



**HAL**  
open science

## The origins of marine fishes endemic to subtropical islands of the Southwest Pacific

André Samayoa, J. David Aguirre, Erwan Delrieu-Trottin, Libby Liggins

### ► To cite this version:

André Samayoa, J. David Aguirre, Erwan Delrieu-Trottin, Libby Liggins. The origins of marine fishes endemic to subtropical islands of the Southwest Pacific. *Journal of Biogeography*, 2023, 50 (8), pp.1388-1401. 10.1111/jbi.14579 . hal-04902906

**HAL Id: hal-04902906**

**<https://hal.science/hal-04902906v1>**

Submitted on 21 Jan 2025

**HAL** is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.



Distributed under a Creative Commons Attribution - NonCommercial 4.0 International License



# The origins of marine fishes endemic to subtropical islands of the Southwest Pacific

André P. Samayoa<sup>1</sup> | J. David Aguirre<sup>1</sup> | Erwan Delrieu-Trottin<sup>2</sup> | Libby Liggins<sup>1,3</sup>

<sup>1</sup>School of Natural Sciences, Massey University, Auckland, New Zealand

<sup>2</sup>CEFE, Univ Montpellier, CNRS, EPHE-PSL University, IRD, Montpellier, France

<sup>3</sup>Natural Sciences, Auckland Museum Tāmaki Paenga Hira, Auckland, New Zealand

## Correspondence

André P. Samayoa, School of Natural Sciences, Massey University, Auckland 0745, New Zealand.

Email: [a.p.samayoa@massey.ac.nz](mailto:a.p.samayoa@massey.ac.nz)

## Funding information

Royal Society of New Zealand Te Apārangi Marsden Grant

**Handling Editor:** Giacomo Bernardi

## Abstract

**Aim:** Lineages colonizing subtropical oceanic islands often have to overcome geographical isolation and novel climate stressors to found new populations. Historical and ecological factors influence the success of colonization and subsequent diversification, leaving a signal in the genetic constitution of the diverged, range-restricted taxa. Here, we examined the historical biogeography of endemic marine fishes to quantify the role of geographical proximity and climate differences in determining colonization, and the underlying mechanisms of speciation.

**Location:** Subtropical islands of the Southwest Pacific.

**Taxa:** Thirty endemic marine ray-finned fishes from 17 genera.

**Methods:** Using parametric biogeographical history models, we estimated the ancestral geographical ranges for 144 species based on time-calibrated phylogenies that included endemic species and their closest sister taxa, linking terminal nodes with geographical distribution classified into 14 biogeographical areas.

**Results:** Ancestral range estimations revealed most species originated in Australia (66%), while only 10% and 7% originated in northern tropical Pacific locations and the East Pacific respectively, with 17% of species-range estimations being inconclusive. Vicariant events alone were identified as the most likely process shaping range evolution in 57% of the 14 best-fitting models, dispersal alone was favoured for 14% of species and both processes had a role for the remaining 29% of species. Across all phylogenies, likelihood-ratio tests confirmed that geographical distance and climatic differences constrained dispersal in 73% and 33% of species, respectively.

**Main Conclusions:** Marine fishes endemic to subtropical islands of the Southwest Pacific originated by vicariant and jump-dispersal events mainly from ancestral populations in mainland Australia. Geographical distance and climatic differences are significant taxon-specific factors influencing the dispersal of marine fishes in the region.

## KEYWORDS

biogeography, climate zone, endemism, geographical distance, jump-dispersal, marine fishes, oceanic islands, Pacific Ocean, vicariance

This is an open access article under the terms of the [Creative Commons Attribution](https://creativecommons.org/licenses/by/4.0/) License, which permits use, distribution and reproduction in any medium, provided the original work is properly cited.

© 2023 The Authors. *Journal of Biogeography* published by John Wiley & Sons Ltd.

## 1 | INTRODUCTION

The relative importance of vicariance and dispersal have been at the centre of vigorous debates regarding the major forces governing geographical patterns of marine biodiversity and endemism (Wiens & Donoghue, 2004). In the sea, vicariance arises when populations become geographically isolated due to not only 'hard' barriers like landmasses, but also from 'soft' barriers such as long oceanic distances between isolated areas (Cowman & Bellwood, 2013; Luiz et al., 2012). Whereas vicariance focuses on the severance of connectivity across a species range, the movement of individuals from a source population to new habitats (jump-dispersal) can also lead to species divergence. In the last two decades, empirical evidence shows that jump-dispersal is a more prevalent process in marine biogeography than considered during the late 20th century (Sanmartín, 2012; Waters, 2008; Waters & Roy, 2004). Jump-dispersal is followed by limited gene flow between the original and new populations leading to founder-event speciation (Matzke, 2014; Templeton, 2008). Consequently, any factor affecting dispersal is likely to alter the probabilities of colonization and eventual diversification of range-restricted biodiversity. Geographical isolation (Dalongeville et al., 2018; Jones et al., 2009) and environmental conditions (Stuart-Smith et al., 2017) can influence dispersal and hence the geographical distribution of marine organisms (Lester et al., 2007). Determining the prevalence of jump-dispersal and factors affecting dispersal probabilities in the sea is essential to better understand marine biogeographical patterns.

An approach that has significantly improved our understanding of processes shaping biogeographical patterns is the combination of geographical information and molecular phylogenies. For instance, parametric methods can explicitly test models representing distinct biogeographical hypotheses (Lamm & Redelings, 2009). Each model includes parameters that correspond to anagenetic (dispersal, extinction) and cladogenetic (vicariance, founder event) processes along a time-calibrated tree (Sanmartín, 2012), as well as the effect of geographical and environmental distances (Van Dam & Matzke, 2016). In this framework, parameters are estimated using maximum-likelihood (Ree et al., 2005) resulting in the inference of ranges at each ancestral node (Ree & Smith, 2008). A number of studies have used this methodology to reveal the origin of evolutionary lineages in marine fishes: the ancestral Tethys Sea for syngnatharians (Santaquiteria et al., 2021), the Central Pacific for pomacanthids (Baraf et al., 2019), and the Indo-Pacific for the clade Gobiidae-Gobionellidae (Thacker, 2015) and for herbivorous groups in the Atlantic (Siqueira et al., 2019). Further evidence shows that both vicariance and jump-dispersal are relevant in the biogeography of marine fishes (Delrieu-Trottin et al., 2019; Piñeros et al., 2019). However, few of these studies have integrated distance parameters in their range evolution models to explicitly test the effect of geographical distance and climatic differences.

The Indo-Australian Archipelago (IAA) harbours the greatest levels of marine fish biodiversity in the Indo-Pacific (Hoeksema, 2007), whereas peripheral isolated islands have low species richness but

high endemism (Allen, 2008; Hughes et al., 2002). Phylogenetic evidence suggests that high species richness in the centre of the IAA is the result of multiple independent events (Cowman et al., 2017), in keeping with our understanding that the biogeographical history of this fauna has been long and complex (Miller et al., 2018). However, relatively less is known about the origins of endemism at the periphery of the Indo-Pacific (Bowen et al., 2013). In the Central Pacific, biogeographical analyses suggest a complex scenario where endemic lineages of Hawaii derive from allopatric speciation mainly driven by geologic events that influenced ocean circulation, with endemism generated in two temporal waves (0–3 and 8–12 Ma) from various geographical origins (Hodge et al., 2014). In the Southeast Pacific, seven of nine parametric models for reef fishes endemic to Easter Island (Rapa Nui), point to jump-dispersal as the main driver of endemism, in some cases via stepping-stone colonization routes within the Pacific (Delrieu-Trottin et al., 2019). Nevertheless, Hawaii and Rapa Nui are the most peripheral of the island groups in the Indo-Pacific, and the likely drivers of endemism in other peripheral islands remain unexplored.

The islands of Lord Howe, Norfolk and Rangitāhua (the Kermadec Islands) lie in subtropical (29°S–31°S) waters of the Southwest Pacific (Francis, 1993), and have been relatively understudied by evolutionary biologists and biogeographers (Liggins et al., 2022). Of volcanic origin, the islands are geologically young with estimated ages of 6.9 Ma for Lord Howe (McDougall et al., 1981), 3.05 Ma for Norfolk (Jones & McDougall, 1973) and 2.58 Ma for Rangitāhua (Brook, 1998). The overall endemism rate for marine fishes across all three island groups is 4.6% (Francis, 1993), with local communities including a mixture of tropical, subtropical and temperate species (Francis & Duffy, 2015). In the last decade, regional expeditions and collections of specimens have improved our knowledge of the regional fish fauna (Duffy & Ahyong, 2015; Francis, 2019), and the recent reconstruction of multi-locus time-calibrated phylogenies for 34 endemic marine fishes suggests that the emergence of these three island groups had a role in generating novel biodiversity in the region (Samayoa et al., 2022).

Here, we examine the historical biogeography of 30 marine fishes endemic to the subtropical islands of the Southwest Pacific using a parametric modelling approach based on previously inferred molecular phylogenies. First, we determine the prevalence of vicariance and jump-dispersal as biogeographical processes shaping endemism patterns in the region. Given the suggested role of the oceanic islands in generating biodiversity (Samayoa et al., 2022), we hypothesized that founder-event speciation models would be favoured. Second, we examine the influence of geographical proximity of source locations and their climate on dispersal probabilities. We hypothesized that lineages from geographically close locations of similar climate would be more successful in colonizing the subtropical islands of the Southwest Pacific prior to taxonomic diversification. Third, we use the estimated biogeographical origins and ancestral ranges to infer colonization routes for each endemic species, and then, using a comparative approach, we describe common patterns of origination and range evolution in the region. We

hypothesized that marine fish endemism in the relatively young and volcanic subtropical islands of the Southwest Pacific originated mainly from old continental islands in the region, such as Australia. Our study proposes a biogeographical scenario where multiple independent processes shaped contemporary endemism patterns in marine fishes of the Southwest Pacific, and where dispersal may be constrained by geographical distances and climate zone differences depending on the taxa examined, providing a new understanding of the evolution of marine fishes endemic to the peripheral islands of the Indo-Pacific.

