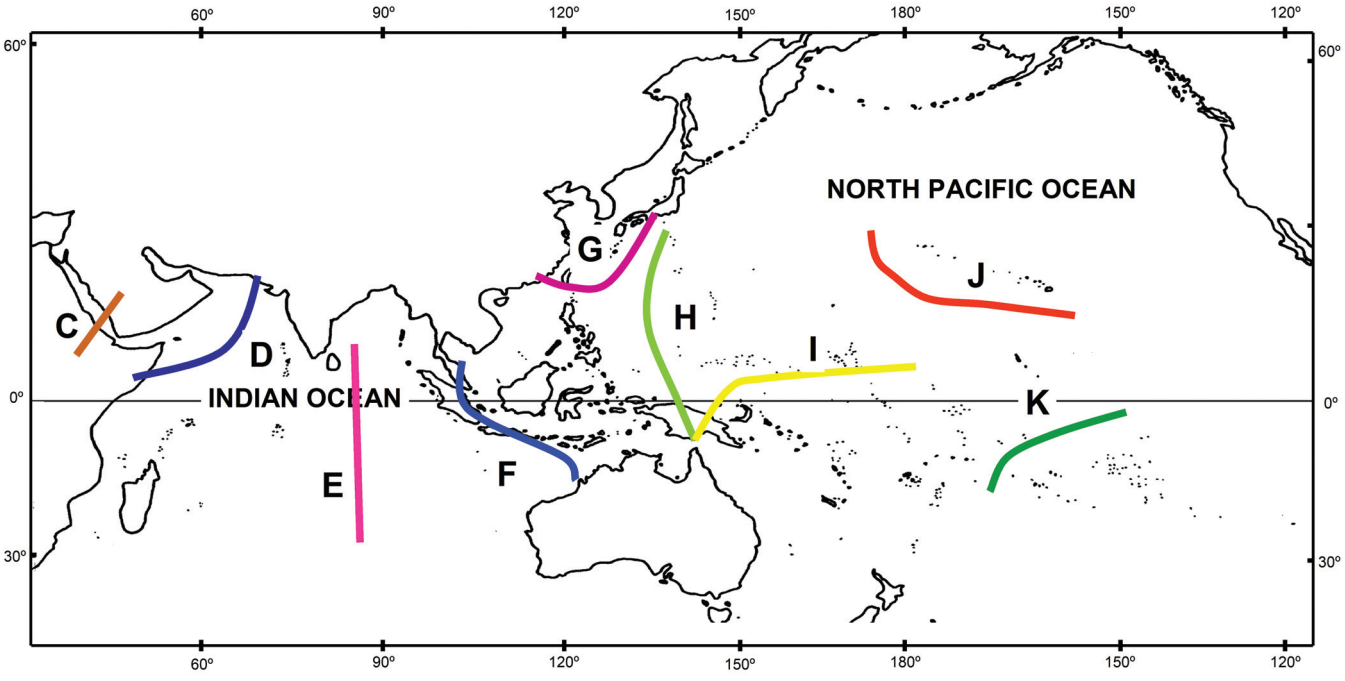

Fig. 4. Barriers (C – K) inferred from the distribution of species within complexes by Blum (1989).

