Molecular Phylogenetics and Evolution 176 (2022) 107584

Contents lists available at ScienceDirect

Molecular Phylogenetics and Evolution

journal homepage: www.elsevier.com/locate/ympev

Molecular phylogenetics reveals the evolutionary history of marine fishes (Actinopterygii) endemic to the subtropical islands of the Southwest Pacific

André P. Samayoa[a,*], Carl D. Struthers[b], Thomas Trnski[c], Clive D. Roberts[b], Libby Liggins[a,c]

[a] School of Natural Sciences, Massey University, Auckland 0745, New Zealand
[b] Museum of New Zealand Te Papa Tongarewa, P.O. Box 467, Wellington, New Zealand
[c] Natural Sciences, Auckland Museum Tāmaki Paenga Hira, Auckland 1010, New Zealand

ARTICLE INFO

Original content:
Alignments_TreeFiles_MolPhyloSWPEndemism (Original data)
Alignments_TreeFiles_MolPhyloSWPEndemism (Original data)

Keywords:
Ray-finned fishes
Time-calibrated phylogeny
Endemism
Southwest Pacific
Oceanic islands
Multi-locus data

ABSTRACT

Remote oceanic islands of the Pacific host elevated levels of actinopterygian (ray-finned fishes) endemism. Characterizing the evolutionary histories of these endemics has provided insight into the generation and maintenance of marine biodiversity in many regions. The subtropical islands of Lord Howe, Norfolk, and Rangitāhua (Kermadec) in the Southwest Pacific are yet to be comprehensively studied. Here, we characterize the spatio-temporal diversification of marine fishes endemic to these Southwest Pacific islands by combining molecular phylogenies and the geographic distribution of species. We built Bayesian ultrametric trees based on open-access and newly generated sequences for five mitochondrial and ten nuclear loci, and using fossil data for time calibration. We present the most comprehensive phylogenies to date for marine ray-finned fish genera, comprising 34 species endemic to the islands, including the first phylogenetic placements for 11 endemics. Overall, our topologies confirm the species status of all endemics, including three undescribed taxa. Our phylogenies highlight the predominant affinity of these endemics with the Australian fish fauna (53%), followed by the East Pacific (15%), and individual cases where the closest sister taxon of our endemic is found in the Northwest Pacific and wider Indo-Pacific. Nonetheless, for a quarter of our focal endemics, their geographic affinity remains unresolved due to sampling gaps within their genera. Our divergence time estimates reveal that the majority of endemic lineages (67.6%) diverged after the emergence of Lord Howe (6.92 Ma), the oldest subtropical island in the Southwest Pacific, suggesting that these islands have promoted diversification. However, divergence ages of some endemics pre-date the emergence of the islands, suggesting they may have originated outside of these islands, or, in some cases, ages may be overestimated due to unsampled taxa. To fully understand the role of the Southwest Pacific subtropical islands as a 'cradle' for diversification, our study advocates for further regional surveys focused on tissue collection for DNA analysis.

1. Introduction

Marine biodiversity is being lost at an accelerated rate, driven by anthropogenic pressures such as overharvesting (Costello et al., 2010), habitat fragmentation (Lotze et al., 2006), and climate change (Beaugrand et al., 2015). In particular, endemic species are more vulnerable to extinction than widespread taxa because of their restricted spatial distribution (Norman, 2003) and specialist niche requirements (Le Feuvre et al., 2021). For this reason, areas of high endemism have been of great research interest, shedding light on how biodiversity is generated and maintained (Bowen et al., 2013), and therefore how it is best conserved (Moritz, 2002). Molecular phylogenetics has been crucial in characterizing the evolutionary history of endemic species in these regions, and, consequently, the macroevolutionary processes that have given rise and maintained unique regional biodiversity (Bellwood and Meyer, 2009). Still, many regions are yet to be comprehensively researched (Hortal et al., 2015), and the evolutionary history of their endemic species characterized. Such understanding is increasingly urgent as pressures on biodiversity continue to escalate (Nunez et al., 2019).

Marine actinopterygians (ray-finned fishes) have been extensively studied in the last decade, providing comprehensive knowledge of their evolutionary relationships and temporal diversification patterns

---

* Corresponding author.
E-mail addresses: A.P.Samayoa@massey.ac.nz (A.P. Samayoa), CarlS@tepapa.govt.nz (C.D. Struthers), ttrnski@aucklandmuseum.com (T. Trnski), CliveR@tepapa.govt.nz (C.D. Roberts), L.Liggins@massey.ac.nz (L. Liggins).

https://doi.org/10.1016/j.ympev.2022.107584
Received 25 February 2022; Received in revised form 3 May 2022; Accepted 12 May 2022

globally (Alfaro et al., 2018, Betancur-R et al., 2013, Betancur-R et al., 2015, Betancur-R et al., 2017, Hughes et al., 2018, Matschiner et al., 2017, Mirande, 2017, Near et al., 2013, Near et al., 2012, Rabosky, 2020, Rabosky et al., 2018). In particular, the deep nodes in the phylogeny of all marine ray-finned fishes have been the subject of intense research; however, towards the tips of the phylogeny, our understanding of the evolutionary relationships among finer taxonomic units (e.g. genera and species) diminishes, and a bias toward the most speciose, conspicuous, and easy to study taxa, increases (Cowman, 2014). This phylogenetic knowledge gap is exacerbated by survey efforts that tend to favor easy-to-access regions with substantial research resources, delaying comprehensive species inventories and associated phylogenetic research in more isolated areas (Hortal et al., 2015). However, it is in isolated oceanic islands that the highest proportions of endemic marine fishes are observed (Van Der Meer et al., 2015), highlighting the importance of turning our attention to these remote centers of endemism to better understand the evolutionary and ecological processes shaping their biodiversity.

Within the Indo-Pacific Ocean, the tropical latitudes are widely recognized as hosting the highest levels of marine biodiversity, where the Coral Triangle represents the region's "bullseye" (Hoeksema, 2007). Synthetic literature has discussed the evolutionary and ecological processes that have formed and maintained this unique pattern (Cowman, 2014, Gaboriau et al., 2018), including the role of diversification in peripheral regions (Bowen et al., 2013). At the periphery of the Pacific Ocean, there tend to be geographically isolated islands characterized by elevated rates of endemism and low species richness (Cowman et al., 2017). For instance, Juan Fernández and Desventuradas in the East Pacific have the highest proportion of endemic coastal marine fishes (61.5%, Friedlander et al., 2016), followed by the Hawaiian archipelago in the Central North Pacific (25%, Randall, 2007), Rapa Nui (Easter Island) (21.7%, Randall and Cea, 2011), and the Marquesas Islands in the Central Pacific (13.7%, Delrieu-Trottin et al., 2015). Studies that combine the checklist of a region's marine fish fauna and the available molecular data in a phylogenetic framework have unveiled biogeographic patterns and the origin of species and biogeographic patterns in tropical areas (e.g. Hodge and Bellwood, 2016), including in the peripheral islands of Hawaii (Hoban and Williams, 2020, Hodge et al., 2014) and Rapa Nui (Delrieu-Trottin et al., 2019) providing a general understanding of the evolutionary histories of the fish fauna.

The Southwest Pacific harbors peripheral oceanic islands where the fish fauna is relatively understudied from an evolutionary perspective (Liggins et al., 2021). Lord Howe Island and Norfolk Island of Australia, and Rangitāhua (the Kermadec Islands) of New Zealand (Aotearoa) straddle subtropical waters (29–32°S, Francis, 1993) (Fig. 1) and are geologically young. Their estimated age of emergence due to intense volcanic activity is 6.92 Ma for Lord Howe (McDougall et al., 1981), 3.05 Ma for Norfolk (Jones and McDougall, 1973), and the Early Pleistocene (i.e. no older than 2.58 Ma) for Rangitāhua (Brook, 1998b). Because of their geographic position, the marine fish fauna is characterized by a mixture of tropical, subtropical, and temperate species (Francis and Duffy, 2015) and the overall rate of endemism for coastal fishes across the three island groups has been reported at 4.6% with individual rates in the 1.2–2.1% range (Francis, 1993). However, since these early studies, there have been several surveys and expeditions to the islands, resulting in the addition of further taxa to the fish checklists and the description of new endemic species (Francis, 2019). In particular, Rangitāhua has been the subject of intense sampling which has led to recent comprehensive checklists for its marine fauna and flora (Duffy and Ahyong, 2015), and its coastal fishes with an updated endemicity rate of 4.6% for this group (Trnski et al., 2015).

As a result of decades of field work, specimen curation, and laboratory work, we live in an age where it is easy to access large volumes of molecular, biological, and ecological information (Allendorf et al., 2010, Hoban et al., 2021). Nevertheless, the rates at which data are being generated, and ulteriorly analyzed differ considerably, increasing the amount of information yet to be evaluated (Hortal et al., 2015). Phylogenetic analyses are one means to summarize molecular data that has been accumulated across disparate taxa and research groups, helping to quickly contextualize new information that becomes available (Antonelli et al., 2017). Therefore, based on recent updates to species checklists for Lord Howe, Norfolk, and Rangitāhua, the acquisition of specimens endemic to the islands, the accumulation of open-access molecular sequences, and phylogenetic knowledge of marine fishes, we are well-placed to improve our understanding of how marine endemism has evolved in these peripheral islands of the Pacific.

In this study, we build the most comprehensive multi-locus time-calibrated phylogenies to date for marine ray-finned fishes occurring in the Southwest Pacific and endemic to the islands of Lord Howe Island, Norfolk Island, and/or Rangitāhua, to shed light on their evolutionary history. Firstly, we aimed to corroborate the species status of each endemic by phylogenetically placing it alongside its presumed sister taxa using molecular sequences. Secondly, we characterized the geographic origin of marine fish endemism in the Southwest Pacific region by associating each taxon to its range within each topology. Finally, we provide a temporal context for diversification in the Southwest Pacific region by estimating the divergence times of these endemic lineages using fossil-based calibration points. Our study generates new knowledge regarding the phylogenetic relationships of fishes endemic to the subtropical islands of the Southwest Pacific, and helps to highlight

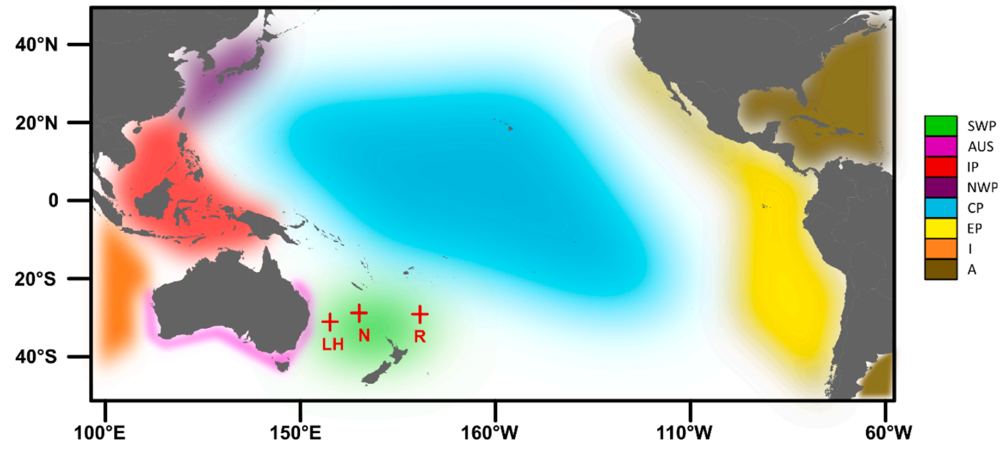

Fig. 1. Geographic location of Lord Howe Island (LH), Norfolk Island (N), and Rangitāhua (the Kermadec Islands) (R) in the Southwest Pacific region. The color palette illustrates the geographic regions defined in our study to classify the origin of sampled taxa: SWP: Southwest Pacific; AUS: Australian coasts; IP: Indian Ocean and West Pacific, centred in the Coral Triangle; NWP: Northwest Pacific; CP: Central Pacific; EP: East Pacific; I: Indian Ocean only; A: Atlantic Ocean.

persisting gaps in taxonomic sampling that preclude definitive dating, the inference of their biogeographic origins, and modes by which endemic species evolved in the Southwest Pacific.

2. Materials and methods

2.1. Taxonomic sampling and study approach

We focused on marine fish taxa that are distributed in at least one of the three oceanic islands of Lord Howe, Norfolk, and Rangitāhua, and occurring in areas within the Southwest Pacific region (i.e. not occurring beyond Aotearoa New Zealand mainland, New Caledonia, and the east coast of Australia; Fig. 1), but not being widespread throughout the Southwest Pacific. We assembled a list of 60 taxa based on this criterion (hereafter “endemic taxa”) using checklists from recent expeditions to Rangitāhua (Duffy and Ahyong, 2015, Francis and Duffy, 2015, Roberts et al., 2015, Trnski et al., 2015), and the most up-to-date checklist for Lord Howe, Norfolk, and Rangitāhua (Francis, 2019). This initial list was further refined according to DNA sequence data availability to select the endemics for subsequent analysis. For instance, we only kept endemics for which we had access to either DNA sequences or tissue for DNA extraction and sequencing. For the selected endemics, we also required that several sister taxa within their genus (and/or family) had either molecular sequences or tissues available, that could be used to phylogenetically place the endemics. The list of all sister taxa was built based on searches through World Register of Marine Species (WoRMS; WoRMS Editorial Board, 2021). To retrieve existing sequences for endemics and sister taxa, we searched the open-access sequence repositories of the National Center for Biotechnology Information (NCBI; https://www.ncbi.nlm.nih.gov/) and the Barcode of Life Data System (BOLD; https://www.barcodinglife.org). For endemics and sister taxa without available sequence information, we requested tissues from existing collections to generate novel sequences. For each endemic, the current valid scientific and vernacular names, and geographic distribution information were retrieved from the Eschmeyer’s Catalog of Fishes (Fricke et al., 2021), Fishbase (Froese and Pauly, 2021), and faunal checklists of the Southwest Pacific (Francis, 2019, Duffy and Ahyong, 2015, Roberts et al., 2015, Roberts et al., 2020) and Australia (Fishes of Australia; https://fishesofaustralia.au.net). For each sister taxon, taxonomic validity and their geographic range were verified using the Eschmeyer’s Catalog of Fishes and FishBase. The final list of endemics for subsequent analysis included 34 taxa (Table 1).

The selection of outgroup taxa for phylogenetic analyses included representatives of related genera within the family of the endemics, related families of the same order, and members of related orders, based on the hypothetical evolutionary relationships among species in published literature and global phylogenies of ray-finned fishes (Alfaro et al., 2018, Betancur-R et al., 2017, Hughes et al., 2018, Matschiner et al., 2017, Near et al., 2013, Near et al., 2012, Rabosky et al., 2018). We specifically aimed for taxa for which fossil data are available according to Matschiner et al. (2017) to inform inference of time-calibrated trees. When possible, endemic and sister taxa that shared fossil calibration points with other endemic and sister taxa were grouped in a single phylogeny to use the same outgroup taxa. Consequently, the 34 endemics were analyzed in 13 phylogenies (Table 1). In our taxonomic sampling strategy, *Notocirrhitus splendens* represented an exceptional case within the NEMOGOCHIA phylogeny (naming as described below), since it is both a monotypic genus that we identify as an endemic – distributed in East Australia, Lord Howe, Norfolk, and Rangitāhua – and a representative of the family Cirrhitidae which has been consistently used as an outgroup to root trees of Aplodactylidae (Burridge, 2000), Chironemidae (Burridge et al., 2006), and Cheilodactylidae/Latridae (Burridge and Smolenski, 2004, Ludt et al., 2019). Given that we had no available sequences for all the presumed sister taxa of *N. splendens*, and that missing species overestimate divergence times among phylogenetic congeners (Hodge and Bellwood, 2016), we opted to use this endemic as an outgroup in the phylogeny and exclude it from our final list of focal endemics.

