

*Pleistocene diversification of the  
Pomacentrus coelestis species complex  
(Pisces: Pomacentridae): historical  
biogeography and species boundaries*

**Laurie Sorenson, Gerald R. Allen, Mark  
V. Erdmann, Chang-Feng Dai & Shang-  
Yin Vanson Liu**

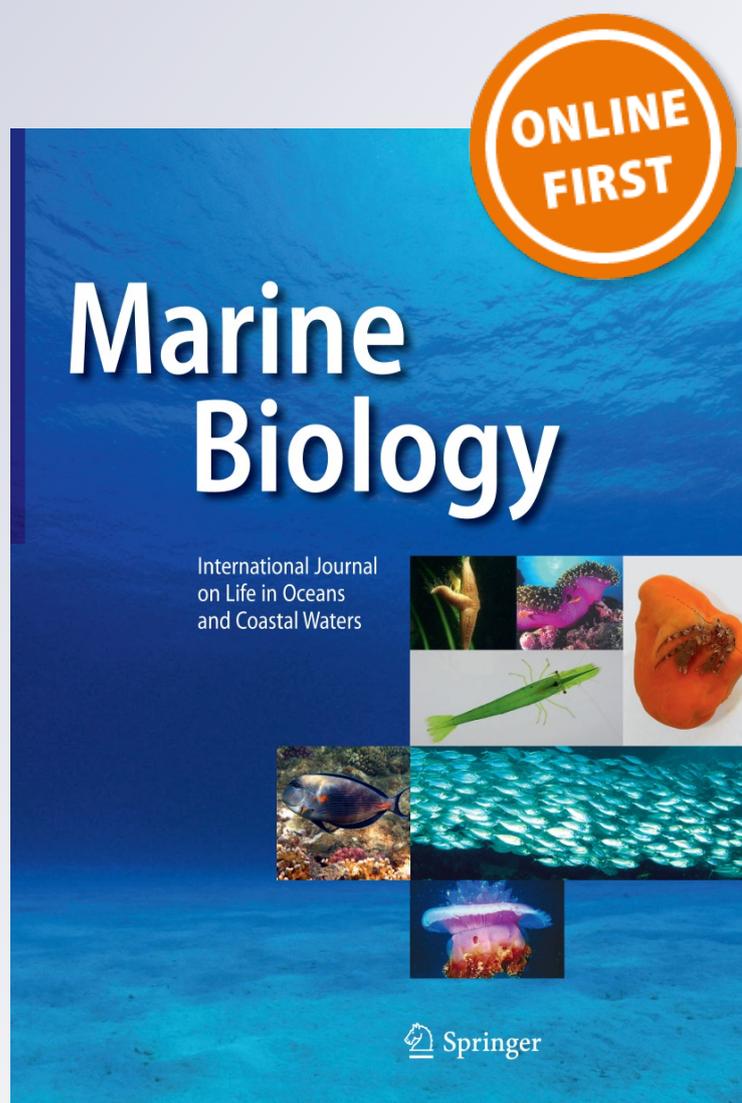
**Marine Biology**

International Journal on Life in Oceans  
and Coastal Waters

ISSN 0025-3162

Mar Biol

DOI 10.1007/s00227-014-2521-8



**Your article is protected by copyright and all rights are held exclusively by Springer-Verlag Berlin Heidelberg. This e-offprint is for personal use only and shall not be self-archived in electronic repositories. If you wish to self-archive your article, please use the accepted manuscript version for posting on your own website. You may further deposit the accepted manuscript version in any repository, provided it is only made publicly available 12 months after official publication or later and provided acknowledgement is given to the original source of publication and a link is inserted to the published article on Springer's website. The link must be accompanied by the following text: "The final publication is available at [link.springer.com](http://link.springer.com)".**

# Pleistocene diversification of the *Pomacentrus coelestis* species complex (Pisces: Pomacentridae): historical biogeography and species boundaries

Laurie Sorenson · Gerald R. Allen · Mark V. Erdmann · Chang-Feng Dai · Shang-Yin Vanson Liu

Received: 2 April 2014 / Accepted: 8 August 2014  
© Springer-Verlag Berlin Heidelberg 2014

**Abstract** Pleistocene eustatic changes in sea level are often invoked to explain genetic divergence among marine organisms. However, molecular phylogenies have revealed relatively few examples of speciation events dating to the Pleistocene. We present a species-level hypothesis of the timing of evolution for the *Pomacentrus coelestis* species complex (Pomacentridae), based on the nuclear S7 intron and the mitochondrial Cytb gene, and reconstruct ancestral range distributions across the timetree. Ancestral range reconstruction suggests the complex originated in the Coral Triangle and East Indian Ocean, with subsequent range expansion outward from this region. We suggest that land barriers to dispersal (e.g., Indo-Pacific barrier) may be responsible for the divergence between Indian (*P. alleni*, *P. similis*, *P. caeruleopunctatus*, and *P. caeruleus*) and Pacific

(*P. micronesicus*, *P. auriventris*, and *P. coelestis*) species groups, and subsequent isolation by Pleistocene sea-level fluctuations in certain areas of the Coral Triangle (glacial refugia) may play an important role in the diversification of this species complex. Additionally, our analyses show cryptic lineages within *P. micronesicus* and highlight the need for comprehensive sampling within and among species to reveal recent speciation events.

## Introduction

Allopatric speciation is the most widely accepted mode of generating new species; however, physical barriers dividing populations are relatively scant in the marine environment. Furthermore, marine organisms often disperse as larvae, reducing the potential for geographic isolation and subsequent differentiation (e.g., Palumbi 1994; Mora and Sale 2002). Several marine barriers to dispersal have been identified, however, that often define well-established biogeographic provinces (see review Briggs and Bowen 2012). An obvious example is the Isthmus of Panama, which divided the East Pacific and western Atlantic marine fauna with its closing approximately 3.1 Ma (Coates and Obando 1996). Allopatry has also been proposed as an important mechanism for speciation in the Indo-Pacific (Randall 1998; Bellwood and Wainwright 2002).

The Indo-Malay-Philippine (IMP) biodiversity hot spot, stretching from northern Australia and New Guinea, to Malaysia, and west to the Philippines, has long been considered the center of marine biodiversity (Allen 2008). A complex geologic history, as well as climatic and oceanographic changes, has resulted in vicariance events and the formation of geographic barriers over the last 16 million years. Land barriers and shallow seas were formed

Communicated by M. Taylor.

**Electronic supplementary material** The online version of this article (doi:10.1007/s00227-014-2521-8) contains supplementary material, which is available to authorized users.

L. Sorenson · S.-Y. V. Liu (✉)  
Department of Ecology and Evolutionary Biology, University of California Los Angeles, 610 Charles E Young Drive South, Los Angeles, CA 90095, USA  
e-mail: oceandiver6426@gmail.com

G. R. Allen  
Department of Aquatic Zoology, Western Australian Museum, Locked Bag 49, Welshpool DC, Perth 6986, Australia

M. V. Erdmann  
Conservation International, Indonesia Marine Program, Jl. Drive Muwardi No. 17, Bali, Indonesia

C.-F. Dai  
Institute of Oceanography, National Taiwan University, Taipei 10617, Taiwan, ROC

in the region during the Middle Miocene (~8–16 Ma) when the Australian and Eurasian plates collided (Kennett et al. 1985). This event also restricted flow between the Indian and Pacific Oceans. The IMP region was also greatly impacted by eustatic changes in sea levels due to glacial cycles since the Pliocene (Naish et al. 2009). Three to six glacial cycles occurred during the Pleistocene, which lowered sea level up to 130 m below present levels (e.g., Chappell 1981; Potts 1983; Voris 2000), and significantly decreased available habitat on continental shelves. Although reproductive isolation may not be generated during short inter-glacial periods, alternate genotypes between temporally separated groups can become fixed, which may lead to accumulated genetic differences even with ongoing gene flow (Rocha and Bowen 2008). These processes are often invoked to explain, at least in part, the high biodiversity within the IMP region (Woodland 1983; Bellwood and Wainwright 2002; Gaither and Rocha 2013).

Despite the potential for vicariance events in the recent past (especially during the Last Glacial Maximum throughout the Pleistocene), molecular studies suggest most divergence events among lineages occurred during the Pliocene or Miocene for most extant reef fauna (e.g., Barber and Bellwood 2005; Renema et al. 2008; Bellwood et al. 2010). Relatively few contradictory examples to this temporal pattern exist: Some sister species splits in chaetodontids (e.g., McMillan and Palumbi 1995; Fessler and Westneat 2007; Bellwood et al. 2010), wrasses (e.g., Choat et al. 2012), *Dascyllus* (Leray et al. 2010), Tetraodontiformes (e.g., Santini et al. 2013a; 2013b; 2013c), and acanthurids (e.g., Sorenson et al. 2013) date to the Pleistocene. Instead, the timing of Pleistocene events appears to be more important in driving intraspecific differentiation (e.g., see Rocha and Bowen 2008). This pattern of prevalent diversification within the Miocene and Pliocene and not the Pleistocene, however, may be due to the paucity of well-sampled phylogenies.