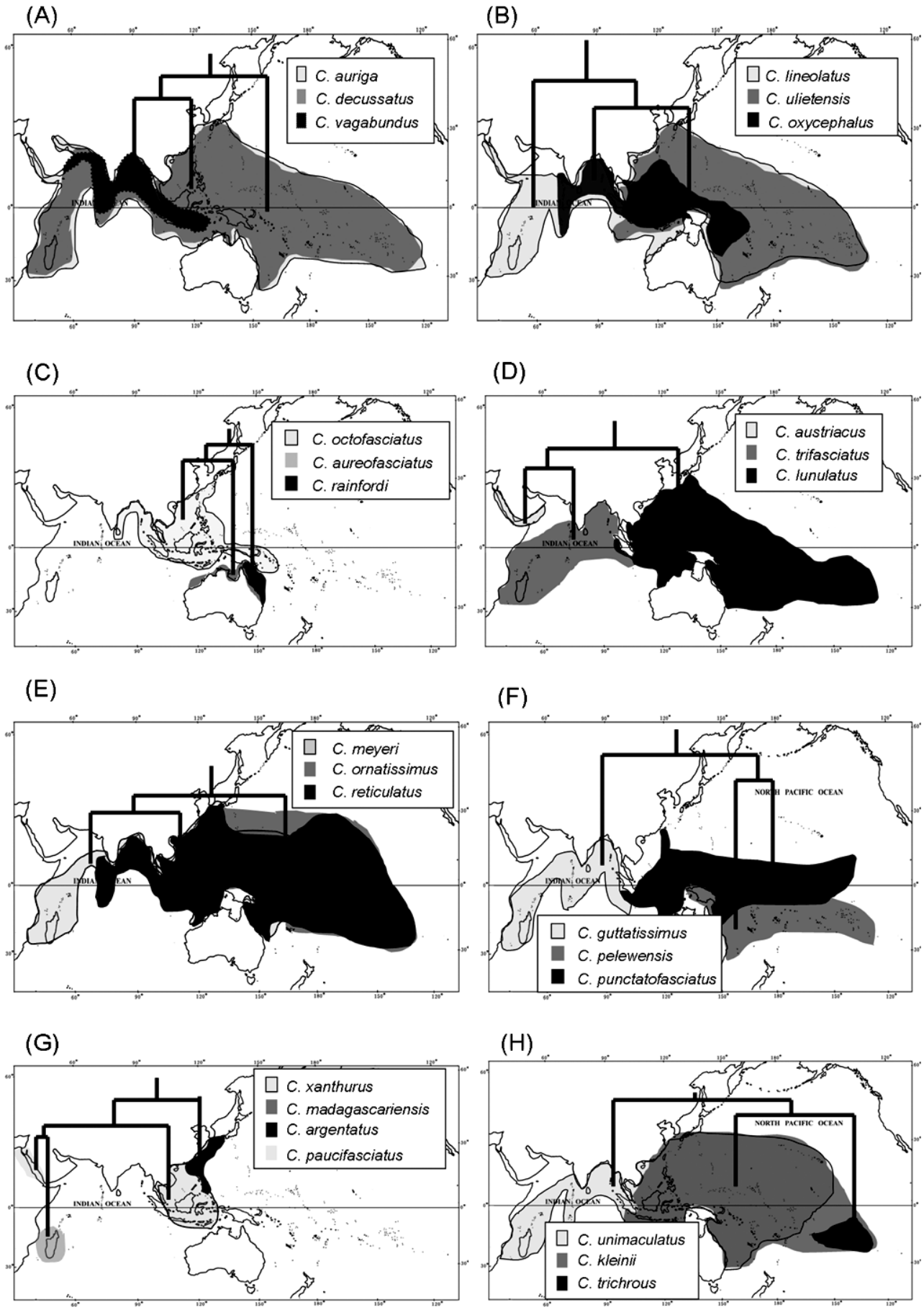

Fig. 5. Distributions and phylogenetic relationships based on the combined tree of the *Chaetodon* species A) *C. auriga* spp. group, B) *C. lineolatus* complex, C) *Discochaetodon* subgenus, D) *Corallochaetodon* subgenus, E) *Citharoedus* subgenus, F) *C. punctatofasciatus* spp. group, G) *Rhombochaetodon* spp. group, H) *Lepidochaetodon* subgenus.

may be different between clades and the evolutionary rate of the mitochondrial D-loop of *Chaetodon* is rapid (McMillan & Palumbi, 1997). Hence, we could not confirm that the times estimated were valid. However, in terms of the 12 barriers (Fig. 4; Blum, 1989), we find different barriers affect different species. For example, barrier F limits the distribution of *C. baronessa* and *C. unimaculatus* while barriers G, H and I do not. On the other hand, the distribution of *C. speculum* was limited by barriers F, G, H and I. According to these results, we suggest that barrier F formed before *C. speculum*, *C. baronessa* and *C. unimaculatus* speciated and barriers G, H and I were formed before *C. speculum* speciated but after *C. baronessa* and *C. unimaculatus* speciated and dispersed. These results also suggest that lowering of sea levels during Pleistocene glaciations may have occurred many times for different barriers and speciations.

Cases of allopatric sister species in which their ranges are strongly asymmetric in size and separated by wide areas of inhospitable habitats are examples of 'classical founder speciation' (Paulay & Meyer, 2002). Endemic species on islands are likely examples of founder speciation. In the Indo-West Pacific, these include *C. austriacus*, *C. madagascariensis*, *C. paucifasciatus* and *C. trichrous* (Fig. 5D, G & H). In the Indo-West Pacific, the geography of speciation events has been repeatedly discussed in explanations of the focus of diversity in the East Indies Triangle between the Philippines, Indonesia and New Guinea (Palumbi, 1997). According to the center-of-origin hypothesis, species originate within this central region and then expand their range outwards (Briggs, 1999, 2000, 2003). In contrast, the central region can be viewed as a 'center of accumulation' for species that originated at the periphery of the Indo-West Pacific (Jokiel & Martinelli, 1992). These two hypotheses make different predictions about the location of recently-formed species and their direction of spread.