2.2. DNA sequence retrieval and generation

We extracted DNA sequences from NCBI and BOLD for 389 of the 401 species that comprise ingroups and outgroups for the 13 phylogenies. To do this, an automated sequence search was facilitated through the ‘regPhylo’ package (Eme et al., 2019) for R v.3.6.3 (R Core Team, 2020) using RStudio v.1.2.5033 (RStudio Team, 2019) following the functions ‘GetSeqInfo_NCBI_taxid’ to search NCBI and ‘GetSeq_BOLD’ to search BOLD. As a general rule, a locus was retained for phylogenetic analysis when it was represented in at least 50% of the ingroup and outgroup species list per phylogeny (see Tables 1S-13S). Loci were exceptionally retained for phylogenetic analysis when they covered unique ingroup species despite not reaching the 50% threshold, or when they were the only nuclear loci and nearly covered 50%. Overall, we downloaded sequences for five mitochondrial markers (the 12S and 16S non-coding ribosomal RNA regions, the non-coding D-loop, and the two coding COI and Cytb) and ten nuclear loci (protein-coding regions RAG1, RAG2, Tmo-4C4, H3, MYH6, PLAGL2, ENC1, GLYT, SH3PX3, and Ptr) to be used in the 13 phylogenies (Table 2). Previous studies have revealed that some fish sequences sourced from NCBI and BOLD are assigned to species identities incorrectly (e.g. Liggins et al., 2021, Tang et al., 2021) which can impact phylogenetic inferences. In our study, we minimized the probability of incorporating sequences from misidentified taxa by: not using “blacklisted” sequences identified in previous studies (Eme et al., 2020, Eme et al., 2019, Tang et al., 2021); ensuring most accessed sequences were supported by vouchers and/or isolate numbers where these metadata were provided; verifying that taxa fell within expected taxonomic groups and positions during phylogenetic inference; and where we encountered discrepancy, we preferentially used sequences included in peer-reviewed reputable phylogenies, or those we knew had corresponding voucher specimens under the authors’ care (e.g. Eme et al., 2020, Eme et al., 2019).

For species that had no sequence information available in NCBI and BOLD, or from collaborators, we accessed tissue from within curated collections of the Museum of New Zealand Te Papa Tongarewa (NMNZ, The National Fish Collection), Auckland Museum Tāmaki Paenga Hira (AIM), Massey University, and the Australian Museum (AMS) to generate novel sequences. All DNA extraction, PCR, and sequencing preparation for the focal taxa was carried out at Massey University Auckland. Genomic DNA was extracted using the DNeasy Blood and Tissue kit (Qiagen, Valencia, CA). To amplify a portion of the COI gene region, we used the primer combination named Fish COI-2 cocktail (Ivanova et al., 2007) or FishF1 and FishR1 or FishR2 (Ward et al., 2005); for the RAG1 gene region, we used Rag1F1 and Rag1R1 (López et al., 2004); and for the 12S gene region, we used 12S53F and 12S613R, and 12S489F and 12S991R to amplify two fragments (McCord and Westneat, 2016). All PCRs were conducted using either the MyTaq^{TM} or MyFi^{TM} DNA polymerase kits (Bioline, Australia Pty Ltd, Alexandria, NSW) as per the kit instructions. For the Fish COI-2 primer cocktail, PCR was performed with a denaturation at 94 °C for 1 min, followed by an initial 5 cycles (94 °C for 30 secs, 50 °C for 40 secs, 72 °C for 1 min), followed by 35 cycles (94 °C for 30 secs, 54 °C for 40 secs, 7 °C for 1 min), then a final extension at 72 °C for 10 mins (as per Ivanova et al., 2007). Using FishF1 and FishR1 or FishR2, PCR was performed with a denaturation at 95 °C for 2 mins, followed by 35 cycles (95 °C for 45 secs, 54 °C for 45 secs, 72 °C for 1 min), then a final extension at 72 °C for 5 mins. For the rag 1 and 12S gene regions, PCR was performed with a denaturation at 94 °C for 5 mins, followed by 40 cycles (94 °C for 30 secs, 50 °C for 30 secs, 72 °C for 45 secs), then a final extension at 72 °C for 3 mins. Following PCR, a 1% agarose gel was run using 2 μL of PCR product and 1 μL of GelRed, alongside a BioLabs Quick Loading DNA Ladder to ensure PCR products were of the right size and sufficient concentration. PCR products were then purified using the ExoSap

Table 1

The 34 marine fish taxa included in our study that are endemic to the subtropical islands of Lord Howe, Norfolk, and/or Rangitāhua (Kermadec) in the Southwest Pacific. Current valid scientific and vernacular names; associated range (AUS: eastern coast of Australia; NC: New Caledonia; LH: Lord Howe; N: Norfolk; R: Rangitāhua - Kermadec Islands; NZ: mainland New Zealand); genus assessed and, in brackets, the number of taxa represented in the phylogeny over the total number of species in the genus; and the name of the phylogeny it is included in. The GAMS phylogeny includes *Girella*, *Atypichthys*, *Microcanthus*, and *Scorpis*. The HL phylogeny includes *Hypoplectrodes* and *Lepidoperca*. The NEMOGOCHIA phylogeny includes *Nemadactylus*, *Morwong*, *Goniistius*, *Chironemus*, and *Aplodactylus*. *Lepidoperca inornata* was selected as an endemic associated with Rangitāhua as it occurs along the Kermadec Ridge. *Girella fimbriata*, *Morwong fuscus*, and *Upeneus francisi* were not considered in mainland New Zealand, as their occurrence in northern locations is rare (Middleton et al., 2021, Middleton et al., In review).

Endemic species	Range						Genus	Phylogeny name
	AUS	NC	LH	N	R	NZ		
Arripis trutta (Forster, 1801) - Kahawai	X		X	X	X	X	*Arripis* (4/4)	Arripis
Arripis xylabion Paulin, 1993 - Northern kahawai			X	X	X	X		
Capromimus abbreviatus (Hector, 1875) - Capro dory					X	X	*Capromimus* (1/1)	Capromimus
Chromis abyssicola Allen & Randall, 1985 - Deepwater demoiselle				X	X	X	*Chromis* (85/110*-59/84**)	Chromis
Chromis dispila Griffin, 1923 - Twospot demoiselle					X	X		
Chromis hypsilepis (Günther, 1867) - Onespot demoiselle	X		X	X		X		
Chromis kennensis Whitley, 1964 ^^ - Yellowspot Puller	X	X	X	X	X			
Chromis nitida (Whitley, 1928) - Barrier reef chromis	X		X					
Enneapterygius kermadecensis Fricke, 1994 - Kermadec triplefin					X		*Enneapterygius* (27/63)	Enneapterygius
Eviota kermadecensis Hoese & Stewart, 2012 - Feathery goby					X		*Eviota* (62/117)	Eviota
Flexor incus Conway, Stewart & Summers 2018 - Kermadec clingfish		X	X		X		*Flexor* (1/1)	Flexor
Girella cyanea Macleay, 1881 - Bluefish	X		X	X	X	X	*Girella* (16/17)	GAMS
Girella fimbriata (McCulloch, 1920) - Caramel drummer					X			
Atypichthys latus McCulloch & Waite, 1916 - Mado			X	X	X	X	*Atypichthys* (2/2)	
Microcanthus joyceae Whitley, 1931 - East-Australian stripey	X	X	X	X			*Microcanthus* (2/2)	
Scorpis violacea (Hutton, 1873) - Blue maomao	X		X	X	X	X	*Scorpis* (4/5)	
Hypoplectrodes sp. A (*sensu* Roberts et al., 2015:1189) - Eyebrow perch	X		X	X	X	X	*Hypoplectrodes* (9/11^)	HL
Hypoplectrodes sp. C (*sensu* Roberts et al., 2015:1192) - Kermadec halfbanded perch					X			
Lepidoperca inornata Regan 1914 - Plain perch					X	X	*Lepidoperca* (8/10)	
Kathetostoma binigrasella Gomon & Roberts, 2011 - Banded stargazer					X	X	*Kathetostoma* (7/8)	Kathetostoma
Nemadactylus douglasii (Hector, 1875) - Porae	X				X	X	*Nemadactylus* (7/8)	NEMOGOCHIA
Nemadactylus n. sp. (*sensu* Roberts et al., 2015:1357) - King tarakihi	X		X	X	X	X		
Morwong ephippium (McCulloch & Waite, 1916) - Painted moki	X		X	X	X	X	*Morwong* (2/2)	
Morwong fuscus (Castelnau, 1979) - Red morwong	X		X					
Goniistius francisi (Burridge, 2004) - Masked morwong		X	X	X	X		*Goniistius* (9/9)	
Goniistius vestitus (Castelnau, 1879) - Crested morwong	X	X	X	X				
Chironemus marmoratus Günther, 1860 - Hiwihiwi	X		X			X	*Chironemus* (6/6)	
Chironemus microlepis Waite, 1916 - Northern kelpfish			X	X	X			
Aplodactylus etheridgii (Ogilby, 1889) - Notch-head marblefish			X	X	X	X	*Aplodactylus* (5/5)	
Optivus agastos Gomon, 2004 - Violet roughy	X	X	X				*Optivus* (3/3)	Optivus
Optivus elongatus (Günther, 1859) - Slender roughy					X	X		
Parma alboscapularis Allen & Hoese, 1975 - Black angelfish			X	X	X	X	*Parma* (7/10)	Parma
Parma kermadecensis Allen, 1987 - Kermadec scalyfin					X			

(*continued on next page*)

Table 1 (*continued*)

Endemic species	Range						Genus	Phylogeny name
	AUS	NC	LH	N	R	NZ		
Upeneus francisi Randall & Guézé, 1992 - Bartail goatfish			X	X	X		*Upeneus* (29/37)	Upeneus

Chromis sensu lato* and ***Chromis sensu stricto* as in Tang et al. (2021); ^includes the eight taxa from WoRMS and the unnamed species A, B, and C; ^^found also in Tonga (Iwatsubo and Motomura, 2018).

Table 2
Gene regions used for the inference of the 13 phylogenies. The GAMS phylogeny includes *Girella*, *Atypichthys*, *Microcanthus*, and *Scorpis*; the HL phylogeny includes *Hypoplectrodes* and *Lepidoperca*; the NEMOGOCHIA phylogeny includes *Nemadactylus*, *Morwong*, *Goniistius*, *Chironemus*, and *Aplodactylus*. Sizes (bp) of mitochondrial (italic) and nuclear (regular text) loci are provided in brackets. The total length of the concatenated sequences used in phylogenetic analysis is provided in the last column.

Phylogeny name	Gene regions (bp)	Concatenated alignments (bp)
Arripis	*16S (1643), COI (651), Cytb (1138)*	3432
Capromimus	*12S (561), 16S (501), COI (651)*, GLYT (843), H3 (310), MYH6 (777), PLAGL2 (789), SH3PX3 (708)	5140
Chromis	*12S (652), 16S (550), COI (539), Cytb (576)*, RAG1 (897)	3214
Enneapterygius	*12S (167), COI (649)*	816
Eviota	*12S (147), COI (1548)*, Ptr (594)	2289
Flexor	*12S (350), COI (648)*, ENC1 (741), MYH6 (658)	2397
GAMS	*16S (479), COI (651), Cytb (420)*, RAG1 (1371), RAG2 (765), Tmo-4C4 (477)	4163
HL	*COI (650)*	650
Kathetostoma	*16S (530), COI (650), Cytb (368), D-loop (357)*, ENC1 (747), RAG1 (1257)	3909
NEMOGOCHIA	*16S (571), COI (650), Cytb (327)*, RAG1 (1410)	2958
Optivus	*COI (651), Cytb (1134)*, MYH6 (729)	2514
Parma	*16S (544), COI (648), Cytb (699)*, RAG1 (1457), Tmo-4C4 (511)	3859
Upeneus	*12S (170), COI (651)*	821

reagents and protocol (Thermo Fisher Scientific, West Palm Beach, FL) and sent for forward and reverse sequencing (Macrogen, Korea). Quality control of the received sequences was carried out using Geneious v.9.0.5 (https://www.geneious.com). All sequence chromatographs were inspected by eye and poor-quality nucleotide bases and primer sequences were trimmed before the alignment of the forward and reverse sequence (where available) to check for consensus. All generated sequences are deposited in NCBI (Accessions: ON368279-91 [CO1], ON387613-7 [12S]) and metadata uploaded to the Genomic Observatories MetaDatabase (GEOME; Deck et al., 2017, Riginos et al., 2020; accessioned at https://n2t.net/ark:/21547/ECL2 [CO1] and https://n2t.net/ark:/21547/EBx2 [12S]).

2.3. Maximum likelihood phylogenetic inference

The 'regPhylo' package was used to perform sequence alignment, trimming, partitioning, and to build a guide tree for a maximum likelihood (ML) inference in RAxML (Stamatakis, 2014). The best sequence per locus per taxon was selected using the median length option of the 'SelBestSeq' function. The absence of stop codons in protein-coding sequences and general alignment quality were checked in Geneious v.9.1.8. Multiple sequence alignments were run by calling MAFFT (Katoh et al., 2002) from 'regPhylo'. Alignments were trimmed with the 'Filtering.align.Trimal' function that calls the program TrimAl (Capella-Gutiérrez et al., 2009) after which we opted for the output from the automated1 option that selects the best-trimmed alignment based on a heuristic approach. All alignments were visually inspected, and manual trimming was required for Cytb and 12S sequences. To aid in the alignment, trimming, and sequence selection: for Cytb, we used a reference sequence that includes adjacent gene regions (accession number GU135519, recommended by Li et al., 2018); and for 12S, we used a reference sequence (accession number AP019333) to help identify the two subregions of the gene region, choosing the subregion with the maximum overlap among taxa for subsequent analysis. The accession numbers for all sequences (retrieved and novel) used in subsequent phylogenetic analyses are included as Supplementary Material.

For each phylogeny, aligned sequences were concatenated in multilocus supermatrices using the 'Align.Concat' function (exceptionally not applied for the COI-based HL phylogeny, see Table 2), after which the program PartitionFinder2 (Lanfear et al., 2017) was called from 'regPhylo' to partition our alignments with the following options: branch lengths linked; greedy search; all models; and AICc model selection criterion. The option RAxML was additionally set to 'True' as this fits the best partitioning scheme according to the substitution models implemented in the program. The function 'ConstraintTaxo2newick' was used to build a multifurcating tree based on soft constraints derived from previously published phylogenies and ready to use in a rapid bootstrap analysis in RAxML (as in Eme et al., 2019, Eme et al., 2020). The program RAxML was run as a command-line interface selecting the GTRGAMMA model of substitution with the autoMRE option which selects the best number of bootstrap replicates to compute bootstrap values for each node. The inference included the partitioning scheme and the guide tree from the previous steps. The best tree was rerooted in Dendroscope v.3.5.9 (Huson and Scornavacca, 2012) using the hypothetical most distantly related taxon per phylogeny. The final rooted tree and bootstrap support values per node were visualized in FigTree v.1.4.4.

2.4. Bayesian estimates of divergence times

BEAST2 (Bouckaert et al., 2019) was used to infer the 13 time-calibrated trees based on the same partitioning scheme we employed during the ML inference. A GTR + G substitution model with four categories of evolutionary rates was set for each subset after selecting the Standard GTR model in the SSM package for BEAUTi (Bouckaert and Xie, 2017). Trees and an uncorrelated relaxed clock model were linked across subsets. A Birth-Death speciation model was assumed in all cases and hard constraints were introduced based on the nodes that displayed 100 bootstrap support on the ML tree with a log-normal distribution prior on each one.

