Zootaxa 779: 1–20 (2004)
www.mapress.com/zootaxa/

ISSN 1175-5326 (print edition)
ZOOTAXA
ISSN 1175-5334 (online edition)

779

Molecular phylogenetics of *Chaetodon* and the Chaetodontidae (Teleostei: Perciformes) with reference to morphology

D. TIMOTHY J. LITTLEWOOD[1*], SARAH M. MCDONALD[1†], ANTHONY C. GILL[2] & THOMAS H. CRIBB[3]

[1]*Department of Zoology, The Natural History Museum, Cromwell Road, London SW7 5BD, England, UK*
[2]*School of Life Sciences, PO Box 874501, Arizona State University, Tempe, AZ 85287-4501, USA*
[3]*Department of Microbiology & Parasitology, The University of Queensland, Brisbane, Australia 4072*
†present address: Marine Botany Laboratory, Stazione Zoologica A. Dohrn, Villa Comunale 1, 80121 Napoli, Italy
**Author for correspondence (E-mail: T.Littlewood@nhm.ac.uk)*

Abstract

Butterflyfish are colourful, pan-tropical coastal fish that are important and distinctive members of coral reef communities. A successful systematic scheme and a robust phylogeny is considered essential in understanding further their biogeography and ecology, although recent cladistic treatments of butterflyfish phylogeny, based on soft tissue and bone morphology and coded at the generic and subgeneric levels, differ in character coding and subsequently tree topology. This study provides an independent test of the morphologically based hypotheses, using molecular systematic data from two partial mitochondrial gene fragments, cytochrome b (cytb) and small subunit rRNA (rrnS), for 52 ingroup chaetodontids and seven pomacanthids used to root the molecular trees. Individual gene trees were largely compatible and a combined molecular phylogeny, inferred from Bayesian analysis, was used to test alternative hypotheses suggested by morphological analyses. The tree was also used to map the latest morphological matrix in order to evaluate potential synapomorphies for various nodes defining butterflyfish interrelationships. A clade comprised of *Chelmon* and *Coradion* was sister group to other chaetodontids. *Heniochus* and *Hemitaurichthys* were each resolved as monophyletic groups, and as sister taxa Of the taxa sampled, *Prognothodes* was resolved as the sister genus to Chaeotodon. Of the ten *Chaetodon* subgenera sampled, all were monophyletic but their interrelationships differed significantly from that inferred from morphological characters. *Lepidochaetodon* was the most basal subgenus followed by *Exornator* and the remaining subgenera. Molecular data support the sister group relationship between *Corallochaetodon* and *Citharoedus* suggested by morphology, but major differences occur among the remaining more derived taxa. *Chaetodon trifascialis* and *C. oligacanthus* were resolved as sister taxa adding weight to the inclusion of the latter in *C. Megaprotodon*. Of those pairs of taxa known to hybridize and sampled with molecular data, all were closely related phylogenetically, except those hybrids known to occur in the *Rabdophorus* subgenus. Two base changes separated *C. pelewensis* from *C.*

Accepted by L. Page: 7 Dec. 2004; published: 20 Dec. 2004

paucifasciatus which have been regarded previously as a single species. Cytb provided greater resolution than rrnS and will likely provide additional resolution with greater taxon sampling.

Key words: butterflyfish, phylogeny, molecular systematics, mitochondrial gene

Introduction

The Chaetodontidae, or butterflyfishes, contain 11 genera with over 130 species (Kuiter, 2002), of which the majority occur in tropical reef environments across the globe. The group are characteristically colourful and get their name (chaetodont = "bristle-tooth") from the distinctive fine jaw teeth. Other defining features include a deeply compressed body and a small protractile mouth with brush-like teeth. Although individual species are often very distinctive, hybrids are known to occur (Randall *et al*., 1977; Allen *et al*. 1998). The interrelationships of the family have been estimated cladistically based on osteological and internal soft anatomy by Blum (1988, 1989) and other workers recoding or revising Blum's original matrix (Ferry-Graham *et al.*, 2001a; Smith *et al.*, 2003). As a pantropical family, closely associated with reef environments, frequently territorial and 'home-ranging' with male-female pair bonds that may last for life in some species (e.g. Driscoll and Driscoll, 1988), larval strategies that allow varying degrees of dispersal and with jaw morphologies that appear to reflect diet (e.g. Motta, 1988, 1989; Ferry-Graham *et al.*, 2001b) and feeding habit (Sano, 1989), the group has fascinated ichthyologists, ecologists, behavioural biologists and biogeographers. A robust phylogeny for the group has been recognized as an important foundation for comparative studies and an extensive morphological assessment was undertaken by Blum (1988) who developed a number of species groups (originally proposed by Burgess, 1978) based on shared characters. Here we take the opportunity to test phylogenies based on morphology with molecular systematic data derived from two mitochondrial genes that have been determined for larval identification (GenBank, unpublished sequences from J.S. Nelson and colleagues, University of Singapore). The genes characterized for larval identification are both mitochondrial; partial small subunit (12S) ribosomal RNA (rrnS) genes and partial cytochrome b (cytb) genes. These same markers have been used widely for phylogeny in establishing relationships among a variety of fish taxa and for a variety of divergence times; e.g. Perdices *et al.* (2002), Chen *et al.* (2003), Near *et al.* (2003), Ruber *et al.* (2003). Sampling in the present study is not exhaustive but the aims include determining the utility of the mitochondrial genes for a possible wider phylogenetic assessment of the Chaetodontidae, the validity (monophyly) of subgenera suggested by Burgess (1978) and developed by Blum (1988, 1989), the sister group relationships of these subgenera suggested by Blum's cladistic analysis (Blum, 1988), and the reinterpretation of these data by Ferry-Graham *et al.* (2001a) and Smith *et al.* (2003). Figure 1 illustrates the strict-consensus solutions of Blum's analysis [Fig. 1a] and that of Ferry-Graham *et al.* [Fig. 1b] after character modifi-

cation (see Ferry-Graham *et al.* 2001a). The latter solution is only marginally more resolved but the overall topologies are very similar. Detailed discussions on character coding differences between these studies and the latest assessment [Fig. 1c] may be found in Smith *et al.* (2003).

