Phylogenomic Analysis of Concatenated Ultraconserved Elements Reveals the Recent Evolutionary Radiation of the Fairy Wrasses (Teleostei: Labridae: *Cirrhilabrus*)

YI-KAI TEA[1,2,*], XIN XU[1,3], JOSEPH D. DIBATTISTA[2], NATHAN LO[1], PETER F. COWMAN[4,5,†], AND SIMON Y. W. HO[1,†]

[1]*School of Life and Environmental Sciences, University of Sydney, New South Wales 2006, Australia*

[2]*Australian Museum Research Institute, Australian Museum, 1 William St, Sydney, New South Wales 2010, Australia*

[3]*College of Life Sciences, Hunan Normal University, Changsha, Hunan 410081, China*

[4]*ARC Centre of Excellence for Coral Reef Studies, James Cook University, Townsville, Queensland, Australia*

[5]*Biodiversity and Geosciences Program, Museum of Tropical Queensland, Queensland Museum, Townsville, Queensland 4810, Australia*

[*]*Correspondence to be sent to: School of Life and Environmental Sciences, Edgeworth David Building A11, University of Sydney, New South Wales 2006, Australia;*

Email: yi-kai.tea@sydney.edu.au

[†]*Authors contributed equally to this article.*

Running title: PHYLOGENOMICS OF THE FAIRY WRASSES

© The Author(s) 2021. Published by Oxford University Press, on behalf of the Society of Systematic Biologists. All rights reserved. For Permissions, please email: journals.permissions@oup.com

Downloaded from https://academic.oup.com/sysbio/advance-article/doi/10.1093/sysbio/syab012/6146943 by guest on 01 March 2021

Abstract.— The fairy wrasses (genus *Cirrhilabrus*) are among the most successful of the extant wrasse lineages (Teleostei: Labridae), with their 61 species accounting for nearly 10% of the family. Although species complexes within the genus have been diagnosed on the basis of coloration patterns and synapomorphies, attempts to resolve evolutionary relationships among these groups using molecular and morphological data have largely been unsuccessful. Here we use a phylogenomic approach with a data set comprising 991 ultraconserved elements (UCEs) and mitochondrial *COI* to uncover the evolutionary history and patterns of temporal and spatial diversification of the fairy wrasses. Our analyses of phylogenetic signal suggest that most gene-tree incongruence is caused by estimation error, leading to poor resolution in a summary-coalescent analysis of the data. In contrast, analyses of concatenated sequences are able to resolve the major relationships of *Cirrhilabrus*. We determine the placements of species that were previously regarded as *incertae sedis* and find evidence for the nesting of *Conniella,* an unusual, monotypic genus, within *Cirrhilabrus*. Our relaxed-clock dating analysis indicates that the major divergences within the genus occurred around the Miocene-Pliocene boundary, followed by extensive cladogenesis of species complexes in the Pliocene-Pleistocene. Biogeographic reconstruction suggests that the fairy wrasses emerged within the Coral Triangle, with episodic fluctuations of sea levels during glacial cycles coinciding with shallow divergence events but providing few opportunities for more widespread dispersal. Our study demonstrates both the resolving power and limitations of UCEs across shallow timescales where there is substantial estimation error in individual gene trees.

KEYWORDS

UCEs, molecular dating, gene trees, summary coalescent, concatenation, gene genealogy interrogation, biogeography

Downloaded from https://academic.oup.com/sysbio/advance-article/doi/10.1093/sysbio/syab012/6146943 by guest on 01 March 2021

The wrasses (family Labridae) are among the largest and most successful of the marine teleostean radiations. With over 600 valid species (Parenti and Randall 2000; Fricke et al. 2020), the family is second only to the gobies (family Gobiidae) in marine species richness (Randall et al. 1990; Fricke et al. 2020). Despite their cosmopolitan distribution, it is within coral reefs that the wrasses attain maximum diversity. Here they are known not only for their impressive diversity in coloration and form, but also in functional specialization and ability to exploit novel trophic guilds (Cowman et al. 2009; Evans et al. 2019; Huertas and Bellwood 2020).