The CA package of Matschiner et al. (2017) for BEAUti was used to include fossil constraints based on the fossil records of the selected taxa from the initial taxonomic sampling and the parameters indicated by the authors (diversification rate of 0.041–0.081; turnover rate of 0.0011–0.37; sampling rate of 0.0066–0.01806). The maximum and minimum ages per fossil (Table 14S) were set according to the first occurrence of each fossil as indicated in the supporting material of Matschiner et al. (2017). In addition, we opted to include two of the fossils used by Frédérich et al. (2013) to calibrate their pomacentrid tree (*Morone* †sp. considered a relevant upper boundary at 74 Ma to root the tree; and *Chromis* †*savornini* to date the origin of the genus *Chromis*, using the age range 5.332–7.246 Ma as indicated in Fossilworks at https://fossilworks.org). The use of fossils as primary calibration points

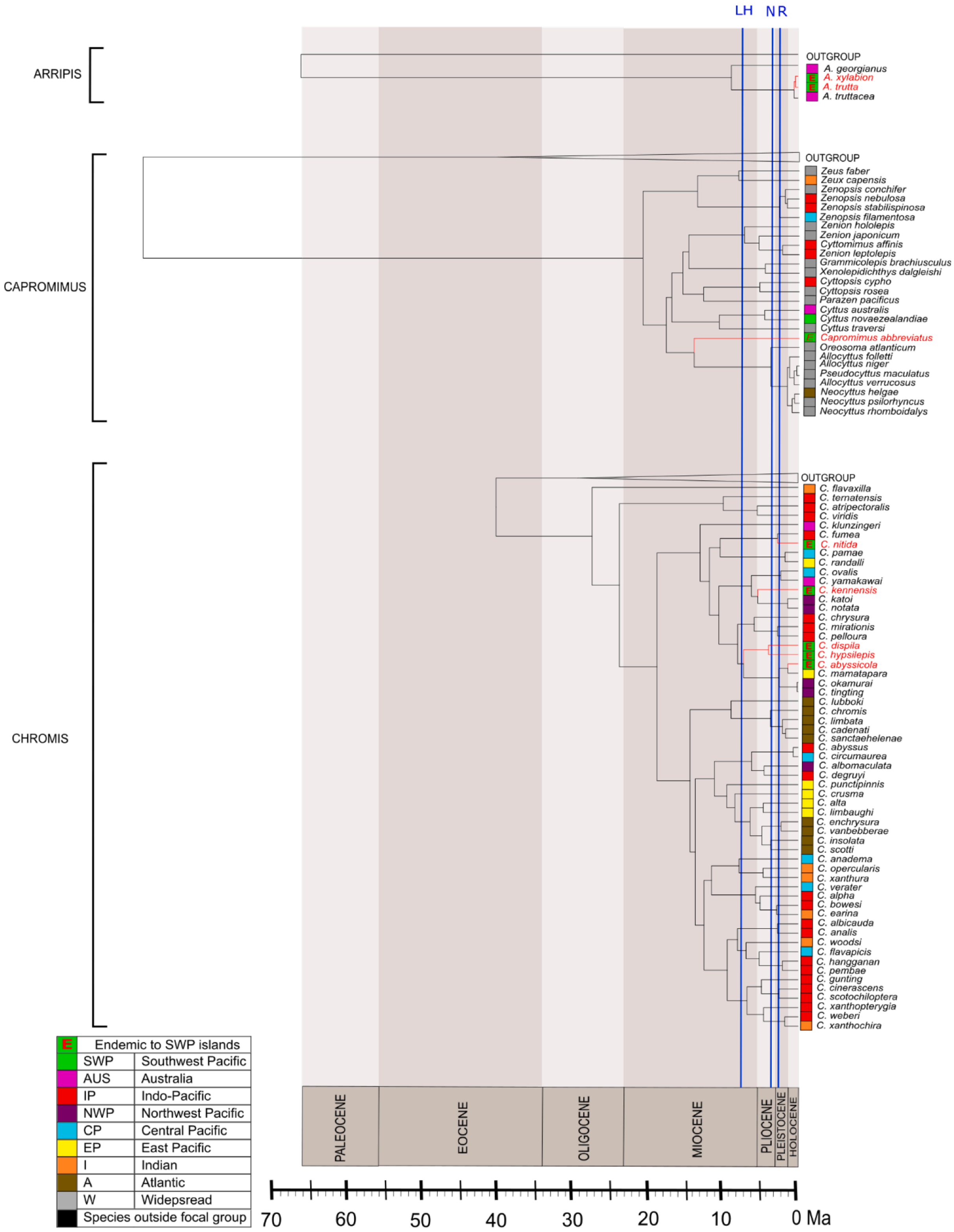

Fig. 2. Molecular placing of 34 marine fish species endemic to the Southwest Pacific islands across 13 time-calibrated Bayesian trees. An endemic's phylogenetic position is displayed as a red branch. Phylogenies and the ages of focal islands (vertical blue lines: LH: Lord Howe; N: Norfolk; R: Rangitāhua - the Kermadecs) are placed on a temporal axis in million years (Ma). Colored squares at the tips of the trees illustrate the range of each taxon following the colored geographic regions of Fig. 1, except for grey (widespread species) and black (outside focal taxonomic group). Squares for the focal endemic taxa are indicated with a red 'E' over the SWP green background. Species omnipresent in the SWP, or restricted to New Zealand mainland, New Caledonia, and/or the east/south of Australia retain the SWP green color. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

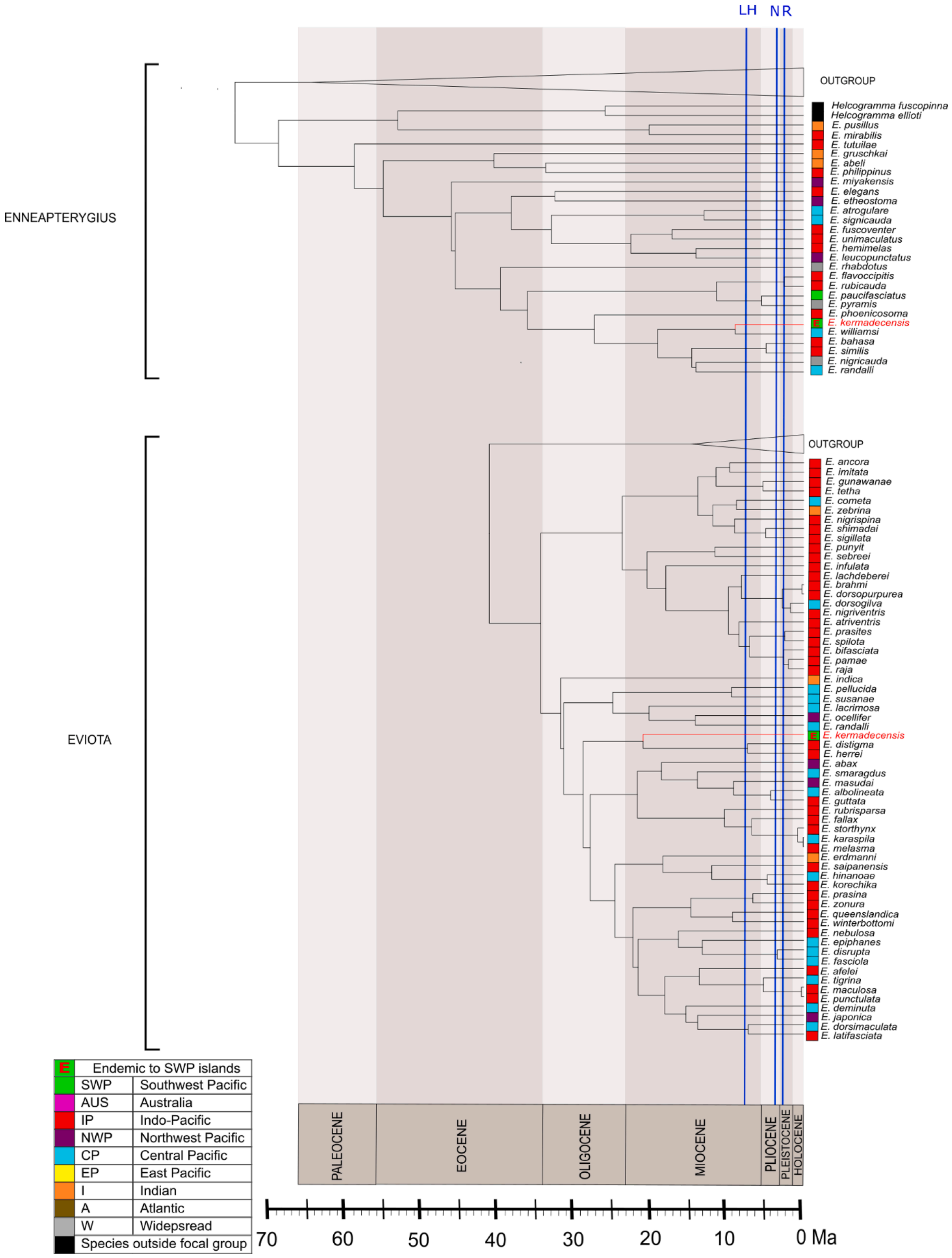

Fig. 2. (*continued*).

was preferred over other methods for divergence time estimations, as this approach avoids the additional errors from the use of secondary calibrations and the circularity of using geological calibration events (Powell et al., 2020, Sauquet, 2013).

For each phylogeny, three independent BEAST2 analyses were run with the BEAGLE package (Ayres et al., 2012) for performance improvement. The chain length of every run was set to 70,000,000 steps, with samples taken every 7,000 iterations, and discarding the first 10% as burn-in. The program Tracer v.1.7.1 (Rambaut et al., 2018) was used to verify effective mixing of the chains, appropriate burn-in, and that

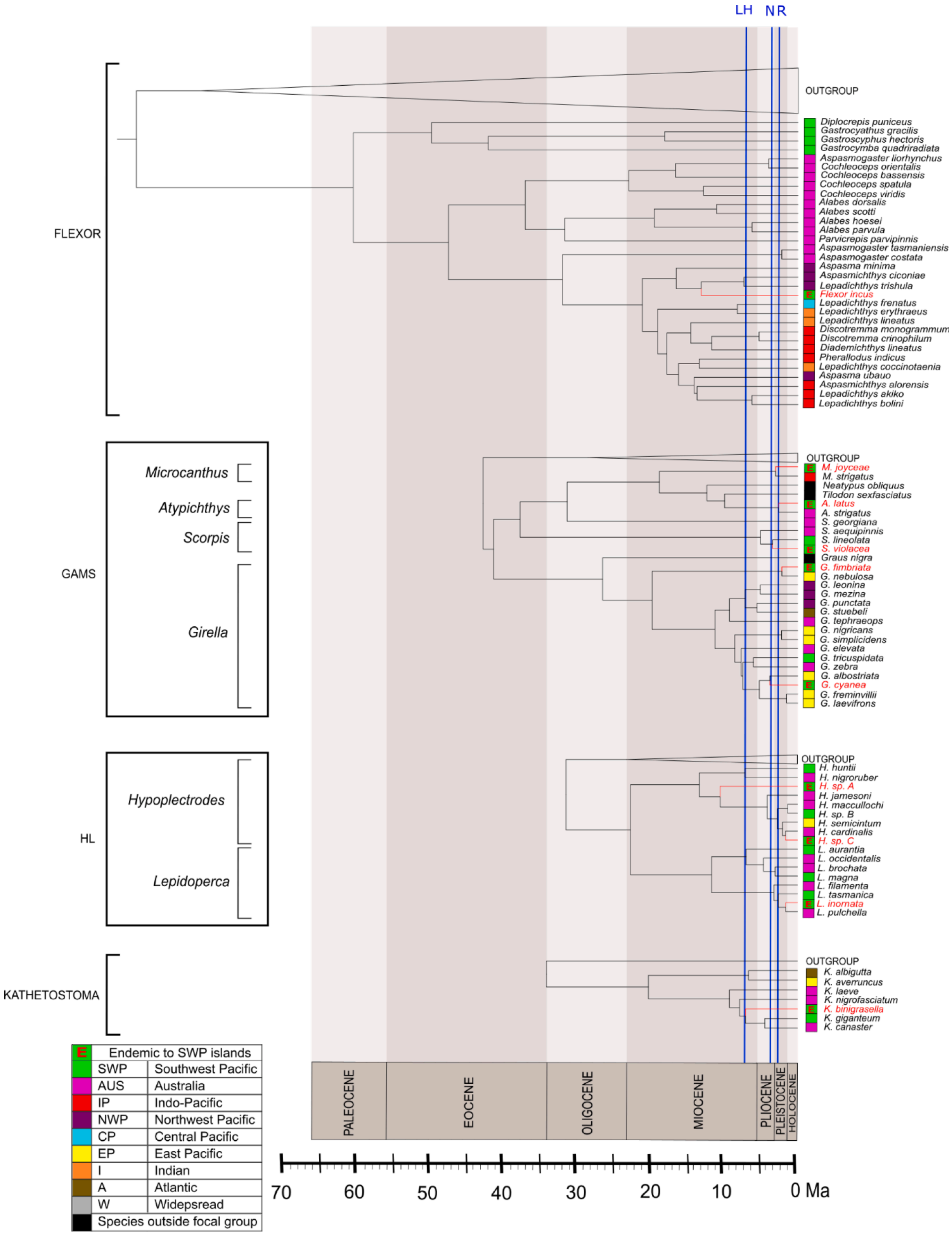

Fig. 2. (*continued*).

each run reached convergence by ensuring that the effective sample size (ESS) computed for every parameter yielded values of over 200. The log and tree files of the three BEAST2 analyses were combined with LogCombiner v.2.6.3 (Bouckaert et al., 2019). The posterior probability of the trees was summarized in a Maximum Clade Credibility Tree in TreeAnnotator v.2.6.3 (Bouckaert et al., 2019). The mean node heights and 95% highest probability density were visualized with FigTree v.1.4.4.

3. Results

Evolutionary relationships were inferred for 34 ray-finned fish species endemic to the Southwest Pacific islands of Lord Howe, Norfolk, and Rangitāhua. Novel sequences were generated for seven species

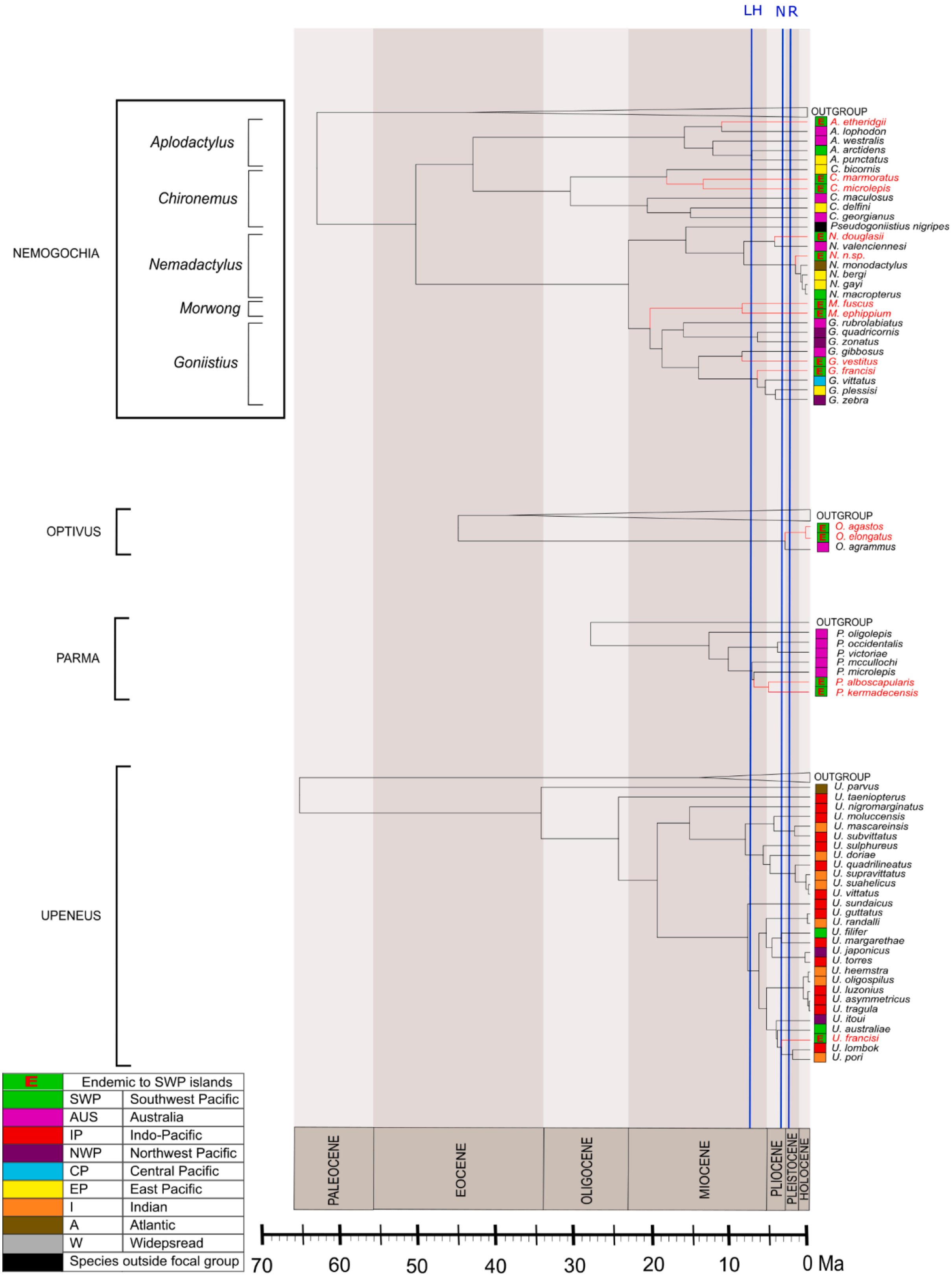

Fig. 2. (*continued*).

previously unrepresented in open-access sequence repositories (*Enneapterygius kermadecensis*, *Eviota kermadecensis*, *Hypoplectrodes huntii*, *Hypoplectrodes* sp. C, *Optivus elongatus*, *Parma kermadecensis*, and *Upeneus francisi*) based on tissues within New Zealand collections, and for three species (*Hypoplectrodes maccullochi*, *Hypoplectrodes nigroruber* and *Lepidoperca brochata*) based on the tissue loans from the Australian Museum, six of which were endemics of the Southwest Pacific islands. We generated additional 12S sequences for *Chromis abyssicola*, *Chromis dispila*, and *Chromis hyspilepis*, and COI for *Flexor incus* and *Parma alboscapularis* to inform phylogenetic analyses. Except for the COI-based

"HL" phylogeny, all phylogenetic inferences were based on concatenated multi-locus sequences between 816 and 5140 bp (Table 2). The evolutionary relationships among species within each endemic's genus are summarized in 13 phylogenies spanning 21 genera (Fig. 2, Fig. 1S-13S). Two phylogenies correspond to the monotypic genera *Capromimus* and *Flexor*, and include all sister taxa with available molecular data at the order and subfamily levels, respectively. Eight genera (*Aplodactylus*, *Arripis*, *Atypichthys*, *Chironemus*, *Goniistius*, *Microcanthus*, *Morwong*, and *Optivus*) are fully resolved with all current valid species included in each phylogeny. Four genera are almost fully resolved with all species per genus sampled except for one, due to missing sequences and unavailable tissues: the Atlantic representatives *Nemadactylus vemae* and *Kathetostoma cubana*, and the Pacific taxa *Girella simplex* and *Scorpis chilensis* are absent within their genus' phylogenies. The remaining genera are missing two or more species. We also report the geographic distribution for each species included in the phylogenetic trees (Fig. 2, Tables 1Sa-13Sa).