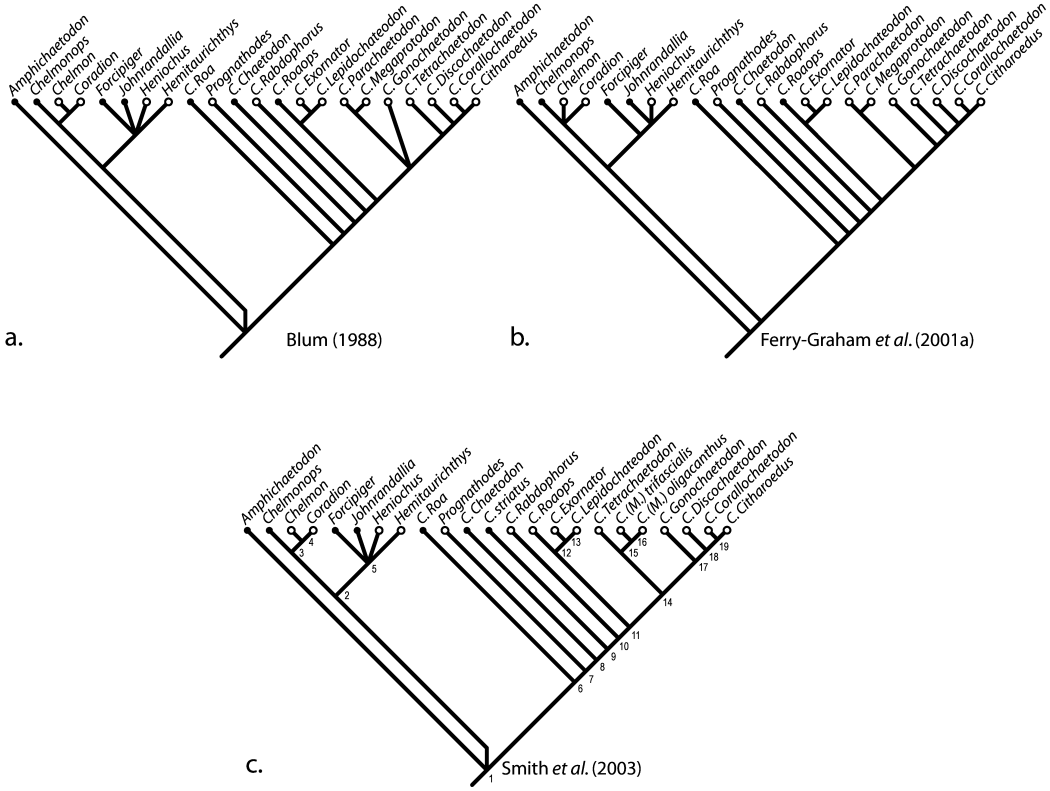

FIGURE 1. Phylogenies of the Chaetodontidae determined using morphological characters. a. Blum (1988) produced the first analysis based largely on osteology and included representatives of all nominal genera and subgenera established prior to 1988; b. Ferry-Graham *et al.* (2001a) presented a new analysis based on a modification of Blum's matrix; c. Smith *et al.* (2003) revised the original matrix of Blum (1988) in the light of Ferry-Graham *et al.* (2001a) and provided 41 osteological and soft-tissue characters – numbered nodes are those used in Table 3. Open circles indicate taxa sampled in the present study.

Materials and Methods

Data from partial fragments of two mitochondrial genes and from 59 taxa were used. Partial small subunit ribosomal (rrnS) and cytochrome b (cytb) sequences were downloaded from GenBank for 43 ingroup taxa, and the outgroups. Although available on GenBank, these sequences do not appear to have been used for systematic studies or used in the liter-

ature. Freshly collected and archival material was also sequenced; see Table 1 for full details. Sequences from 52 species of Chaetodontidae were utilized and were rooted against seven selected sequences of Pomacanthidae, the family considered to be the sister group to the Chaetodontidae (Burgess, 1978). An additional nine species were sampled and sequenced (AJ748297-AJ748314) as described below.

TABLE 1. Taxonomic listing of species analysed. Sequences new to the present study are indicated by '§'. Accession numbers for rrnS sequences for *Heniochus acuminatus* and *Chaetodon collare* have been transposed in the Table and are emboldened; original GenBank accessions are AF108543 and AF108511 for *H. acuminatus* and *C. collare* respectively (see text for explanation).

			mt DNA	
Classification		Species and Authority	cytb	rrnS
outgroup Pomacanthidae				
		Chaetodontoplus caeruleopunctatus Yasuda & Tominaga, 1976	AF108640	AF108565
		Chaetodontoplus duboulayi (Günther, 1867)	AF108641	AF108566
		Chaetodontoplus melanosoma (Bleeker, 1853)	AF108642	AF108567
		Chaetodontoplus meredithi Kuiter, 1989	AF108643	AF108568
		Pomacanthus imperator (Bloch, 1787)	AF108647	AF108572
		Centropyge bispinosus (Günther, 1860)	AF108627	AF108552
		Genicanthus melanospilos (Bleeker, 1857)	AF108645	AF108570
ingroup Chaetodontidae				
Genus	Subgenus[1]			
Chelmon		*Chelmon rostratus* (Linnaeus, 1758)	AF108612	AF108537
Coradion		*Coradion altivelis* McCulloch, 1916	AF108613	AF108538
		Coradion chrysozonus (Cuvier, 1831)	AF108614	AF108539
Hemitaurichthys		*Hemitaurichthys polylepis* (Bleeker, 1857)	AF108616	AF108541
		Hemitaurichthys zoster (Bennett, 1831)	AF108617	AF108542
Heniochus		*Heniochus acuminatus* (Linnaeus, 1758)	AF108618	**AF108511**
		Heniochus chrysostomus Cuvier, 1831	AF108619	AF108544
		Heniochus pleurotaenia Ahl, 1923	AF108620	AF108545
		Heniochus varius (Cuvier, 1829)	AF108621	AF108546
Prognathodes		*Prognathodes aculeatus* Poey, 1860	AF108579	AF108504
Chaetodon	*Chitharoedus*	*Chaetodon meyeri* Bloch & Schneider, 1801	AF108597	AF108522
		Chaetodon ornatissimus Cuvier, 1831	AF108600	AF108525
		Chaetodon reticulatus Cuvier, 1831	AJ748297§	AJ748306§
	Corallochaetodon	*Chaetodon austriacus* Rüppell, 1836	AF108583	AF108508
		Chaetodon lunulatus Quoy & Gaimard, 1825	AJ748298§	AJ748307§
		Chaetodon trifasciatus Park, 1797	AF108608	AF108533

.....continued on the next page

TABLE 1 continied

Classification		Species and Authority	mt DNA	
			cytb	rrnS
	Discochaetodon	*Chaetodon aureofasciatus* Macleay, 1878	AF108581	AF108506
		Chaetodon octofasciatus Bloch, 1787	AF108599	AF108524
		Chaetodon rainfordi McCulloch, 1923	AF108605	AF108530
	Exornator	*Chaetodon argentatus* Smith & Radcliffe, 1911	AF108580	AF108505
		Chaetodon citrinellus Cuvier, 1831	AF108585	AF108510
		Chaetodon guttatissimus Bennett, 1833	AF108590	AF108515
		Chaetodon madagascariensis Ahl, 1923	AJ748299§	AJ748308§
		Chaetodon paucifasciatus Ahl, 1923	AJ748300§	AJ748309§
		Chaetodon pelewensis Kner, 1868	AJ748301§	AJ748310§
		Chaetodon punctatofasciatus Cuvier, 1831	AF108603	AF108528
		Chaetodon quadrimaculatus Gray, 1831	AJ748302§	AJ748311§
		Chaetodon xanthurus Bleeker, 1857	AF108611	AF108536
	Gonochaetodon	*Chaetodon baronessa* Cuvier, 1829	AF108584	AF108509
		Chaetodon larvatus Cuvier, 1831	AF108592	AF108517
	Lepidochaetodon	*Chaetodon kleinii* Bloch, 1790	AF108591	AF108516
		Chaetodon trichrous Günther, 1874	AJ748303§	AJ748312§
		Chaetodon unimaculatus Bloch, 1787	AJ748304§	AJ748313§
	Megaprotodon	*Chaetodon trifascialis* Quoy & Gaimard, 1825	AF108607	AF108532
	*Parachaetodon**	*Chaetodon oligacanthus* Bleeker, 1850	AF108622	AF108547
	Rabdophorus	*Chaetodon auriga* Forsskål, 1775	AF108582	AF108507
		Chaetodon collare Bloch, 1787	AF108586	**AF108543**
		Chaetodon decussatus Cuvier, 1829	AF108587	AF108512
		Chaetodon ephippium Cuvier, 1831	AF108588	AF108513
		Chaetodon flavirostris Günther, 1874	AF108589	AF108514
		Chaetodon lineolatus Cuvier, 1831	AF108593	AF108518
		Chaetodon lunula (Lacepède, 1802)	AF108594	AF108519
		Chaetodon melannotus Bloch & Schneider, 1801	AF108595	AF108520
		Chaetodon mesoleucos Forsskål, 1775	AF108596	AF108521
		Chaetodon ocellicaudus Cuvier, 1831	AF108598	AF108523
		Chaetodon oxycephalus Bleeker, 1853	AF108601	AF108526
		Chaetodon rafflesi Bennett, 1830	AF108604	AF108529
		Chaetodon semilarvatus Cuvier, 1831	AJ748305§	AJ748314§
		Chaetodon ulietensis Cuvier, 1831	AF108609	AF108534
		Chaetodon vagabundus Linnaeus, 1758	AF108610	AF108535
	Tetrachaetodon	*Chaetodon plebeius* Cuvier, 1831	AF108602	AF108527
		Chaetodon speculum Cuvier, 1831	AF108606	AF108531