The limited evidence for Pleistocene lineage diversification events may also be, in part, due to our tendency to define species based on discernible morphological traits (Rocha and Bowen 2008; Gaither and Rocha 2013). These morphological characters may be very subtle, especially between marine fishes where species distinctions are frequently based on coloration or meristic traits. Increasing genetic evidence indicates cryptic lineages are commonly found in reef fishes, especially those with broad geographic ranges, which may or may not be accompanied by morphological variation (Craig et al. 2007; Steinke et al. 2009; Drew and Barber 2009; Drew et al. 2010; Liu et al. 2012). Because currently unrecognized species (e.g., cryptic species) may not be considered in molecular phylogenetic studies, inferred sister species divergence times will be older and very recent cladogenesis events will remain undetected (Gaither and Rocha 2013).

Damselfishes (Pomacentridae) are a morphologically and ecologically diverse group of fishes that include many conspicuous members of coral reef communities. Many species within this family have broad geographic distributions, and cryptic lineages have been reported in several of these species (Drew et al. 2008, 2010; Steinke et al. 2009; Liu et al. 2012). One widespread species, *P. coelestis*, was recently shown to represent at least two distinct species (*P. coelestis* and *P. micronesicus*; Liu et al. 2012, 2013) based on morphology and genetics. Mitochondrial data suggest that the *P. coelestis* lineage underwent rapid expansion during the Pleistocene, suggesting that eustatic changes in sea level may have initiated speciation (Liu et al. 2012).

*Pomacentrus coelestis* and *P. micronesicus* are both parts of the *P. coelestis* species complex (Liu et al. 2013) that includes five additional species: *P. alleni*, *P. auriventris*, *P. caeruleopunctatus*, *P. caeruleus* and *P. similis*. All of the species within the complex are morphologically similar and difficult to distinguish from *P. coelestis* in situ, except for *P. alleni* and *P. auriventris*, which have distinct color morphs. Detailed morphological examination of *P. coelestis* specimens suggested that an additional cryptic lineage may exist (Liu et al. 2013). The presence of cryptic lineages and species descriptions based on color pattern variations within the complex make it an ideal model to determine whether Pleistocene vicariance events play an important role in reef fish evolution.

Here, we present a biogeographic study of the *P. coelestis* species complex, including samples from all seven species from multiple localities across their ranges. We use mitochondrial Cytb and ribosomal intron S7 sequence to infer the relationships among species within the complex and time-calibrate our dataset to estimate divergence times between species. Our results reveal cryptic lineages within the species complex and suggest that Pleistocene sea-level changes and glacial refugia provided a mechanism for speciation in this fish group.

## Materials and methods

### Sample collection

A total of 145 individuals including all members of the *P. coelestis* complex was collected for this study. In addition, a *P. moluccensis* sample was collected for use as an outgroup. Fishes were collected using either a hand net or small spear (Table 1), and one *P. coelestis* and three *P. micronesicus* samples were purchased from a Taiwanese aquarium shop. Individuals were collected from multiple locations within each species' range when possible (Table 1): *P. similis* ( $N = 17$ ), *P. caeruleus* ( $N = 17$ ), *P. auriventris* ( $N = 44$ ), *P. alleni* ( $N = 12$ ), *P. caeruleopunctatus* ( $N = 2$ ), *P. coelestis*

**Table 1** List of species, sampling localities, and biogeographic regions used in this study

Taxon name	Sampling location	Sample code	N	Region
<i>Pomacentrus alleni</i>	Narcondam Island, Andaman Sea	all_ana	4	E
	Cinque Islands, Andaman Sea	all_and	7	
	Nusa Penida, Bali, Indonesia	all_nus	1	
	Species total		12	
<i>Pomacentrus auriventris</i>	Banda	aur_ba	6	E, T
	Kofiau Island, South Raja Ampat	aur_kaf	3	
	Kawe Island, North Raja Ampat	aur_kaw	14	
	Boo Kecil, Raja Ampat	aur_kec	8	
	Nusa Laut	aur_la	2	
	Nusa penida, Bali, Indonesia	aur_nu	1	
	Silang, Bali, Indonesia	aur_silang	10	
Species total		44		
<i>Pomacentrus coelestis</i>	Bali, Indonesia	coe_balis	4	E, T, P, N
	Australia	coe_cap	9	
	Palawan	coe_palawan	7	
	Taiwanese aquarium shop	coe_twsp	1	
Species total		21		
<i>Pomacentrus micronesicus</i>	Cenderawasih Bay, West Papua, Indonesia	mic_ce	14	T, P
	Chuuk Atoll, Caroline Islands, Micronesia	mic_ch	8	
	Bali, Indonesia	mic_balis	2	
	Taiwanese aquarium shop	mic_twsp	3	
	Palawan	mic_palawan	1	
Species total		28		
<i>Pomacentrus caeruleopunctatus</i>	Madagascar	cap_mad	2	
Species total			2	
<i>Pomacentrus caeruleus</i>	Madagascar	cae_pca	17	W
	Species total		17	
<i>Pomacentrus similis</i>	Mergui, Myanmar	sim_sim	3	E
	Narcondam Island, Andaman Sea	sim_sna	14	
	Species total		17	

Biogeographic region information was used in the ancestral range reconstruction analysis. *Pomacentrus caeruleopunctatus* was not run separately in the range reconstruction analysis, so biogeographic range coding is not provided. Region abbreviations are as follows: *W* West Indian, *E* East Indian, *T* Coral Triangle, *P* Central Pacific, *N* North Pacific

( $N = 21$ ), and *P. micronesicus* ( $N = 28$ ). Fin clips or whole fishes were preserved in ethanol until use.

#### Gene amplification and sequencing

Genomic DNA was extracted from preserved tissues using the Genomic DNA extraction kit (Genomics BioScience and Technology, New Taipei City, Taiwan) following the

manufacturer's protocol. A partial Cytb fragment and the S7 ribosomal protein intron were amplified using the polymerase chain reaction (PCR) using the following primers: GluDG-L and H16460 (Palumbi 1996) for Cytb; S7RPEX1F and S7RPEX2R primers (Chow and Hazama 1998) for S7. PCRs were carried out in 25  $\mu$ l reactions containing 5–10 ng DNA, 10 mM Tris HCl (pH 8.3), 1.5 mM MgCl<sub>2</sub>, 50 mM KCl, 0.2 mM dNTPs, 1 U Taq

DNA polymerase (Mdbio, Taipei, Taiwan), and 0.3 mM each of the appropriate forward and reverse primers. PCR cycling consisted of 2 min at 94 °C for the first cycle, followed by 34 cycles of 95 °C for 30 s, 55 °C for 30 s, and 70 °C for Cytb. The S7 intron was amplified by the same thermocycling profile except with an annealing temperature of 64 °C. The nucleotide sequences of the PCR products were determined using an ABI 377-automated sequencer (Applied Biosystems). Additionally, Cytb sequence data were downloaded from GenBank for three outgroups used in this study: *Chromis atrilobata*, *C. multilineata*, and *P. moluccensis*.

### Phylogenetic analyses

Chromatograms were manually edited and assembled using Sequencher 4.2 (Gene Code, Ann Arbor, MI, USA). Gene sequences were aligned using the MUSCLE software (Edgar 2004) in Geneious 5.3 (Biomatters Ltd.). Alignments were checked by eye and trimmed to minimize missing characters. The final gene alignments consisted of 1,080 bp for Cytb and 707 bp for S7. All generated sequences were deposited in GenBank (Table S1).

We used jModelTest (Posada 2008) to select the best-fitting substitution model for Cytb and S7 using Akaike Information Criterion (AIC). We did not consider models with the invariant sites parameter since this is accounted for by the gamma parameter. The HKY + G model was selected as the most appropriate model for both loci.

Phylogenetic trees were generated under maximum likelihood (ML) using RAxML (Stamatakis 2006). We ran analyses for two datasets: (1) Cytb only, which included 99 individuals and (2) a complete dataset containing 50 individuals for which there were Cytb and S7 sequence data. Because there were no *P. micronesicus* (Chuuk) individuals with sequence for both genes, we combined data for ch2 (S7) and ch1 (Cytb) to include a representative from this clade in the tree. We assigned the GTR + G model, the closest model to HKY + G that can be implemented in RAxML, to each gene and ran 500 fast bootstrap replicates using the GTR + CAT model. For the complete dataset with Cytb and S7, each gene was partitioned separately and 1,000 fast bootstrap replicates were used with the GTR + CAT model.