The generation of the *Arripis* time-calibrated tree involved five primary calibration points located at deep evolutionary positions (Table 14S) given that the genus is placed at basal nodes within Scombriformes (Matschiner et al., 2017, Sanciangco et al., 2016): the three most closely related orders (Ophidiiformes, Batrachoidiformes, and Gobiiformes; Alfaro et al., 2018, Betancur-R et al., 2017, Hughes et al., 2018, and Matschiner et al., 2017), and two more distant orders (Lampriformes, the only outgroup lacking fossils, and Myctophiformes; Matschiner et al., 2017) were required to correctly root and calibrate the final version of our tree (Fig. 1S). The monophyly of *Arripis* is strongly supported by both ML and Bayesian inferences (BI) (BS = 100; PP = 1; Fig. 1S) based on a 3432 bp concatenated sequence of three mitochondrial markers (Table 2). The endemics *A. trutta* and *A. xylabion* sit in a well-supported clade (BS = 97; PP = 1) that includes *A. truttacea* (Australia). All species are distributed between Australia and New Zealand. With all congeners sampled, the first lineage split is estimated to have occurred 8.87 Ma, and the divergence of both endemics from *A. truttacea* at 0.332 Ma.

The monotypic *Capromimus* lies outside its presumed confamilials within Zeniontidae appearing as the sister taxon of all members of the Oreosomatidae family (Fig. 2S). The phylogenetic position of the endemic *C. abbreviatus* is consistent in all our tree inferences with strong support in the BI (PP > 0.99) and is based on the 5140 bp concatenated sequence of three mitochondrial and five nuclear markers. The estimated divergence time of the endemic's lineage is 14.18 Ma based on our list of sampled taxa.

The "Chromis" phylogeny was inferred using a 3214 bp concatenated sequence of four mitochondrial loci and one nuclear locus (Table 2). Our initial phylogenies included all species named under the *Chromis* genus prior to Tang et al. (2021) (i.e. *Chromis sensu lato*) (results not shown). Based on their study and our data, we observed that all our focal endemics were placed within the *Chromis sensu stricto* clade, thus opting to update our phylogenies by including only strict *Chromis* taxa. Final results (Fig. 3S) show highly-supported sister taxa relationships at terminal positions for four endemics: *C. nitida* (endemic)/*C. fumea* (Indo-Pacific) (BS = 100; PP = 1); *C. dispila*/*C. hypsilepis* (both endemic) (BS = 66; PP > 95); and *C. abyssicola* (endemic)/*C. mamatapara* (East Pacific) (BS = 92; PP > 99). The position of *C. kennensis* was poorly supported (BS = 56; PP > 0.16). All endemics consistently belong to a clade of 19 taxa across our inferences with strong support in the BI tree (PP > 0.81). Based on our genus coverage (59/84 species), the estimated divergence ages are 1.33 Ma for *C. abyssicola*, 2.72 Ma for *C. nitida*, 3.94 Ma for the clade of endemics *C. dispila*/*C. hypsilepis*, and 5.34 for *C. kennensis*.

The phylogeny of *Enneapterygius* is based on two mitochondrial loci (COI and Cytb) analyzed as a supermatrix of 816 bp (Table 2), including 27 of 63 described species to date. The ML phylogeny (Fig. 4S) shows extremely weak support at deep and intermediate nodes. However, the nodes at the terminal clades *Enneapterygius similis*/*E. bahasa*, *Enneapterygius rubicauda*/*E. flavoccipitis*, *Enneapterygius paucifasciatus*/*E. pyramis*, and *Enneapterygius kermadecencis*/*E. williamsi*/*E. nigricauda*/*E. randalli* show support values of over 89. The BI chronogram supports the monophyly of the members of the Tripterygiidae family (PP = 1) which includes the 27 species of *Enneapterygius* and the two taxa of *Helcogramma*. Only four nodes are supported by PP values of over 0.70: *Enneapterygius gruschkai*/*E. abeli*/*E. philippinus* (0.7655); *Enneapterygius flavoccipitis*/*E. rubicauda* (0.7977); *Enneapterygius paucifasciatus*/*E. pyramis* (0.7062); and *Enneapterygius kermadecensis*/*E. williamsi* (0.7103). In all our trees, the endemic *Enneapterygius kermadecensis* is grouped with *E. williamsi* (Central Pacific), and *E. pusillus* is consistently placed closer to the genus *Helcogramma* than to other *Enneapterygius* taxa. The lineage including *Enneapterygius kermadecensis* is estimated to have diverged 8.93 Ma from its closer congeners within our dataset.

Phylogenetic relationships within *Eviota* were inferred based on 62 of its 117 valid species to date. We used the concatenation (2289 bp) of two mitochondrial and one nuclear locus (Table 2). The monophyly of the genus is supported in our Bayesian phylogeny (PP > 0.96; Fig. 5S). Intermediate nodes are characterized by low BS and PP values, both increasing when approaching terminal positions. The endemic *Eviota kermadecensis* is consistently located within its genus as the sister taxon of *E. distigma* and *E. herrei* (both from the Indo-Pacific), albeit in poorly supported phylogenetic positions. Given our taxonomic sampling, the lineage represented by *Eviota kermadecensis* is estimated to have diverged 20.97 Ma.

Evolutionary relationships of *Flexor* were inferred from a supermatrix of 2397 bp encompassing two mitochondrial and two nuclear loci (Table 2). The monotypic genus is placed within its own family in a clade supported by a PP of 0.99 (Fig. 6S). In both the ML and BI topologies, the endemic *F. incus* is grouped with *Lepadichthys trishula*, *Aspasmichthys ciconiae*, and *Aspasma minima* (BS = 32; PP = 0.72), all Northwest Pacific taxa. In both inferences, the monophyly of the Diademichthyinae subfamily *sensu* Conway et al. (2020) is supported by BS = 100 and PP = 1 where the pair *Aspasmogaster tasmaniensis* and *Aspasmogaster ciconiae* are the sister taxa of the subfamily members with the same support values. Based on the taxa sampled in our analysis, the lineage represented by *F. incus* is estimated to have diverged 13.46 Ma.

Our GAMS phylogeny includes the genera *Girella*, *Atypichthys*, *Microcanthus*, and *Scorpis*, and is based on the concatenated sequence (4163 bp) of three mitochondrial and three nuclear loci (Table 2). All current valid species within each genus were sampled except in *Girella* (16/17) and *Scorpis* (4/5). *Girella*, *Microcanthus*, and *Atypichthys* appear monophyletic across all analyses (Fig. 7S) with BS values of 95–99 and PP of 0.97–1. *Scorpis* is a strongly supported clade (BS = 100; PP = 1) when *S. aequipinnis*, *S. lineolata*, and *S. violacea* are considered. *S. georgiana* is consistently positioned outside its presumed congeners. Given our taxonomic coverage, the endemics *G. cyanea*, *G. fimbriata*, *A. latus*, *M. joyceae*, and *S. violacea* are located within their genus members in respective sister taxa relationships with *G. albostriata* (Southeast Pacific), *G. nebulosa* (Southeast Pacific), *A. strigatus* (Australia), *M. strigatus* (Indo-Pacific), and *S. lineolata* (Southwest Pacific), and estimated divergence times of 3.77 Ma, 2.12 Ma, 2.56 Ma, 2.99 Ma, and 3.35 Ma.

The HL phylogeny includes *Hypoplectrodes* and *Lepidoperca* using a 650 bp sequence of COI (Table 2). Both genera are nearly-fully resolved, and their monophyly is supported in our ML and BI analyses (BS = 100 both cases; PP > 0.93 *Hypoplectrodes*; PP = 1 *Lepidoperca*; see Fig. 8S). All species occur in waters of Australia or the Southwest Pacific region, except *H. semicinctum* that inhabits remote oceanic islands of the Southeast Pacific. Species sampled within *Hypoplectrodes* include six of the eight described species and the three undescribed taxa: *H.* sp. A (not valid as *H. coronatus* in Roberts et al., 2015), *H.* sp. B (not valid as *H. dimidius* in Roberts et al., 2015), and *H.* sp. C (not valid as *H. igneus* in Roberts et al., 2015). The endemic *H.* sp. A has a weakly supported placement (BS = 46; PP = 0.4909) as the ancestor of the monophyletic clade (BS = 100; PP = 1) formed by *H. maccullochi*, *H.* sp. B, *H. jamesoni*, *H.* sp. C, *H. cardinalis*, and *H. semicinctum*. Within this clade, the endemic

H. sp. C displays a variable position across our trees. The phylogenetic relationships among *Lepidoperca* congeners are constant in our analyses, with the endemic *L. inornata* and *L. pulchella* (Australia) forming a highly supported clade (BS = 80; PP > 0.98) nested in a larger well-supported clade (BS = 100; PP = 1) that includes *L. tasmanica* and *L. filamenta*. Based on our species coverage, the estimated divergence times are 10.26 Ma for *H.* sp. A, 1.58 Ma for *H.* sp. C, and 1.56 for *L. inornata*.

The monophyly of the genus *Kathetostoma* is well supported in our analyses (BS = 77; PP = 1; see Fig. 9S). There are no differences among the ML and BI topologies which are based on the concatenated sequence (3909 bp) of four mitochondrial and two nuclear markers (Table 2), and seven of the eight current valid species. Phylogenies are characterized by a clade of two taxa (*K. albiguttata* and *K. averruncus*) geographically distributed in the Atlantic and East Pacific regions, and a clade of five species found in Australia and Southwest Pacific (Fig. 2). In the second group, the Australian *K. laeve* is the sister taxon of the remaining four taxa. In our dataset, the endemic lineage representing *K. binigrasella* is estimated to have diverged 6.98 Ma from its congeners.

Our NEMOGOCHIA phylogeny includes the closely related genera *Nemadactylus*, *Morwong*, *Goniistius*, *Chironemus*, and *Aplodactylus* based on a concatenated sequence (2958 bp) of four mitochondrial loci and one nuclear locus. Our initial taxonomic search included all *Cheilodactylus* species within Cheilodactylidae prior to the nomenclature revision of Ludt et al. (2019). Based on their work, our final species list included taxa reassigned to Latridae, specifically the genera *Pseudogoniistius*, *Goniistius*, and *Morwong*. Our focal endemics spanned five genera, of which four were fully resolved and one (*Nemadactylus*) missed one species. Overall, the monophyly of *Aplodactylus*, *Chironemus*, *Morwong*, and *Nemadactylus* is highly supported in our ML and BI trees (BS = 96–100; PP = 0.99–1; see Fig. 10S). The genus *Goniistius* forms a supported monophyletic clade in our ML tree only (BS = 89; PP = 0.48 respectively). The two valid *Morwong* species are both considered endemics and are included in our study with an estimated divergence time of 8.51 Ma. Within *Chironemus*, *C. microlepis* is consistently grouped with *C. marmoratus*, both of which are closely related to *C. bicornis* from the Southeast Pacific (BS = 75; PP > 0.99). Based on all sampled congeners, the endemics are estimated to have diverged 13.61 Ma. Within the fully resolved *Aplodactus*, *A. etheridgii* is constantly coupled with *A. lophodon* from Australia (BS = 56; PP > 0.94) with a divergence time of 11.18 Ma. Within *Goniistius*, *G. vestitus* forms a moderately supported sister taxa relationship with *G. gibbosus* from Australia (BS = 73; PP = 0.69). The position of *G. francisi* varies across our analyses but remains closely related to a clade formed by *G. vestitus*, *G. vittatus*, *G. zebra*, and *G. plessisi*. The estimated divergence times of *G. vestitus* and *G. francisi* are 8.53 Ma and 6.54 Ma respectively. For the near complete phylogeny of *Nemadactylus*, *Pseudogoniistius nigripes* consistently appears as the sister taxon (BS = 58; PP > 0.99). *Nemadactylus douglasii* (estimated age: 4.29 Ma) forms a strongly supported clade (BS = 96; PP = 1) with *N. valenciennesi* (Australia). *Nemadactylus* n. sp. (not valid as *N. rex* in Roberts et al., 2015, and Ludt et al., 2019; estimated age: 1.56 Ma) is the ancestor in a well-supported clade (BS = 97; PP = 1) that includes *N. monodactylus*, *N. bergi*, *N. gayi*, and *N. macropterus*.

The fully-resolved genus *Optivus* is monophyletic in our phylogenies (BS = 99; PP = 1) with its three valid species displaying the same positions in the ML and BI topologies (Fig. 11S), based on a 2514 bp concatenated sequence of two mitochondrial loci and one nuclear locus (Table 2). The endemics *O. elongatus* and *O. agastos* form a monophyletic clade (BS = 99; PP > 0.99) with an estimated divergence time of 0.589 Ma. Both occur in the Southwest Pacific while the basal *O. agrammus* occurs in Australia.

The *Parma* phylogeny includes 7 of the 10 recognized *Parma* species, all of which are geographically restricted to Australia and the Southwest Pacific. The monophyly of the genus is highly supported (BS = 100; PP = 1) with *Mecaenichthys immaculatus* appearing as the sister taxon (Fig. 12S). ML and BI trees were inferred from a sequence of 3859 bp that concatenates three mitochondrial and two nuclear loci (Table 2). Based on our taxonomic sampling, the endemics *P. alboscapularis* and *P. kermadecensis* consistently form a clade across all our inferences, albeit with low node support (BS = 40; PP = 0.33). The estimated age of both endemics is 5.1 Ma.

The relationships within *Upeneus* are based on 29 of 37 described species, and the 821 bp concatenated sequence of two mitochondrial loci (Table 2). Based on the number of sampled taxa, our ML and BI results indicate that the genus is monophyletic (BS = 100; PP = 1) with *U. parvus* (Atlantic) as the sister taxon of its remaining congeners (Fig. 13S). In our dataset, the endemic *U. francisi* (estimated age: 3.65 Ma) is weakly (BS = 44; PP = 0.41) but consistently positioned within its genus as the base of the clade *U. lombok*/*U. pori*, all three forming a weak but consistent clade (BS = 44; PP = 0.49) that diverges from *U. australiae* (Australia).

4. Discussion

Understanding the evolutionary history of endemic species in remote islands has provided insight as to how marine biodiversity is generated and maintained (Bowen et al., 2013), and, in particular, molecular phylogenies have helped illuminate the evolutionary trajectories of endemic marine fishes (Bellwood and Meyer, 2009). Here we present the first comprehensive molecular phylogenies for 21 marine fish genera that encompass 34 taxa endemic to the subtropical islands of Lord Howe, Norfolk, and Rangitāhua in the Southwest Pacific, helping to refine the recent evolutionary histories and patterns for marine ray-finned fishes in the region. Our analyses confirm the species status of all focal endemics by placing them in time-calibrated phylogenies including all relevant sister taxa where DNA sequences could be retrieved or generated. Below we discuss the spatial and temporal diversification patterns of marine fish endemics in the Southwest Pacific based on our results.