1 – subgenus according to Blum (1988, 1989)

* – subgenus *Megaprotodon* according to Smith *et al*. (2003)

Collection of taxa

Heart muscle samples were collected in Moorea, French Polynesia, during November and December 1999 from six *Chaetodon* species, namely *C. lunulatus* Quoy & Gaimard, 1824, *C. pelewensis* Kner, 1868, *C. quadrimaculatus* Gray, 1831, *C. reticulatus* Cuvier, 1831, *C. trichrous* Günther, 1874 and *C. unimaculatus* Bloch, 1787. The samples were fixed and stored in 100% ethanol. Small sections of the samples were taken for DNA analysis using a sterilised scalpel and forceps. Archival samples held in the Natural History Museum, were sampled by removing small fin clippings from the following three species: *C. madagascariensis* Ahl, 1923, *C. paucifasciatus* Ahl, 1923 and *C. semilarvatus* Cuvier, 1831. The sample from *C. madagascariensis* was collected from the Indian Ocean, off the coast of Kenya (1994) and the samples from *C. paucifasciatus* and *C. semilarvatus* were collected from the Red Sea, off the coast of Egypt (1997). The sample from *C. madagascariensis* contained some muscle that was used for the analysis, as it was not likely to be contaminated by parasites. For the other two species, small sections of the fin clipping were taken from the edge nearest the body and washed twice in 70% ethanol to remove any external parasites and debris. Details of samples, including GenBank accession numbers and classification are given in Table 1.

The use of the species name *Chaetodon oligacanthus* Bleeker, 1850 is used here, rather than *Parachaetodon ocellatus* (Cuvier, 1831), as the species is clearly a member of the *Chaetodon* clade (Smith *et al*., 2003; see also www.fishbase.org). There is some debate as to which subgenus the species belongs (either *Megaprotodon* or its own *Parachaetodon*) but the present study tests further the need to recognize this additional subgenus.

DNA extraction, gene amplification and sequencing

The samples were soaked in TE buffer (1M Tris-EDTA, pH 8.0) overnight to remove the ethanol. The samples were digested to extract the DNA by grinding the material in TE with 1% SDS and then by incubating with proteinase K (approx. 0.2 mg/ml) at 37°C for 24 hours. Phenol extraction was used to remove proteins and the DNA was concentrated using Millipore® Microcon-100™ columns. PCR amplifications were carried out using 1µl of extracted DNA (20-100 ng gDNA) and of the 5' and 3' primers at 10mM concentration (for primers used see Table 2), one Ready-To-Go PCR bead (Amersham Pharmacia Biotech) and 22µl of ddH_2O making a reaction volume of 25µl. The PCR profile was as follows: 94°C 5mins (hot start), followed by 35 cycles of 94°C 30s, 46°C 30s, 72°C 1min, and a final extension period of 7mins at 72°C. 1µl of each PCR product was visualised on 1% TAE agarose gels and successful PCR products were purified using QIAGEN®QIAquick™ purification columns according to the manufacturer's protocols. Cycle sequencing using ABI Big Dye™ chemistry was carried out using manufacturer's protocols. Dye terminators were removed through precipitation with alcohol and 3M Sodium Acetate (pH 5.2) and the resultant pellets were dried at 95°C for 2mins, resuspended in dH_2O and sequenced using a ABI Prism 377™ automated sequencer. The

sequences from the new species were edited and contigs built using Sequencher[a] version 3.1.1 (GeneCodes Corp.).

TABLE 2. Primers used to amplify mitochondrial rrnS and cytochrome b from species of Chaetodontidae. The primers are based on those suggested by Palumbi (1996).

Gene	Primer	Sequence (5'-3')
rrnS	12SA-5'	AAA CTG GGA TTA GAT ACC CCA CTA T
	12SB-3'	GAG GGT GAC GGG CGG TGT GT
cytb	CB1-5'	CCA TCC ACC ATC TCA GCA TGA AA
	CB2-3'	CCC TCA GAA TGA TAT TTG TCC TCA

Sequence alignment and phylogenetic analysis

Sequences from the new species and from GenBank were aligned by eye using MacClade (Maddison & Maddison, 2000). Neither gene was difficult to align and no insertions or deletions were detected. The nine newly sequenced cytb fragments were 75bp (25 amino acids) shorter than the published sequences at the 3'-end of the fragment. We coded these sites as N (unknown), rather than missing. Each molecular data partition was analysed by parsimony (MP) and minimum evolution (ME) using PAUP* version 4.0b10 (Swofford, 2002), and Bayesian inference (BI) using MrBayes version 3.0b4 (Huelsenbeck, 2000). Posterior probabilities were approximated over 1,000,000 generations (ngen=1,000,000) via 4 simultaneous Markov Chain Monte Carlo chains (nchains=4) with every 1000th tree saved (samplefreq=1000). Default values were used for the MCMC parameters. Consensus trees with mean branch lengths were constructed using the 'sumt' command with the 'contype=allcompat' option and ignoring the initial topologies saved during 'burn in' (300 each for rrnS and cytb and 400 for the combined data set); the initial *n*-generations before log-likelihood values and substitution parameters plateau (see Huelsenbeck and Ronquist, 2001). MP analyses were run using a heuristic search strategy with tree bisection-reconnection branch swapping, equally weighted unordered characters and gaps treated as missing data. Trees were rooted against the pomacanthid sequences, and bootstrap support was estimated (1,000 replicates). ME analyses were based on genetic distances estimated by maximum likelihood using a general time-reversible (GTR) model of nucleotide evolution incorporating estimates of invariant (I) sites and among-site rate (G) variation; Modeltest (Posada and Crandall, 1998) was used to estimate the best substitution model for each data partitition and in each case this was GTR+I+G. BI analyses were conducted using a GTR+I+G model for each data partition independently and

also individually in a combined rrnS+cytb analysis, thus allowing separate estimates for each model parameter per data set.

Assessment of morphology

Species groupings (subgenera) and interrelationships between subgenera suggested by Blum (1988, 1989) were scored for each of the molecular phylogenetic solutions. Initially we treated *C. trifascialis* and *C. oligacanthus* as being members of separate subgenera, *Megaprotodon* and *Parachaetodon* respectively, in order to test whether they were sister groups and likely both members of the same subgenus *Megaprotodon*, as suggested by Smith *et al*. (2003). For each individual, and for the combined gene partition, constraint analyses with ME using a maximum likelihood model (incongruence length difference tests as implemented in PAUP*) were employed to determine whether the molecular data were compatible with the latest morphological phylogenetic assessment of relationships between genera and between *Chaetodon* subgenera (Smith *et al*., 2003); only those relationships among taxa where representative sequences were available, were tested. Constraint analyses, enforcing particular nodes to reflect morphological hypotheses were run using ME; resulting constraint trees were saved and were tested against unconstrained tree solutions using the Shimodaira-Hasegawa (S-H) test under maximum likelihood. Additionally, characters detailed in Smith *et al*. (2003) were mapped onto the combined molecular phylogenetic tree using MacClade (Maddison and Maddison, 2000) to determine which, if any, apomorphies changed or support new nodes.