While there have been several efforts to resolve the evolutionary relationships among the wrasses (Westneat and Alfaro 2005; Cowman et al. 2009; Cowman and Bellwood 2011), contention persists at almost all levels of the labrid phylogeny, ranging from contested monophyly (Kaufman and Liem 1982; Schultz 1985; Bellwood 1994; Parenti and Randall 2000) to widespread systematic disarray at the intrafamilial level (Westneat and Alfaro 2005). One group in critical need of systematic revaluation is the pseudocheilines, an informal group of wrasses comprising around 100 species from six genera: *Cirrhilabrus*, *Conniella*, *Paracheilinus*, *Pseudocheilinus*, *Pseudocheilinops*, and *Pteragogus*.

Among the pseudocheilines, the fairy wrasses (*Cirrhilabrus*) are of particular systematic interest. With 61 species, they account for nearly 10% of all extant labrids (second only to the polyphyletic *Halichoeres*; see Barber and Bellwood 2005). The genus consists of small, brightly colored fishes found mostly on low-complexity rubble slopes adjacent to coral reefs, at depths of 10 to 250 m. Although sequential hermaphroditism and sexual dimorphism are common among labrids (Warner 1984; Kuiter 2010), it is in the fairy wrasses that these traits are most conspicuous. In contrast to the females, males are very colorful and are known for their exuberant nuptial displays. Their iridescent brilliance and underwater courtship performances have attracted considerable interest, not only among labrid systematists, but

Downloaded from https://academic.oup.com/sysbio/advance-article/doi/10.1093/sysbio/syab012/6146943 by guest on 01 March 2021

also among underwater photographers and aquarium fish collectors (Kuiter 2010; Tea et al. 2016).

The evolutionary relationships of the fairy wrasses have been difficult to resolve, leaving open questions concerning their diversification and evolutionary success. Previous phylogenetic studies have relied on nucleotide sequences of the mitochondrial *COI* gene (Tea et al. 2016; Victor et al. 2016; Tea et al. 2018b; Tea et al. 2020). However, these have been met with limited success, in part due to the extensive sharing of *COI* haplotypes among the fairy wrasses, even between species that are morphologically or geographically separated. Furthermore, morphological characters diagnosing species are often phylogenetically uninformative, with several of these characters being autapomorphic or homoplasious, or showing substantial intraspecific variation (Tea et al. 2016; Tea et al. 2020). These factors have made *Cirrhilabrus* one of the more phylogenetically recalcitrant groups of labrids, with scrutiny of color patterns or diagnoses based on a combination of characters being the only way to separate certain members of the genus.

The genus attains a large but asymmetric geographic distribution spanning the Indo-Pacific, with over half of the known species restricted to the Coral Triangle. For many species of *Cirrhilabrus*, analyses of mitochondrial DNA have inferred relationships that are incongruent with patterns of biogeographic distribution, species-level coloration, and morphological characteristics (Victor and Randall 2014; Allen et al. 2015; Tea et al. 2016). This situation, where high levels of species diversity are not reflected in mitochondrial DNA, is typical of taxa that have either evolved rapidly or are undergoing large-scale introgression (Harris et al. 2018; Hench et al. 2019). However, neither of these hypotheses has been thoroughly investigated, and little is known regarding the spatio-temporal evolution of the fairy wrasses.