4.1. Phylogenetic placement of endemic taxa

We provide the first published placement within a molecular phylogeny for 11 species endemic to the subtropical islands of the Southwest Pacific and smaller areas within (*Arripis xylabion*, *Enneapterygius kermadecensis*, *Eviota kermadecensis*, *Atypichthys latus*, *Lepidoperca inornata*, *Hypoplectrodes* sp. A, *Hypoplectrodes* sp. C, *Nemadactylus* n.sp., *Optivus elongatus*, *Parma kermadecensis*, and *Upeneus francisi*), and the first comprehensive molecular phylogenies for four of the studied genera (*Enneapterygius*, *Lepidoperca*, *Hypoplectrodes*, and *Upeneus*). Our molecular results and phylogenetic analyses for three endemics (*Hypoplectrodes* sp. A, *Hypoplectrodes* sp. C, and *Nemadactylus* n. sp.) will support their formal species description, and for *Arripis xylabion* (Paulin, 1993), we provide new knowledge of its evolutionary relationship to other taxa, previously undescribed based on either molecular or morphological studies.

For some endemic species already described using phenotypic traits, we found that their suggested sister taxon relationships were not completely corroborated by our molecular and phylogenetic inference. For instance, *U. francisi* and 16 other congeners belong to the "*japonicus* group" (Uiblein and Motomura, 2021) based on morphometric, meristic, and color characters (Uiblein and Heemstra, 2010). In our ML and BI phylogenies, however, the 10 "*japonicus* group" species with available sequences (*U. asymmetricus*, *U. australiae*, *U. francisi*, *U. guttatus*, *U. itoui*, *U. japonicus*, *U. lombok*, *U. parvus*, *U. pori*, and *U. torres*) do not cluster, and *U. francisi* has a constant close kinship with only three of its presumed sister taxa (*U. australiae*, *U. lombok*, and *U. pori*). Similarly for *Eviota kermadecensis*, while Hoese and Stewart (2012) related the endemic to *E. abax* and *E. masudai* within the group I defined by Lachner and Karnella (1980), our phylogenies reveal a closer relationship with different congeners (*E. distigma* and *E. herrei*) within group I (Fig. 5S). For *Enneapterygius kermadecensis*, we found complete disagreement between taxon relationships suggested by our results and former meristic studies (Fricke, 1994): the endemic was originally placed within a complex that included *E. hemimelas*, *E. bichrous* (synonym of *E. flavoccipitis*), and *E. niger*, whereas our analyses do not show a phylogenetic closeness between any of these taxa and *E. kermadecensis*. As the original taxonomic relationships for these taxa were described based on phenotypic similarities, these incongruences could be indicative of homoplasy, rather than homology in the meristic traits among the originally proposed sister taxa. However, particularly for the *Enneapterygius* and *Eviota* genera for which taxonomic sampling was not comprehensive, the inclusion of further taxa and more gene regions to support some of the weakly supported nodes may change the topology of the phylogeny, and the divergence date of the endemic lineages.

Twenty-three of the endemic taxa included in our study had previously been included in a molecular phylogeny. For two taxa (*Kathetostoma binigrasella* and *Microcanthus joyceae*), we present their accepted binomial name in a time-calibrated tree for the first time, as they were officially named, or resurrected, following their previous inclusion in a phylogeny: *K. binigrasella* was named the “banded giant stargazer of New Zealand” in the phylogenetic inference of Smith et al. (2006) before its formal description by Gomon and Roberts (2011), and; *M. joyceae* was resurrected by Tea and Gill (2020) after providing molecular and morphological evidence that the Southwest population of *M. strigatus* deserved distinct species status. For these species, and the remaining 21 focal endemics, our phylogenetic inferences are congruent with previously reported topologies. Nonetheless, our analyses have extended previous phylogenetic studies through the inclusion of more taxa for several taxonomic groups, including *Arripis* (previously studied by Sanciangco et al., 2016), *Chromis sensu lato* and *Parma* (most recently studied by McCord et al., 2021, and Tang et al., 2021), *Girella* and *Scorpis* (previously studied by Beldade et al., 2021, and Knudsen et al., 2019), *Nemadactylus* (previously studied by Ludt et al., 2019), *Optivus* (previously studied by Ghedotti et al., 2021), the subfamily Diademichthyinae (including *Flexor incus*, previously studied by Conway et al., 2020), and the order Zeiformes (including *Capromimus abbreviatus*, previously studied by Grande et al., 2018). Furthermore, our time-calibrated phylogenies based on fossils provide a temporal context for the evolutionary history of several taxonomic groups previously without any divergence time estimates (*Goniistius* and *Morwong*, Ludt et al., 2019; *Aplodactylus*, Burridge, 2000), and corroborate divergence time estimates based on other methods (*Chironemus*, Burridge et al., 2006; *Nemadactylus*, Papa et al., 2021). Overall, our standardized methodological approach to phylogenetic analysis has made use of all available data and collections to present the most comprehensive understanding of the evolutionary history of the endemic Southwest Pacific fish fauna to date. Our research has also highlighted persisting gaps in taxonomic sampling, representation, and consequently our knowledge of this regional fish fauna.

4.2. *Resolution and accuracy of evolutionary relationships*

The 13 phylogenies presented in this study recover the molecular data and metadata available in open-access repositories for 389 of our total set of 401 fish species. As we endeavored to include all known valid taxa, our results directly reflect the sequence availability in these groups and portray the repercussions of taxonomic gaps in analysis. In our case, we acknowledge three inherent caveats that affect the resolution and robustness of our inferred phylogenies.

Firstly, the taxonomic coverage is disparate across the 21 genera. Eight are fully sampled (*Aplodactylus*, *Arripis*, *Atypichthys*, *Chironemus*, *Goniistius*, *Microcanthus*, *Morwong*, and *Optivus*), and four are nearly fully sampled missing one species each: *Kathetostoma* and *Nemadactylus* are missing their Atlantic member (likely not affecting our inferred patterns within the Pacific), and *Girella* and *Scorpis* are missing one Pacific member (which should not significantly impact our inferences, but the possibility cannot be ruled out), whereas coverage is only partial for *Hypoplectrodes* (82%), *Lepidoperca* (80%), *Upeneus* (78%), *Chromis* (70%), *Parma* (70%), *Eviota* (53%), and *Enneapterygius* (43%). Having included all species that have sequences in the open-access repositories, our results highlight the need to increase the collection and preservation of specimens for DNA analysis, in particular for the most speciose genera. Missing taxa are a source of divergence time overestimations (Hodge and Bellwood, 2016), so great care should be taken in the interpretation of our incompletely sampled phylogenies.

Secondly, each phylogeny was inferred with a variable number of gene regions (Table 2), ranging between one (“HL”) and eight (“Capromimus”) with a mean of four. Four phylogenies were based on mitochondrial loci only, and the rest in a combination of both mitochondrial and nuclear loci, illustrating the traditional prevalence of mitochondrial markers due to their evolutionary properties (Wilson et al., 1985) and the tendency to examine multiple loci over this single locus (Brito and Edwards, 2009). Within mitochondrial loci, COI occurs in all inferences, followed by Cytb and 16S in seven phylogenies, 12S in six, and the D-loop in one. COI has been the marker of choice for species delimitation, including fishes, owing to its almost-universal resolution power for closely allied taxa (Hebert et al., 2003). For nuclear loci, RAG1 was most commonly used (four phylogenies), followed by MYH6 (three), Tmo-4C4 (two), and seven other regions were used in one phylogeny each. In contrast to phylogenies based on a single locus or the mitochondrial locus, our inferences that include multiple loci are expected to provide better inferences for closely related species given that: a multilocus approach increases species delimitation success (Dupuis et al., 2012); concatenation counteracts the heterogeneity of single-locus inferences (Roe et al., 2010); and the integrative analysis of nuclear and mitochondrial markers leads to more informative and better well-resolved topologies (Rubinoff and Holland, 2005).

Thirdly, we report missing data values per phylogeny between 0% (HL) and 43% (*Enneapterygius*) with an average of 29% across all phylogenies. We generally included loci that covered a minimum of 50% of species within a phylogeny to avoid elevating levels of missing data. However, adding genes can improve poorly supported nodes, even when that gene region is not represented in a high proportion of taxa (Jiang et al., 2014). As a consequence, some of the low support values, particularly at intermediate nodes of our ML topologies, might have been improved with the inclusion of further gene regions albeit with an increase in the percent missing data across our multilocus supermatrices. Nonetheless, our inferred trees largely agree with previously published phylogenies, and divergence time estimates using an uncorrelated log-normal clock in BEAST, as is our case, are relatively insensitive to varying degrees of missing data (Zheng and Wiens, 2015).

4.3. *Geographic affinities of endemic taxa*

Despite contrasting topologies and ranges of included taxa (Fig. 2), the majority of our focal endemics are most closely related to Australian taxa, followed by East Pacific species (Table 3). We also found individual cases where an endemic is closely related to taxa found throughout the broader Indo-Pacific region and the Northwest Pacific. However, the geographic affinity of a quarter of the endemics analyzed here remains unresolved.

4.3.1. *Australia*

Our findings indicate a geographic affinity between 18 of our endemics and Australia mainland. Nine species are distributed in six genera that include sampled taxa only found in the Southwest Pacific and the larger Australia region. For fully sampled genera (*Arripis*, *Atypichthys*, and *Optivus*), the geographic affinity is clearly illustrated by the divergence of the endemics from taxa distributed in Australia, including its western coasts (Fig. 2; Tables 1S,7S,11S). For the partially sampled *Parma*, despite missing two taxa, all described members are restricted to New Zealand and Australia (Tang et al., 2021), and *P. oligolepis* (East Australia) appears as the sister taxon of all congeners. For the partially sampled *Lepidoperca* and *Scorpis*, missing taxa are distributed in other oceanic regions (*L. coatsii*: Atlantic and Indian; *S. chilensis*: East Pacific), adding a small degree of uncertainty to our inferences. However, we still

Table 3

Geographic affinity of the 34 endemics based on our study. Number, percentage, and taxa allocated to each presumed geographic region following those shown in Fig. 1 and Fig. 2.

Geographic affinity	Number of taxa	Taxa
Australia	18/34–53%	_Aplodactylus etheridgii, Arripis trutta, Arripis xylabion, Atypichthys latus, Goniistius francisi, Goniistius vestitus, Hypoplectrodes_ sp.A, _Hypoplectrodes_ sp.C, _Kathetostoma binigrasella, Lepidoperca inornata, Nemadactylus douglasii, Nemadactylus_ n.sp., _Optivus agastos, Optivus elongatus, Parma alboscapularis, Parma kermadecensis, Scorpis violacea, Upeneus francisi_
East Pacific	5/34–15%	_Chironemus marmoratus, Chironemus microlepis, Chromis abyssicola, Girella cyanea, Girella fimbriata_
Indo-Pacific	1/34–3%	_Chromis nitida_
Northwest Pacific	1/34–3%	_Flexor incus_
Not yet resolved*	9/34–26%	_Capromimus abbreviatus, Chromis dispila, Chromis hypsilepis, Chromis kennensis, Enneapterygius kermadecensis, Eviota kermadecensis, Microcanthus joyceae, Morwong ephippium, Morwong fuscus_

*Geographic origin not yet resolved due to incomplete taxonomic sampling.

detect an affinity with Australia as *L. inornata* forms a strongly supported clade (BS = 80; PP > 0.98) with *L. pulchella* from New South Wales and Victoria, and *S. violacea* descends from congeners of Australia and New Zealand mainland.

Seven further endemics are proximal to Australian taxa, but the ranges of sampled congeners extend to other oceanic regions. The phylogenetic placements of *A. etheridgii*, *G. vestitus*, and *N. douglasii* are consistent across our analyses forming well-supported clades with *A. lophodon*, *G. gibbosus*, and *N. valenciennesi* respectively, all three restricted to Australian coasts. The placement of *K. binigrasella* persists also in all our analyses, deriving from *K. nigrofasciatum* found in Australia mainland. A more subtle affinity is detected for three endemics that represent consistent single lineages that descend from Australia and/or Southwest Pacific congeners: *U. francisi* splits from *U. australiae* (Australia and New Caledonia); *N.* n.sp. from the clade *N. valenciennesi* (Australia)/*N. douglasii* (Southwest Pacific); and *H.* sp. A from the clade *H. nigroruber* (Australia)/*H. huntii* (New Zealand). An even more subtle affinity is observed for two endemics placed within clades that descend from Australian taxa, but their phylogenetic position within the clade is variable: the species *G. rubrolabiatus*, found in Australia, appears in our ML phylogeny as the sister taxon of all its congeners, including the endemic *G. francisii*, and is the oldest *Goniistius* lineage in our BI tree; and the clade *H.* sp. C/ *H. jamesoni*/*H. cardinalis*/*H. semicinctum*/*H. macculochi*/*H.* sp. B diverges from *H.* sp. A, for which we have previously inferred an Australian affinity.

The strong geographic affinity between our focal endemics and Australia is likely determined by the regional oceanography, climate similarity, and past geological events. Our results show that in at least 15 of the 18 endemics with an Australian affinity, there is a direct or indirect evolutionary divergence from a taxon restricted to mainland Australia, suggesting eastward dispersal. It has been proposed that this movement is facilitated by the Tasman Front, a branch of the East Australia Current, which likely transports larvae in an eastward flow, connecting East Australia, Lord Howe, Norfolk, and Rangitāhua (Trnski and de Lange, 2015). This route has been previously suggested to influence the biogeography and species distributions of the region's marine fish fauna (Francis, 1993), as well as the region's coastal corals (Brook, 1999), echinoderms (Bronstein et al., 2019), and marine mollusks (Brook, 1998a). Additionally, environmental factors could also explain affinities between Australia and the islands of the Southwest Pacific. For instance, Wicks et al. (2010) reported similarities in the hydrodynamics (e.g. wave energy) of coastal sites in Southeast Australia and Rangitāhua, and suggested it had a role in shaping the similar coral assemblages they observed. Furthermore, sea surface temperature is known to shape the range boundaries of marine fauna (Stuart-Smith et al., 2017) and is often approximated by latitude. In the Southwest Pacific, our focal islands sit in a subtropical belt shared with Australia mainland and no other major landmass regionally, resulting in similar marine fish biota within this same marine climate zone. Lastly, historical events have likely also influenced the current regional biogeography. For 15 of the 18 endemics with an Australian affinity, the closest sister taxon inhabits temperate waters which clearly points to the temperate Australian coasts as a significant origin of biodiversity. Divergence time estimates for these taxa range from 11.17 Ma (*Aplodactylus etheridgii*) to 0.332 Ma (*Arripis trutta* and *Arripis xylabion*) (Table 4), a timing that coincides with the global cooling phase that started in the Middle Miocene (ca. 15 Ma), witnessing a major influence of cooler waters at mid-latitudes in the Southwest Pacific (Nelson and Cooke, 2001) and the receding of tropical and subtropical marine taxa to lower latitudes in both hemispheres (Flower and Kennett, 1994). The expansion of cold marine conditions to lower latitudes would have facilitated the northward input of cold-water temperate Australian fishes into modern-day subtropical latitudes where they would have found suitable habitats, a process that would have resulted in independent lineage diversification events at different time scales. Two other endemics (*Hypoplectrodes* sp. C and *Atypichthys latus*) are closely related to subtropical Australian species, and one (*Upeneus francisi*) to a tropical Australian taxon. All three have recent diversifications (3.65–1.58 Ma) (Table 4), exemplifying that diversification from warmer Australian latitudes has been occurring over more contemporary time scales.

Table 4

Mean node age and confidence interval (Ma) for the 34 endemics investigated in this study. The dashed line separates species based on our threshold of 6.92 Ma for the maximum geological age of our focal islands of the Southwest Pacific.

Species	Ma	Confidence intervals
_Arripis trutta_*	0.3316	0.0276–0.7643
_Arripis xylabion_*	0.3316	0.0276–0.7643
_Optivus agastos_*	0.5894	0.025–1.447
_Optivus elongatus_*	0.5894	0.025–1.447
_Chromis abyssicola_ˆˆ	1.3287	0.4658–2.3425
Nemadactylus n.sp.ˆ	1.5559	0.5458–2.8226
_Lepidoperca inornata_ˆˆ	1.5646	0.3202–3.1078
Hypoplectrodes sp. Cˆˆ	1.5831	0.3218–3.1659
_Girella fimbriata_ˆ	2.1243	0.6036–3.9862
_Atypichthys latus_*	2.5647	0.172–6.2026
_Chromis nitida_ˆˆ	2.7212	1.2932–4.4505
_Microcanthus joyceae_*	2.9897	0.8873–5.6315
_Scorpis violacea_ˆ	3.3464	1.1055–5.9461
_Upeneus francisi_ˆˆ	3.6506	1.1718–6.6257
_Girella cyanea_ˆ	3.7686	1.1038–6.7333
_Chromis dispila_ˆˆ	3.9381	1.8694–6.1324
_Chromis hypsilepis_ˆˆ	3.9381	1.8694–6.1324
_Nemadactylus douglasii_ˆ	4.2883	1.9409–7.0924
_Parma alboscapularis_ˆˆ	5.1040	2.2551–8.1674
_Parma kermadecensis_ˆˆ	5.1040	2.2551–8.1674
_Chromis kennensis_ˆˆ	5.3413	2.6805–8.4153
_Goniistius francisi_*	6.5410	2.2662–12.2732
_Kathetostoma binigrasella_ˆ	6.9795	4.3927–11.4104
_Morwong ephippium_*	8.5109	2.6548–15.9602
_Morwong fuscus_*	8.5109	2.6548–15.9602
_Goniistius vestitus_*	8.5287	3.7424–14.2155
_Enneapterygius kermadecensis_ˆˆ	8.9317	4.3217–14.3042
Hypoplectrodes sp. Aˆˆ	10.2637	3.7005–18.1031
_Aplodactylus etheridgii_*	11.1763	5.6481–18.0152
Flexor incus‡	13.4632	6.2634–21.1712
_Chironemus marmoratus_*	13.6110	5.2852–22.8507
_Chironemus microlepis_*	13.6110	5.2852–22.8507
Capromimus abbreviatus‡	14.1837	5.2581–23.5608
_Eviota kermadecensis_ˆˆ	20.9723	11.3006–31.5149

*Fully sampled genera; ˆNearly-fully sampled genera (one species missing); ˆˆPartially sampled genera (two or more missing taxa); ‡ Monotypic.