## 2 | MATERIALS AND METHODS

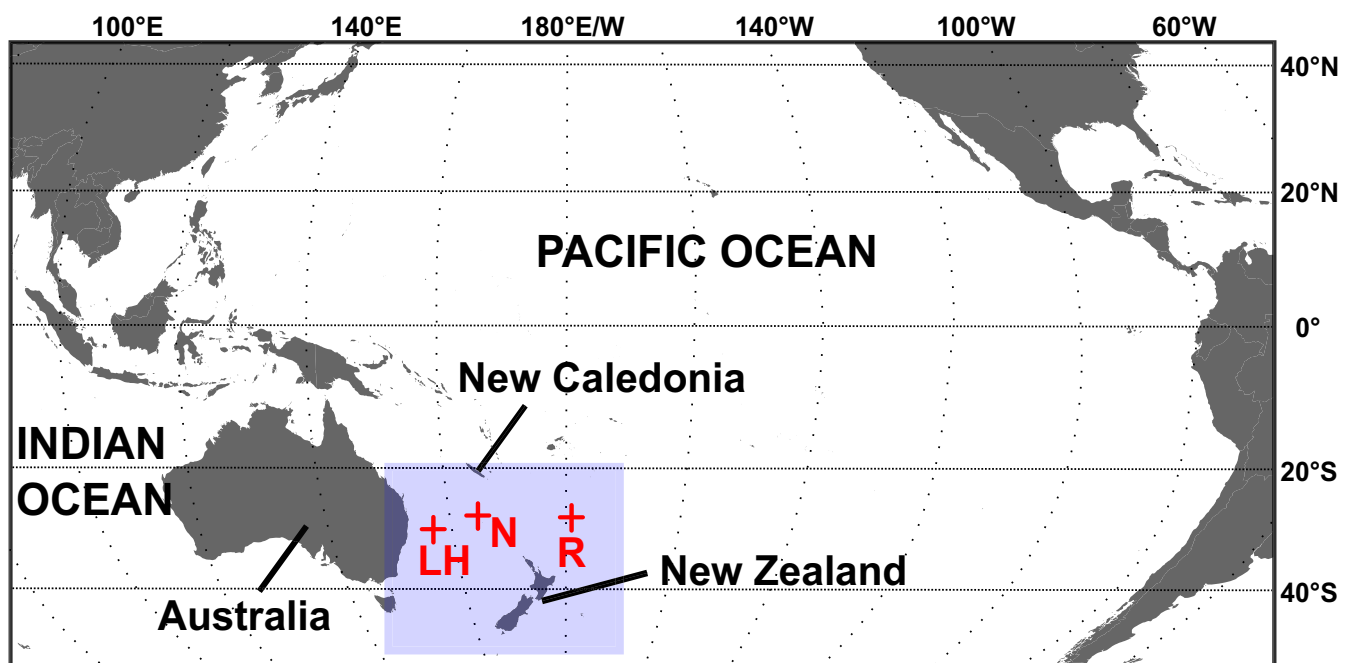
### 2.1 | Time-calibrated phylogenies

Our analyses were based on the time-calibrated trees and sequence data presented in Samayoa et al. (2022). Briefly, the authors defined marine fish taxa as endemic to the Southwest Pacific when they occurred in at least one of the subtropical islands of Lord Howe, Norfolk and Rangitāhua, without extending beyond mainland Aotearoa (New Zealand), New Caledonia and the East coast of Australia (Figure 1). Sequence data (12S, 16S, COI, CytB, D-Loop, ENC1, GLYT, H3, MYH6, PLAGL2, PTR, RAG1, RAG2, SH3PX3 and/or TMO-4C4) were available for 34 endemics and their congeners, representing 21 genera. In total, 13 independent Bayesian time-calibrated phylogenies were inferred across the tip of the actinopterygian tree using fossil calibrations (Bouckaert et al., 2019). In our study, we selected 17 of the 21 genera analysed in Samayoa et al. (2022) to examine the

historical biogeography of endemic taxa from the Southwest Pacific across the best-sampled trees. We excluded two monotypic genera and two genera that only included around 50% of known congeners to minimize any inaccuracies introduced by missing taxa (Hodge & Bellwood, 2016). Most genera were from discontinuous locations of the actinopterygian tree, and so were pruned (and hereafter referred to as clades) for analysis using the 'ape' package (Paradis & Schliep, 2019) for R 3.6.3 (R Core Team, 2020) in RStudio 1.2.5033 (RStudio Team, 2019). In three cases, when published evidence confirmed that the genera were sister groups to each other and/or that we had all intermediate taxa in our phylogeny, the clades used for analysis included two genera. Specifically, we kept *Chironemus* and *Aplodactylus* (Ludt et al., 2019; Sanciangco et al., 2016) in one clade, *Goniistius* and *Morwong* (Ludt et al., 2019) in another clade, and *Microcanthus* and *Atypichthys* in the same clade along with *Neatypus obliquus* and *Tilodon sexfasciatus*, which all form a consistent group of taxa within Microcanthidae (Tea & Gill, 2020). Consequently, our study is based on 14 clades (Table 1; Figures S2.15–S285) which encompass 30 endemics across 17 genera.

### 2.2 | Geographical distributions

We defined 14 areas based on marine biogeographical regionalizations suggested by Kulbicki et al. (2013) and Bowen et al. (2016), distinguishing each of the three subtropical islands of the Southwest Pacific and the subtropical and temperate climate zones for mainland Aotearoa and Australia (according to Commonwealth of Australia, 2012, and New Zealand Department



**FIGURE 1** Location of the Southwest Pacific (blue square), including East Australia, New Caledonia and Aotearoa New Zealand, as well as three subtropical oceanic islands (red crosses: Lord Howe—LH; Norfolk Island—N and Rangitāhua—R), following Samayoa et al. (2022). Map projection: Mollweide.



Genus/genera included	Taxa	Taxa endemic to the Southwest Pacific
<i>Optivus</i>	3	<i>Optivus agastos</i> , <i>O. elongatus</i>
<i>Arripis</i>	4	<i>Arripis trutta</i> , <i>A. xylabion</i>
<i>Scorpis</i>	4	<i>Scorpis violacea</i>
<i>Atypichthys</i> – <i>Microcanthus</i>	6	<i>Atypichthys latus</i> , <i>Microcanthus joyceae</i>
<i>Kathetostoma</i>	7	<i>Kathetostoma binigrasella</i>
<i>Nemadactylus</i>	8	<i>Nemadactylus douglasii</i> , <i>N. n. sp.</i>
<i>Lepidoperca</i>	8	<i>Lepidoperca inornata</i>
<i>Parma</i>	8	<i>Parma alboscapularis</i> , <b><i>P. kermadecensis</i></b>
<i>Hypoplectrodes</i>	9	<i>Hypoplectrodes</i> sp. A, <b><i>H. sp. C</i></b>
<i>Chironemus</i> – <i>Aplodactylus</i>	11	<i>Chironemus marmoratus</i> , <b><i>C. microlepis</i></b> , <b><i>Aplodactylus etheridgii</i></b>
<i>Goniistius</i> – <i>Morwong</i>	11	<i>Goniistius francisi</i> , <i>G. vestitus</i> , <i>Morwong ephippium</i> , <i>M. fuscus</i>
<i>Girella</i>	17	<i>Girella cyanea</i> , <b><i>G. fimbriata</i></b>
<i>Chromis</i>	19 <sup>a</sup>	<i>Chromis abyssicola</i> , <i>C. dispila</i> , <i>C. hypsilepis</i> , <i>C. kennensis</i> , <i>C. nitida</i>
<i>Upeneus</i>	29	<b><i>Upeneus francisi</i></b>

Note: Species in bold are restricted to the subtropical islands (Lord Howe, Norfolk and/or Rangitāhua).

<sup>a</sup>*Chromis* sensu stricto as in Tang et al. (2021).

Area	Code	Locations included
Australia temperate	AUST	New South Wales (North boundary: Bermagui), Victoria, Tasmania, South Australia, lower Western Australia
Australia subtropical	AUSST	South Queensland, New South Wales (South boundary: Bermagui)
Lord Howe	LH	Lord Howe Island
Norfolk	N	Norfolk Island
Rangitāhua	R	Rangitāhua (Kermadec Islands)
New Zealand subtropical	NZST	North Island limited by East Cape in the East and Cape Egmont in the West
New Zealand temperate	NZT	Remaining North Island and South Island
West and South-Central Pacific	WCP	West Pacific, South Pacific from New Caledonia to French Polynesia
Northwest Pacific	NWP	Japan, Korea, China
Hawaii	H	Hawaii
Rapa Nui/Pitcairn	RNP	Rapa Nui (Easter Island), Pitcairn
Eastern Pacific	EP	Tropical East and Central Pacific (California, Mexico, Galápagos), Southeast Pacific (Juan Fernández, Desventuradas, Peru, Chile)
Indian	I	Indian Ocean
Atlantic	A	Atlantic Ocean

of Conservation, 2019) (Table 2). To define the area/s that each taxon occupied, we referenced their occurrence data in OBIS (<https://obis.org>), Fishes of Australia (<http://fishesofaustralia.au.net>), the Catalogue of Fishes (Fricke et al., 2021), Fishbase (Froese & Pauly, 2021), faunal checklists of the Southwest Pacific (Duffy

& Ah Yong, 2015; Francis, 2019) and expert consultation. Within clades, taxa were coded as present or absent across areas to generate geographical matrices. Overall, 24 of the 30 endemic species had a continental island (either Australia, or Aotearoa or both) in their geographical range, and only six taxa were found exclusively

TABLE 1 Name of clades, number of taxa included per clade and endemic taxa examined in this study.

TABLE 2 The areas (biogeographical regions) defined to accommodate taxa across the 14 clades used for biogeographical analyses in 'BioGeoBEARS'.

associated with the subtropical islands (Lord Howe, Norfolk and/or Rangitāhua, Table 1).