Downloaded from https://academic.oup.com/sysbio/advance-article/doi/10.1093/sysbio/syab012/6146943 by guest on 01 March 2021

The challenges in reconstructing the evolutionary history of the fairy wrasses can potentially be overcome by using a phylogenomic approach (Zhang et al. 2019; Cowman et al. 2020), such as the targeted capture of ultraconserved elements (UCEs) (Faircloth et al. 2012). Analyses of UCE data have shown great promise in resolving evolutionary relationships across vertebrates, including fishes at both shallow (DiBattista et al. 2018; Ludt et al. 2019) and deep evolutionary scales (Faircloth et al. 2013; Longo et al. 2017; Alfaro et al. 2018; Faircloth et al. 2020). However, the low information content of individual UCEs and incongruence between the UCE gene trees can pose problems for phylogenomic analyses (Gatesy and Springer 2014; Xi et al. 2015; Meiklejohn et al. 2016) and other downstream analyses that utilize the resulting trees such as estimation of divergence times using molecular dating.

Using an integrative phylogenomic approach that combines UCEs and mitochondrial *COI*, we explore the efficacy of UCEs in resolving the evolutionary history of the fairy wrasses. Our data set includes all pseudocheiline genera, including the monotypic *Pseudocheilinops* and *Conniella*, the latter a rare and enigmatic taxon investigated here in a phylogenetic context for the first time. We compare methods of phylogenomic inference based on the summary coalescent versus concatenation and investigate gene-tree estimation error caused by low phylogenetic information content. Our study provides new insights into the drivers of species diversity in a challenging taxonomic group, demonstrating both the resolving power and limitations of UCEs and their ability to capture the details of recent evolutionary history.

MATERIALS AND METHODS

Taxon Sampling, and DNA Sequencing

Downloaded from https://academic.oup.com/sysbio/advance-article/doi/10.1093/sysbio/syab012/6146943 by guest on 01 March 2021

We sampled 39 of the 61 valid species of *Cirrhilabrus* (64% taxon coverage) for target capture of UCEs. For 28 of these species, along with 10 additional species not included in our UCE sampling, we obtained publicly available sequences of mitochondrial *COI*. In combination, the UCE and mitochondrial data provided 80% taxon coverage across the genus *Cirrhilabrus*. We included the following pseudocheiline genera as outgroup taxa, following the relationships proposed by Westneat (1993): *Paracheilinus*, *Pseudocheilinus*, *Pseudocheilinops*, and *Pteragogus*. We also include the monotypic *Conniella* for complete representation of the pseudocheilines.

Tissue samples were preserved in 80–100% ethanol and stored at −20 °C prior to extraction using a DNeasy Blood and Tissue kit (Qiagen, Hildren, Germany). Library preparation, enrichment, and sequencing were performed by Arbor Biosciences (Ann Arbor, Michigan, USA) following methods described by Quattrini et al. (2018). Following target-capture enrichment, target-enriched libraries were sequenced on one lane of Illumina HiSeq 2500 (100 bp PE reads). A detailed description of taxon sampling and library preparation, as well as a complete list of tissues, sampling locations, and museum voucher information, is presented in the Supplementary Material.

De-multiplexed Illumina reads were processed using PHYLUCE (Faircloth 2016). To investigate the potential impacts of missing data, we assembled two UCE supermatrices based on loci that were represented in at least 75% of all sampled taxa, as well as a more conservative threshold of 95%. The 75%-occupancy matrix consists of 991 UCE loci (330,550 bp) and the 95%-occupancy matrix consists of 47 UCE loci (14,920 bp) from 38 ingroup taxa and 11 outgroup taxa. To test for model violations, which can cause systematic errors in gene-tree inference, we tested for compositional heterogeneity and substitution saturation using PhyloMAd (Duchêne et al. 2018). Further details of filtering, trimming, UCE assembly statistics, and model testing are in the Supplementary Material.