4.3.2. East Pacific

Several of our Southwest Pacific endemics had a geographic affinity with remote insular territories of Chile in the Southeast Pacific, namely Juan Fernández and Desventuradas (*C. marmoratus*, *C. microlepis*, *G. cyanea*) and Rapa Nui (*C. abyssicola*, *G. fimbriata*). For instance, *C. abyssicola* and *G. cyanea* are consistently grouped with East Pacific congeners, however, we currently lack additional congenerics to confidently infer the East Pacific as their geographic origin in both cases. Using meristic traits, the pomacentrid was initially allied with species from West Australia, Japan, and Hawaii (Allen and Randall, 1985), but molecular evidence brought by Tang et al. (2021) and our study (Fig. 3S; Table 3S) groups it with its congener from Rapa Nui (*C. mamatapara*). Since around one-third of *Chromis sensu stricto* species are missing in our phylogeny, however, we refrain from further biogeographic inferences. For the girellid, we present the most comprehensively sampled time-calibrated molecular phylogeny to date for its genus (16/17 taxa), finding that *G. cyanea* is grouped with *G. albostriata* (Juan Fernández and Desventuradas), a similar result to the 16S topology of Beldade et al. (2021). All our topologies show that this clade is in a sister-taxa relationship with the clade *G. freminvillii*/*G. laevifrons*, both from the East Pacific, and that both clades descend from Southwest Pacific taxa (*G. tricuspidata*, *G. zebra*, and *G. elevata*). Our current data is insufficient to confidently establish if *G. cyanea* diverged from its closest phylogenetic neighbors from the Southeast, or its Southwest ancestors giving rise to its Southeast congeners. For the remaining endemics with East Pacific affinities, our time-calibrated trees show that they are chronological descendants of Southeast Pacific taxa (Fig. 2): *C. marmoratus* and *C. microlepis* diverge from *C. bicornis* (Juan Fernández and Desventuradas), and the clade *G. fimbriata* (endemic)/*G. nebulosa* (Rapa Nui) from *Graus nigra* (Southeast Pacific). Our confidence in these inferred patterns are supported by: previous phylogenies of *Chironemus* (Burridge et al., 2006) and *Girella* (Knudsen et al., 2019, Knudsen and Clements, 2016); *Chironemus* is fully sampled; and *Girella* is missing only one species.

The presence of congeners at the longitudinal extremes of the Pacific implies trans-oceanic dispersal events, a pattern reported in other marine fishes (Robertson et al., 2004, Rosenblatt and Waples, 1986). In their phylogenetic analysis of the *Cirripictes alboapicalis* complex distributed across the South Pacific, Delrieu-Trottin et al. (2018) describe similar patterns. Their results highlight evolutionary proximity between Rangitāhua and Rapa Nui clades, despite the presence of other clades within the species complex occupying intervening islands of French Polynesia. Such patterns indicate that chance colonization among these island groups, followed by allopatry, may have occurred in the evolutionary history of several fishes. For *C. marmoratus*, *C. microlepis*, and *G. fimbriata*, our analysis suggests that these Southwest Pacific endemic lineages originated in the East Pacific: their divergence postdates the estimated ages of eastern taxa, implying westward dispersal via the warm low latitude South Equatorial Current, and eventual spreading along the East Australian Current. Endemics with an East Pacific affinity and their closest sister taxon diverged between 13.61 Ma (*C. marmoratus* and *C. microlepis*) and 1.33 Ma (*C. abyssicola*), and are mostly subtropical, implying in our case that organisms with similar climatic adaptations to the subtropics have dispersed since the Mid-Miocene across the Pacific as independent events at different times. This inferred pattern is counter to the eastward colonization routes inferred for the same trans-Pacific congeners within *Chironemus* (Burridge et al., 2006), but also *Nemadactylus* (Burridge, 1999), the flat oyster (Ó Foighil et al., 1999), and other marine taxa (Waters, 2008), presumably facilitated by the West Wind Drift that flows in a clockwise rotation around Antarctica and dominates at higher latitudes. Determining the prevalence of westward dispersal of subtropical taxa, the species groups that are advected, and the periods when this dispersal has been the most influential in shaping the biodiversity in the Southwest Pacific requires further examination across broader taxonomic groups.

4.3.3. Indo-Pacific, Northwest Pacific, and unresolved affinities

The only endemic that displays a geographic affinity with the Indo-Pacific belongs to *Chromis* (Table 3), a genus characterized by its reef-associated species mostly found in the tropical latitudes of the Indo-Pacific. We recover the monophyletic clade (BS = 99; PP = 1) between *C. nitida* and *C. fumea* (Indo-Pacific) from previous molecular phylogenies (Frédérich et al., 2013, McCord et al., 2021, Tang et al., 2021), pointing to a geographic closeness between our Southwest Pacific endemic and its tropical congener. Although there is dispersal and ongoing colonization of the subtropical Southwest islands by tropical fishes (Duffy and Ahyong, 2015, Liggins et al., 2020), the speciose genus *Chromis* is poorly sampled in our analysis (Table 1). For this reason, we only provisionally associate a tropical affinity to the *Chromis* endemic, anticipating that the inclusion of further taxa may alter this inference.

The sole endemic of our dataset with an anti-equatorial affinity is *Flexor incus*, showing a proximal association with temperate taxa from the Northwest Pacific (Fig. 2; Table 6S). Previously known as *Aspasmogaster* sp. (Francis and Duffy, 2015, Roberts et al., 2009, Stewart, 2015, Trnski et al., 2015), the endemic was presumed linked to Indo-Pacific taxa based on its formal description (Conway et al., 2018), but was later grouped with *Aspasmichthys ciconiae* (Northwest Pacific) within the Diademichthyinae subfamily using molecular evidence (Conway et al., 2020). Our results recover this trend with the association of *F. incus* with two further Northwest Pacific taxa, suggesting an anti-equatorial origin and the dispersal of organisms from the Northern to the Southern Hemisphere. This trans-equatorial north-to-south pathway has been previously reported in girellids (Beldade et al., 2021), and the opposite south-to-north in microcanthids (Tea et al., 2019). Similar to these studies, we presume that oceanic connectivity between northern and southern regions happened during periods of global climatic change. However, whereas Beldade et al. (2021) and Tea et al. (2019) describe connectivity within the Pleistocene, our estimated age of *F. incus* (ca. 13.5 Ma) suggests that the Mid-Miocene, when there was a major global cooling phase, may have also been a period of trans-equatorial connectivity (i.e. the Middle Miocene Climatic Transition, Crame, 2018).

Lastly, we were unable to address a biogeographic pattern for nine endemics due to unclear evolutionary relationships (Table 3). The monotypic *Capromimus* is placed among Zeiformes species, missing six taxa required to fully sample the Order. The endemic is the sole sister taxon of Oreosomatidae outside its presumed confamilials (Fig. 2S), as already reported by Grande et al. (2018), with no distinct closeness to other species, nor a clear divergence from a specific geographic region (Fig. 2). In the case of *Chromis dispila*, *Chromis hypsilepis*, *Morwong ephippium*, *Morwong fuscus*, and *Microcanthus joyceae*, each forms a monophyletic clade with a congener, but there is still insufficient ancestral evidence to infer their geographic affinities. Both *Chromis* endemics are sister taxa to each other, also reported by McCord et al. (2021) and Tang et al. (2021), but there is no clear ancestor to associate them with (Fig. 2 and Fig. 3S). Similarly, both *Morwong* endemics form a sister taxa relationship (BS = 100; PP > 0.98; Fig. 10S) within the resurrected genus (Ludt et al., 2019), but with unclear ancestral taxa to relate with. *Microcanthus joyceae* is the assumed ancestor within its clade with *M. strigatus* (Tea et al., 2019), but our topologies show *Microcanthus* as the basal genus within our fully sampled Microcanthidae. Since we did not aim for inter-genus relationships, we lack evidence for the ancestry of the family, the genus, and ultimately, the endemic. Lastly, the three endemics *Chromis kennensis*, *Enneapterygius kermadecensis*, and *Eviota kermadecensis* appear as single lineages within poorly sampled genera, impeding the inference of a clear geographic affinity (Table 1). Nonetheless, our time-calibrated phylogeny for the *Enneapterygius* genus shows that *E. kermadecensis* forms a moderately supported clade with *E. williamsi* (PP = 0.71), suggesting an association with a tropical species found in Tonga, Vanuatu, and New Caledonia. The geographic affinity of *Eviota kermadecensis* is more diffuse as its closest taxa vary across our trees, and although it consistently appears as the sister taxon of a clade of Indo-Pacific congeners (*E. distigma* and *E. herrei*), we encourage conservative interpretation of these patterns until further taxonomic sampling can be undertaken to resolve its geographic affinities.

4.4. Temporal diversification of endemic taxa

Based on our analyses, 23 of our 34 (67.6%) endemics are neoendemics estimated to have diverged at, or after, the emergence of the oldest island in the Southwest Pacific (6.92 Ma for Lord Howe in McDougall et al., 1981; Table 4). A divergence age that postdates the emergence of an island is indicative of a diversification event after the formation of the oceanic element, suggesting that the remote volcanic islands of the Southwest Pacific have provided a suitable habitat for the generation of unique evolutionary lineages of marine fishes. Furthermore, for 19 of the 34 endemics analyzed, their clades were not fully sampled and so our divergence time estimates may be overestimated, meaning the first island emergence may be even more important in driving species origination in the region, or that the younger islands are more important than currently suggested by our time estimates. Our conclusions add the Southwest Pacific islands to the list of peripheral islands in the Pacific that export marine biodiversity (Bowen et al., 2013, Cowman et al., 2017), confirming the role of these topographical features in providing the opportunity for origination and hosting unique taxa found nowhere else. Moreover, the Southwest Pacific islands likely continue to promote contemporary diversification in “Oceanic” *Chrysiptera* (*sensu* Tang et al., 2021). The significant morphological divergence of the “Rangitāhua demoiselle” (formerly recognized as *C. rapanui*) from *C. rapanui* of Rapa Nui, and congeners of neighboring regions, suggest that this taxon should be considered a very recent neoendemic of Rangitāhua (Liggins et al., 2021).

For 32.4% of our focal endemics, their divergence times pre-date the emergence of the subtropical Southwest Pacific islands. As suggested, these cases could be due to the overestimation of node ages caused by missing taxa within genera (*Hypoplectrodes*: 18%; *Eviota*: 47%; *Enneapterygius*: 57%), the subfamily Diademichthyinae for *F. incus* (missing six of 23 members), and the Zeiformes for *C. abbreviatus* (missing six of 27 members). However, it is also plausible that our divergence estimates are close to real ages, and that these species could have found suitable habitats in either emerged islands (that are now submerged or on a mainland) or former seamounts in the vicinity of our focal islands, resulting in older endemic taxa than the islands they are endemic to (Heads, 2011). In the case of the fully sampled *Morwong* and *Goniistius*, the closeness of the endemics’ estimated age (ca. 8.5 Ma) to our Lord Howe’s threshold (6.92 Ma) might indicate that their representatives could have settled in nearby oceanic formations before the actual emergence of the islands. For *F. incus* (estimated age:13.5 Ma), it was initially described only in Rangitāhua (Conway et al., 2018) of ca 2.5 Ma age, but is recently reported in Lord Howe and New Caledonia (Fujiwara et al., 2021) of much older ages of 6.92 Ma and 37 Ma respectively (Grandcolas et al., 2008), suggesting that the endemic originated outside our focal islands. A similar scenario might be hypothesized for endemics whose range also includes New Zealand mainland (*A. etheridgii* and *C. abbreviatus*) and Australia mainland (*Hypoplectrodes* sp. A), with these two locations acting as the main center of origination in each case. Finally, if we consider our estimates close to real ages for the fully sampled *Chironemus*, our much older estimated ages imply that *C. marmoratus* and *C. microlepis* diverged well before the emergence of the islands and that their current ranges are relicts of a much wider past range. Since both taxa seem to have diverged from a Southeast congener, our results suggest the splitting of an ancestral chironemid lineage, with a trans-Pacific distribution rather than the westward colonization route solely hypothesized on geographic affinities. A similar example of paleoendemism is reported for the Kermadec Islands giant limpet (*Scutellastra kermadecensis*), in which case fossil evidence suggested that climatic fluctuations altered its former widespread range, reducing it to its current restricted distribution in Rangitāhua (Fleming, 1973).

5. Conclusions

Our molecular phylogenetic analysis provides the most comprehensive understanding of the evolutionary histories of endemic ray-finned fishes in the Southwest Pacific islands to date. From our general work, further research can be undertaken to understand the biogeography and evolution of marine ray-finned fishes at finer spatial and temporal scales. While we confirm the species status of our focal endemics by phylogenetically positioning them within their sister taxa, and have fully sampled eight genera, presenting the first inferences on the geographic origin and diversification timings of endemism in the Southwest Pacific islands, we also highlight some limitations. Multiple genera have a large number of missing taxa, although some require only a single species to achieve complete taxonomic representation. This limitation would benefit from the provisioning of suitable DNA samples in specimen collections, and increased collection effort, in particular in Norfolk and Lord Howe islands. Additionally, we find that four of our phylogenies are solely covered by mitochondrial loci, emphasizing the necessity to include more nuclear markers, and the importance of high-throughput technologies that increasingly enable the parallel sequencing of thousands of homologous gene regions across taxa (Faircloth et al., 2020, Tea et al., 2022). By closing the divide between scientific expeditions, specimen collection, DNA sample preparation, and increasing locus coverage, we will learn more about the formation and maintenance of marine ray-finned fish diversity in the Southwest Pacific.

CRediT authorship contribution statement

André P. Samayoa: Investigation, Data curation, Formal analysis, Validation, Writing – original draft, Writing – review & editing. **Carl D. Struthers:** Resources, Writing – review & editing. **Thomas Trnski:** Resources, Writing – review & editing. **Clive D. Roberts:** Writing – review & editing. **Libby Liggins:** Conceptualization, Data curation, Funding acquisition, Project administration, Supervision, Resources, Validation, Writing – review & editing.

Declaration of Competing Interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Data availability-

Alignments_TreeFiles_MolPhyloSWPEndemism (Original data) (Mendeley Data)

Acknowledgments

We are grateful to the following colleagues for their role in this research: Giacomo Bernardi (University of California Santa Cruz), Kendall Clements (University of Auckland), and co-authors for sharing 16S sequences for *Girella*; Amanda Hay (Australian Museum) for facilitating tissue loans; Diane Pitassy (Smithsonian Institution), Richard Pyle (Bishop Museum), and Luis Rocha (California Academy of Sciences) for providing information on specimens within their institute’s fish collections and for searching potentially usable tissues; José I. Carvajal for laboratory assistance and analysis advice; and J. David Aguirre (Massey University), Erwan Delrieu-Trottin (Université Montpellier II), and Adam N. H. Smith (Massey University) for helpful feedback during discussions of the project. We additionally thank two anonymous reviewers and the editor for their constructive insights. This research was a continuation of the collaboration between Massey University and Te Papa Tongarewa founded by Marti J. Anderson and Clive D. Roberts and funded by a Royal Society of New Zealand Te Apārangi Marsden Grant (15-MAU-132). We wish to thank and acknowledge mana whenua (the traditional owners) of Rangitāhua, the Māori iwi of Ngāti Kuri for their support, and the New Zealand Department of Conservation for sampling permission (Authorisation number: 47976 MAR). Expeditions to Rangitāhua were made possible by the crew of the RV Braveheart, Stoney Creek Shipping Company. These expeditions were supported by Natural History New Zealand and the New Zealand on Air Platinum Fund, the Auckland Museum Institute, Auckland Museum, C&L Gregory Trust, the Pew Charitable Trusts, the School of Natural Sciences Massey University, and the National Institute of Water and Atmospheric Research Ltd Core Funded Coasts & Oceans Programme 2: Biological Resources subcontract for fundamental knowledge of marine fish biodiversity with the Museum of New Zealand Te Papa Tongarewa, Museum of New Zealand Te Papa Tongarewa Collection Development fund: AP's 4788 and 4862. André P. Samayoa is supported by a Massey University Doctoral Scholarship; Libby Liggins is supported by a Rutherford Discovery Fellowship (RDF-20-MAU-001). Ongoing biodiversity research of Rangitāhua is supported by a New Zealand Ministry of Business, Innovation, and Employment Endeavour Research Fund "Te Mana o Rangitāhua" (69495-ENDRP-AWMMU2001).