Phylogenetic relationships were also inferred under a Bayesian framework using MrBayes 3.2 (Ronquist et al. 2012) for both datasets. The HKY + G substitution model was used, and the complete dataset was partitioned by gene. Two independent analyses were run for 30 million generations each using four chains (one cold, three heated), sampling every 3,000 generations. Trace files were checked using Tracer 1.5 (Drummond and Rambaut 2007) to ensure convergence. Analyses were combined, and the first 25 %

of trees were discarded as burnin. Post-burnin trees were combined to obtain a 50 % majority rule consensus tree.

We calculated mean genetic distance among described species and clades recovered in our phylogenetic analyses using the cytochrome *b* alignment. We assumed a Kimura 2-parameter model and ran the analysis using MEGA 5.2 (Tamura et al. 2011).

### Dating analysis

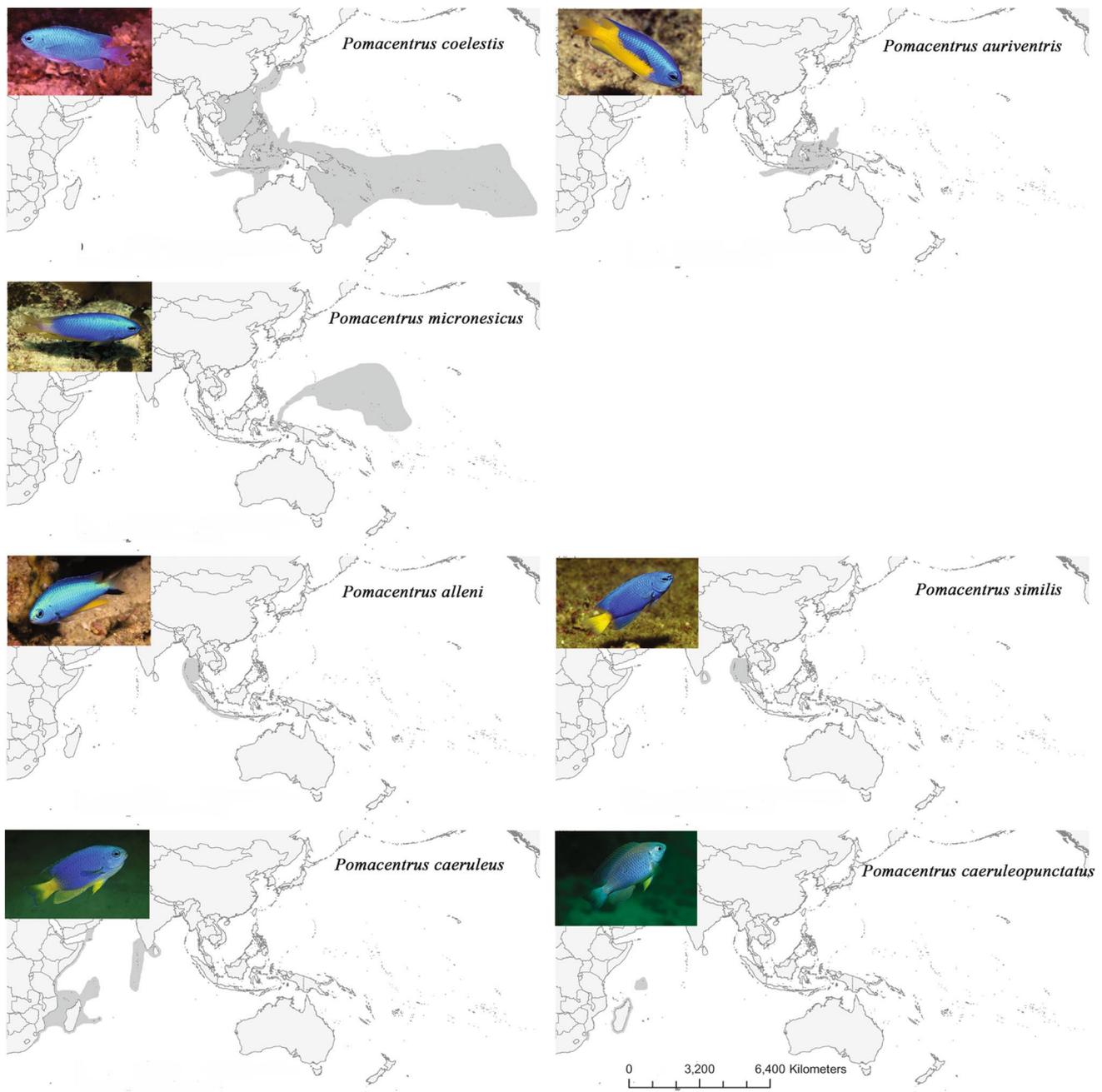
For the dating analysis, we selected a representative taxon from each of the species (except *P. caeruleopunctatus*) in the *P. coelestis* species complex that had both Cytb and S7 sequences. We used one *P. caeruleus* individual to represent *P. caeruleus* + *P. caeruleopunctatus* since we were unable to distinguish the two using our genetic data. Because there were no *P. micronesicus* (Chuuk) individuals with sequence for both genes, we used the same individuals from the RAxML complete dataset analysis that combined ch2 and ch1. We used *P. moluccensis*, *C. atrilobata*, and *C. multilineata* as outgroups; the latter two species were represented only by Cytb sequence data in the matrix.

We used the program BEAST 1.7 (Drummond et al. 2012) to infer divergence times for the eight lineages. An HKY + G nucleotide substitution model was assigned to each gene partition, and we used an uncorrelated lognormal relaxed molecular clock. A birth–death prior was assigned to cladogenesis rates. We used the timing of the rise of the Isthmus of Panama (3.1–3.5 Ma; Coates and Obando 1996) to date the split between the Atlantic *C. multilineata* and the Pacific *C. atrilobata*. We set an exponential prior with an offset of 3.0 and a mean of 0.17. Analyses were run for 40 million generations, sampling every 4,000 generations. Convergence was confirmed using Tracer 1.5 (Drummond and Rambaut 2007) and to ensure ESS values exceeded 200. We discarded the first 10 % of trees as burnin and used TreeAnnotator 1.7 (Drummond et al. 2012) to generate the maximum clade credibility timetree.

### Ancestral range reconstruction

The Indo-Pacific was divided into five broad regions: (1) West Indian Ocean, (2) East Indian Ocean, (3) Coral Triangle, (4) North Pacific, and (5) Central Pacific. Each species represented in the timetree was coded as either being present or absent from each of the five regions based on published records (Allen 1991) and personal field observations of G. Allen and M. Erdmann (Fig. 1).

We reconstructed ancestral geographic ranges using a stochastic model of geographic range evolution incorporating dispersal, extinction, and cladogenesis (DEC model). Reconstruction was inferred using a maximum likelihood approach implemented in the package



**Fig. 1** Species distributions of taxa in the *Pomacentrus coelestis* species complex based on Allen (1991) and the personal communication with Dr. Gerald Allen and Dr. Mark Erdmann

BioGeoBEARS 0.2 (Matzke 2013) in R 2.15 (R Development Team 2012). We tested the fit of two different models: (1) a model where dispersal among all adjacent regions was equally probable and (2) a model where the probability of dispersal between the Coral Triangle (T) and the East Indian Ocean (E) was constrained to be 0.05, reflecting limited dispersal due to the Indo-Pacific Barrier during the Pleistocene. Model fit was assessed using Akaike Information Criterion (AIC).