## 2.3 | Biogeographical analysis

Ancestral range estimates were inferred with 'BioGeoBEARS' (Matzke, 2013) in R using RStudio. The package allows the comparison between three of the most commonly used models in historical biogeography that differ in their assumptions on range evolution during cladogenetic events: the likelihood-based Dispersal–Extinction–Cladogenesis (DEC) model (Ree & Smith, 2008) as implemented in LAGRANGE, which assumes that each descendant lineage inherits a single-area range via sympatry or vicariance within a subset of the ancestral range; the likelihood interpretation of the Dispersal–Vicariance Analysis (DIVA) (Ronquist, 1997), called DIVALIKE, assuming that descendants inherit one or more areas via vicariance; and the likelihood interpretation of the BAYAREA model (Landis et al., 2013), called BAYAREALIKE, which assumes no cladogenesis with each ancestral range inherited by both daughter lineages. Besides the anagenetic parameters for range expansion ("d") and extinction ("e") estimated by default for all standard models, we tested the three parameters "j" (jump-dispersal weight), "x" (geographical distance) and "n" (distance based on climate zone) during the ancestral range estimations by setting them from fixed (default value of 0, equivalent to standard models) to free parameters. For each clade, we tested the three standard models, and, in each case, seven more complex models based on the combination of tested parameters, for a total of 336 biogeographical models across the 14 clades. All runs used the optimization routine 'optimx' under the 'speedup' option set on FALSE, which avoids shortcuts during the full maximum-likelihood search. We assessed optimization issues for each run, making sure that the most complex models resulted in equal or higher log-likelihood values compared to nested models. When 'optimx' yielded lower log-likelihood, we replaced it with the optimizer 'GenSA', which performs better for models with a high number of parameters. As 'GenSA' did not automatically improve log-likelihood for all four and five parameter-based models during our initial runs, we proceeded with the replacement only when 'optimx' performed poorly.

Our BioGeoBEARS analyses included time constraints to account for geological events (i.e. island emergence and formation of a land barrier) likely to have affected dispersal among areas. Because the package requires a specific age value to set time constraints, we considered emergence timings of oceanic islands at 2.5 Ma for Rapa Nui (Clouard & Bonneville, 2005), 2.58 Ma for Rangitāhua (Brook, 1998), 3.05 Ma for Norfolk (Jones & McDougall, 1973), 6 Ma for Juan Fernández (Burrige et al., 2006) and 6.9 Ma for Lord Howe (McDougall et al., 1981), and for the Central American land barrier at 3.1 Ma (Cowman & Bellwood, 2013). Furthermore, as oceanic islands emerge from multiple processes (Neill & Trewick, 2008), including a seamount origin likely to be colonized before its surfacing, we selected the oldest estimated ages of contemporary surfaced

islands as a standard across our focal islands. Similarly, the closure of the Central American Seaway involved complex oceanographic processes between 12 and 3 Ma (Kamikuri et al., 2009) where gene flow between Atlantic and Pacific marine populations could have been interrupted before, during and after the final closure ca. 3 Ma (Lessios, 2008; O'Dea et al., 2016). Following previous biogeographical analyses of marine fishes (Cowman & Bellwood, 2013; Siqueira et al., 2019; Thacker, 2015), we used 3.1 Ma as the Central American Seaway closure event.

For each clade, we partitioned the evolutionary time-scale in time slices using the ages of the geological events that influenced the geographical distribution of taxa included in the phylogeny. For clades with taxa found in the East Pacific, we included Juan Fernández as a time constraint only when it was the sole location covered within our East Pacific area. Conversely, when additional locations were covered within the East Pacific, we did not include Juan Fernández as a time constraint, which was the case for the *Girella* and *Chironemus–Aplodactylus* clades. Across all clades, we applied constraints on dispersal within each time slice using the 'areas\_allowed' option which removes the possibility of dispersal (value of 0) to oceanic islands, when these were not yet emerged while moving backward in time, and between areas in the Pacific and the Atlantic separated by the closure of the Panama Isthmus.

For clades with Atlantic representatives, we set a value of 0 among Atlantic and Pacific areas in time slices where the Isthmus of Panama was closed (i.e. younger than 3.1 Ma), whereas the values were set to 1 (areas allowed) for older time slices. We allowed connectivity between Atlantic and specific Pacific areas in younger time slices depending on the geographical distribution of particular taxa. For *Nemadactylus*, taxa are found in cold southern waters of the Atlantic and the Pacific, for which reason we allowed the connection between the Atlantic and temperate Australia to simulate the connectivity of these waters along the West Wind Drift (Waters, 2008). For *Girella*, given that the sole Atlantic member is closely related to congeners found in the Northwest Pacific (Beldade et al., 2021; Samayoa et al., 2022), we allowed connectivity among both areas. For *Upeneus*, the Atlantic and the Pacific were connected through the Indian Ocean where most taxa occur.

## 2.4 | Geographical distances and climate zones

The estimation of "x" (geographical distance) and "n" (climate zone distance) required the generation of distance matrices. For "x", we measured great-circle distances (i.e. the shortest distance between two points along a sphere's surface), following oceanic routes between midpoints of areas (the centre of an area or archipelago, the mid-point of a coastline and/or oceanic islands) using the 'sf' package (Pebesma, 2018) for R. For each clade, distances were kept constant across time slices except between areas of the Atlantic and the Pacific which were separated by the Panama Isthmus at 3.1 Ma. For areas of the Atlantic and Pacific separated 3.1 Ma, distances were calculated westwards, whereas for areas separated by more

than 3.1 Ma, measurements were performed eastwards if they were the shortest distances among areas across the open channel. For “n”, we allocated distances on a scale from 1 to 3: a value of 1 was set between areas within the same climate (subtropical, tropical or temperate); a value of 2 was set for adjacent climates (temperate/subtropical and subtropical/tropical); and a value of 3 was set for non-adjacent climates (tropical/temperate). Distances in both geographical and climate matrices were divided by the highest value to obtain relative distances. When distance parameters are set free, the parameters “d” and “j” are multiplied by distance to a power (e.g. “x” and “n”) (Matzke, 2013; Van Dam & Matzke, 2016): when “x” (or “n”) equals 0, distance has no effect, corresponding to the default setting; when “x” (or “n”) is negative, dispersal rate decreases when distance increases. To run our BioGeoBEARS analyses, we used the default limits of  $-2.5$  and  $0.0$  from the package for “x” and “n”.

## 2.5 | Best-fitting model selection

We selected best-fitting models using the sample-size corrected Akaike information criterion (AICc) preferentially, which corrects AIC for small sample sizes. However, given that we estimated up to five parameters (“d”, “e”, “j”, “x”, “n”) and that the sample sizes should be over  $k + 1$ , wherein  $k$  is the number of estimated parameters, AICc was not computable for clades with six taxa or less (i.e. *Optivus*, *Arripis*, *Scorpis* and *Atypichthys-Microcanthus*). In these cases, we used AIC scores as guidance for model selection. Because AIC is not intended for very small-sized samples, these AIC results (Tables S1.1S–S1.8S; Figures S2.1S–S2.8S) should be interpreted with caution. After selecting a best-fitting model based on the lowest AICc/AIC score, and when the model included at least one of the three free parameters we were interested in (“j”, jump-dispersal; “x”, geographic distance; “n”, climate zone distance), we confirmed that adding the parameter(s) significantly increased the log-likelihood of the standard model (DEC, DIVALIKE or BAYAREALIKE) using a log-likelihood-ratio test (LRT). Accordingly, we determined the parameters that significantly improved model fit by performing LRTs between nested models and their more complex counterparts, using a threshold of 0.05 for the  $p$ -value per comparison. For each clade, we performed 57 LRTs, yielding a total of 798 tests across all clades (Tables S1.1S–S1.28S).

## 3 | RESULTS

### 3.1 | Optimization of parameter estimates

We compared values for log-likelihood along 798 pairs of models for 336 models to estimate the ancestral range of our 30 endemic species. In 60 of the 798 pairwise comparisons, the more complex models yielded lower log-likelihood than their respective nested model. In 38 of the 60 comparisons, the complex models involved four or five free parameters, and replacing ‘optimx’ with ‘GenSA’

increased log-likelihood for 17% (17/99) and 63% (21/23) of four- and five-parameter models, respectively (Table S1.29S). In 22 of the 60 comparisons, the average log-likelihood difference was 0.0024 and changing the optimizer did not increase log-likelihood. As more complex models were expected to yield either higher or the same log-likelihood, and given the small difference detected for the 22 tests, we considered in these cases that the more complex models displayed the same log-likelihood value.