Appendix A. Supplementary material

Supplementary data to this article can be found online at https://doi.org/10.1016/j.ympev.2022.107584.

References

Alfaro, M.E., Faircloth, B.C., Harrington, R.C., Sorenson, L., Friedman, M., Thacker, C.E., Oliveros, C.H., Černý, D., Near, T.J., 2018. Explosive diversification of marine fishes at the Cretaceous-Palaeogene boundary. Nat. Ecol. Evol. 2, 688–696.

Allen, G.R., Randall, J.E., 1985. A new species of damselfish (Pomacentridae) from eastern Australia and the Norfolk Island Ridge. Rec. Western Aust. Museum 12, 241–245.

Allendorf, F.W., Hohenlohe, P.A., Luikart, G., 2010. Genomics and the future of conservation genetics. Nat. Rev. Genet. 11, 697–709.

Antonelli, A., Hettling, H., Condamine, F.L., Vos, K., Nilsson, R.H., Sanderson, M.J., Sauquet, H., Scharn, R., Silvestro, D., Töpel, M., 2017. Toward a self-updating platform for estimating rates of speciation and migration, ages, and relationships of taxa. Syst. Biol. 66, 152–166.

Ayres, D.L., Darling, A., Zwickl, D.J., Beerli, P., Holder, M.T., Lewis, P.O., Huelsenbeck, J.P., Ronquist, F., Swofford, D.L., Cummings, M.P., 2012. BEAGLE: an application programming interface and high-performance computing library for statistical phylogenetics. Syst. Biol. 61, 170–173.

Beaugrand, G., Edwards, M., Raybaud, V., Goberville, E., Kirby, R.R., 2015. Future vulnerability of marine biodiversity compared with contemporary and past changes. Nat. Clim. Change 5, 695–701.

Beldade, R., Longo, G.C., Clements, K.D., Robertson, D.R., Perez-Matus, A., Itoi, S., Sugita, H., Bernardi, G., 2021. Evolutionary origin of the Atlantic Cabo Verde nibbler (*Girella stuebeli*), a member of a primarily Pacific Ocean family of antitropical herbivorous reef fishes. Mol. Phylogenet. Evol. 156, 107021.

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

Betancur-R, R.R.E., Broughton, E.O., Wiley, K., Carpenter, J.A., López, C., Li, N.I., Holcroft, D., Arcila, M., Cureton II, Sanciangco J.C., 2013. The tree of life and a new classification of bony fishes. PLoS Currents 5 ecurrents. tol.53ba26640df0ccaee75bb165c8c26288.

Betancur-R, R., Ortí, G., Pyron, R.A., 2015. Fossil-based comparative analyses reveal ancient marine ancestry erased by extinction in ray-finned fishes. Ecol. Lett. 18, 441–450.

Betancur-R, R., Wiley, E.O., Arratia, G., Acero, A., Bailly, N., Miya, M., Lecointre, G., Orti, G., 2017. Phylogenetic classification of bony fishes. BMC Evol. Biol. 17, 1–40.

Bouckaert, R., Xie, D., 2017. Standard nucleotide substitution models v1. 0.1. https://doi.org/10.5281/zenodo.995740.

Bouckaert, R., Vaughan, T.G., Barido-Sottani, J., Duchêne, S., Fourment, M., Gavryushkina, A., Heled, J., Jones, G., Kühnert, D., de Maio, N., 2019. BEAST 2.5: An advanced software platform for Bayesian evolutionary analysis. PLoS Comput. Biol. 15, e1006650.

Bowen, B.W., Rocha, L.A., Toonen, R.J., Karl, S.A., 2013. The origins of tropical marine biodiversity. Trends Ecol. Evol. 28, 359–366.

Brito, P.H., Edwards, S.V., 2009. Multilocus phylogeography and phylogenetics using sequence-based markers. Genetica 135, 439–455.

Bronstein, O., Kroh, A., Miskelly, A.D., Smith, S.D., Dworjanyn, S.A., Mos, B., Byrne, M., 2019. Implications of range overlap in the commercially important pan-tropical sea urchin genus Tripneustes (Echinoidea: Toxopneustidae). Mar. Biol. 166, 34.

Brook, F., 1998a. The coastal molluscan fauna of the northern Kermadec Islands, Southwest Pacific Ocean. J. Royal Soc. N.Z. 28, 185–233.

Brook, F., 1998b. Stratigraphy and paleontology of Pleistocene submarine volcanic-sedimentary sequences at the northern Kermadec Islands. J. Roy. Soc. N.Z. 28, 235–257.

Brook, F., 1999. The coastal scleractinian coral fauna of the Kermadec Islands, southwestern Pacific Ocean. J. Royal Soc. N.Z. 29, 435–460.

Burridge, C.P., Roberto, M.C., Dyer, B.S., 2006. Multiple origins of the Juan Fernández kelpfish fauna and evidence for frequent and unidirectional dispersal of cirrhitoid fishes across the South Pacific. Syst. Biol. 55, 566–578.

Burridge, C.P., Smolenski, A.J., 2004. Molecular phylogeny of the Cheilodactylidae and Latridae (Perciformes: Cirrhitoidea) with notes on taxonomy and biogeography. Mol. Phylogenet. Evol. 30, 118–127.

Burridge, C., 1999. Molecular phylogeny of *Nemadactylus* and *Acantholatris* (Perciformes: Cirrhitoidea: Cheilodactylidae), with implications for taxonomy and biogeography. Mol. Phylogenet. Evol. 13, 93–109.

Burridge, C.P., 2000. Molecular phylogeny of the Aplodactylidae (Perciformes: Cirrhitoidea), a group of Southern Hemisphere marine fishes. J. Nat. Hist. 34, 2173–2185.

Capella-Gutiérrez, S., Silla-Martínez, J.M., Gabaldón, T., 2009. trimAl: a tool for automated alignment trimming in large-scale phylogenetic analyses. Bioinformatics 25, 1972–1973.

Conway, K.W., Stewart, A.L., Summers, A.P., 2018. A new genus and species of clingfish from the Rangitāhua Kermadec Islands of New Zealand (Teleostei, Gobiesocidae). ZooKeys 786, 75–104.

Conway, K.W., King, C.D., Summers, A.P., Kim, D., Hastings, P.A., Moore, G.I., Iglésias, S. P., Erdmann, M.V., Baldwin, C.C., Short, G., Fujiwara, K., Trnski, T., Voelker, G., Rüber, L., 2020. Molecular phylogenetics of the clingfishes (Teleostei: Gobiesocidae)—implications for classification. Copeia 108 (4).

R Core Team, 2020. R: A Language and Environment for Statistical Computing. R Foundation for Statistical Computing, Vienna, Austria. https://www.R-project.org/.

Costello, M.J., Coll, M., Danovaro, R., Halpin, P., Ojaveer, H., Miloslavich, P., 2010. A census of marine biodiversity knowledge, resources, and future challenges. PLoS ONE 5, e12110.

Cowman, P.F., Parravicini, V., Kulbicki, M., Floeter, S.R., 2017. The biogeography of tropical reef fishes: endemism and provinciality through time. Biol. Rev. 92, 2112–2130.

Cowman, P.F., 2014. Historical factors that have shaped the evolution of tropical reef fishes: a review of phylogenies, biogeography, and remaining questions. Front. Genet. 5, 394.

Crame, J.A., 2018. Key stages in the evolution of the Antarctic marine fauna. J. Biogeogr. 45, 986–994.

Deck, J., Gaither, M.R., Ewing, R., Bird, C.E., Davies, N., Meyer, C., Riginos, C., Toonen, R.J., Crandall, E.D., 2017. The Genomic Observatories Metadatabase (GeOMe): A new repository for field and sampling event metadata associated with genetic samples. PLoS Biol. 15 (8), e2002925.

Delrieu-Trottin, E., Brosseau-Acquaviva, L., Mona, S., Neglia, V., Giles, E.C., Rapu-Edmunds, C., Saenz-Agudelo, P., 2019. Understanding the origin of the most isolated endemic reef fish fauna of the Indo-Pacific: Coral reef fishes of Rapa Nui. J. Biogeogr. 46 (4), 723–733.

Delrieu-Trottin, E., Williams, J., Bacchet, P., Kulbicki, M., Mourier, J., Galzin, R., de Loma, T.L., Mou-Tham, G., Siu, G., Planes, S., 2015. Shore fishes of the Marquesas Islands, an updated checklist with new records and new percentage of endemic species. Check List 11, 1758.

Delrieu-Trottin, E., Liggins, L., Trnski, T., Williams, J.T., Neglia, V., Rapu-Edmunds, C., Planes, S., Saenz-Agudeulo, P., 2018. Evidence of cryptic species in the blenniid *Cirripectes alboapicalis* species complex, with zoogeographic implications for the South Pacific. Zookeys 810, 127–138.

Duffy, C.A., Ahyong, S.T., 2015. Annotated checklist of the marine flora and fauna of the Kermadec Islands Marine Reserve and northern Kermadec Ridge, New Zealand. Bullet. Auckland Museum 20, 19–124.

Dupuis, J.R., Roe, A.D., Sperling, F.A., 2012. Multi-locus species delimitation in closely related animals and fungi: one marker is not enough. Mol. Ecol. 21, 4422–4436.

Eme, D., Anderson, M.J., Struthers, C.D., Roberts, C.D., Liggins, L., 2019. An integrated pathway for building regional phylogenies for ecological studies. Glob. Ecol. Biogeogr. 28, 1899–1911.

Eme, D., Anderson, M., Myers, E., Roberts, C., Liggins, L., 2020. Phylogenetic measures reveal eco-evolutionary drivers of biodiversity along a depth gradient. Ecography 43, 689–702.

Faircloth, B.C., Alda, F., Hoekzema, K., Burns, M.D., Oliveira, C., Albert, J.S., Melo, B.F., Ochoa, L.E., Roxo, F.F., Chakrabarty, P., 2020. A target enrichment bait set for studying relationships among ostariophysan fishes. Copeia 108, 47–60.

Fleming, C., 1973. Kermadec Island giant limpet occurring fossil in New Zealand, and relict distributions in the tropics. N. Z. J. Mar. Freshwater Res. 7, 159–164.

Flower, B.P., Kennett, J.P., 1994. The middle Miocene climatic transition: East Antarctic ice sheet development, deep ocean circulation and global carbon cycling. Palaeogeogr. Palaeoclimatol. Palaeoecol. 108, 537–555.

Ó Foighil, D., Marshall, B.A., Hilbish, T.J., Pino, M.A., 1999. Trans-Pacific range extension by rafting is inferred for the flat oyster *Ostrea chilensis*. Biol. Bullet. 196, 122–126.

Francis, M.P., Duffy, C.A., 2015. New records, checklist and biogeography of Kermadec Islands' coastal fishes. Bullet. Auckland Museum 20, 481–495.

Francis, M., 2019. Checklist of the coastal fishes of Lord Howe, Norfolk, and Kermadec Islands, southwest Pacific Ocean. https://doi.org/10.6084/m9.figshare.c.4428305.v1.

Francis, M.P., 1993. Checklist of the coastal fishes of Lord Howe, Norfolk, and Kermadec Islands, Southwest Pacific Ocean. Pac. Sci. 47, 136–170.

Frédérich, B., Sorenson, L., Santini, F., Slater, G.J., Alfaro, M.E., 2013. Iterative ecological radiation and convergence during the evolutionary history of damselfishes (Pomacentridae). Am. Nat. 181, 94–113.

Fricke, R., Eschmeyer, W.N., Van Der Laan, R., 2021. Eschmeyer's Catalog of Fishes: Genera, Species, References. http://researcharchive.calacademy.org/research/ichthyology/catalog/fishcatmain.asp (accessed 8 November 2021).

Fricke, R., 1994. Tripterygiid Fishes of Australia, New Zealand and the Southwest Pacific Ocean: With Descriptions of 2 New Genera and 16 New Species (Teleostei). Lubrecht & Cramer Limited.

Friedlander, A.M., Ballesteros, E., Caselle, J.E., Gaymer, C.F., Palma, A.T., Petit, I., Varas, E., Muñoz Wilson, A., Sala, E., Fulton, C.J., 2016. Marine biodiversity in Juan Fernández and Desventuradas Islands, Chile: global endemism hotspots. PLoS ONE 11 (1), e0145059.

Froese, R., Pauly, D., 2021. FishBase. World Wide Web electronic publication. http://www.fishbase.org, version (06/2021) (accessed 8 November 2021).

Fujiwara, K., Conway, K.W., Motomura, H., 2021. First record of the Kermadec Clingfish, *Flexor incus* Conway, Stewart & Summers, 2018 (Gobiesocidae), from New Caledonia and Australia. Check List 17, 769.

Gaboriau, T., Leprieur, F., Mouillot, D., Hubert, N., 2018. Influence of the geography of speciation on current patterns of coral reef fish biodiversity across the Indo-Pacific. Ecography 41, 1295–1306.

Ghedotti, M.J., Dekay, H.M., Maile, A.J., Smith, W.L., Davis, M.P., 2021. Anatomy and evolution of bioluminescent organs in the slimeheads (Teleostei: Trachichthyidae). J. Morphol. 282, 820–832.

Gomon, M.F., Roberts, C.D., 2011. A second New Zealand species of the stargazer genus *Kathetostoma* (Trachinoidei: Uranoscopidae). Zootaxa 2776, 1–12.

Grandcolas, P., Murienne, J.M., Robillard, T., Desutter-Grandcolas, L., Jourdan, H., Guilbert, E., Deharveng, L., 2008. New Caledonia: a very old Darwinian island? Philos. Trans. Roy. Soc. B: Biol. Sci. 363, 3309–3317.

Grande, T.C., Borden, W.C., Wilson, M.V., Scarpitta, L., 2018. Phylogenetic relationships among fishes in the order Zeiformes based on molecular and morphological data. Copeia 106, 20–48.

Heads, M., 2011. Old taxa on young islands: a critique of the use of island age to date island-endemic clades and calibrate phylogenies. Syst. Biol. 60, 204–218.

Hebert, P.D., Ratnasingham, S., de Waard, J.R., 2003. Barcoding animal life: cytochrome c oxidase subunit 1 divergences among closely related species. Proc. R. Soc. Lond. B Biol. Sci. 270, S96–S99.

Hoban, S., Bruford, M., Funk, W.C., Galbusera, P., Griffith, M.P., Grueber, C.E., Heuertz, M., Hunter, M.E., Hvilsom, C., Kalamujic, S.B., 2021. Global commitments to conserving and monitoring genetic diversity are now necessary and feasible. Bioscience 71, 964–976.

Hoban, M.L., Williams, J.T., 2020. *Cirripectes matatakaro*, a new species of combtooth blenny from the Central Pacific, illuminates the origins of the Hawaiian fish fauna. PeerJ 8, e8852.

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

Hodge, J.R., van Herwerden, L., Bellwood, D.R., 2014. Temporal evolution of coral reef fishes: global patterns and disparity in isolated locations. J. Biogeogr. 41, 2115–2127.

Hoeksema, B.W., 2007. Delineation of the Indo-Malayan Centre of Maximum Marine Biodiversity: The Coral Triangle. In: Renema, W. (Ed.), Biogeography, Time, and Place: Distributions, Barriers, and Islands. Topics In Geobiology, vol. 29. Springer, Dordrecht.

Hoese, D.F., Stewart, A.L., 2012. A new species of the gobiid genus *Eviota* (Teleostei: Gobioidei) from the Kermadec Islands, New Zealand. Zootaxa 3418, 61–67.