Results

A total of 362bp of rrnS were available for alignment. No insertions or deletions were needed to align the genes unambiguously. Of these sites, 225 were invariant and 106 informative under the principles of parsimony. A total of 399bp of cytb were available for alignment. No insertions or deletions were needed to align the genes unambiguously. Of these sites, 210 were invariant and 177 (129 from coding position 3) were informative under the principles of parsimony. Alignments of both gene partitions have been submitted to EBI/EMBL under the following accession numbers: ALIGN_000802 for partial rrnS and ALIGN_000803 for partial cytb; both are available from http://www3.ebi.ac.uk/Services/webin/help/webin-align/align_SRS_help.html

Mitochondrial rrnS

With rrnS MP found 32 equally parsimonious trees; length=436, CI=0.389, RC=0.312. The strict consensus of these trees was not in conflict with either the minimum evolution tree (ME score= 1.483) or the BI solution [Fig. 2a]. For both ME and BI, Modeltest showed that a general-time reversible substitution model estimating both invariant sites and gamma distribution (GTR+I+G) was significantly better than other models ($P<0.001$).

All analyses using Bayesian inference included the following parameters: nst=6, rates=invgamma, ncat=4, shape=estimate, inferrates=yes, and basefreq=empirical, that corresponds to the best substitution model estimated (general-time-reversible including estimates of invariant sites and gamma distributed among-site rate variation). Log likelihood scores had 'plateaued' after approximately 200,000 replicates and we estimated tree parameters and posterior probabilities (pp) for the final 700,000 results (burnin=300). Poor nodal support, and low posterior probabilities, particularly within and between the *Chaetodon* species reflects the relatively few variable base positions in this gene fragment. Most genera were resolved as monophyletic, but within *Chaetodon* only the subgenus *Lepidochaetodon* was strongly supported (pp=98). *Chelmon* and *Coradion* were resolved as sister taxa and these in turn were resolved as a sister group to a paraphyletic *Heniochus* and *Hemitauricthys* clade. Unexpectedly, the sequence for *Chaetodon collare* Bloch, 1787 (AF108511) was resolved in this latter clade, thus rendering *Chaetodon* a non-monophyletic clade with this gene. Similarly, the sequence for *Heniochus acuminatus* (Linnaeus, 1758) (AF108543) was resolved within the *C. Rabdophorus* subgenus. It seems likely that the rrnS sequences for these taxa have been transposed and we have chosen to analyse the rrnS data on this basis (see Table 1). With the sequences correctly transposed, the relative position of these taxa matches more closely the cytb solution (Fig. 2b) and the taxonomy. For confirmation, rrnS for *C. collare* and *H. acuminatus* needs to be (re-) determined.

Mitochondrial cytb

With cytb MP found 21 equally parsimonious trees; length=1280, CI=0.243, RC=0.147. The strict consensus of these trees was not in conflict with either the minimum evolution tree (ME score=5.321) or the BI solution [Fig. 2b]. ME and BI analyses were conducted using a GTR+I+G model. For BI, log likelihood scores had plateaued after approximately 250,000 replicates and we estimated tree parameters and posterior probabilities for the final 700,000 results (burnin=300). This gene fragment provided considerably greater resolution than rrnS, reflected in longer internal branches and higher nodal support. All genera were well supported and within *Chaetodon*, where there was more than one representative for each subgenus sampled, all the subgenera were strongly resolved as monophyletic (pp=100). An analysis of nonsynonymous positions only gave identical tree topologies for each method used (data not shown). The position of *Chaetodon collare* fell within the *Rabdophorus* subgenus as might be predicted, lending further doubt on the veracity of the *C. collare* rrnS sequence. *Heniochus* and *Hemitauricthys* were resolved as sister taxa. A monophyletic *Chelmon* and *Coradion* clade was sister group to the remaining species (*Prognathodes* + *Chaetodon*).

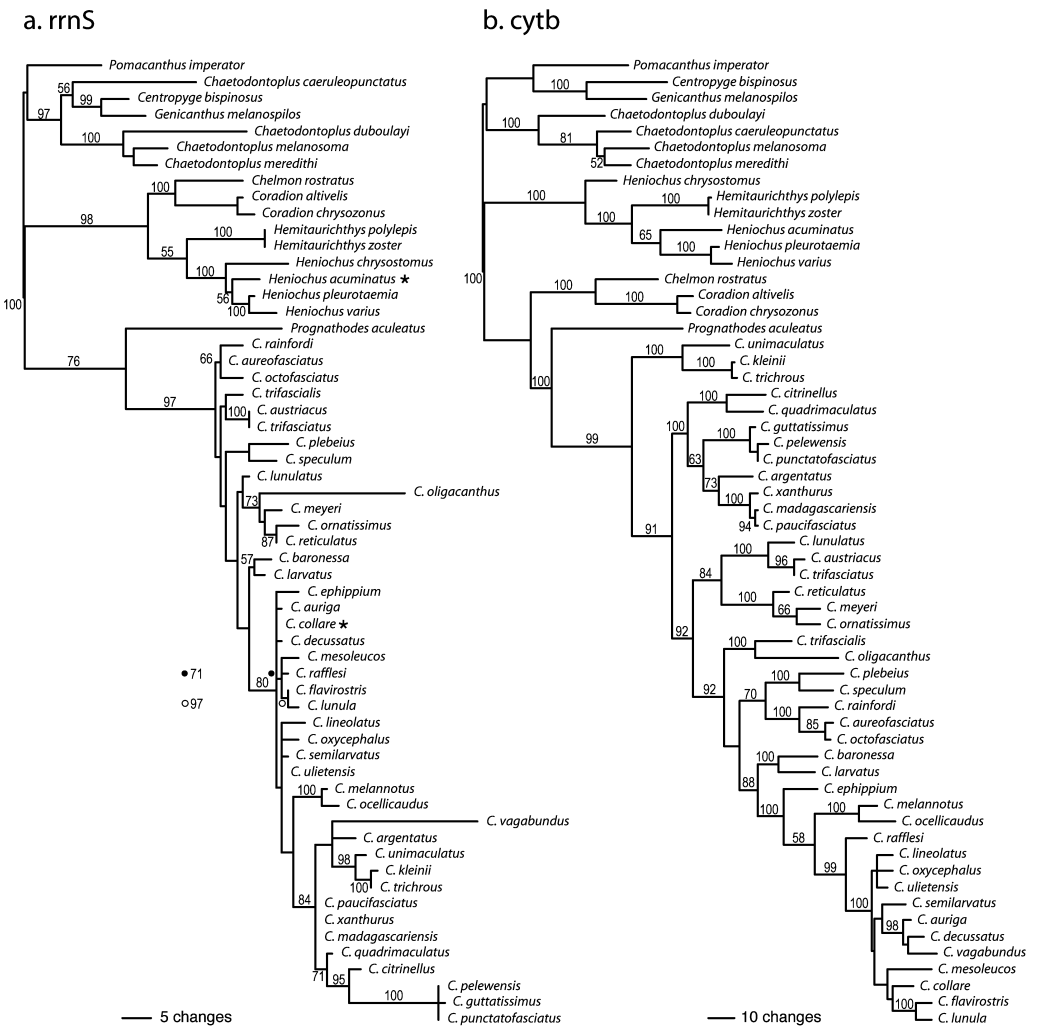

FIGURE 2. Molecular phylogenies of the Chaetodontidae, rooted against representative pomacanthids based on a. partial small subunit ribosomal RNA genes, and b. partial mitochondrial cytochrome b. Both trees are solutions provided by Bayesian inference with posterior probabilities indicated at the nodes. Asterisks indicate the likely transposition of the original sequences submitted to GenBank; see text for further details.