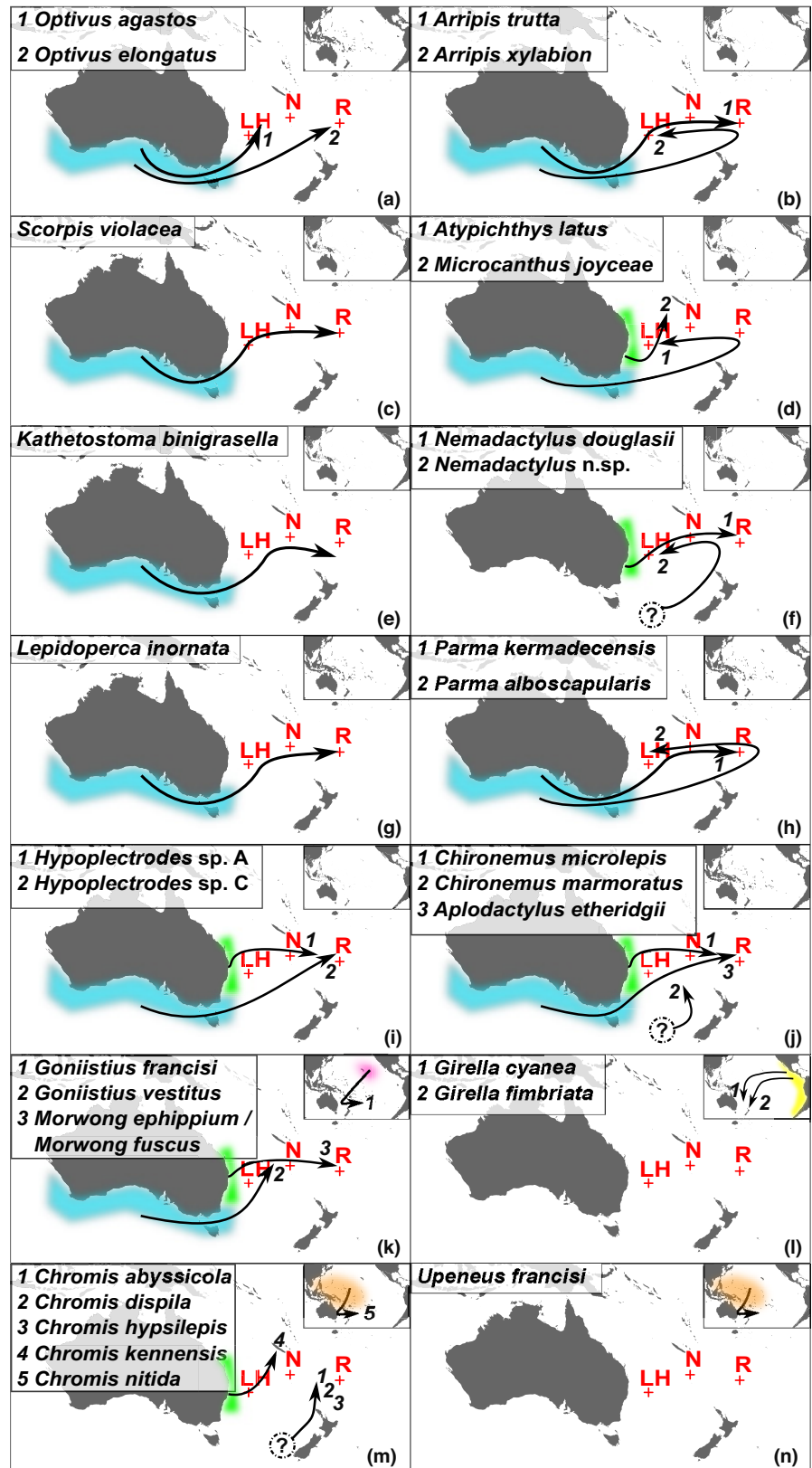
Among the parameters estimated across the 336 models, jump-dispersal estimates (“j”) reached their maximum value in 23 models using the default BioGeoBEARS settings, an indication of optimization issues. Maximum “j” estimates were mainly reported for BAYAREALIKE (18/23 cases; max = 0.9999), followed by DIVALIKE (5/23 cases; max = 1.9999), with no reports for DEC models. High “j” values were consistently yielded under different ‘BioGeoBEARS’ settings: trials with three optimizers; unconstrained analyses; setting the re-scaling option ‘rescale\_params’ on TRUE when free parameters are in different scales; and modifying limits for “d”, “e”, and “j”. Optimization issues were minimized by verifying we did not include phylogenetic trees with ultra-short branches; rescaling distances; checking log-likelihood differences between complex and nested models; using ‘GenSA’ when appropriate and verifying that convergence was achieved after optimization. Since we tried all available optimizers in the package, we likely reached the best estimates possible under the current ‘BioGeoBEARS’ version, and given the available data for our focal genera. Estimates appeared unrelated to sample size since our largest clade (*Upeneus*, 29 taxa) showed low “j” values similar to one of the smallest clades (*Parma*, 8 taxa).

For *Optivus*, values over 1 for range expansion (“d”) were recovered in three of the eight BAYAREALIKE models, and for extinction (“e”), values over 1 were recovered in seven of the eight BAYAREALIKE models (Table S1.1S). Since values of  $e < 1$  indicate less than one anagenetic event (either range expansion or extinction) per million years (Matzke, 2013), the high values in *Optivus* suggest more frequent anagenetic events for this clade. However, this clade also exhibited the lowest taxonomic coverage in our dataset (Table 1), requiring caution in interpreting estimates for anagenetic processes.

### 3.2 | Ancestral range estimations

Inferred biogeographical origins and colonization routes for endemic taxa revealed that most very likely originated in Australia (Figure 2; Figures S2.1S–S2.28S). In all, 13 endemic lineages likely originated from Temperate Australia (*Optivus agastos*, *O. elongatus*, *Arripis trutta*, *A. xylabion*, *Scorpis violacea*, *Atypichthys latus*, *Kathetostoma binigrasella*, *Lepidoperca inornata*, *Parma alboscapularis*, *P. kermadecensis*, *Hypoplectrodes* sp. C, *Chironemus marmoratus* and *Goniistius vestitus*), while seven likely originated from Subtropical Australia (*Microcanthus joyceae*, *Nemadactylus douglasii*, *Hypoplectrodes* sp. A, *Aplodactylus etheridgii*, *Morwong ephippium*, *M. fuscus* and *Chromis kennensis*). Our analyses also showed that two endemic

**FIGURE 2** Inferred biogeographical origin and colonization pathways for 30 marine fishes endemic to the subtropical Southwest Pacific based on the best-fitting models in 'BioGeoBEARS'. Colours indicate biogeographical regions included in the analysis: blue is Temperate Australia, green is Subtropical Australia, pink is Hawaii, yellow is East Pacific and orange is West/Central Pacific. Circles with a question mark indicate an unresolved origin. Red crosses indicate the position of the Southwest Pacific islands of Lord Howe (LH), Norfolk (N) and Rangitāhua (R). Arrows indicate proposed colonization routes of an endemic lineage from its most likely geographical origin, as estimated in BioGeoBEARS. (a) *Optivus*; (b) *Arripis*; (c) *Scorpis*; (d) *Atypichthys*-*Microcanthus*; (e) *Kathetostoma*; (f) *Nemadactylus*; (g) *Lepidoperca*; (h) *Parma*; (i) *Hypoplectrodes*; (j) *Chironemus*-*Aplodactylus*; (k) *Goniistius*-*Morwong*; (l) *Girella*; (m) *Chromis*; (n) *Upeneus*.



lineages likely originated from the East Pacific (*Girella cyanea* and *G. fimbriata*), two in the West and Central Pacific (*Chromis nitida* and *Upeneus francisi*), and one in Hawaii (*Goniistius francisi*). Five lineages had unresolved origins (*Nemadactylus* n.sp., *Chironemus microlepis*, *Chromis abyssicola*, *C. dispila* and *C. hypsilepis*). Furthermore, endemic

species restricted to the subtropical islands (Table 1) did not appear to have a common origin.

DEC was the most common biogeographical model supported by our data (7/14), with the geographical distance parameter ("x") significantly improving the model in six cases, the climate parameter

("n") in three cases, and jump-dispersal ("j") in two cases (Table 3; Tables S1.1S–S1.28S). DIVALIKE was the second-best standard model supported by our data (5/14), with geographical distance ("x") and jump-dispersal ("j") significantly improving the model in two cases each, with no significant improvement by any parameter in two cases. BAYAREALIKE was the best-supported model in 2-of-14 clades, always in combination with jump-dispersal ("j"), with geographic distance ("x") and climate ("n") significantly improving one model each.

Overall, jump-dispersal was included in 6-of-14 best-fitting models (Table 3). Ancestral range estimates show that 8-of-30 endemic species emerged through jump-dispersal (Figures S2.1S–S2.28S): *A. etheridgii*, *C. microlepis*, *C. nitida*, *G. cyanea*, *G. fimbriata*, *G. francisi*, *Hypoplectrodes* sp. C and *U. francisi*. When contrasting this result with the divergence timing of the endemic lineages, we found that both old and young endemics have emerged from founder-event speciation (Figure 3), suggesting that this mode of speciation is unrelated to the time of divergence. Of the eight endemic taxa that likely diverged from a founder-event, five are restricted to the subtropical islands (Figure 3).

TABLE 3 Best-fitting biogeographical model per clade, contrasted with the number of taxa sampled.

Name of clade	Number of taxa	Best-fitting model
<i>Optivus</i> *	3	DEC + x
<i>Arripis</i> *	4	DEC + x + n
<i>Scorpiis</i> *	4	DEC + x
<i>Atypichthys</i> – <i>Microcanthus</i> *	6	DEC + x
<i>Kathetostoma</i>	7	DIVALIKE + x
<i>Nemadactylus</i>	8	DEC + n
<i>Lepidoperca</i>	8	DIVALIKE
<i>Parma</i>	8	DIVALIKE
<i>Hypoplectrodes</i>	9	DIVALIKE + j
<i>Chironemus</i> – <i>Aplodactylus</i>	11	BAYAREALIKE + j + x
<i>Goniistius</i> – <i>Morwong</i>	11	DIVALIKE + j + x
<i>Girella</i>	17	DEC + j + x
<i>Chromis</i>	19	DEC + j + x + n
<i>Upeneus</i>	29	BAYAREALIKE + j + n

Note: Model selection was based on the corrected version of the Akaike information criterion (AICc) for small sample sizes, except for clades with an asterisk (\*) for which AICc was not computable owing to their small sample sizes. In these cases, AIC-based best-fitting models for such clades are presented, but caution should be taken in their interpretation. Models based on Matzke (2013).

Abbreviations: BAYAREALIKE, likelihood implementation of the BAYAREALIKE model; DEC, Dispersal-Extinction-Cladogenesis model; DIVALIKE, likelihood implementation of the Dispersal-Variance Analysis model; +j, model with significant jump-dispersal parameter "j"; +n, model with significant environmental distance parameter "n", based on climate zone; +x, model with significant geographic distance parameter "x".

## 4 | DISCUSSION

Our analyses of the historical biogeography of marine fishes endemic to subtropical islands of the Southwest Pacific highlight vicariance and jump-dispersal as significant actors shaping endemism in the region, with geographical distance and climate zones significantly limiting dispersal in particular marine fish groups. We present hypotheses on the origin of endemic lineages and subsequent colonization routes, where eastward colonization from mainland Australia is the dominant pattern, followed by rare southward pathways from tropical areas, and westward routes from the East Pacific. Finally, by acknowledging caveats in our analysis, we highlight our study's contributions to understanding the historical biogeography of marine fish endemism in the Southwest Pacific.