In *Chaetodon*, we have found evidence of speciation in both the Central Indo-West Pacific and at its periphery. A number of narrowly-endemic species were distributed around the periphery of the Indo-West Pacific, but not all the species can be attributed to recent speciation events. Some species (*C. pelewensis*, Fig. 5F; *C. trichrous*, Fig. 5H) are terminal branches of the phylogeny and are therefore the products of relatively recent divergence, as predicted by the center-of-accumulation hypothesis. However, the others are relatively basal branches of the phylogeny (*C. auriga*, Fig. 5A; *C. lineolatus*, Fig. 5B; *C. lunulatus*, Fig. 5D; *C. unimaculatus*, Fig. 5H) and represent ancient speciation events or even peripheral relicts (as predicted by the center-of-origin idea). We have no evidence of range expansion by peripheral species toward the center of the Indo-West Pacific. Endemic species are also present in the Central region of Indo-West Pacific and again, some are relatively basal (*C. xanthurus*, Fig. 5G) as predicted by the center-of-origin hypothesis, whereas others are terminal (*C. octofasciatus*, Fig. 5C). In summary, evidence from *Chaetodon* adds to that from other studies and reviews (Palumbi, 1996, 1997; Palumbi et al., 1997; Paulay, 1997; Bellwood & Wainwright, 2002; Bernardi et al., 2004; Williams & Reid, 2004) that suggest that speciation occurred throughout the Indo-West Pacific, as a result of diverse mechanisms.

The processes in evolutionary history are not only speciations, but also extinctions. In the absence of a good fossil record, there is no direct evidence of extinction in the genus *Chaetodon*. In general, post-Miocene extinction is believed to have been lower in the Indo-West Pacific than in the other regions of the marine tropics (Vermeij, 1989; Paulay, 1997; Meyer, 2003). Pliocene extinctions were most severe in the Western Atlantic and Eastern Pacific where 32% and 15% of genera went extinct, respectively, compared to none in the Indo-West Pacific (Vermeij, 1989).

However, at least one pattern within the subgenus *Exornator* (*C. fremblii* spp. group.) does imply extinction. There is great geographical distance between *C. fremblii* in Hawaii and its sister clade, *C. blackburnii*, in East-South Africa and the distance is too far to be explained by dispersal. Another similar pattern happens within the subgenus *Chaetodon*. The great distance between *C. marleyi* in East-South Africa and its closely-related species in Western Atlantic also implies extinction. According to these two patterns described above, the distance between the two sister species are too great for dispersal. Thus, we suggest that there were extinction events that happened in the genus *Chaetodon*.

ACKNOWLEDGMENTS

This research was supported by grants from Academia Sinica. The authors would like to thank the members of the laboratory of Fish Ecology & Evolution, Research Center for Biodiversity, Academia Sinica for help in sampling. We would also like to thank the two anonymous reviewers for comments that greatly improved the manuscript.

LITERATURE CITED

Allen, G. R., 1979. *Butterfly and angelfishes of the world*. Vol. 2. J. Wiley, New York & Toronto. 352 pp.

Allen, G. R., R. Steene & M. Allen, 1998. *A guide to angelfishes and butterflyfishes*. Odyssey, WA. 250 pp.

Bellwood, D. R. & P. C. Wainwright, 2002. The history and biogeography of fishes on coral reefs. In: Sale, P. F. (ed.), *Coral reef fishes: dynamics and diversity in a complex ecosystem*. Academic Press, San Diego. Pp. 5-32.

Benzie, J. A. H., 1999. Genetic structure of coral reef organisms: ghosts of dispersal past. *American Zoology*, **39**: 131-145.

Bernardi, G., G. Bucciarelli, D. Costagliola, D. R. Robertson, & J. B. Heiser, 2004. Evolution of coral reef fish *Thalassoma* spp. (Labridae). 1. Molecular phylogeny and biogeography. *Marine Biology*, **144**: 369-375.

Blum, S. D., 1988. *The Osteology and Phylogeny of the Chaetodontidae (Teleostei: Perciformes)*. PhD. Dissertation. University of Hawaii, Manoa. 365 pp. Available from: http://www.calacademy.org/research/informatics/sblum/pub/.