Combined mitochondrial rrnS and cytb

The combined data, with or without the *Chaetodon collare* rrnS sequence, yielded tree topologies almost identical to the cytb solution. However, MP analysis, which found 21 equally parsimonious solutions (length=177, CI=0.283, RC=0.175), yielded relatively poor nodal support. Figure 3 shows the BI solution, in which a GTR+I+G model was estimated for each data partition independently (burnin=400). Although the MP and ME solutions were almost identical in topology throughout there were some minor differences. A comparison of results between the BI solution and MP and ME trees is shown in Table 3,

in the context of which subgenera were supported. Within clade 7 of Smith *et al*.'s (2003) solution (Fig. 1c), only the monophyly of clades 10 and 19 were confirmed by all phylogenetic reconstruction methods, and clade 16 by ME and BI. Again, strong nodal support grouped *Chelmon* and *Coradion* as sister taxa (pp=100), and *Heniochus* and *Hemitauricthys* as sister taxa, (pp=100) but in this case with *Chelmon* and *Coradion* as sister group to the remaining taxa. *Prognathodes* was resolved as the sister group to the *Chaetodon* clade. Of the subgenera, *Lepidochaetodon* was the most basal and *Rabdophorus* the most derived. Constraint analyses indicated that the interrelationships between chaetodontid genera were not significantly different from Blum's analysis (Blum, 1988, 1989; see Fig. 1a), but the interrelationships of the subgenera of *Chaetodon* were significantly different (S-H test; $P<0.0001$).

Characters described in Smith *et al.* (2003) were transcribed into a matrix using MacClade for those taxa sampled in this study, and mapped onto the combined molecular phylogenetic solution shown in Figure 3 with reference to the species groups (genera and subgenera of *Chaetodon*) only. The coding for the hypothetical ancestor used in Smith *et al.* (2003) was retained for the outgroup and for mapping. Figure 4 shows the distribution of unambiguously mapped characters as they were optimised by MacClade. Unreversed synapomorphies, non-unique synapomorphies and autapomorphies are all shown on the tree as they were in the morphological analysis by Smith *et al.* (2003, their Figure 8; see also their Appendixes A-C). All character numbers and coding, and taxon delineation follow Smith *et al.* (2003). Not surprisingly, the different topology found by the molecular analysis yields significantly different character optimizations at a variety of nodes, when compared to the tree based on morphology. Notwithstanding the incomplete coverage of species groups, the morphological consequences of accepting the molecular solution are worth outlining, as they allow an assessment of the utility of the genes chosen and an independent test of morphology.

Mapping morphological characters onto the molecular phylogeny

Character 29 (second circumorbital excluded from margin of orbit), was not illustrated by Smith *et al.* (2003) on their tree, although with a recorded consistency index of 1.0 on their solution, it is consistent with the present analysis as a synapomorphy for the *Prognathodes+Chaetodon* clade (*Roa* is included in this grouping in Smith *et al.*, 2003). In Smith *et al.* (2003) character 29 was not an unambiguous unreversed synapomorphy for this clade as optimization could easily have made the 0 state a synapomorphy for the sister clade (from *Chelmonops* to *Heniochus* in their Figure 8). The *Prognathodes+Chaetodon* clade is supported in the present analysis by an additional synapomorphy, character 34 (separated palato-vomerine ligaments). It is unambiguous in the present analysis because *Amphichaetodon* was not included in the present analysis.

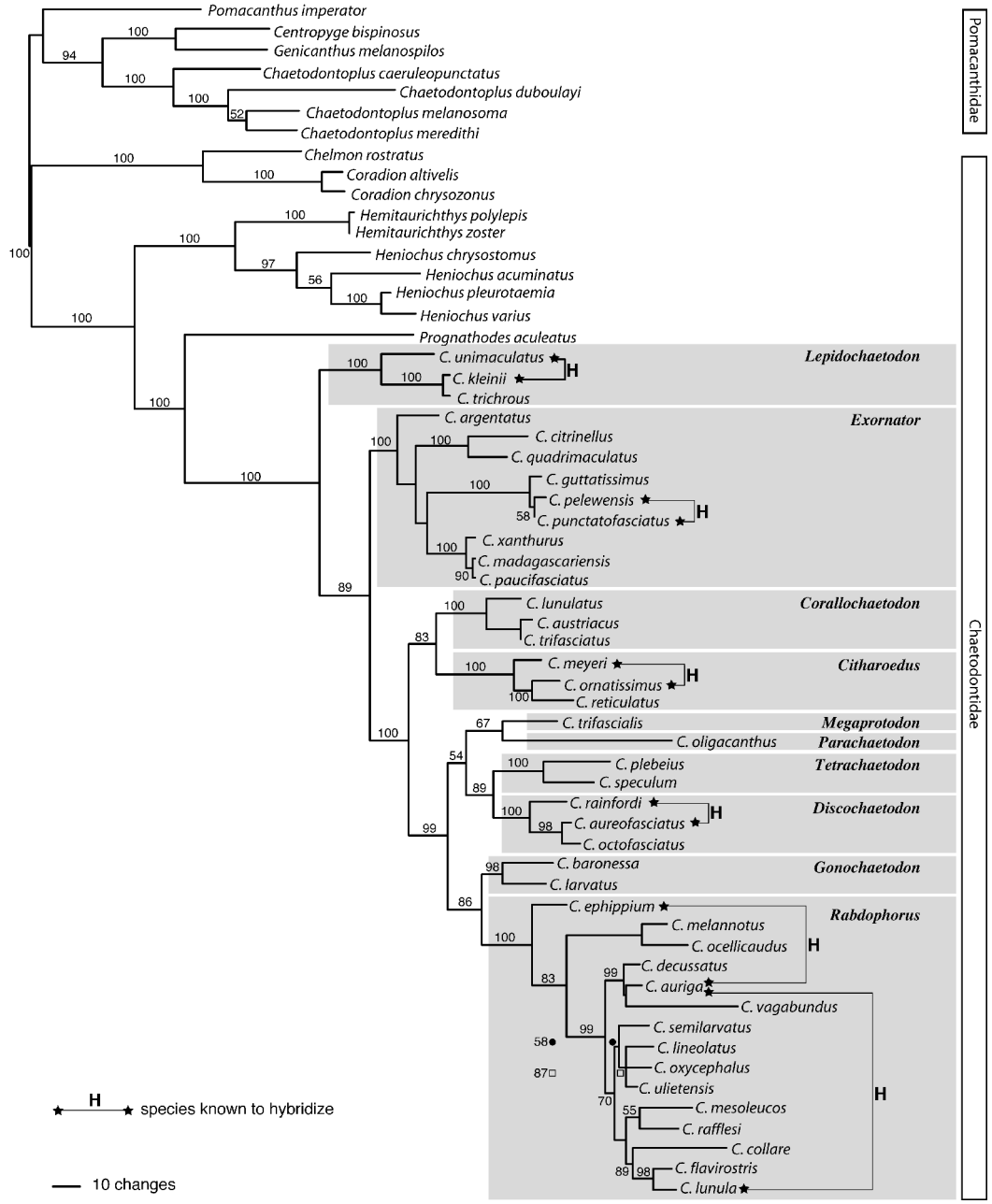

FIGURE 3. A combined molecular phylogenetic analysis of the Chaetodontidae from a Bayesian inference of rrnS and cytb, each partition modelled independently. Names of *Chaetodon* subgenera are indicated in the shaded boxes and nodal support values are posterior probabilities. Members of subgenera sampled that are known to hybridize with one another are indicated by stars; other known hybrid pairs are mentioned in the text.

Two synapomorphies support the clade uniting *C. trifascialis (*subgenus *Megaprotodon*), *C. oligacanthus* (subgenus *Parachaetodon*) and the subgenera *Tetrachaetodon*, *Discochaetodon, Gonochaetodon* and *Rabdophorus*: character 36 (state 1, direct laterophysic connection), and character 37 (state 1, swim bladder unattached to the peritoneum). *C. trifascialis* and *C. oligacanthus* are supported as sister groups by a weakly developed anterior laminae (character 4, state 1) although in the present study it is shown as an unreversed synapomorphy since characters were not ordered; character 4 was ordered in Smith *et al.* (2003) and state 1 was deemed intermediate. The molecular data and these characters perhaps strengthen the case for incorporating *C. oligacanthus* into the *Megaprotodon* clade, hence supporting the nomenclature used in Figure 4.