### 4.1 | Biogeographical processes: Vicariance and jump-dispersal

Our results indicate that vicariance is a prevalent driver of allopatric speciation in the Southwest Pacific. The selection of DEC for most clades (Table 3) implies that range evolution in these cases has been driven by cladogenetic processes such as peripatry and within-area vicariance (Ree & Smith, 2008), although careful interpretation is required for four clades due to their small sample sizes (*Optivus*, *Arripis*, *Scorpiis* and *Atypichthys*–*Microcanthus*). Interestingly, for *Microcanthus*, a previous assessment of the historical biogeography of Microcanthidae retrieved DIVALIKE as the best-fitting model (Tea et al., 2019). Despite fewer sister taxa sampled in the previous investigation, both results show vicariance as a significant driver of range evolution in microcanthids. The selection of five DIVALIKE-based models in our analyses (Table 3) indicates that classic vicariance has also been significant in the range evolution of many taxa in our study (Ronquist & Sanmartin, 2011). Our results show that, over time, populations residing on volcanic islands can become isolated via 'soft' vicariant barriers (Luiz et al., 2012), caused by geographical or environmental distance or climatic difference, facilitating their speciation as more range-restricted lineages. Nonetheless, for our best-fitting models based on BAYAREALIKE (*Chironemus*–*Aplodactylus* and *Upeneus*), the simulations did not include cladogenetic processes (i.e. peripatry and vicariance), implying that vicariance is not the only determinant of range evolution in fishes of the Southwest Pacific.

Marine island biogeography theory predicts that recently emerged oceanic islands are colonized by organisms with high dispersal capacities, followed by an increase in speciation and endemism richness (Dawson, 2016). Our results support this statement, where almost half of our models included jump-dispersal events, highlighting the relevance of founder-event speciation in the emergence of endemic species restricted to these subtropical, oceanic islands (5/8; Table 3; Figure 2). A similar prevalence of jump-dispersal in the range evolution has been found for surgeonfishes and parrotfishes (Siqueira et al., 2019), kyphosids (Knudsen et al. 2019), and for two taxa endemic to Rapa Nui (*Myripristis tiki* and *Chrysiptera*,



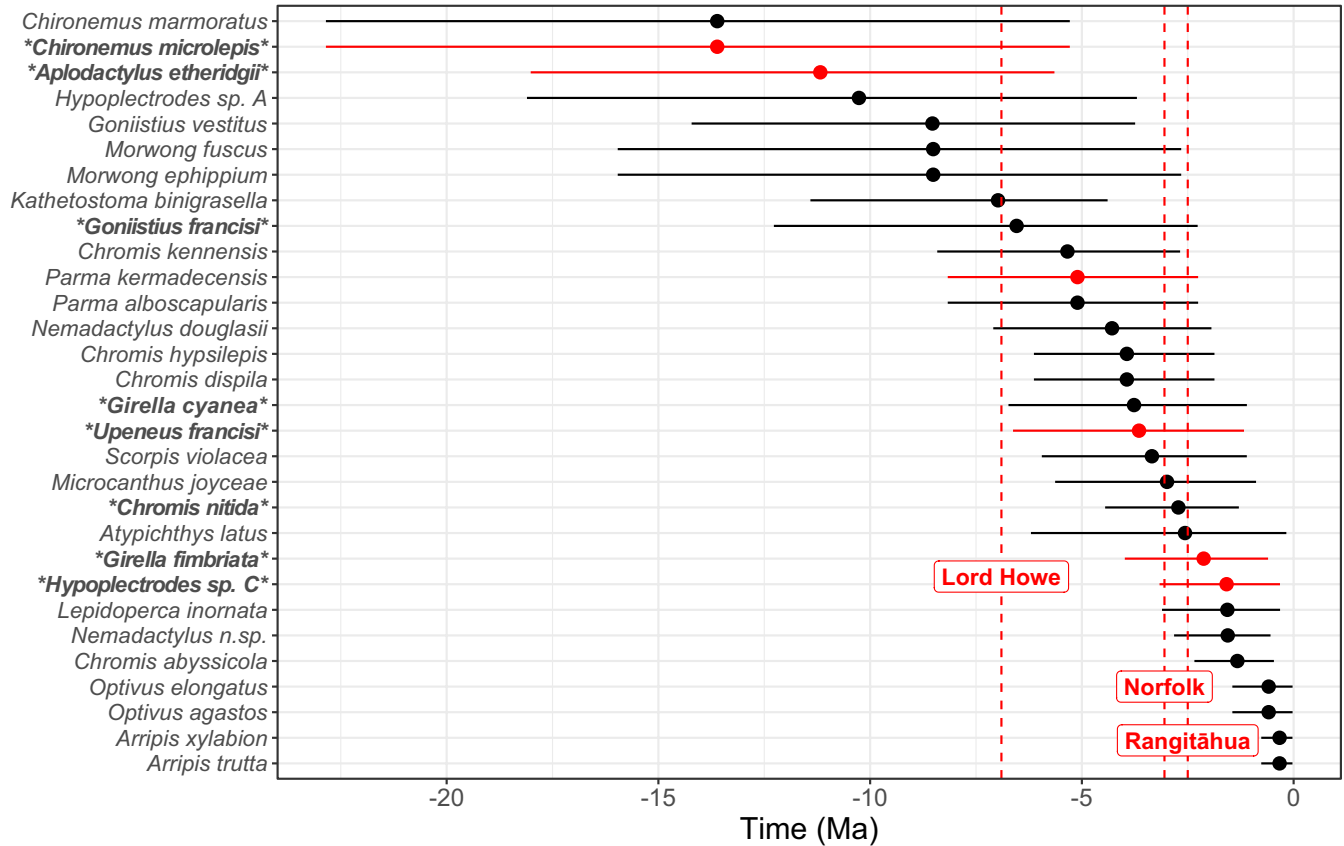


FIGURE 3 Divergence time estimates in million years (Ma) for 30 marine fish taxa endemic to the Southwest Pacific, based on Samayoa et al. (2022). Mean divergence age (dots) and 95% posterior distribution (horizontal bars) are represented. Red bars indicate endemic taxa restricted to the subtropical islands of Lord Howe, Norfolk and/or Rangitāhua, and vertical dotted lines show the estimated age of the islands: Lord Howe (6.9 Ma); Norfolk (3.05 Ma) and Rangitāhua (2.58 Ma). Asterisk-framed taxa in bold indicate endemics with strong support for jump-dispersal speciation according to best-fitting biogeographical models.

Delrieu-Trotin et al., 2019). In contrast, for endemic species present in both continental and oceanic islands of the Southwest Pacific, we found jump-dispersal events were more rarely included in the best-fitting models (3/24). Furthermore, for several taxa within *Hypoplectrodes*, *Goniistius-Morwong*, *Girella* and *Chromis*, best-fitting models indicated that both vicariance and jump-dispersal had a role in driving range evolution (i.e. DEC/DIVALIKE plus “J”). Overall, our biogeographical results demonstrate that jump-dispersal is significant in most but not all of our models, and instead show that vicariance and jump-dispersal act with different weights depending on taxonomic and spatiotemporal scales (Sanmartín, 2012), a conclusion that has been reported in marine crustaceans (Liu et al., 2018) and molluscs (Cunha et al., 2019), as well as in freshwater stingrays (Fontenelle et al., 2021).

#### 4.2 | Effect of geographical distance and climate zone differences on dispersal

Geographical distance is known to potentially limit dispersal in the sea (Cowen & Sponaugle, 2009). In our study, we found evidence for this assertion, as best-fitting models for 64% of clades included

a significant negative effect of geographical distance in dispersal probabilities. Affected taxa vary considerably in their ecology, evolutionary history and biology, including various: diet types, such as herbivory (e.g. *Girella fimbriata*), omnivory (e.g. *G. cyanea*) and zooplanktivory (e.g. *S. violacea*) (Knudsen et al., 2019); patterns of originations (Figure 2); and divergence timings (Figure 3). Early life-history data are only available for five endemics (*C. abyssicola*, *C. dispila*, *C. hypsilepis*, *C. nitida* and *Girella fimbriata*) in Fishbase (Froese & Pauly, 2021). The four *Chromis* taxa lay demersal eggs, supporting other studies that suggest that the dispersal of taxa with benthic eggs is more affected by geographical distance than in those with pelagic eggs (Riginos et al., 2011). Interestingly, in four clades (*Chironemus-Aplodactylus*, *Goniistius-Morwong*, *Girella* and *Chromis*), range evolution is driven by both geographical distance constraints and jump-dispersal events. This result suggests long-distance dispersal events can occur even in marine fishes where colonization mainly occurs among geographically proximal locations (Benestan et al., 2021).

We hypothesized that lineages from climatically similar locations would be more successful in colonizing the subtropical islands of the Southwest Pacific, and would then speciate to become endemic species. We found that climate zone differences significantly



limited dispersal in four clades (Table 3; Tables S1.3S, S1.11S, S1.25S, S1.27S). These groups comprise endemic lineages that display relatively recent diversification timings (5.34 Ma [2.68–8.41, 95% height posterior density (HPD)] for *C. kennensis* to 0.33 Ma [0.02–0.76, 95% HPD] for *Arripis trutta*/*A. xylabion*; Figure 3), and include the temperate taxa within *Arripis* and *Nemadactylus* and the only two cases of tropical origination among our dataset (*Chromis* and *Upeneus*; Figure 2m–n). Our results suggest that climate has been a relevant factor in marine fish speciation within the last 5 Ma, a period characterized by rapid glacial/interglacial periods with latitudinal shifts in subtropical and temperate oceanic boundaries in the Southwest Pacific (McClymont et al., 2016; Nelson & Cooke, 2001). This geological setting is known to have influenced the biogeography of *Arripis* taxa which are highly sensitive to climate changes (Moore & Chaplin, 2014), and the longitudinal distribution of *Nemadactylus* members across Australia (Burrige, 2000). For endemic species within *Chromis* and *Upeneus*, biogeographical patterns could be partially explained by poleward shifts of subtropical marine boundaries in the Southwest Pacific during interglacial periods (McClymont et al., 2016), facilitating rare colonization of higher latitude islands from taxonomic groups mainly inhabiting tropical waters.