## Results

### Phylogenetic analyses

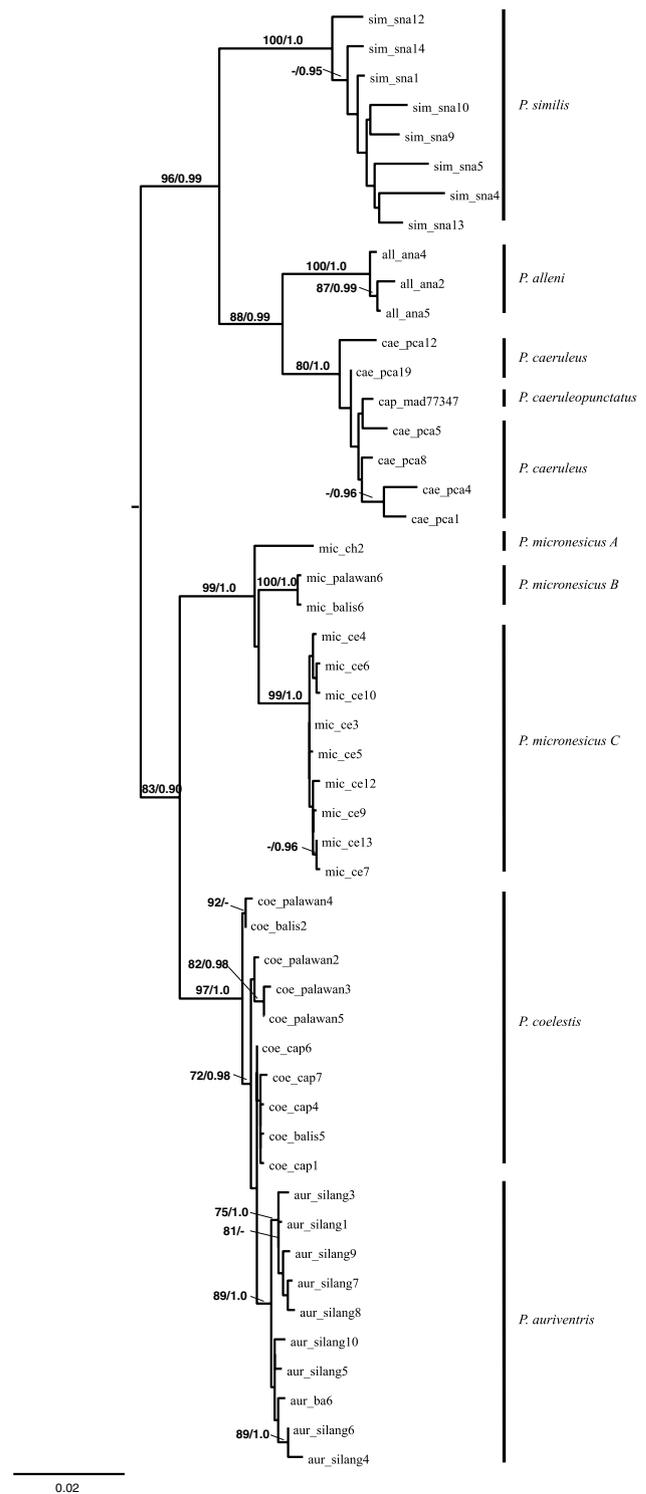
The relationships among taxa based on the ML and Bayesian Cytb analyses were identical, with high support for most of the groups (Figure S1 and S2, respectively). Two major clades were inferred; one composed of species found in the Indian Ocean (91 bootstrap proportions, bsp; 1.0

posterior probability, pp) and another composed of species distributed in the Pacific Ocean (63 bsp, 0.95 pp). Within the Indian Ocean subclade, *P. similis* (98 bsp, 1.0 pp) is sister to a clade including *P. caeruleus*, *P. caeruleopunctatus*, and *P. alleni* (96 bsp, 1.0 pp). The relationships among *P. caeruleus*, *P. caeruleopunctatus*, and *P. alleni* could not be resolved, but *P. alleni* monophyly was highly supported (99 bsp, 1.0 pp) and formed a distinct clade from *P. caeruleus* + *P. caeruleopunctatus*. Within the Pacific clade, two well-supported subclades were inferred. The first included *P. coelestis* sampled from Indonesia, the Philippines, Australia, and one individual purchased at a Taiwanese aquarium shop, and *P. auriventris* (95 bsp, 1.0 pp). *Pomacentrus coelestis* sequences were not monophyletic, but *P. auriventris* sequences formed a well-supported group (91 bsp, 1.0 pp). The second Pacific subclade was composed of three distinct lineages, all with bsp = 100 and pp = 1.0: clade A of *P. micronesicus* sampled from Chuuk Atoll, Micronesia; clade B of *P. micronesicus* sampled from Bali, Indonesia, Palawan, Philippines, and individuals from a Taiwanese aquarium shop (TWSP 1-3); and clade C of *P. micronesicus* sampled from Cenderawasih Bay, West Papua, Indonesia. The relationships among these three clades, however, could not be resolved.

Species relationships based on the ML and Bayesian analyses of the complete dataset were identical, and this topology was highly congruent with the cytochrome *b* tree (Fig. 2). Indian (96 bsp, 0.99 pp) and Pacific (83 bsp, 0.90 pp) Ocean clades were again recovered. *Pomacentrus similis* (96 bsp, 1.0 pp), *P. alleni* (100 bsp, 1.0 pp), and *P. caeruleus* + *P. caeruleopunctatus* (80 bsp, 1.0 pp) were all monophyletic, but unlike the Cytb trees, *P. alleni* was sister to the *P. caeruleus* + *P. caeruleopunctatus* clade (88 bsp, 0.99 pp). In the Pacific clade, the first subclade (83 bsp, 0.90 pp) contained *P. coelestis* individuals, as well as a strongly supported *P. auriventris* clade (97 bsp, 1.0 pp). The second Pacific subclade (99 bsp, 1.0 pp) included the three lineages of *P. micronesicus*, all with  $\geq 99$  bsp, 1.0 pp).

### Genetic divergence

*Pomacentrus coelestis* individuals were combined into one group because of the low support for their non-monophyly recovered by our phylogenetic analyses. We also separated *P. caeruleus* and *P. caeruleopunctatus* into two groups because they are previously described species. Finally, we were interested in assessing the divergence among the three subclades of *P. micronesicus*, so we defined the subclades as three different groups. Mean genetic divergence between

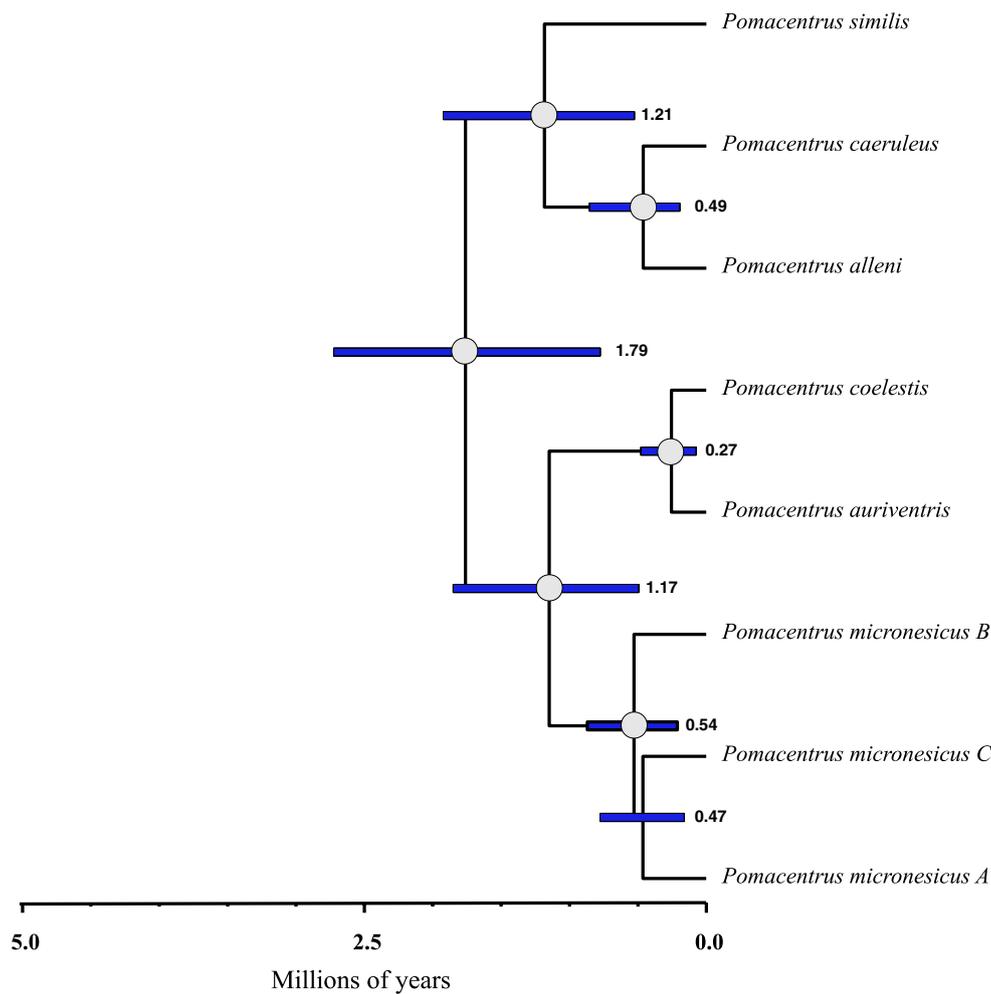


**Fig. 2** Maximum likelihood hypothesis of relationships within the *Pomacentrus coelestis* species complex based on the complete dataset. Only bootstrap proportions  $\geq 70\%$  and posterior probability values  $\geq 0.95$  based on the Bayesian analysis are shown preceding the nodes. The outgroup has been removed