TABLE 3. Clades suggested by Blum (1989) and Smith *et al.* (2003) on the basis of morphology, supported by combined rrnS and cytb molecular systematic data; + supported, - unsupported.

Blum (1989) Table 1 clades	Smith *et al*. (2003) Fig. 8 clades (see Fig. 1c)	Analysis		
		MP	ME	BI
1 [= *Amphichaetodon*]		not tested	not tested	not tested
2 [= *Chelmonops*]		not tested	not tested	not tested
3 [= *Chelmon*]		not tested	not tested	not tested
4 [= *Coradion*]		+	+	+
5 [= *Forcipiger*]		not tested	not tested	not tested
6 [= *Heniochus*]		-	+	+
7 [= *Hemitaurichthys polylepis* spp. gp]		+	+	+
8 [= *Hemitaurichthys thompsoni* spp. gp]		not tested	not tested	not tested
9 [= *Prognathodes*]		not tested	not tested	not tested
10 [= *Roa*]		not tested	not tested	not tested
11 [= *C. Chaetodon*]		not tested	not tested	not tested
12 [= *C. Rabdophorus*; *Chaetodontops* spp. gp]		-	-	-
13 [= *C. Rabdophorus*; *C. auriga* spp. gp]		+	+	+
14 [= *C. Rabdophorus*; *C. ephippium* spp. gp]		not tested	not tested	not tested
15 [= *C. Rabdophorus*; *C. falcula* spp. gp]		not tested	not tested	not tested
16 [= *C. Rabdophorus*; *C. lineolatus* spp. gp]		-	+	-
17 [= *C. Rabdophorus*; *C. lunula* spp. gp]		not tested	not tested	not tested
18 [= *C. Rabdophorus*; *C. melannotus* spp. gp]		+	+	+
19 [= *C. Rabdophorus*; *C. selene* spp. gp]		not tested	not tested	not tested

.....continued on the next page

TABLE 3 continied.

Blum (1989) Table 1 clades	Smith *et al*. (2003) Fig. 8 clades (see Fig. 1c)	Analysis		
		MP	ME	BI
20 [= *C. Roaops*; *C. tinkeri* spp. gp]		not tested	not tested	not tested
21 [= *C. Exornator*; *Rhombochaetodon* spp. gp]		-	+	+
22 [= *C. Exornator*; *C. fremblii* spp. gp]		not tested	not tested	not tested
23 [= *C. Exornator*; *C. miliaris* spp. gp]		not tested	not tested	not tested
24 [= *C. Exornator*; *C. punctatofasciatus* spp. gp]		+	+	+
25 [= *C. Lepidochaetodon*; *C. kleinii* spp. gp]		+	+	+
26 [= *C. Megaprotodon+C. Parachaetodon*]		-	+	+
27 [= *C. Gonochaetodon*]		+	+	+
28 [= *C. Tetrachaetodon*]		+	+	+
29 [= *C. Discochaetodon*]		+	+	+
30 [= *C. Corallochaetodon*]		+	+	+
31 [= *C. Citharoedus*]		+	+	+
	1	not tested	not tested	not tested
	2	+	-	-
	3	not tested	not tested	not tested
	4	+	+	+
	5	+	+	+
	6	not tested	not tested	not tested
	7	+	+	+
	8	not tested	not tested	not tested
	9	not tested	not tested	not tested
	10	+	+	+
	11	-	-	-
	12	-	-	-
	13	-	-	-
	14	-	-	-
	15	-	-	-
	16	-	+	+
	17	-	-	-
	18	-	-	-
	19	+	+	+

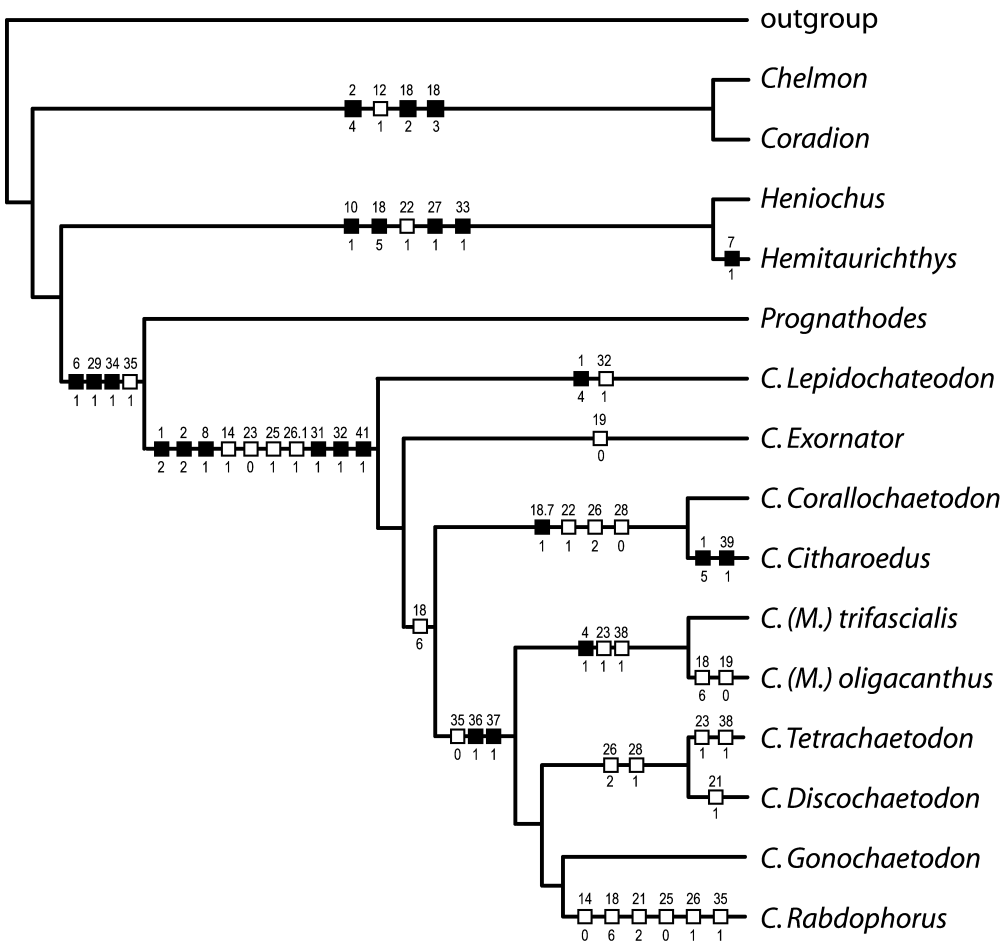

FIGURE 4. The Bayesian inference solution shown in Fig. 3, but with Smith *et al.*'s (2003) character matrix mapped onto the combined molecular phylogeny. Synapomorphies are listed before the node; unreversed synapomorphies are depicted as black squares and non-unique synapomorphies are depicted as open squares. Autapomorphies in terminal taxa are also shown optimized on the cladogram. Character numbers (above each square) represent the character numbers listed in Appendix C of Smith *et al.* (2003) with character states below the square. *C.* (*M.*) *trifascialis* and *C.* (*M.*) *oligacanthus* are equivalent to *Chaetodon* subgenera *Megaprotodon* and *Parachaetodon* respectively (see text).