### 4.3 | Origin and colonization routes of the Southwest Pacific islands

Similar to terrestrial fauna (Goldberg et al., 2008), our historical biogeographical scenarios (Figure 2) revealed mainland Australia as the inferred origin for 66% (20/30) of endemic lineages, representing the major source of marine fish endemism in the Southwest Pacific. Of the 20 lineages, 13 likely originate in Temperate Australia and 7 in Subtropical Australia. These results support the qualitative inferences from Samayoa et al. (2022) where 53% of Southwest Pacific endemics were most closely related to Australian fauna. Our data also show that colonization routes departing from Temperate Australia involve an intermediate subtropical range, and that all routes leaving Australia follow an eastward pathway likely facilitated by the predominant eastward oceanic circulation of the region (Trnski & de Lange, 2015). Lineages originating in Australia include diversification timings between 13.61 Ma (5.28–22.85, 95% HPD) (*C. microlepis*) and 0.33 Ma (0.02–0.76, 95% HPD) (*Arripis trutta*/*A. xylabion*; Figure 3), implying multiple independent colonization events and processes over a long temporal scale. In some cases, the diversification of endemic taxa could predate the emergence of islands they are endemic to (paleoendemism, Stebbins & Major, 1965; e.g. *C. microlepis*), potentially indicating the importance of these islands in providing refuge for relict lineages extirpated from other parts of their range. We note that BioGeoBEARS considers mean divergence ages (dots, Figure 3) from the Bayesian phylogenies, excluding the 95% posterior distribution per estimated age (horizontal bars, Figure 3) in the analyses. Therefore, although a mean age older than an oceanic island supports paleoendemism, a 95% HPD that includes the island's age does not rule out in situ diversification.

Low-latitude, tropical regions are generally more biodiverse than high-latitude regions, and tend to export marine fish biodiversity to subtropical and temperate regions (Bowen et al., 2013). Our results suggest that three of the endemic lineages of the Southwest Pacific likely originated in tropical areas (Figure 2). For *C. nitida*, results point to an origin in the West/Central Pacific, with an intermediate range in Subtropical Australia, and eventual expansion to Lord Howe (Figure S2.25S). For *G. francisi*, ancestral range estimation (Figure S2.21S) shows Hawaii as the potential origin, subsequent migration to New Caledonia 6.54 Ma, and the eventual colonization of the Southwest Pacific islands. However, whereas the importance of New Caledonia in the speciation process has previously been reported in Delrieu-Trottin et al. (2019), a Hawaiian origin is dependent on the placement of *G. vitattus* in the phylogeny, which may require more investigation (see Ludt et al., 2019). For *U. francisi*, our data indicate a most likely origin in the West/Central Pacific, a migration to Lord Howe 3.65 Ma, and an expansion to Norfolk and Rangitāhua (Figure S2.27S). In all three cases, jump-dispersal is a significant driver of range evolution (Table 3). Colonization from warmer tropical regions is similarly implicated over recent time-scales in Rangitāhua (Liggins et al., 2020), and temperate Aotearoa (Middleton et al., 2021), and has been described in Eastern and Western Australia (Stuart-Smith et al., 2018; Wernberg et al., 2016). These immigration and colonization events are becoming more frequent as the subtropical and even temperate regions warm over contemporary time-scales due to climate change (Hastings et al., 2020).

Molecular evidence has also highlighted the connection between marine fish fauna on both sides of the Pacific (Burrige et al., 2006; Delrieu-Trottin et al., 2018; Liggins et al., 2022), a pattern revealed by our study in two recently diversified lineages within *Girella*, where colonization is inferred to originate in the East Pacific (Figure 2; Figure S2.23S), and driven by vicariant and jump-dispersal events (Table 3). For five lineages, we were unable to determine their origin and colonization pathways. In two cases, the ancestral ranges before divergence are considerably wide: for *Nemadactylus* n. sp., the ancestral range includes Australia, Aotearoa and the Atlantic (Figure S2.11S); and for *C. microlepis*, the ancestor is estimated to have occurred in Australia mainland and Subtropical Aotearoa, a range from which it would have colonized Lord Howe (Figure S2.19S). The other three lineages belong to *Chromis*, and all display unresolved origins at the most proximate ancestral nodes (Figure S2.26S).

Oceanic islands are key actors in actinopterygian evolution in the Indo-Pacific (Bowen et al., 2013), and in our study, Lord Howe, Norfolk, and Rangitāhua are included in ancestral range estimations after their emergence (dashed lines in Figures S2.1S–S2.28S), highlighting their role in range evolution for endemic lineages (Samayoa et al., 2022). Lord Howe appears as the first stepping-stone in the northern pathway of *Optivus agastos* (Figure 2a; Figure S2.1S), and the eastward routes of *S. violacea* (Figure 2c; Figure S2.5S), and *U. francisi* (Figure 2n; Figure S2.27S). Its oldest emergence and closest proximity to Australia among the three island groups (Francis, 1993) might have facilitated its leading role in shaping endemism patterns

in the region. Our analyses also highlight the role of Subtropical Aotearoa as an intermediate stepping-stone for seven endemic lineages: in three cases (*Optivus elongatus*, *K. binigrasella* and *L. inornata*), the area appears as a stepping-stone before colonization of Rangitāhua in the north; however, in four cases (*A. latus*, *Arripis xylobion*, *Nemadactylus* n.sp. and *P. alboscapularis*), analyses indicate a westward pathway from Aotearoa towards Lord Howe, a rare route inferred for plants (Birch & Keeley, 2013) and freshwater gastropods (Zielske et al., 2017). While a westward route seems counter-intuitive for the region's marine fishes as it implies active dispersal against the predominant eastward oceanic currents, it has also been suggested for three morphologically distinct *Chrysiptera* taxa thought to have originated in the Central Pacific, and later colonizing the Southwest Pacific (Liggins et al., 2022).

## 5 | CONCLUSIONS

In our study, we investigated the historical biogeography of 30 marine fish taxa endemic to subtropical islands of the Southwest Pacific, characterized by a dominant eastward colonization from mainland Australia and rare routes from northern tropical and eastern Pacific areas, involving independent colonization events across a wide temporal scale. Vicariance and jump-dispersal were not mutually exclusive processes in the evolution of endemism in the region, and both geographical distance and climate zone differences were relevant but not dominant factors determining dispersal. Our inferences are limited by small sample sizes in four clades, and by the necessity of grouping small marine areas into larger regions (e.g. clumping small islands into 'East Pacific' obscuring the exact origin of endemic *Girella* lineages) to ensure that our analyses were computationally feasible. Additionally, by keeping constant climatic distances along time, our analyses simplify the effect of past latitudinal shifts in climate (McClymont et al., 2016) on range evolution. Integrating interdisciplinary evidence from biological and oceanographic sources to better understand dispersal processes (Cowen & Sponaugle, 2009) could potentially clarify the modes of speciation and dispersal routes unveiled by our analyses, including those which are most relevant to contemporary evolution. In particular, the role of processes operating at smaller spatial scales, such as the interaction between organisms (e.g. priority effects, competition, predation) are not included in our modelling approach, but are important factors known to shape the biogeography of marine organisms (Waters et al., 2013) and marine fishes (Hodge & Price, 2022). Nonetheless, our study has unveiled the origins and the role of multiple processes contributing to marine fish endemism in subtropical islands of the Southwest Pacific, improving our understanding of the historical biogeography of marine fishes in this understudied region of the Pacific.

## ACKNOWLEDGEMENTS

We are grateful to two reviewers, the handling editor and the chief editor for their comments that substantially improved the

manuscript. We thank Nicholas J. Matzke for providing constructive insights on BioGeoBEARS, as well as Adam N. H. Smith (Massey University) and José I. Carvajal for the early feedback on this study. 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 (Authorization 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. 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). André P. Samayoa is supported by a Massey University Doctoral Scholarship. Libby Liggins is supported by a Rutherford Discovery Fellowship (RDF-20-MAU-001). Open access publishing facilitated by Massey University, as part of the Wiley - Massey University agreement via the Council of Australian University Librarians.

## CONFLICT OF INTEREST STATEMENT

The authors declare no conflict of interest.

## DATA AVAILABILITY STATEMENT

The original molecular data (gene regions and accession numbers) used for phylogenetic inference can be found as supplementary material at <https://doi.org/10.1016/j.jympev.2022.107584>. Alignment and tree files are available in Mendeley at <https://doi.org/10.17632/cmz9gjd97g.1>. The BioGeoBEARS files generated in this study are available in Mendeley at <https://doi.org/10.17632/dydcnc77b5.2>.

## ORCID

André P. Samayoa  <https://orcid.org/0000-0003-3852-7935>

J. David Aguirre  <https://orcid.org/0000-0001-7520-441X>

Erwan Delrieu-Trottin  <https://orcid.org/0000-0002-4120-9316>

Libby Liggins  <https://orcid.org/0000-0003-1143-2346>

## REFERENCES

- Allen, G. R. (2008). Conservation hotspots of biodiversity and endemism for Indo-Pacific coral reef fishes. *Aquatic Conservation: Marine and Freshwater Ecosystems*, 18(5), 541–556.