Blum, S. D., 1989. Biogeography of Chaetodontidae: an analysis of allopatry among closely related species. *Environmental Biology of Fishes*, **25**: 9-31.

Briggs, J. C., 1999. Coincident biogeographic patterns: Indo-West Pacific Ocean. *Evolution*, **53**: 326-335.

Briggs, J. C., 2000. Centrifugal speciation and centres of origin. *Journal of Biogeography*, **27**: 1183-1188.

Briggs, J. C., 2003. Marine centres of origin as evolutionary engines. *Journal of Biogeography*, **30**: 1-18.

Burgess, W. E., 1978. *Butterflyfishes of the world, a monograph of the family Chaetodontidae*. T. F. H. Publications, Neptune City. 832 pp.

Chen, C. L., K. H. Pollock & J. M. Hoenig, 1998. Combining change-in-ratio, index-removal, and models for estimating population size. *Biometrics*, **54**:815-827.

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

Colborn, J., R. E. Crabtree, J. B. Shaklee, E. Pfeiler & B. W. Bowen, 2001. The evolutionary enigma of bonefishes (*Albula* spp.): cryptic species and ancient separations in a globally distributed shorefish. *Evolution*, **55**: 807-820.

Dettman, J. R., D. J. Jacobson & J. W. Taylor, 2003. A multilocus genealogical approach to phylogenetic species recognition in the model eukaryote *Neurospora*. *Evolution*, **57**: 2703-2720.

Duda, T. & S. Palumbi, 1999. Population structure of the black tiger prawn, *Penaeus monodon*, among western Indian Ocean and western Pacific population. *Marine Biology*, **134**: 705-710.

Felsenstein, J., 1985. Confidence limits on phylogenies: an approach using the bootstrap. *Evolution*, **39**: 783-791.

Felsenstein, J., 1993. *PHYLIP (Phylogeny Inference Package) Version 3.6a2*. Department of Genetics, University of Washington, Seattle.

Harrison, J. S., 2004. Evolution, biogeography and the utility of mitochondrial 16S and COI genes in phylogenetic analysis of the crab genus *Austinixa* (Decapoda: Pinnotheridae). *Molecular Phylogenetics and Evolution*, **30**: 743-754.

Håstad, O. & M. Björklun, 1998. Nucleotide substitution models and estimation of phylogeny. *Molecular Biology and Evolution*, **15**: 1381-1389.

Hellberg, M. E., 1998. Sympatric seashells along the sea's shore: the geography of speciation in the marine gastropod *Tegula*. *Evolution*, **52**: 1311-1324.

Hillis, D. M. & J. J. Bull, 1993. An empirical test of bootstrapping as a method for assessing confidence in phylogenetic trees. *Systematic Biology*, **42**: 182-192.

Higgins, D. G., A. J. Bleasby & R. Fuchs, 1992. CLUSTAL V: improved software for multiple sequence alignment. *CABIOS*, **8**: 189-191.

Johns, G. C. & J. C. Avise, 1998. A comparative summary of genetic distances in the vertebrates from the mitochondrial cytochrome b gene. *Molecular Biology Evolution*, **15**: 1481-1490.

Jokiel, P. & F. J. Martinelli, 1992. The vortex model of coral reef biogeography. *Journal of Biogeography*, **19**: 449-458.

Kimura, M., 1980. A simple method for estimating evolutionary rate of base substitution through comparative studies of nucleotide sequences. *Journal of Molecular Evolution*, **16**: 111-120.

Kuiter, R. H., 1995. *Chaetodon lunulatus*, a sibling species of *C. trifasciatus*, with observations on other sibling species of butterflyfish (Chaetodontidae). *Reviews Françoise Aquariology*, **21**(3-4): 105-106.

Kumar, S., K. Tamura, I. B. Jakobsen & M. Nei, 2001. *MEGA2: Molecular Evolutionary Genetics Analysis Software*. Arizona State University, Tempe.

Lessios, H. A., B. D. Kessing & J. S. Pearse, 2001. Population structure and speciation in tropical seas: global phylogeography of the sea urchin *Diadema*. *Evolution*, **55**: 955-975.

Littlewood, D. T. J., S. M. McDonald, A. C. Gill & T. H. Cribb, 2004. Molecular phylogenetics of *Chaetodon* and the Chaetodontidae (Teleostei: Perciformes) with reference to morphology. *Zootaxa*, **779**: 1-20.