**Table 2** Mean genetic divergence among *Pomacentrus coelestis* species complex taxa and subclades

	cap	coe	mic A	aur	all	mic C	mic B	cae	sim
cap									
coe	0.0600								
mic A	0.0690	0.0480							
aur	0.0640	0.0100	0.0480						
all	0.0190	0.0660	0.0770	0.0710					
mic C	0.0660	0.0480	0.0220	0.0480	0.0720				
mic B	0.0740	0.0490	0.0230	0.0490	0.0810	0.0290			
cae	0.0060	0.0620	0.0700	0.0680	0.0210	0.0660	0.0740		
sim	0.0480	0.0630	0.0660	0.0690	0.0530	0.0730	0.0750	0.0490	

Genetic divergence calculations were computed using MEGA 5.2 (Tamura et al. 2011) assuming a Kimura 2-parameter model. *Pomacentrus micronesicus* subclades A–C were grouped separately. See Table 1 for taxon abbreviations



**Fig. 3** Time-calibrated phylogeny of the *Pomacentrus coelestis* species complex based on BEAST 1.7 analysis. Node bars indicate 95 % highest posterior density, and white circles on nodes indicate pos-

terior probability values  $\geq 0.99$ . The outgroups have been removed. Inferred evolutionary timing indicates a Pleistocene origin and diversification of the entire clade

groups (Table 2) ranged from 0.006 (*P. caeruleus* vs. *P. caeruleopunctatus*) to 0.077 (*P. allenii* vs. *P. micronesicus* A). Mean divergence between *P. micronesicus* subclades ranged from 0.022 to 0.029 (Table 2).

#### Divergence time estimation

The BEAST topology (Fig. 3) was identical to the ML and Bayesian trees based on the complete dataset, and we

recovered  $pp \geq 0.99$  for all nodes except for the sister relationship between *P. micronesicus* A and C (0.92 pp).

The *P. coelestis* species complex arose in the Early to Middle Pleistocene; we inferred a crown age of  $\sim 1.8$  Ma (0.77–2.72 Ma 95 % highest posterior density, HPD) for the group. The Indian and Pacific Ocean subclades arose contemporaneously  $\sim 1.2$  Ma. The split between *P. alleni* and *P. caeruleus* + *P. caeruleopunctatus* occurred  $\sim 0.5$  Ma (0.19–0.85 Ma 95 % HPD). Within the first Pacific subclade, *P. auriventris* split from *P. coelestis*  $\sim 0.3$  Ma (0.07–0.48 Ma 95 % HPD). Divergence times among lineages in the second Pacific subclade significantly overlapped. We inferred a crown age of  $\sim 0.5$  Ma (0.21–0.87 Ma 95 % HPD) for the *P. micronesicus* subclade containing lineages A–C, with only a slightly younger 95 % HPD range for the split between *P. micronesicus* A and C (0.16–0.78 Ma 95 % HPD).

### Historical biogeographic ranges

For our model comparison, the fit of model 1, where dispersal among all adjacent regions was equally probable, was preferred over model 2 ( $\Delta AIC = 9.6$ ). We reconstructed the ancestral ranges under conditions of model 1 and inferred that the most recent common ancestor of the *P. coelestis* species complex most likely originated in the East Indian Ocean/Coral Triangle (Fig. 4; Table S2) with subsequent expansion outward from this region. The ancestor to the Indian Ocean clade occurred in the East Indian Ocean region. *Pomacentrus similis* maintained this ancestral range, while the ancestor to *P. alleni* and *P. caeruleus* expanded its range to include the West Indian Ocean. The split between *P. alleni* and *P. caeruleus* divided the EW range, with *P. alleni* in the East Indian, and *P. caeruleus* into the West Indian Ocean.

The ancestor to the Pacific Ocean clade was most likely broadly distributed across the East Indian, Coral Triangle, and Central and North Pacific regions. This distribution was maintained in the palawan4 lineage, but the silang3 lineage originated in the East Indian and Coral Triangle regions. The *P. micronesicus* clade arose in the Coral Triangle, and *P. micronesicus* subclade A expanded its range to include the Central Pacific.

## Discussion

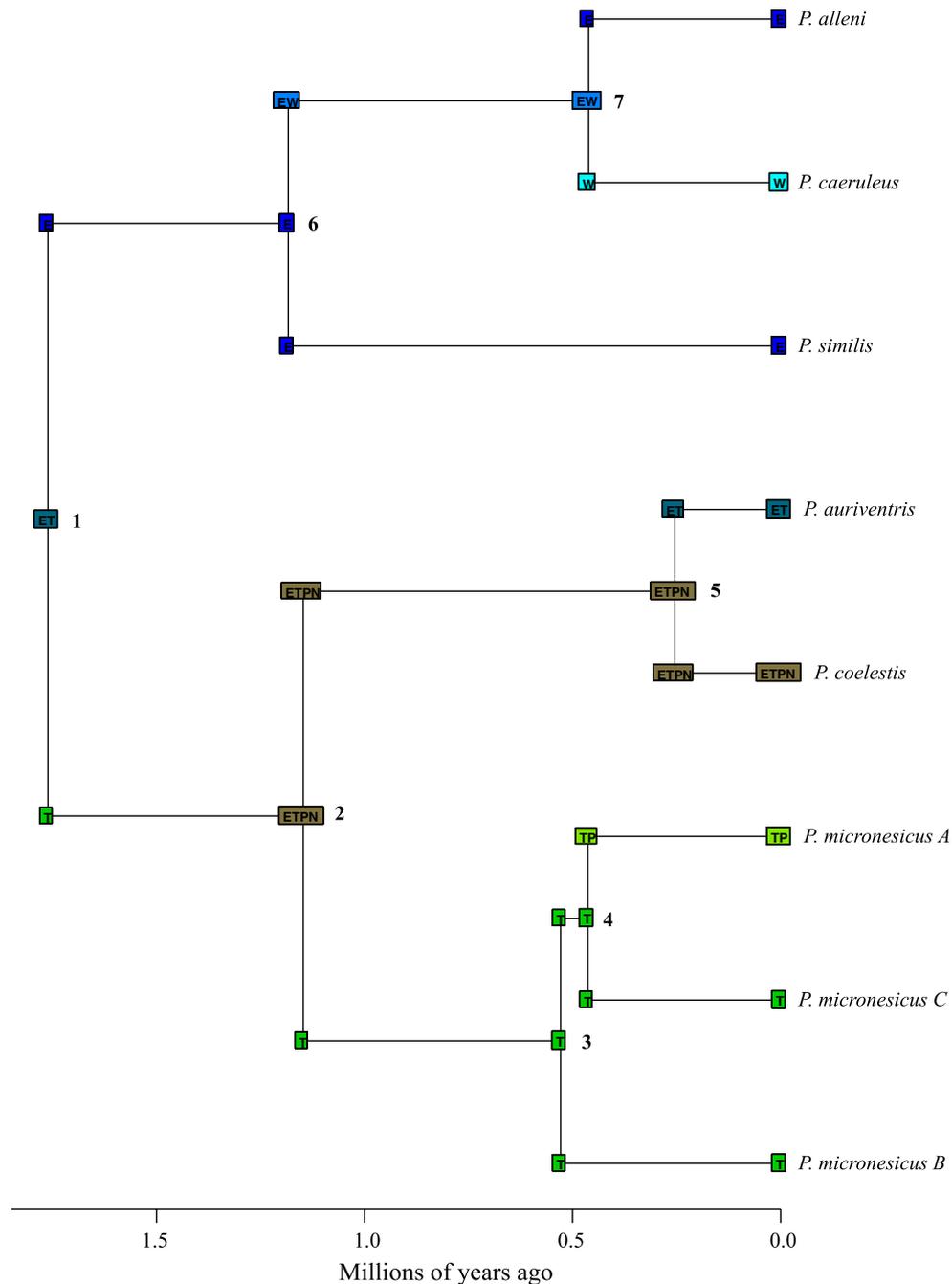
### Speciation process and historical biogeography

Our study provides some of the first evidence of origin and subsequent diversification of an entire species complex during the Pleistocene. Although eustatic sea levels during the Pleistocene have been implicated in the diversification

of many reef groups, evidence is scant based on previously published studies. Most timing of origin and divergence time estimates for reported lineages tend to date to the Pliocene and/or Miocene (e.g., McCafferty et al. 2002; Barber and Bellwood 2005; Bellwood et al. 2010; Cowman and Bellwood 2011; Santini et al. 2013a, 2013b, 2013c; Sorenson et al. 2013), calling into question the impact of Pleistocene sea levels on reef fish speciation. The importance of Pleistocene speciation may be revealed, however, with better sampled phylogenies. Increasing the sampling coverage will include more sister taxa comparisons that may uncover very recent speciation events, whereas poorly sampled trees will inevitably elevate divergence time estimates (Gaither and Rocha 2013). Therefore, generating fully sampled phylogenies that also include multiple taxa per species (that could represent cryptic lineages) is vital to understanding the evolutionary histories of reef fish fauna.