- Baraf, L. M., Pratchett, M. S., & Cowman, P. F. (2019). Ancestral biogeography and ecology of marine angelfishes (F: Pomacanthidae). *Molecular Phylogenetics and Evolution*, 140, 106596. <https://doi.org/10.1016/j.ympev.2019.106596>
- 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. *Molecular Phylogenetics and Evolution*, 156, 107021. <https://doi.org/10.1016/j.ympev.2020.107021>
- Benestan, L., Fietz, K., Loiseau, N., Guerin, P.-E., Trofimenko, E., Rühls, S., Schmidt, C., Rath, W., Biastoch, A., & Pérez-Ruzafa, A. (2021). Restricted dispersal in a sea of gene flow. *Proceedings of the Royal Society B*, 288(1951), 20210458. <https://doi.org/10.1098/rspb.2021.0458>
- Birch, J. L., & Keeley, S. C. (2013). Dispersal pathways across the Pacific: The historical biogeography of *Astelia* s.l. (Asteliaceae, Asparagales). *Journal of Biogeography*, 40(10), 1914–1927.
- 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 Computational Biology*, 15(4), e1006650.
- Bowen, B. W., Gaither, M. R., DiBattista, J. D., Iacchi, M., Andrews, K. R., Grant, W. S., Toonen, R. J., & Briggs, J. C. (2016). Comparative phylogeography of the ocean planet. *Proceedings of the National Academy of Sciences of the United States of America*, 113(29), 7962–7969.
- Bowen, B. W., Rocha, L. A., Toonen, R. J., & Karl, S. A. (2013). The origins of tropical marine biodiversity. *Trends in Ecology & Evolution*, 28(6), 359–366.
- Brook, F. (1998). Stratigraphy and paleontology of Pleistocene submarine volcanic-sedimentary sequences at the northern Kermadec Islands. *Journal of the Royal Society of New Zealand*, 28(2), 235–257.
- Burridge, C. P. (2000). Biogeographic history of geminate cirrhitoids (Perciformes: Cirrhitidae) with east–west allopatric distributions across southern Australia, based on molecular data. *Global Ecology and Biogeography*, 9(6), 517–525.
- 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. *Systematic Biology*, 55(4), 566–578.
- Clouard, V., & Bonneville, A. (2005). Ages of seamounts, islands, and plateaus on the Pacific plate. In G. R. Foulger (Ed.), *Plates, plumes, and paradigms* (pp. 71–90). Geological Society of America.
- Commonwealth of Australia. (2012). *Marine bioregional plan for the temperate east marine region*. Department of Sustainability, Environment, Water, Population and Communities, Australian Government. <https://www.dcceew.gov.au/environment/marine/marine-bioregional-plans/temperate-east>
- Cowen, R. K., & Sponaugle, S. (2009). Larval dispersal and marine population connectivity. *Annual Review of Marine Science*, 1, 443–466. <https://doi.org/10.1146/annurev.marine.010908.163757>
- Cowman, P. F., & Bellwood, D. R. (2013). Vicariance across major marine biogeographic barriers: Temporal concordance and the relative intensity of hard versus soft barriers. *Proceedings of the Royal Society B: Biological Sciences*, 280(1768), 20131541. <https://doi.org/10.1098/rspb.2013.1541>
- Cowman, P. F., Parravicini, V., Kulbicki, M., & Floeter, S. R. (2017). The biogeography of tropical reef fishes: Endemism and provinciality through time. *Biological Reviews*, 92(4), 2112–2130.
- Cunha, T. J., Lemer, S., Bouchet, P., Kano, Y., & Giribet, G. (2019). Putting keyhole limpets on the map: Phylogeny and biogeography of the globally distributed marine family Fissurellidae (Vetigastropoda, Mollusca). *Molecular Phylogenetics and Evolution*, 135, 249–269.
- Dalongeville, A., Andreollo, M., Mouillot, D., Lobreaux, S., Fortin, M. J., Lasram, F., Belmaker, J., Rocklin, D., & Manel, S. (2018). Geographic isolation and larval dispersal shape seascape genetic patterns differently according to spatial scale. *Evolutionary Applications*, 11(8), 1437–1447.
- Dawson, M. N. (2016). Island and Island-like marine environments. *Global Ecology and Biogeography*, 25(7), 831–846.
- Delrieu-Trottin, E., Liggins, L., Trnski, T., Williams, J. T., Neglia, V., Rapu-Edmunds, C., Planes, S., & Saenz-Agudelo, P. (2018). Evidence of cryptic species in the blennioid *Cirripectes alboapicalis* species complex, with zoogeographic implications for the South Pacific. *Zookeys*, 810, 127–138.
- 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. *Journal of Biogeography*, 46(4), 723–733.
- Duffy, C. A., & Ah Yong, S. T. (2015). Annotated checklist of the marine flora and fauna of the Kermadec Islands Marine Reserve and northern Kermadec Ridge, New Zealand. *Bulletin of the Auckland Museum*, 20, 19–124.
- Fontenelle, J. P., Marques, F., P. L., Kolmann, M. A., & Lovejoy, N. R. (2021). Biogeography of the Neotropical freshwater stingrays (Myliobatiformes: Potamotrygoninae) reveals effects of continent-scale paleogeographic change and drainage evolution. *Journal of Biogeography*, 48(6), 1406–1419.
- Francis, M. P. (1993). Checklist of the coastal fishes of Lord Howe, Norfolk, and Kermadec Islands, Southwest Pacific Ocean. *Pacific Science*, 47(2), 136–170.
- 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., & Duffy, C. A. (2015). New records, checklist and biogeography of Kermadec Islands' coastal fishes. *Bulletin of the Auckland Museum*, 20, 481–495.
- Fricke, R., Eschmeyer, W.N. & Van der Laan, R. (2021). Eschmeyer's catalog of fishes: Genera Species, References. Retrieved 8 November 2021, from <http://researcharchive.calacademy.org/research/ichthyology/catalog/fishcatmain.asp>
- Froese, R. & Pauly, D. (2021). *FishBase*. World wide web electronic publication, version (06/2021). Retrieved 8 November 2021, from <http://www.fishbase.org>
- Goldberg, J., Trewick, S. A., & Paterson, A. M. (2008). Evolution of New Zealand's terrestrial fauna: A review of molecular evidence. *Philosophical Transactions of the Royal Society B: Biological Sciences*, 363(1508), 3319–3334.
- Hastings, R. A., Rutterford, L. A., Freer, J. J., Collins, R. A., Simpson, S. D., & Genner, M. J. (2020). Climate change drives poleward increases and equatorward declines in marine species. *Current Biology*, 30(8), 1572–1577.
- 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. *Journal of Biogeography*, 43(7), 1324–1335.
- Hodge, J. R., & Price, S. A. (2022). Biotic interactions and the future of fishes on coral reefs: The importance of trait-based approaches. *Integrative and Comparative Biology*, 62(6), 1734–1747.
- Hodge, J. R., Van Herwerden, L., & Bellwood, D. R. (2014). Temporal evolution of coral reef fishes: Global patterns and disparity in isolated locations. *Journal of Biogeography*, 41(11), 2115–2127.
- Hoeksema, B. W. (2007). Delineation of the Indo-Malayan Centre of maximum marine biodiversity: The coral triangle. In W. Renema (Ed.), *Biogeography, time, and place: Distributions, barriers, and islands* (pp. 117–178). Springer.