Discussion

The interrelationships of the Chaetodontidae have been resolved with some success using morphology alone. Reasonably well-resolved phylogenies have been produced with a suite of characters coded at the generic and species group level (as recognized by Blum, 1988). Nevertheless, even with a morphologically based character set, which itself has evolved depending on character coding and interpretation (Ferry-Graham *et al.*, 2001a; Smith *et al.*, 2003), there remain unresolved and poorly supported nodes. The present study pro-

vides additional phylogenetic resolution with two short mitochondrial gene fragments. The molecular data, collected originally for larval identification, provides convincing support for all the species groups sampled and supports many, but also challenges some, of the interrelationships suggested by morphology. Mitochondrial cytochrome b (cytb) provides resolution throughout the tree and, to some extent, within the species groups. The partial gene fragment used appears to be a suitable marker worth pursuing for additional taxon sampling within the family Chaetodontidae. Ribosomal small subunit RNA genes (rrnS) provided considerably less resolution with poorer nodal support and relatively shorter internal branch lengths, and it is unlikely that the gene will provide a cost-effective marker in further resolving chaetodontid relationships. The analysis also highlights a likely misnumbering on GenBank. The rrnS sequences for *Chaetodon collare* (AF108511) and *Heniochus acuminatus* (AF108543) appear to have been transposed; this needs to be verified. Certainly, until the GenBank entry is verified, or until additional rrnS sequences for these taxa are characterized, the sequences should not be used for purposes of larval identification, the purpose for which it was originally determined.

The molecular data support the integrity of the species groups sampled in this study, although it should be noted that various authors differ in their circumscription of subgenera (e.g. compare Kuiter, 2000). Species groups used in this study were originally defined on the basis of osteological characters that defined 21 morphologically distinct groups of chaetodontids (Blum, 1988). Notwithstanding the differences in coding that have evolved since the original treatment by Blum (1988), all morphological analyses recognize the monophyly, or putative monophyly, of these groups and many studies have used them as operational taxonomic units when discussing broader patterns of relatedness or biology and biogeography. The monophyly of the groups, as tested here with molecular data, is well worth extending with greater taxon sampling (additional subgenera and denser sampling within subgenera). Until this is done, we believe it would be premature to investigate too deeply the hypotheses concerning, for example, biogeography or the adaptive radiation of jaw structure and feeding ecology. Thus, the present study is restricted to discussing overall patterns and differences in phylogeny between the latest morphological assessment (Smith *et al.* 2003) and the solutions provided by mitochondrial genes.

Looking towards the base of the chaetodontid phylogeny, the molecular solution supports a sister group relationship between *Chelmon* and *Coradion*, but not between this clade and a clade containing *Heniochus* and *Hemitaurichthys* as found with morphology. Instead, *Chelmon* and *Coradion* are the most basal chaetodontid genera in the present study. However, poor nodal support between these clades suggests that the molecular data is not seriously challenging the results from morphology. Although resolved as monophyletic genera, *Heniochus* and *Hemitaurichthys* are not well differentiated genera morphologically. Blum (1988) listed an elongate fourth dorsal spine as the sole autapomorphy of *Heniochus*, and Smith et al. (2003) recognized only one character that differentiated the genera; character 7, less than 60 lateral line scales in *Heniochus* and more than 60 in *Hemi-*

taurichthys. Far greater nodal support is found among the remaining deeper branches and the data set proves more useful in establishing relationships within and between the subgenera of *Chaetodon*.

In the absence of members of *C. Roa* (recognized as a separate genus *Roa* containing three species by Smith *et al.*, 2003), *Prognathodes* was resolved as the sister group to a monophyletic clade comprising all *Chaetodon* species. The subgenus *Lepidochaetodon* is the most basal *Chaetodon* clade and it confirms the close relationship between between *Chaetodon kleinii* Bloch, 1790 and *C. trichrous* Günther, 1874 discussed by Blum (1988, 1989). No samples were available for *C. striatus* or *Roaops*, but the next clade to branch off includes the species in *Exornator*. The close relationships between *Exornator* and *Lepidochaetodon* is also well established through morphological analyses (they appear as sister taxa), albeit without any unique unreversed synapomorphies; character 12 of Smith *et al.* (2003), branchiostegal rays reduced to five, unites these taxa, but also unites *Chelmonops*, *Chelmon* and *Coradion*.

The sister group relationship between the monophyletic subgenera *Corallochaetodon* Burgess, 1978 and *Citharoedus* Kaup, 1860, is confirmed here with molecular data. Blum (1989) discussed the morphological similarities between these taxa including jaw structure, and colour patterning and, as with Smith *et al.*'s (2003) character mapping, the present analysis confirmed a unique unreversed synapomorphy for the grouping; character 18, state 7, teeth of nearly equivalent length, coalesced into brush with increased number of bands.

As with morphology, *Megaprotodon* and *Parachaetodon* were resolved as sister taxa, supported by one unreversed synapomorphy; character 4, state 1, pleural rib laminae with weakly developed anterior laminae. This evidence supports Smith *et al.* (2003) synonymization of *Parachaetodon* and *Megaprotodon* as illustrated in Figure 4 although ideally, additional species of *Megaprotodon* need to be sampled for at least cytb to test this. *Tetrachaetodon* and *Discochaetodon* were resolved as sister taxa, with two non-unique synapomorphies (character 26, state 2, and character 28, state 1). This is in contrast to the morphological analysis which had *Discochaetodon* as sister group to *Corallochaetodon+Citharoedus*. Indeed, it is largely the relationships among the more derived *Chaetodon* subgenera (*Corallochaetodon+Citharoedus*, *Megaprotodon+Parachaetodon*, *Tetrachaetodon*, *Discochaetodon*, *Gonochaetodon* and *Rabdophorus*) where the major differences between the molecular and morphological treatments appear. Nevertheless, although the topologies are significantly different, two unique unreversed morphological synapomorphies support the clade (*Megaprotodon*, *Tetrachaetodon*, *Discochaetodon*, *Gonochaetodon*, *Rabdophorus*); character 36, state 1, a direct laterophysic connection and character 37, state 1, swim bladder unattached to the peritoneum. Neither of these characters appears as unreversed synapomorphies in the morphological analysis of Smith *et al.* (2003), but importantly, they provide a resolution to the ambiguity found in the morphological analysis. Smith *et al.* (2003) appealed for additional species-level phylogenetic

analyses to resolve apparent anomalies concerning the evolution laterophysic connections. The molecular analyses yielded a different topology that obviates such additional sampling, although indeed, wider taxonomic and gene sampling remains preferable.

Hybridization between closely related *Chaetodon* species has long been suspected, and in a few cases proven and employed for taxonomic revision. For example, McMillan *et al*. (1999) considered that overwhelming evidence from behavioural, genetic, and phenotypic (colour) patterns between *C. pelewensis* and *C. punctatofasciatus* made it impossible to justify their continued species-level status. Smith *et al*. (2003) synonymized these two taxa and evidence from the current study lends some support to this; the rrnS sequence for each taxon was identical and there were only two base differences between the cytb sequences for each taxon, one synsonymous and the other nonsynonymous (CGA coding for arginine in *C. pelewensis* and CAA coding for glutamine). A further use for the phylogeny is evaluating the relative position of taxa thought to hybridize; members of subgenera known to hybridize are indicated in Fig. 3 and are listed below. It is not surprising that cases of probable hybridization involve species in the same subgenus, and often closely related taxa, e.g. *C. aureofasciatus* x *C. rainfordi*, *C. pelewensis* x *C. puntatofasciatus*, and *C. kleini* x *C. unimaculatus* (Randall *et al*., 1977). However, in listing all known hybrids among *Chaetodon* species Gill (1999) noted that only five pairings (of 12 known from the literature) involved sister parent species according to the classification in Blum (1989: Table 1). The known cases of natural hybridization in Chaetodontidae include: *Chaetodon aureofasciatus* x *C. rainfordi*, *C. auriga* x *C. ephippium*, *C. auriga* x *C. fasciatus*, *C. auriga* x *C. lunula*, *C. ephippium* x *S. semeion*, *C. ephippium* x *C. xanthocephalus*, *C. kleini* x *C. unimaculatus*, *C. meyeri* x *C. ornatissimus*, *C. miliaris* x *C. tinkeri*, *C. ocellatus* x *C. striatus* and *C. pelewensis* x *C. punctatofasciatus* (Burgess, 1974, 1978; Randall *et al*., 1977; Allen, 1981; Randall and Fridman, 1981; Clavijo, 1985). It remains to be seen whether other hybrid pairs reflect these close phylogenetic relationships. Attempts to hybridize *C. kleinii* x *C. trichrous* and *C. aureofasciatus* x *C. octofasciatus*, each pair representing sister taxa, would appear to be justified on the basis of phylogeny and history of hybridization of one member of each these pairs. Hybridization among member pairs of *Rabdophorus* subgenus were more distantly related and this may reflect poorer resolution within this clade, or a greater propensity to hybridize within the group. Widespread distribution, colour pattern variability, the opportunity to hybridize and the difficulty in differentiating taxa based on morphology alone, all suggest that a molecular phylogenetic approach has much to offer.