McCartney, M. A., G. Keller & H. A. Lessios, 2000. Dispersal barriers in tropical oceans and speciation in Atlantic and eastern Pacific sea urchins of the genus *Echinometra*. *Molecular Ecology*, **9**: 1391-1400.

McMillan, W. O. & S. R. Palumbi, 1995. Concordant evolutionary patterns among Indo-west Pacific butterflyfishes. *Proceeding of the Royal Society B-Biological Sciences,* **260**: 229-236.

McMillan, W. O. & S. R. Palumbi, 1997. Rapid rate of control-region evolution in Pacific butterflyfishes (Chaetodontidae). *Journal of Molecular Evolution,* **45**: 473-484.

McMillan, W. O., L. A. Weigt & S. R. Palumbi, 1999. Color pattern evolution, assortative mating and genetic differentiation in brightly colored butterflyfishes (Chaetodontidae). *Evolution,* **53**: 247-260.

Meyer, C. P., 2003. Molecular systematics of cowries (Gastropoda: Cypraeidae) and diversification patterns in the tropics. *Biological Journal of the Linnaean Society*, **79**: 401-459.

Nicholas, J. S., D. C. Hoyle & P. G. Higgs, 2001. RNA sequence evolution with secondary structure constraints: comparison of substitution rate models using maximum-likelihood methods. *Genetics*, **157**: 399-411.

Palumbi, S. R., 1996. What can molecular genetics contribute to marine biogeography? An urchin's tale. *Journal of Experimental Marine Biology and Ecology*, **203**: 75-92.

Palumbi, S. R., 1997. Molecular biogeography of the Pacific. *Coral Reefs*, **16**(Supplement 1): S47-S52.

Palumbi, S. R., G. Grabowsky, T. Duda, L. Geyer & N. Tachino, 1997. Speciation and population genetic structure in tropical Pacific sea urchins. *Evolution*, **51**: 1506-1517.

Paulay, G., 1997. Diversity and distribution of reef organisms. In: Birkeland, C. (ed.), *Life and death of coral reefs*. Chapman and Hall, New York. Pp. 298-353.

Paulay, G. & C. Meyer, 2002. Diversification in the tropical Pacific: comparisons between marine and terrestrial systems and the importance of founder speciation. *Integral Computation Biology and Chemistry, Biology*, **42**: 922-934.

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

Rosenberg, N. A., 2002. The probability of topological concordance of gene trees and species trees. *Theoretical Population Biology*, 61: 225-247.

Sanderson, M. J., 2002. r8s: *Analysis of rates (''r8s'') of evolution (and other stuff), version 1.5*. Section of Evolution and Ecology, University of California, Davis. Distributed at http://ginger.ucdavis.edu/r8s/.

Slatkin, M. & J. L. Pollack, 2006. The concordance of gene trees and species trees at two linked loci. *Genetics*, **172**: 1979-1984.

Smith, W. L., J. F. Webb & S. D. Blum, 2003. The evolution of the laterophysic connection with a revised phylogeny and taxonomy of butterflyfishes (Teleostei : Chaetodontidae). *Cladistics*, **19**: 287-306.

Steene, R. C., 1978. *Butterfly and angelfishes of the world*. Vol. 1. Mergus Verlag, Melle. 144 pp.

Tamura, K. & M. Nei, 1993. Estimation of the number of nucleotide substitutions in the control region of mitochondrial DNA in humans and chimpanzees. *Molecular Biology and Evolution*, **10**: 512-516.

Tajima, F., 1989. Statistical method for testing the neutral mutation hypothesis by DNA polymorphism. *Genetics*, **123**: 585-595.

Vacquier, V. D., 1998. Evolution of gamete recognition proteins. *Science*, **281**: 1995-1998.

Vermij, G. J., 1989. Interoceanic differences in adaptation: effects of history and productivity. *Marine Ecology-progress Series*, **57**: 293-230.

Williams, S. T., 2000. Species boundaries in the starfish genus *Linckia*. *Marine Biology*, **136**: 137-148.

Williams, S. T. & J. A. H. Benzie, 1998. Evidence of a biogeographic break between populations of a high dispersal starfish: congruent regions within the Indo-West Pacific defined by color morphs, mtDNA and allozyme data. *Evolution*, **52**: 87-99.

Williams, S. T. & D. G. Reid, 2004. Speciation and diversity on tropical rocky shores: a global phylogeny of snails of the genus *Echinolittorina*. *Evolution*, **58**: 2227-2251.