- Hughes, T. P., Bellwood, D. R., & Connolly, S. R. (2002). Biodiversity hotspots, centres of endemism, and the conservation of coral reefs. *Ecology Letters*, 5(6), 775–784.
- Jones, G., Almany, G., Russ, G., Sale, P., Steneck, R., Van Oppen, M., & Willis, B. (2009). Larval retention and connectivity among populations of corals and reef fishes: History, advances and challenges. *Coral Reefs*, 28(2), 307–325.
- Jones, J., & McDougall, I. (1973). Geological history of Norfolk and Philip islands, Southwest Pacific Ocean. *Journal of the Geological Society of Australia*, 20(3), 239–254.
- Kamikuri, S., Motoyama, I., Nishi, H., & Iwai, M. (2009). Evolution of eastern Pacific warm Pool and upwelling processes since the middle Miocene based on analysis of radiolarian assemblages: Response to Indonesian and central American seaways. *Palaeogeography, Palaeoclimatology, Palaeoecology*, 280(3–4), 469–479.
- Knudsen, S. W., Choat, J. H., & Clements, K. D. (2019). The herbivorous fish family Kyphosidae (Teleostei: Perciformes) represents a recent radiation from higher latitudes. *Journal of Biogeography*, 46(9), 2067–2080.
- Kulbicki, M., Parravicini, V., Bellwood, D. R., Arias-González, E., Chabanet, P., Floeter, S. R., Friedlander, A., McPherson, J., Myers, R. E., & Vigliola, L. (2013). Global biogeography of reef fishes: A hierarchical quantitative delineation of regions. *PLoS One*, 8(12), e81847. <https://doi.org/10.1371/journal.pone.0081847>
- Lamm, K. S., & Redelings, B. D. (2009). Reconstructing ancestral ranges in historical biogeography: Properties and prospects. *Journal of Systematics and Evolution*, 47(5), 369–382.
- Landis, M. J., Matzke, N. J., Moore, B. R., & Huelsenbeck, J. P. (2013). Bayesian analysis of biogeography when the number of areas is large. *Systematic Biology*, 62(6), 789–804.
- Lessios, H. A. (2008). The great American schism: Divergence of marine organisms after the rise of the central American isthmus. *Annual Review of Ecology, Evolution, and Systematics*, 39, 63–91.
- Lester, S. E., Ruttenberg, B. I., Gaines, S. D., & Kinlan, B. P. (2007). The relationship between dispersal ability and geographic range size. *Ecology Letters*, 10(8), 745–758.
- Liggins, L., Kilduff, L., Trnski, T., Delrieu-Trottin, E., Carvajal, J. I., Arranz, V., Planes, S., Saenz-Agudelo, P., & Aguirre, J. D. (2022). Morphological and genetic divergence supports peripheral endemism and a recent evolutionary history of *Chrysiptera* demoiselles in the subtropical South Pacific. *Coral Reefs*, 41(3), 797–812.
- 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. *Scientific Reports*, 10(1), 1–7.
- Liu, H., Li, S., Ugolini, A., Momtazi, F., & Hou, Z. (2018). Tethyan closure drove tropical marine biodiversity: Vicariant diversification of intertidal crustaceans. *Journal of Biogeography*, 45(4), 941–951.
- Ludt, W. B., BurrIDGE, C. P., & Chakrabarty, P. (2019). A taxonomic revision of Cheilodactylidae and Latridae (Centrarchiformes: Cirrhitidae) using morphological and genomic characters. *Zootaxa*, 4585(1), 121–141.
- Luiz, O. J., Madin, J. S., Robertson, D. R., Rocha, L. A., Wirtz, P., & Floeter, S. R. (2012). Ecological traits influencing range expansion across large oceanic dispersal barriers: Insights from tropical Atlantic reef fishes. *Proceedings of the Royal Society B: Biological Sciences*, 279(1730), 1033–1040.
- Matzke, N. J. (2013). Probabilistic historical biogeography: New models for founder-event speciation, imperfect detection, and fossils allow improved accuracy and model-testing. *Frontiers of Biogeography*, 5(4), 242–248.
- Matzke, N. J. (2014). Model selection in historical biogeography reveals that founder-event speciation is a crucial process in island clades. *Systematic Biology*, 63(6), 951–970.
- McClymont, E. L., Elmore, A. C., Kender, S., Leng, M. J., Greaves, M., & Elderfield, H. (2016). Pliocene-Pleistocene evolution of sea surface and intermediate water temperatures from the Southwest Pacific. *Paleoceanography*, 31(6), 895–913.
- McDougall, I., Embleton, B., & Stone, D. (1981). Origin and evolution of Lord Howe Island, Southwest Pacific Ocean. *Journal of the Geological Society of Australia*, 28(1–2), 155–176.
- Middleton, I., Aguirre, J. D., Trnski, T., Francis, M., Duffy, C., & Liggins, L. (2021). Introduced alien, range extension or just visiting? Combining citizen science observations and expert knowledge to classify range dynamics of marine fishes. *Diversity and Distributions*, 27(7), 1278–1293.
- Miller, E. C., Hayashi, K. T., Song, D., & Wiens, J. J. (2018). Explaining the ocean's richest biodiversity hotspot and global patterns of fish diversity. *Proceedings of the Royal Society B*, 285(1888), 20181314. <https://doi.org/10.1098/rspb.2018.1314>
- Moore, G., & Chaplin, J. (2014). Contrasting demographic histories in a pair of allopatric, sibling species of fish (Arripidae) from environments with contrasting glacial histories. *Marine Biology*, 161(7), 1543–1555.
- Neall, V. E., & Trewick, S. A. (2008). The age and origin of the Pacific islands: A geological overview. *Philosophical Transactions of the Royal Society B: Biological Sciences*, 363(1508), 3293–3308.
- 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. *New Zealand Journal of Geology and Geophysics*, 44(4), 535–553.
- New Zealand Department of Conservation. (2019). *New Zealand marine protected areas. Gaps analysis*. Department of Conservation, Ministry for the Environment, Ministry for Primary Industries. <https://www.doc.govt.nz/globalassets/documents/conservation/marine-and-coastal/marine-protected-areas/mpa-publications/mpa-gaps-analysis-report-2019.pdf>
- O'Dea, A., Lessios, H. A., Coates, A. G., Eytan, R. I., Restrepo-Moreno, S. A., Cione, A. L., Collins, L. S., De Queiroz, A., Farris, D. W., & Norris, R. D. (2016). Formation of the isthmus of Panama. *Science Advances*, 2(8), e1600883.
- Paradis, E., & Schliep, K. (2019). Ape 5.0: An environment for modern phylogenetics and evolutionary analyses in R. *Bioinformatics*, 35(3), 526–528.
- Pebesma, E. J. (2018). Simple features for R: Standardized support for spatial vector data. *R J*, 10(1), 439–446.
- Piñeros, V. J., Beltrán-López, R. G., Baldwin, C. C., Barraza, E., Espinoza, E., Martínez, J. E., & Domínguez-Domínguez, O. (2019). Diversification of the genus *Apogon* (Lacepède, 1801) (Apogonidae: Perciformes) in the tropical eastern Pacific. *Molecular Phylogenetics and Evolution*, 132, 232–242.
- R Core Team. (2020). *R: A language and environment for statistical computing*. R Foundation for Statistical Computing. <https://www.R-project.org/>
- Ree, R. H., Moore, B. R., Webb, C. O., & Donoghue, M. J. (2005). A likelihood framework for inferring the evolution of geographic range on phylogenetic trees. *Evolution*, 59(11), 2299–2311.
- Ree, R. H., & Smith, S. A. (2008). Maximum likelihood inference of geographic range evolution by dispersal, local extinction, and cladogenesis. *Systematic Biology*, 57(1), 4–14.
- Riginos, C., Douglas, K. E., Jin, Y., Shanahan, D. F., & Tremblay, E. A. (2011). Effects of geography and life history traits on genetic differentiation in benthic marine fishes. *Ecography*, 34(4), 566–575.
- Ronquist, F. (1997). Dispersal-vicariance analysis: A new approach to the quantification of historical biogeography. *Systematic Biology*, 46(1), 195–203.
- Ronquist, F., & Sanmartín, I. (2011). Phylogenetic methods in biogeography. *Annual Review of Ecology, Evolution, and Systematics*, 42, 441–464.



- RStudio Team. (2019). *RStudio: Integrated development for R*. RStudio, Inc. <http://www.rstudio.com/>
- Samayoa, A. P., Struthers, C. D., Trnski, T., Roberts, C. D., & Liggins, L. (2022). Molecular phylogenetics reveals the evolutionary history of marine fishes (Actinopterygii) endemic to the subtropical islands of the Southwest Pacific. *Molecular Phylogenetics and Evolution*, 176, 107584. <https://doi.org/10.1016/j.ympev.2022.107584>
- Sanciango, M. D., Carpenter, K. E., & Betancur-R, R. (2016). Phylogenetic placement of enigmatic percomorph families (Teleostei: Percomorphaceae). *Molecular Phylogenetics and Evolution*, 94, 565–576.
- Sanmartín, I. (2012). Historical biogeography: Evolution in time and space. *Evolution: Education and Outreach*, 5(4), 555–568.
- Santaquiteria, A., Siqueira, A. C., Duarte-Ribeiro, E., Carnevale, G., White, W. T., Pogonoski, J. J., Baldwin, C. C., Ortí, G., Arcila, D., & Ricardo, B.-R. (2021). Phylogenomics and historical biogeography of seahorses, dragonets, goatfishes, and allies (Teleostei: Syngnatharia): Assessing factors driving uncertainty in biogeographic inferences. *Systematic Biology*, 70(6), 1145–1162.
- Siqueira, A. C., Bellwood, D. R., & Cowman, P. F. (2019). Historical biogeography of herbivorous coral reef fishes: The formation of an Atlantic fauna. *Journal of Biogeography*, 46(7), 1611–1624.
- Stebbins, G. L., & Major, J. (1965). Endemism and speciation in the California flora. *Ecological Monographs*, 35(1), 2–35.
- Stuart-Smith, J., Pecl, G., Pender, A., Tracey, S., Villanueva, C., & Smith-Vaniz, W. F. (2018). Southernmost records of two *Seriola* species in an Australian ocean-warming hotspot. *Marine Biodiversity*, 48(3), 1579–1582.
- Stuart-Smith, R. D., Edgar, G. J., & Bates, A. E. (2017). Thermal limits to the geographic distributions of shallow-water marine species. *Nature Ecology & Evolution*, 1(12), 1846–1852.
- Tang, K. L., Stiassny, M. L., Mayden, R. L., & DeSalle, R. (2021). Systematics of Damsel-fishes. *Ichthyology & Herpetology*, 109(1), 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(1), 41–60.
- Tea, Y. K., Van Der Wal, C., Ludt, W. B., Gill, A. C., Lo, N., & Ho, S. Y. (2019). Boomeranging around Australia: Historical biogeography and population genomics of the anti-equatorial fish *Microcanthus strigatus* (Teleostei: Microcanthidae). *Molecular Ecology*, 28(16), 3771–3785.
- Templeton, A. R. (2008). The reality and importance of founder speciation in evolution. *BioEssays*, 30(5), 470–479.
- Thacker, C. E. (2015). Biogeography of goby lineages (Gobiiformes: Gobioidaei): Origin, invasions and extinction throughout the Cenozoic. *Journal of Biogeography*, 42(9), 1615–1625.
- Trnski, T., & de Lange, P. J. (2015). Introduction to the Kermadec biodiversity expedition 2011. *Bulletin of the Auckland Museum*, 20, 1–18.
- Van Dam, M. H., & Matzke, N. J. (2016). Evaluating the influence of connectivity and distance on biogeographical patterns in the southwestern deserts of North America. *Journal of Biogeography*, 43(8), 1514–1532.
- Waters, J. M. (2008). Driven by the west wind drift? A synthesis of southern temperate marine biogeography, with new directions for dispersalism. *Journal of Biogeography*, 35(3), 417–427.
- Waters, J. M., Fraser, C. I., & Hewitt, G. M. (2013). Founder takes all: Density-dependent processes structure biodiversity. *Trends in Ecology & Evolution*, 28(2), 78–85.
- Waters, J. M., & Roy, M. S. (2004). Out of Africa: The slow train to Australasia. *Systematic Biology*, 53(1), 18–24.
- Wernberg, T., Bennett, S., Babcock, R. C., De Bettignies, T., Cure, K., Depczynski, M., Dufois, F., Fromont, J., Fulton, C. J., & Hovey, R. K. (2016). Climate-driven regime shift of a temperate marine ecosystem. *Science*, 353(6295), 169–172.
- Wiens, J. J., & Donoghue, M. J. (2004). Historical biogeography, ecology and species richness. *Trends in Ecology & Evolution*, 19(12), 639–644.
- Zielske, S., Ponder, W. F., & Haase, M. (2017). The enigmatic pattern of long-distance dispersal of minute freshwater gastropods (Caenogastropoda, Truncatelloidea, Tateidae) across the South Pacific. *Journal of Biogeography*, 44(1), 195–206.

### BIOSKETCH

**André P. Samayoa** is focused on understanding the biogeography of marine organisms using phylogenetic tools. This study is part of his PhD work within the Marine Evolutionary Ecology research group (Massey University) headed by Assoc. Prof. Libby Liggins. His doctoral research is aimed at understanding the origin, evolution and spatial patterns of marine fish endemism in the Southwest Pacific applying innovative phylogenetic methods.

**Author contributions:** APS and LL conceived the study; APS, EDT and LL designed the sampling; APS and JDA analysed the data; APS wrote the original draft; APS, EDT, JDA and LL reviewed and edited the manuscript; LL supervised the investigation and provided funding.

### SUPPORTING INFORMATION

Additional supporting information can be found online in the Supporting Information section at the end of this article.

**How to cite this article:** Samayoa, A P., Aguirre, J D., Delrieu-Trotin, E., & Liggins, L. (2023). The origins of marine fishes endemic to subtropical islands of the Southwest Pacific. *Journal of Biogeography*, 50, 1388–1401. <https://doi.org/10.1111/jbi.14579>