In conclusion, even with small partial fragments of mitochondrial cytb, and to a lesser extent rrnS, sufficient phylogenetic resolution is achieved to confirm the circumscription of chaetodontid genera and subgenera. The utility of recognizing subgenera comes to the fore only when discussing the evolution of morphological characters that define them, but as can be seen in the case with *Megaprotodon* subsuming *Parachaetodon*, there may need to be some revisions. However, the delineation of subgenera appears rather arbitrary and

may obscure other finer level relationships. Ultimately, a species level phylogeny will be of greater utility. Additionally, the molecular solution strengthens the case for homology among a suite of osteological and soft tissue characters, even demonstrating two as unique unreversed synapomorphies where morphological analyses do not. An extended study sequencing additional cytb and other mitochondrial genes of these and additional taxa will certainly test and corroborate further the characters described by Smith *et al.* (2003). As predicted by Smith *et al.* (2003), combining molecular and morphological characters to resolve a species-level phylogeny of the family is a worthwhile and attainable goal. Such a phylogeny will provide an excellent framework with which the extensive comparative ecological, behavioural and biogeographic data can be examined further in an evolutionary context. A large molecular dataset is already to hand for expansion and supplementation.

Acknowledgements

New sequences for this study were completed by S.M.M. as part of an MSc project. We thank colleagues on the University of Queensland / Université de Perpignan / NHM collecting trip in Moorea for securing specimens of *Chaetodon*. D.T.J.L. was funded by a Wellcome Trust Senior Research Fellowship (043965/Z/95/Z). T.H.C. was supported by funds from the Australian Research Council.

References

Allen, G., Steene, R. & Allen, M. (1998) *A Guide to Angelfishes and Butterflyfishes*, Tropical Reef Research, Perth, 250 pp.

Blum, S.D. (1988) *Osteology and phylogeny of the Chaetodontidae (Pisces: Perciformes).* Ph.D. Dissertation, University of Hawaii, Honolulu, 365 pp.

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

Burgess, W.E. (1974) Evidence for the elevation to family status of the angelfishes (Pomacanthidae) previously considered to be a subfamily of the butterflyfish family, Chaetodontidae. *Pacific Science,* 28, 57—71.

Burgess, W.E. (1978). *Butterflyfishes of the World*, Tropical Fish Hobbyist Publications, Neptune City, NJ, 832 pp.

Chen, W.J., Bonillo, C. & Lecointre, G. (2003) Repeatability of clades as a criterion of reliability: a case study for molecular phylogeny of Acanthomorpha (Teleostei) with larger number of taxa. *Molecular Phylogenetics and Evolution*, 26, 262—288.

Clavijo, I.E. (1985). A probable hybrid butterflyfish from the western Atlantic. *Copeia*, 1985, 235—238.

Driscoll, J.W. & Driscoll, J.L. (1988) Pair behavior and spacing in butterflyfishes (Chaetodontidae). *Environmental Biology of Fishes*, 22, 29—37.

Ferry-Graham, L.A., Wainwright, P.C., Hulsey, C.D. & Bellwood, D.R. (2001a) Evolution and mechanics of long jaws in butterflyfishes (Family Chaetodontidae). *Journal of Morphology,* 248, 120—143.

Ferry-Graham, L.A., Wainwright, P.C. & Bellwood, D.R. (2001b) Prey capture in long-jawed butterflyfishes (Chaetodontidae): the functional basis of novel feeding habits. *Journal of Experimental Marine Biology and Ecology*, 256, 167—184.

Gill, A.C. (1999) Subspecies, geographic forms and widespread Indo-Pacific coral-reef fish species: a call for change in taxonomic practice. *In*: Seret, B. & Sire, J.-Y. (Eds), Proceedings of the 5th Indo-Pacific Fish Conference, Nouméa, November 1997. Société Française d'Ichtyologie and Institut de Recherche pour le Développement, Paris, pp. 79—87.

Huelsenbeck, J. P. (2000) MrBayes: Bayesian inference of phylogeny. *Distributed by the author. Department of Biology, University of Rochester.*

Huelsenbeck, J. P. & Ronquist, F. (2001). MrBayes: Bayesian inference of phylogenetic trees. *Bioinformatics*, 17, 754—755.

Kuiter, R.H. (2002) *Butterflyfishes, Bannerfishes and their Relatives. A Comprehensive Guide to Chaetodontidae and Microcanthidae*, TMC Publishing, Chorleywood, UK, 208 pp.

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

Motta, P.J. (1988) Functional morphology of the feeding apparatus of ten species of Pacific butterflyfishes (Perciformes, Chaetodontidae): an ecomorphological approach. *Environmental Biology of Fishes*, 22, 99—67.

Motta, P. J. (1989) Dentition patterns among Pacific and Western Atlantic butterflyfishes (Perciformes: Chaetodontidae): relationships to feeding ecology and evolutionary history. *Environmental Biology of Fishes*, 25, 159—170.

Maddison, D.R. & Maddison, W.P. (2000) MacClade 4: Analysis of Phylogeny and Character Evolution. Version 4.0. Sinauer Associates, Sunderland, Massachusetts.

Near, T.J., Kassler, T.W., Koppelman, J.B., Dillman, C.B. & Philipp, D.P. (2003). Speciation in North American black basses, *Micropterus* (Actinopterygii : Centrarchidae). *Evolution*, 57, 1610—1621.

Palumbi, S. R. (1996) Nucleic Acids II: the polymerase chain reaction. *In*: Hillis, D.M., Moritz, C. & Mable, B.K. (Eds) *Molecular Systematics*. Sinauer Associates, Sunderland MA, pp. 235—238.

Perdices, A., Bermingham, E., Montilla, A. & Doadrio, I. (2002) Evolutionary history of the genus Rhamdia (Teleostei : Pimelodidae) in Central America. *Molecular Phylogenetics and Evolution*, 25, 172—189.

Posada, D. & Crandall, K.A. (1998) Modeltest: testing the model of DNA substitution. *Bioinformatics*, 14, 817—818.

Randall, J.E., Allen, G.R. & Steene, R.C. (1977) Five probable hybrid butterflyfishes of the genus *Chaetodon* from the Central and Western Pacific. *Records of the Western Australian Museum*, 6, 3—26.

Randall, J.E. & Fridman, D. (1981) *Chaetodon auriga* x *Chaetodon fasciatus*, a hybrid butterflyfish from the Red Sea. *Revue Française d'Aquariologie*, 7, 113—116.

Ruber, L., Van Tassell, J.L. & Zardoya, R. (2003). Rapid speciation and ecological divergence in the American seven-spined gobies (Gobiidae, Gobiosomatini) inferred from a molecular phylogeny. *Evolution*, 57, 1584—1598.

Sano, M. (1989) Feeding habits of Japanese butterflyfishes (Chaetodontidae). *Environmental Biology of Fishes*, 25, 195—205.

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

Swofford, D.L. (2003). PAUP* Phylogenetic Analysis Using Parsimony (*and Other Methods). Version 4. Sinauer Associates, Sunderland, Massachusetts.