Physical barriers and loss of reef habitat due to glacial cycles during the Pleistocene have certainly impacted the evolutionary history of marine fishes (Rocha and Bowen 2008). By the Pleistocene, modern geological and oceanographic features and patterns were mostly established in the Indian Ocean and Indo-West Pacific. Also importantly at this time, eustatic sea levels resulted in exposure of the Sunda Shelf (Voris 2000; Naish et al. 2009), resulting in the Indo-Pacific Barrier (IPB) that decreased genetic flow between the Pacific and Indian Oceans. The IPB appears to be a primary evolutionary driving force early in the history of the *P. coelestis* species complex; we found support for two distinct subclades: one Indian and one Pacific. The divergence time between these two subclades is comparable to previous dating of 17 fish species pairs (0.3–6.6 Ma) across the Indo-Pacific barrier (e.g., Gaither and Rocha 2013).

During times of low sea level, glacial refugia throughout the Indo-West Pacific may have played an important role in generating diversity in *Pomacentrus*. Sea-level fluctuations have had a great influence on marine fish diversification in the Indo-West Pacific (Randall 1998). Over the past 0.5 million years, sea levels have dropped below 100 m than present day levels at least five times (Siddall et al. 2003; Rohling et al. 2010). Several areas in the Coral Triangle have been documented as Pleistocene refugia, such as the Celebes-to-Banda Sea and the Philippine regions (Carpenter and Springer 2005). In the Coral Triangle, a mosaic of deep seas (i.e., the South China, Sulu, Celebes, Molucca, and Banda Seas) maintains highly fragmented coral reef communities during periods of low sea level (Carpenter et al. 2011). During Pleistocene lowstands, Tomini Bay of Sulawesi (Wallace et al. 2002), Cenderawasih Bay of West Papua (Allen and Erdmann 2006; 2012), and the Andaman Sea (Benzie 1999; Satapoomin 2002) were semi-closed from adjacent seas and were proposed to be the driving force behind the high endemism



**Fig. 4** Ancestral range reconstruction for the *Pomacentrus coelestis* species complex. Most likely ranges for nodes and branches are indicated in colored boxes on the timetree. Region abbreviations are as follows: W West Indian, E East Indian, T Coral Triangle, P Cen-

tral Pacific, N North Pacific. Numbers following internal nodes correspond to ancestral range reconstruction probability values provided in Table S2

observed today. Such glacial refugia may have provided a mechanism for the rapid diversification of the *P. coelestis* complex, especially Pacific Ocean lineages, and those in the northeast Indian Ocean, *P. alleni* and *P. similis*. All five Pacific lineages occur in the Coral Triangle, and ancestral range reconstructions suggest origins and population

expansions throughout this region for all lineages, even the widespread *P. coelestis*. *P. micronesicus C* is primarily found in Cenderawasih Bay, and *P. alleni* and *P. similis* are distributed in the Andaman Sea, lending support to the glacial refugia hypothesis (e.g., Benzie 1999; Satapoomin 2002; Allen and Erdmann 2006).

Based on the results of biogeography reconstruction, the *P. coelestis* species complex most likely originated in the eastern Indian Ocean or the Coral Triangle. We inferred population expansion westward in the ancestor of *P. alleni* + *P. caeruleus*, but *P. alleni* continues to reside in the proposed ancestral East Indian Ocean and Coral Triangle portion of the range. Very recent population expansion (since ~0.5 Ma) within the eastern Indian Ocean region has resulted in the current *P. alleni* range stretching north from the Andaman Sea, south to Bali, Indonesia. Most likely, this expansion followed a rise in sea level that increased habitat availability along the shallow continental margin of western Indonesia. *Pomacentrus similis* and *P. alleni* occur in sympatry in the northeastern Indian Ocean. *Pomacentrus alleni* is one of the two species in the species complex with significant color pattern variation (black streak on the lower part of the caudal fin). Taylor and Hellberg (2005) proposed that differentiation in gobies (genus *Elacatinus*) is initiated in allopatry, but maintained by assortative mating based on color differences after secondary contact. A similar mechanism is plausible for *P. alleni*, evolving a distinct color morph and maintaining species identity during secondary contact following glacial retreat.

Following westward dispersal, divergence between *P. alleni* and *P. caeruleus* (+*P. caeruleopunctatus*) may have been established and maintained by restricted connectivity between the East and West Indian Ocean (see review by Ridgway and Sampayo 2005). The surface current regime was suggested to be the main factor constraining marine organism dispersal within the Indian Ocean (Goetze 2011; Vogler et al. 2012). Therefore, we suggest that rare dispersal events from the East Indian to the West Indian Ocean, followed by isolation by oceanographic regimes, may explain the speciation process leading to the separation of *P. alleni* and *P. caeruleus* (+*P. caeruleopunctatus*). The ancestor of *P. caeruleus* and *P. caeruleopunctatus* established populations in the western Indian Ocean, but the evolutionary history of these species is not well understood. Although subtle morphological differences exist between *P. caeruleus* and *P. caeruleopunctatus*, we could not distinguish between the two species based on Cytb and S7 sequence data. Several studies have found that genetic markers, such as mitochondrial genes, may lack the resolution to differentiate color polymorphisms found in some marine fishes (e.g., Messmer et al. 2005; Schultz et al. 2007). Craig et al. (2007) suggested that coloration may evolve faster than molecular rates to explain the incongruence between molecular markers and color morphs, and this may be a possible explanation for the lack of genetic variation between *P. caeruleus* and *P. caeruleopunctatus*. Alternatively, there may not be reproductive isolation between these two species, and *P. caeruleopunctatus* may represent a color morph of *P. caeruleus*. Phylogenomic and

population genomic methods may provide greater resolution and solve this intriguing evolutionary enigma in the future.

### Species boundaries

Our molecular study is the first to include complete sampling across the entire *P. coelestis* species complex, providing valuable information to discern the evolutionary history of this group. Based on our combined dataset analyses, we find support for the monophyly of five of the seven previously described species within the complex based on morphological characters. The exceptions are *P. caeruleus*, rendered paraphyletic by *P. caeruleopunctatus*, and *P. coelestis*, which is paraphyletic with respect to *P. auriventris*.

Within the Indian Ocean clade, our analyses highly supported the monophyly of *P. similis* and *P. alleni*. We were unable to resolve the relationship between *P. caeruleus* and *P. caeruleopunctatus*. The former species is widely distributed throughout the western and north central Indian Ocean; *P. caeruleopunctatus* is distributed sympatrically, but limited primarily to islands (currently known from Madagascar and Seychelles) off the eastern African coast. Though morphologically similar, *P. caeruleopunctatus* achieves greater maximum size compared to *P. caeruleus* and is also distinguished based on body shape, and details in the color pattern along the anal and caudal fins and body scales margins (Allen 2002). The lack of genetic divergence between these two species may be due to *P. caeruleopunctatus* representing a color morph of the more widespread *P. caeruleus*. However, given the additional morphological variation besides coloration between the two species, and our estimate of a very recent stem age of ~493,000 years for *P. caeruleus* + *P. caeruleopunctatus*, *P. caeruleopunctatus* is likely an incipient species, and fixed genetic differences between *P. caeruleus* and *P. caeruleopunctatus* have not had enough time to accumulate. Overlapping geographic distributions of *P. caeruleus* and *P. caeruleopunctatus* suggest that *P. caeruleopunctatus* may have evolved in sympatry or parapatry with *P. caeruleus*.