Hortal, J., de Bello, F., Diniz-Filho, J.A.F., Lewinsohn, T.M., Lobo, J.M., Ladle, R.J., 2015. Seven shortfalls that beset large-scale knowledge of biodiversity. Annu. Rev. Ecol. Evol. Syst. 46, 523–549.

Hughes, L.C., Ortí, G., Huang, Y., Sun, Y., Baldwin, C.C., Thompson, A.W., Arcila, D., Betancur-R, R., Li, C., Becker, L., 2018. Comprehensive phylogeny of ray-finned fishes (Actinopterygii) based on transcriptomic and genomic data. Proc. Natl. Acad. Sci. 115, 6249–6254.

Huson, D.H., Scornavacca, C., 2012. Dendroscope 3: an interactive tool for rooted phylogenetic trees and networks. Syst. Biol. 61, 1061–1067.

Ivanova, N.V., Zemlak, T.S., Hanner, R.H., Hebert, P.D., 2007. Universal primer cocktails for fish DNA barcoding. Mol. Ecol. Notes 7, 544–548.

Iwatsubo, H., Motomura, H., 2018. *Chromis katoi*, a new species of damselfish from the Izu Islands, Japan, with a key to species in the *Chromis notata* species complex (Perciforms: Pomacentridae). Int. J. Ichthyol. 24, 27–34.

Jiang, W., Chen, S.-Y., Wang, H., Li, D.-Z., Wiens, J.J., 2014. Should genes with missing data be excluded from phylogenetic analyses? Mol. Phylogenet. Evol. 80, 308–318.

Jones, J., McDougall, I., 1973. Geological history of Norfolk and Philip islands, southwest Pacific Ocean. J. Geol. Soc. Aust. 20, 239–254.

Katoh, K., Misawa, K., Kuma, K.I., Miyata, T., 2002. MAFFT: a novel method for rapid multiple sequence alignment based on fast Fourier transform. Nucleic Acids Res. 30, 3059–3066.

Knudsen, S.W., Choat, J.H., Clements, K.D., 2019. The herbivorous fish family Kyphosidae (Teleostei: Perciformes) represents a recent radiation from higher latitudes. J. Biogeogr. 46, 2067–2080.

Knudsen, S.W., Clements, K.D., 2016. World-wide species distributions in the family Kyphosidae (Teleostei: Perciformes). Mol. Phylogenet. Evol. 101, 252–266.

Lachner, E.A., Karnella, S.J., 1980. Fishes of the Indo-Pacific genus *Eviota* with descriptions of eight new species (Teleostei, Gobiidae). In: Smithsonian Contributions to Zoology, Smithsonian Insitution Press, Washington.

Lanfear, R., Frandsen, P.B., Wright, A.M., Senfeld, T., Calcott, B., 2017. PartitionFinder 2: new methods for selecting partitioned models of evolution for molecular and morphological phylogenetic analyses. Mol. Biol. Evol. 34, 772–773.

le Feuvre, M.C., Dempster, T., Shelley, J.J., Davis, A.M., Swearer, S.E., 2021. Range restriction leads to narrower ecological niches and greater extinction risk in Australian freshwater fish. Biodivers. Conserv. 30, 2955–2976.

Li, X., Shen, X., Chen, X., Xiang, D., Murphy, R.W., Shen, Y., 2018. Detection of potential problematic Cytb gene sequences of fishes in GenBank. Front. Genet. 9, 30.

Liggins, L., Sweatman, J.A., Trnski, T., Duffy, C.A., Eddy, T.D., Aguirre, J.D., 2020. Natural history footage provides new reef fish biodiversity information for a pristine but rarely visited archipelago. Sci. Rep. 10, 1–7.

Liggins, L., Kilduff, L., Trnski, T., Delrieu-Trottin, E., Carvajal, J.I., Arranz, V., Planes, S., Saenz-Agudelo, P., Aguirre, J.D., 2021. Morphological and genetic divergence supports peripheral endemism and a recent evolutionary history of *Chrysiptera* demoiselles in the subtropical South Pacific. Coral Reefs 1–16.

López, J.A., Chen, W.-J., Ortí, G., 2004. Esociform phylogeny. Copeia 2004, 449–464.

Lotze, H.K., Lenihan, H.S., Bourque, B.J., Bradbury, R.H., Cooke, R.G., Kay, M.C., Kidwell, S.M., Kirby, M.X., Peterson, C.H., Jackson, J.B., 2006. Depletion, degradation, and recovery potential of estuaries and coastal seas. Science 312, 1806–1809.

Ludt, W.B., Burridge, C.P., Chakrabarty, P., 2019. A taxonomic revision of Cheilodactylidae and Latridae (Centrarchiformes: Cirrhitoidei) using morphological and genomic characters. Zootaxa 4585, 121–141.

Matschiner, M., Musilová, Z., Barth, J.M., Starostová, Z., Salzburger, W., Steel, M., Bouckaert, R., 2017. Bayesian phylogenetic estimation of clade ages supports trans-Atlantic dispersal of cichlid fishes. Syst. Biol. 66, 3–22.

McCord, C.L., Westneat, M.W., 2016. Phylogenetic relationships and the evolution of BMP4 in triggerfishes and filefishes (Balistoidea). Mol. Phylogenet. Evol. 94, 397–409.

McCord, C.L., Nash, C.M., Cooper, W.J., Westneat, M.W., 2021. Phylogeny of the damselfishes (Pomacentridae) and patterns of asymmetrical diversification in body size and feeding ecology. PLoS ONE 16, e0258889.

McDougall, I., Embleton, B., Stone, D., 1981. Origin and evolution of Lord Howe Island, southwest Pacific Ocean. J. Geol. Soc. Aust. 28, 155–176.

Middleton, I., Aguirre, J.D., Trnski, T., Francis, M., Duffy, C., Liggins, L., Leroy, B., 2021. Introduced alien, range extension or just visiting? Combining citizen science observations and expert knowledge to classify range dynamics of marine fishes. Divers. Distrib. 27 (7), 1278–1293.

Middleton, I., Francis, M., Aguirre, J.D., Struthers, C., Trnski, T., Duffy, C., Anderson, J., Carney, M., Cox, N., Hemstalk, S., Herbert, B., Johnston, C., Middleton, C., Oldfield, T., Robinson, R., Sheath, C., Skipworth, I., Stancich, J., Whittington, H., Liggins, L., In review. Spatio-temporal variability in occurrences of tropical, subtropical and rare marine fishes in Aotearoa New Zealand. J. Biogeogr.

Mirande, J.M., 2017. Combined phylogeny of ray-finned fishes (Actinopterygii) and the use of morphological characters in large-scale analyses. Cladistics 33, 333–350.

Moritz, C., 2002. Strategies to protect biological diversity and the evolutionary processes that sustain it. Syst. Biol. 51, 238–254.

Near, T.J., Eytan, R.I., Dornburg, A., Kuhn, K.L., Moore, J.A., Davis, M.P., Wainwright, P.C., Friedman, M., Smith, W.L., 2012. Resolution of ray-finned fish phylogeny and timing of diversification. Proc. Natl. Acad. Sci. 109, 13698–13703.

Near, T.J., Dornburg, A., Eytan, R.I., Keck, B.P., Smith, W.L., Kuhn, K.L., Moore, J.A., Price, S.A., Burbrink, F.T., Friedman, M., 2013. Phylogeny and tempo of diversification in the superradiation of spiny-rayed fishes. Proc. Natl. Acad. Sci. 110, 12738–12743.

Nelson, C.S., Cooke, P.J., 2001. History of oceanic front development in the New Zealand sector of the Southern Ocean during the Cenozoic—a synthesis. N. Z. J. Geol. Geophys. 44, 535–553.

Norman, M., 2003. Biodiversity hotspots revisited. Bioscience 53, 916–917.

Nunez, S., Arets, E., Alkemade, R., Verwer, C., Leemans, R., 2019. Assesing the impacts of climate change on biodiversity: is below 2 ^{o}C enough? Clim. Change 154, 351–365.

Papa, Y., Halliwell, A.G., Morrison, M.A., Wellenreuther, M., Ritchie, P.A., 2021. Phylogeographic structure and historical demography of tarakihi (*Nemadactylus macropterus*) and king tarakihi (*Nemadactylus* n. sp.) in New Zealand. N. Z. J. Mar. Freshwater Res. 1–25.

Paulin, C., 1993. Review of the Australian fish family Arripididae (Percomorpha), with the description of a new species. Mar. Freshw. Res. 44, 459–471.

Powell, C.L.E., Waskin, S., Battistuzzi, F.U., 2020. Quantifying the error of secondary vs. distant primary calibrations in a simulated environment. Front. Genet. 11, 252.

Rabosky, D.L., Chang, J., Title, P.O., Cowman, P.F., Sallan, L., Friedman, M., Kaschner, K., Garilao, C., Near, T.J., Coll, M., 2018. An inverse latitudinal gradient in speciation rate for marine fishes. Nature 559, 392–395.

Rabosky, D.L., 2020. Speciation rate and the diversity of fishes in freshwaters and the oceans. J. Biogeogr. 47, 1207–1217.

Rambaut, A., Drummond, A.J., Xie, D., Baele, G., Suchard, M.A., 2018. Posterior summarization in Bayesian phylogenetics using Tracer 1.7. Syst. Biol. 67, 901.

Randall, J.E., Cea, A., 2011. Shore fishes of Easter Island. University of Hawaii Press, Honolulu.

Randall, J., 2007. Reef fishes of Hawaii. University of Hawaii Press, Honolulu.

Riginos, C., Crandall, E.D., Liggins, L., Gaither, M.R., Ewing, R.B., Meyer, C., Andrews, K.R., Euclide, P.T., Titus, B.M., Therkildsen, N.O., Salces-Castellano, A., Stewart, L.C., Toonen, R.J., Deck, J., 2020. Building a global genomics observatory: Using GEOME (the Genomic Observatories Metadatabase) to expedite and improve deposition and retrieval of genetic data and metadata for biodiversity research. Mol. Ecol. Resour. 20 (6), 1458–1469.

Roberts, C.D., Paulin, C.D., Stewart, A.L., Mcphee, R P., Mcdowall, R.M., 2009. Checklist of New Zealand Chordata: Living lancelets, jawless fishes, cartilaginous, and bony fishes. In: Gordon, D.P. (Ed.), New Zealand Inventory of Biodiversity, vol. 1, Kingdom Animalia. Canterbury University Press, Christchurch, pp. 527–536, 566.

Roberts, C.D., Stewart, A.L., Struthers, C.D., Barker, J.J., Kortet, S., 2020. Checklist of the Fishes of New Zealand. Online version 1.2. Museum of New Zealand Te Papa Tongarewa, Wellington, pp. 241. https://collections.tepapa.govt.nz/document/10564.

Roberts, C.D., Stewart, A.L., Struthers, C.D., 2015. The Fishes of New Zealand. Te Papa Press, Wellington.

Robertson, D.R., Grove, J.S., McCosker, J.E., 2004. Tropical transpacific shore fishes. Pac. Sci. 58, 507–565.

Roe, A.D., Rice, A.V., Bromilow, S.E., Cooke, J.E., Sperling, F.A., 2010. Multilocus species identification and fungal DNA barcoding: insights from blue stain fungal symbionts of the mountain pine beetle. Mol. Ecol. Resour. 10, 946–959.

Rosenblatt, R.H., Waples, R.S., 1986. A genetic comparison of allopatric populations of shore fish species from the eastern and central Pacific Ocean: dispersal or vicariance? Copeia 1986 (2), 275.

Rstudio Team, 2019. RStudio: Integrated Development for R. RStudio Inc, Boston, MA. http://www.rstudio.com/.

Rubinoff, D., Holland, B.S., 2005. Between two extremes: mitochondrial DNA is neither the panacea nor the nemesis of phylogenetic and taxonomic inference. Syst. Biol. 54, 952–961.

Sanciangco, M.D., Carpenter, K.E., Betancur-R, R., 2016. Phylogenetic placement of enigmatic percomorph families (Teleostei: Percomorphaceae). Mol. Phylogenet. Evol. 94, 565–576.

Sauquet, H., 2013. A practical guide to molecular dating. C.R. Palevol 12, 355–367.

Smith, P.J., McPhee, R.P., Roberts, C.D., 2006. DNA and meristic evidence for two species of giant stargazer (Teleostei: Uranoscopidae: *Kathetostoma*) in New Zealand waters. N. Z. J. Mar. Freshwater Res. 40, 379–387.

Stamatakis, A., 2014. RAxML version 8: a tool for phylogenetic analysis and post-analysis of large phylogenies. Bioinformatics 30, 1312–1313.

Stewart, A.L., 2015. Family Gobiesocidae. Clingfishes. In: Roberts, C.D., Stewart, A.L., Struthers, C.D. (Eds.), The Fishes of New Zealand, vol. 4. Te Papa Press, Wellington, 2008, pp. 1539–1555.

Stuart-Smith, R.D., Edgar, G.J., Bates, A.E., 2017. Thermal limits to the geographic distributions of shallow-water marine species. Nat. Ecol. Evol. 1, 1846–1852.

Tang, K.L., Stiassny, M.L., Mayden, R.L., Desalle, R., 2021. Systematics of Damselfishes. Ichthyol. Herpetol. 109, 258–318.

Tea, Y.-K., Gill, A.C., 2020. Systematic reappraisal of the anti-equatorial fish genus *Microcanthus* Swainson (Teleostei: Microcanthidae), with redescription and resurrection of *Microcanthus joyceae* Whitley. Zootaxa 4802, 41–60.

Tea, Y.-K., Van Der Wal, C., Ludt, W.B., Gill, A.C., Lo, N., Ho, S.Y.W., 2019. Boomeranging around Australia: Historical biogeography and population genomics of the anti-equatorial fish *Microcanthus strigatus* (Teleostei: Microcanthidae). Mol. Ecol. 28 (16), 3771–3785.

Tea, Y.-K., Xu, X., Dibattista, J.D., Lo, N., Cowman, P.F., Ho, S.Y., 2022. Phylogenomic analysis of concatenated ultraconserved elements reveals the recent evolutionary radiation of the fairy wrasses (Teleostei: Labridae: *Cirrhilabrus*). Syst. Biol. 71, 1–12.

Trnski, T., de Lange, P.J., 2015. Introduction to the Kermadec Biodiscovery Expedition 2011. Bullet. Auckland Museum 20, 1–18.

Trnski, T., Duffy, C.A., Francis, M.P., McGrouther, M.A., Stewart, A.L., Struthers, C.D., Zintzen, V., 2015. Recent collections of fishes at the Kermadec Islands and new records for the region. Bullet. Auckland Museum 20, 463–480.

Uiblein, F., Heemstra, P.C., 2010. A taxonomic review of the Western Indian Ocean goathfishes of the genus *Upeneus* (Family Mullidae), with descriptions of four new species. Smithiana Bullet. 11, 35–71.

Uiblein, F., Motomura, H., 2021. Three new goatfishes of the genus *Upeneus* from the Eastern Indian Ocean and Western Pacific, with an updated taxonomic account for *U. itoui* (Mullidae: *japonicus*-species group). Zootaxa 4938, 298–324.

van der Meer, M.H., Berumen, M.L., Hobbs, J.-P.- A., van Herwerden, L., 2015. Population connectivity and the effectiveness of marine protected areas to protect vulnerable, exploited and endemic coral reef fishes at an endemic hotspot. Coral Reefs 34 (2), 393–402.

Ward, R.D., Zemlak, T.S., Innes, B.H., Last, P.R., Hebert, P.D., 2005. DNA barcoding Australia's fish species. Philos. Trans. Roy. Soc. B: Biol. Sci. 360, 1847–1857.

Waters, J.M., 2008. Driven by the West Wind Drift? A synthesis of southern temperate marine biogeography, with new directions for dispersalism. J. Biogeogr. 35, 417–427.

Wicks, L.C., Gardner, J.P., Davy, S.K., 2010. Spatial patterns and regional affinities of coral communities at the Kermadec Islands Marine Reserve, New Zealand—a marginal high-latitude site. Mar. Ecol. Prog. Ser. 400, 101–113.

Wilson, A.C., Cann, R.L., Carr, S.M., George, M., Gyllensten, U.B., Helm-Bychowski, K.M., Higuchi, R.G., Palumbi, S.R., Prager, E.M., Sage, R.D., 1985. Mitochondrial DNA and two perspectives on evolutionary genetics. Biol. J. Linn. Soc. 26, 375–400.

Worms Editorial Board, 2021. World Register of Marine Species. http://www.marinespecies.org at VLIZ. https://doi.org/10.14284/170 (accessed 8 November 2021).

Zheng, Y., Wiens, J.J., 2015. Do missing data influence the accuracy of divergence-time estimation with BEAST? Mol. Phylogenet. Evol. 85, 41–49.