Within the Pacific Ocean clade, we did not recover monophyly of the widely distributed *P. coelestis*, albeit with low support in nearly all of our analyses, due one lineage within the clade giving rise to *P. auriventris*. Based on our analyses, *P. auriventris* is nested deeply within the *P. coelestis* clade, suggesting that the two species may have evolved in parapatry or sympatry. Poor resolution within *P. coelestis* may be due to recent and rapid population expansion. Liu et al. (2012) performed a mismatch distribution test and coalescence analysis based on mitochondrial control region sequence data, assuming a mutation rate range from 8.2 to 9.3 % per million years, and found evidence for rapid population growth in *P. coelestis* within the Pacific

between ~750,000–860,000 years ago. The estimate for expansion time by Liu et al. (2012) is slightly older than our inferred crown age of *P. coelestis* (0.27 Ma), but still falls well within the stem age (1.2 Ma, 0.49–1.85 Ma 95 % HPD) we inferred for this species. This slight variation in estimates may be due to several factors, including the faster evolutionary rate at the control region locus compared to Cytb, which may lead to older age estimates; variation in sampling between the two studies; and also because we used only one taxon to represent the species in our dating analysis.

Morphologically, *P. coelestis* is more similar to the recently described *P. micronesicus* (Liu et al. 2013), which was split from *P. coelestis* based on CoxI genetic distance and slight variations in anal fin ray counts and number of gill rakers (Liu et al. 2013). Interestingly, however, we did not recover these two species as sister taxa. Instead, we found high support for a clade, sister to *P. coelestis*+ a highly supported monophyletic *P. auriventris*. We recovered three distinct lineages of *P. micronesicus*, supporting the finding of three subclades within the ‘Micronesian clade’ of Liu et al. (2012). One of the subclades includes all Cenderawasih Bay individuals (subclade C); the second one includes all Chuuk Atoll individuals (subclade A); and the third one includes some individuals from Bali (Indonesia) and Palawan (Philippines), and three samples were purchased from a Taiwanese aquarium store, which were imported from the Philippines (TWSP 1-3) (subclade B). Those individuals collected from subclade B revealed that the distribution of this lineage is not only restricted to the area between Bali and Raja Ampat (Liu et al. 2012), but extends northward to the Philippines.

The genetic lineages we recovered within the *P. micronesicus* clade may represent cryptic species. Discovery of genetically distinct, but morphologically indistinguishable organisms, is a common phenomenon in the marine environment (e.g., Knowlton 2000). Based on our analysis of Cytb genetic distance, mean divergence among the three *P. micronesicus* lineages is greater than the mean divergence between comparisons involving four previously described species (*P. caeruleopunctatus* vs. *P. caeruleus*; *P. alleni* vs. *P. coelestis*; *P. alleni* vs. *P. caeruleus*; and *P. alleni* vs. *P. caeruleopunctatus*). Furthermore, mean genetic distance values between each of the *P. micronesicus* subclades to all other species are comparable to mean distances calculated between other species within the complex. Morphological evidence may also exist to support distinct species within the *P. micronesicus* clade. Examination of 23 specimens from Cenderawasih Bay, West Papua, Indonesia (subclade C), showed variation in anal fin ray counts and number of gill rakers compared to the *P. micronesicus*-type specimens from the Kwajalein Atoll, Marshall Islands (Liu et al. 2013). Systematic morphological examinations are now

underway to determine fixed morphological differences between these suspected cryptic lineages.

## Conclusions

Our results are some of the first to reveal Pleistocene origination and diversification of an entire marine fish species complex, lending support to the importance of land barriers and isolation in glacial refugia caused by sea-level fluctuations. Ancestral range reconstruction across the timetree suggests origination in the Coral Triangle and East Indian Ocean, with range expansion outward, resulting in two subclades: one Indian and one Pacific. The IPB therefore appears to be a factor in driving early speciation within the species complex. The presence of three lineages in semi-enclosed regions during Pleistocene lowstands (*P. micronesicus* C in Cenderawasih Bay, and *P. alleni* and *P. similis* in the Andaman Sea) suggests the importance of glacial refugia in the diversification of this fish group. Broad sampling for most species revealed cryptic lineages within *P. micronesicus* that are as genetically distinct from each other as some previously described species comparisons within the complex. Our comprehensive sampling regime, including all described species and multiple individuals per species, highlights the importance of complete sampling to reveal recent speciation events.

**Acknowledgments** We would like to thank Andrew Bentley (University of Kansas Biodiversity Institute) and Ofer Gon (South African Institute for Aquatic Biodiversity) for providing tissue samples of *Pomacentrus caeruleopunctatus* (SAIAB 80854 and SAIAB 77347, respectively), and Bruno Frederich for sharing samples of *Pomacentrus caeruleus* from Madagascar. This study was supported by a grant from the National Science Council of Taiwan ROC (NSC99-2811-M-002-151) to C.F.D.

## References

- Allen GR (1991) Damselfishes of the World. Mergus Publishers, Melle
- Allen GR (2002) Description of two new species of damselfishes (Pomacentridae: *Pomacentrus*) from Madagascar. *Aqua* 6:45–52
- Allen GR (2008) Conservation hotspots of biodiversity and endemism for Indo-Pacific coral reef fishes. *Aquat Conser Mar Freshw Ecosyst* 18:541–556
- Allen GR, Erdmann MV (2006) *Cirrhilabrus cenderawasih*, a new wrasse (Pisces: Labridae) from Papua, Indonesia. *Aqua J Ichthyol Aquat Biol* 11:125–131
- Allen GR, Erdmann MV (2012) Reef Fishes of the East Indies. Tropical Reef Research, Perth
- Barber PH, Bellwood DR (2005) Biodiversity hotspots: evolutionary origins of biodiversity in wrasses (*Halichoeres*: Labridae) in the Indo-Pacific and new World tropics. *Mol Phylogenet Evol* 35:235–253
- Bellwood DR, Wainwright PC (2002) The history and biogeography of fishes on coral reefs. In: Sale PF (ed) *Coral Reef Fishes*:

- Dynamics and Diversity on a Complex Ecosystem. Academic Press, New York
- Bellwood DR, Klanten S, Cowman PF, Pratchett MS, Konow N, van Herwerden L (2010) Evolutionary history of the butterflyfishes (f: Chaetodontidae) and the rise of coral feeding fishes. *J Evol Biol* 23:335–349
- Benzie JAH (1999) Major genetic differences between crown-of-thorns starfish (*Acanthaster planci*) populations in the Indian and Pacific Oceans. *Evolution* 53:1782–1795
- Briggs JC, Bowen BW (2012) A realignment of marine biogeographic provinces with particular reference to fish distributions. *J Biogeogr* 39:12–30
- Carpenter KE, Springer VG (2005) The center of the center of marine shore fish biodiversity: the Philippine Islands. *Environ Biol Fish* 72:467–480
- Carpenter KE, Barber PH, Crandall ED, Ablan-Lagman MCA, Mahardika GN, Ambariyanto et al (2011) Comparative phylogeography of the Coral Triangle and implications for marine management. *J Mar Biol* 2011:1–14
- Chappell J (1981) Relative and average sea level changes, and endo-, epi-, and exogenic processes on the Earth. In: Allison I (ed) Sea level, ice, and climatic change. International Association of Hydrological Sciences, vol 131. International Association of Hydrological Science, United Kingdom
- Choat JH, Klanten OS, van Herwerden L, Robertson DR, Clements KD (2012) Patterns and processes in the evolutionary history of parrotfishes (Family Labridae). *Biol J Linn Soc* 107:529–557
- Chow S, Hazama K (1998) Universal PCR primers for S7 ribosomal protein gene introns in fish. *Mol Ecol* 7:1247–1263
- Coates AG, Obando JA (1996) The geologic evolution of the Central American isthmus. In Jackson J, Budd AF, Coates AG (eds) Evolution and environment in tropical America, University of Chicago Press, Chicago
- Cowman PF, Bellwood DR (2011) Coral reefs as drivers of cladogenesis: expanding coral reefs, cryptic extinction events, and the development of biodiversity hotspots. *J Evol Biol* 24:2543–2562
- Craig MT, Eble JA, Bowen BW, Robertson DR (2007) High genetic connectivity across the Indian and Pacific Oceans in the reef fish *Myripristis berndti* (Holocentridae). *Mar Ecol Prog Ser* 334:245–254
- Drew J, Barber PH (2009) Sequential cladogenesis of the reef fish *Pomacentrus moluccensis* (Pomacentridae) supports the peripheral origin of marine biodiversity in the Indo-Australian archipelago. *Mol Phylogenet Evol* 53:335–339
- Drew J, Allen G, Kaufman L, Barber PH (2008) Endemism and regional colour and genetic differences in five putatively cosmopolitan reef fishes. *Conserv Biol* 22:965–975
- Drew JA, Allen GR, Erdmann MV (2010) Congruence between mitochondrial genes and color morphs in a coral reef fish: population variability in the Indo-Pacific damselfish *Chrysiptera rex* (Snyder, 1909). *Coral Reefs* 29:439–444
- Drummond AJ, Rambaut A (2007) BEAST: Bayesian evolutionary analysis by sampling trees. *BMC Evol Biol* 7:214
- Drummond AJ, Suchard MA, Xie D, Rambaut A (2012) Bayesian phylogenetics with BEAUti and the BEAST 1.7. *Mol Biol Evol* 29:1969–1973
- Edgar RC (2004) MUSCLE: multiple sequence alignment with high accuracy and high throughput. *Nucleic Acids Res* 32:1792–1797
- Fessler JL, Westneat MW (2007) Molecular phylogenetics of the butterflyfishes (Chaetodontidae): taxonomy and biogeography of a global coral reef fish family. *Mol Phylogenet Evol* 45:50–68
- Gaither MR, Rocha LA (2013) Origins of species richness in the Indo-Malay-Philippine biodiversity hotspot: evidence for the centre of overlap hypothesis. *J Biogeogr* 40:1638–1648
- Goetze E (2011) Population differentiation in the open sea: insights from the pelagic copepod *Pleuromamma xiphias*. *Integr Comp Biol* 51:580–597
- Kennett JP, Keller G, Srinivasan MS (1985) Miocene planktonic foraminiferal biogeography and paleogeographic development of the Indo-Pacific region. *Geol Soc Mem* 163:197–236
- Knowlton N (2000) Molecular genetic analyses of species boundaries in the sea. *Hydrobiologia* 420:73–90
- Leray M, Beldade R, Holbrook SJ, Schmitt RJ, Planes S, Bernardi G (2010) Allopatric divergence and speciation in coral reef fish: the three-spot *Dascyllus*, *Dascyllus trimaculatus*, species complex. *Evolution* 64:1218–1230
- Liu SYV, Dai CF, Allen GR, Erdmann MV (2012) Phylogeography of the neon damselfish *Pomacentrus coelestis* indicates a cryptic species and different species origins in the West Pacific Ocean. *Mar Ecol Prog Ser* 458:155–167
- Liu SYV, Ho HCH, Dai CF (2013) A new species of *Pomacentrus* (Actinopterygii: Pomacentridae) from Micronesia, with comments on its phylogenetic relationships. *Zool Stud* 52:6–13
- Matzke NJ (2013) BioGeoBEARS: BioGeography with Bayesian (and likelihood) evolutionary analyses in R scripts. Dissertation, University of California, Berkeley
- McCafferty S, Bermingham E, Quenouille B, Planes S, Hoelzer G, Asoh K (2002) Historical biogeography and molecular systematics of the Indo-Pacific genus *Dascyllus* (Teleostei: Pomacentridae). *Mol Ecol* 11:1377–1392
- McMillan WO, Palumbi SR (1995) Concordant evolutionary patterns among Indo-West Pacific butterflyfishes. *Proc R Soc B-Biol Sci* 260:229–236
- Messmer V, van Herwerden L, Munday PL, Jones GP (2005) Phylogeography of colour polymorphism in the coral reef fish *Pseudochromis fuscus*, from Papua New Guinea and the Great Barrier Reef. *Coral Reefs* 24:392–402
- Mora C, Sale PF (2002) Are populations of coral reef fish open or closed? *Trends Ecol Evol* 17:422–428
- Naish T, Powell R, Levy R, Wilson G, Scherer R, Talarico F, Krissek L, Niessen F, Pompilio M, Wilson T et al (2009) Oblivious-paced Pliocene West Antarctic ice sheet oscillations. *Nature* 458:322–328
- Palumbi SR (1994) Genetic divergence, reproductive isolation, and marine speciation. *Annu Rev Ecol Syst* 25:547–572
- Palumbi SR (1996) What can molecular genetics contribute to marine biogeography? An urchin's tale. *J Exp Mar Biol Ecol* 203:75–92
- Posada D (2008) JModelTest: phylogenetic model averaging. *Mol Biol Evol* 25:123–1256
- Potts DC (1983) Evolutionary disequilibrium among Indo-Pacific corals. *Bull Mar Sci* 33:619–632
- Randall JE (1998) Zoogeography of shore fishes of the Indo-Pacific region. *Zool Stud* 37:227–268
- R Development Team (2012) R: a language and environment for statistical computing. <http://www.r-project.org>
- Renema W, Bellwood DR, Braga JC, Bromfield K, Hall R, Johnson KG, Lunt P, Meyer CP, McMonagle LB, Morley RJ, O'Dea A, Tood JA, Wesselingh FP, Wilson MEJ, Pandolfi JM (2008) Hopping hotspots: global shifts in marine biodiversity. *Science* 321:654–657
- Ridgway T, Sampayo EM (2005) Population genetic status of the western Indian Ocean; What do we know? *West Indian Ocean J Mar Sci* 4:1–9
- Rocha LA, Bowen BW (2008) Speciation in coral-reef fishes. *J Fish Biol* 72:1101–1121
- Rohling EJ, Braun K, Grant K, Kucera M, Robers AP, Siddall M, Trommer G (2010) Comparison between Holocene and Marine Isotope Stage-11 sea-level histories. *Earth Planet Sci Lett* 291:97–105
- Ronquist F, Teslenko M, van der Mark P, Ayres DL, Darling A, Höhna S, Larget B, Liu L, Suchard M, Huelsenbeck JP (2012) MrBayes 3.2: efficient Bayesian phylogenetic inference and model choice across a large model space. *Syst Biol* 61:539–542

- Santini F, Nguyen MTT, Sorenson L, Waltzek TB, Lynch Alfaro JW, Eastman JM, Alfaro ME (2013a) Do habitat shifts drive diversification in teleost fishes? An example from the pufferfishes (Tetraodontidae). *J Evol Biol* 26:1003–1018
- Santini F, Sorenson L, Alfaro ME (2013b) A new multi-locus time-scale reveals the evolutionary basis of diversity patterns of triggerfishes and filefishes (Balistidae, Monacanthidae; Tetraodontiformes). *Mol Phylogenet Evol* 69:165–176
- Santini F, Sorenson L, Marcroft T, Dornburg A, Alfaro ME (2013c) A multilocus molecular phylogeny of boxfishes (Aracnidae, Ostraciidae; Tetraodontiformes). *Mol Phylogenet Evol* 66:153–160
- Satapoomin U (2002) In: Comparative study of reef fish fauna in Thai waters: the Gulf of Thailand versus the Andaman Sea. Technical Paper Phuket Marine Biological Center
- Schultz JK, Pyle RL, DeMartini E, Bowen BW (2007) Genetic connectivity among color morphs and Pacific archipelagos for the flame angelfish, *Centropyge loriculus*. *Mar Biol* 151:167–175
- Siddall M, Rohling EJ, Almogi-Labin A, Hemleben C, Meischner D, Schmelzer I, Smeed DA (2003) Sea-level fluctuations during the last glacial cycle. *Nature* 423:853–858
- Sorenson L, Santini F, Carnevale G, Alfaro ME (2013) A multi-locus timetree of surgeonfishes (Acanthuridae, Percomorpha), with revised family taxonomy. *Mol Phylogenet Evol* 68:150–160
- Stamatakis A (2006) RAxML-VI-HPC: maximum likelihood-based phylogenetic analyses with thousands of taxa and mixed models. *Bioinformatics* 22:2688–2690
- Steinke D, Zemlak TS, Hebert PDN (2009) Barcoding Nemo: DNA-based identifications for the ornamental fish trade. *Plos One*:e6300
- Tamura K, Peterson D, Peterson N, Stecher G, Nei M, Kumar S (2011) MEGA5: molecular evolutionary genetics analysis using maximum likelihood, evolutionary distance, and maximum parsimony methods. *Mol Biol Evol* 28:2731–2739
- Taylor MS, Hellberg ME (2005) Marine radiations at small geographic scales: speciation in neotropical reef gobies (*Elacatinus*). *Evolution* 59:374–385
- Vogler C, Benzie J, Barber PH, Erdmann MV, Ambariyanto, Shepard C, Tenggardjaja K, Gérard K, Wörheide G (2012) Phylogeography of the crown-of-thorns starfish in the Indian Ocean. *Plos One*:e43499
- Voris HK (2000) Maps of Pleistocene sea levels in Southeast Asia: shorelines, river systems and time durations. *J Biogeogr* 27:1153–1167
- Wallace CC, Paulay G, Hoeksema BW, Bellwood DR, Hutchings PA, Barber PH, Erdmann M, Wolstenholme J (2002) Nature and origins of unique high diversity reef faunas in the Bay of Tomini, Central Sulawesi: the ultimate “centre of diversity”? *P Ninth Int Coral Reef Symp* 1:185–192
- Woodland DJ (1983) Zoogeography of the Siganidae (Pisces): an interpretation of distribution and richness patterns. *Bull Mar Sci* 33:713–717