Downloaded from https://academic.oup.com/sysbio/advance-article/doi/10.1093/sysbio/syab012/6146943 by guest on 01 March 2021

Phylogenomic Analyses

In order to improve taxon coverage, we assembled data sets based on concatenated sequences of the 75%-occupancy UCE matrix and mitochondrial *COI* (Supplementary Tables S2–S5). In total, the concatenated sequence alignment had a length of 331,202 bp. To evaluate the effects of missing data, we analyzed four concatenated data sets: 75%-occupancy UCE + *COI* matrix; 75%-occupancy UCE matrix; 95%-occupancy UCE matrix; and the most variable 20% of the UCE loci (variable-loci UCE matrix; see expanded methods in the Supplementary Material). We analyzed all data sets using maximum likelihood in IQ-TREE v1.6.12 (Nguyen et al. 2015) and Bayesian inference in ExaBayes (Aberer et al. 2014) (details provided in the Supplementary Material).

To account for gene-tree incongruence among loci, we analyzed the UCE data set using the summary-coalescent method in ASTRAL-III (Zhang et al. 2018). We inferred the species tree under a range of conditions (see Supplementary Material), including ASTRAL analyses of all 991 UCE gene trees inferred from the 75%-occupancy matrix. We then analyzed the data after contracting branches in gene trees with less than 10%, 20%, and 30% bootstrap support. We performed ASTRAL analyses of a combined data set that comprised gene trees for all 991 UCEs + mitochondrial *COI*, and of the variable-loci UCE matrix. Topological distances between the various trees inferred using summary-coalescent and concatenation approaches were calculated using the normalized Robinson-Foulds metric using the R package phangorn (Schliep 2011). We also examined concordance between the gene trees and the concatenation-based species tree using PhyParts (Smith et al. 2015).

We conducted a gene genealogy interrogation (GGI; Arcila et al. 2017), to test topological hypotheses for five species complexes in *Cirrhilabrus* whose relationships were

Downloaded from https://academic.oup.com/sysbio/advance-article/doi/10.1093/sysbio/syab012/6146943 by guest on 01 March 2021

consistently poorly resolved in all summary-coalescent analyses: the *rubriventralis*, *filamentosus*, *cyanopleura*, *punctatus*, and *temminckii* complexes. In this analysis, we examined all 15 possible unrooted topologies for these five groups, using constrained maximum-likelihood searches for each of the 915 UCE loci (76 UCE loci were excluded from this analysis because of missing data). Site likelihood scores for all of the trees were compared using the approximately unbiased test (Shimodaira, 2002). A full description of the gene genealogy interrogation is given in the Supplementary Material.

Phylogenomic Dating and Biogeographic Analyses

We used MCMCTree in the PAML package (Yang 2007) to perform a Bayesian dating analysis of the concatenated 75%-occupancy UCE + *COI* matrix. Because our aim was to infer divergence times for closely related species, we chose not to employ deep fossil-based calibrations that would require the inclusion of distantly related taxa. Available labrid fossils for calibration include species from several distantly related lineages (summarized by Cowman & Bellwood, 2011; Bellwood et al. 2019), for which no UCE data are available. Instead, we employed secondary calibrations from a well resolved, time-calibrated labrid phylogeny (Cowman et al. 2009). This involved two node calibrations: crown age of the pseudocheilines (*Pteragogus* versus all other taxa, 25.7–49.8 Ma); and the divergence between *Pseudocheilinus* and *Paracheilinus* + *Cirrhilabrus* (10–34 Ma). These were implemented as uniform priors with soft bounds (Yang and Rannala 2006).

Ancestral biogeographic ranges were estimated for the time-calibrated phylogeny using the R package BioGeoBEARS (Matzke 2013). Biogeographic provinces were delineated following those described by Kulbicki et al. (2013), with modifications based on the contemporary distributions of *Cirrhilabrus*. We used the corrected Akaike information criterion to compare three biogeographic models: dispersal, extinction and cladogenesis

Downloaded from https://academic.oup.com/sysbio/advance-article/doi/10.1093/sysbio/syab012/6146943 by guest on 01 March 2021

model (DEC; Ree and Smith 2008); dispersal-vicariance model (DIVA-like; Ronquist 1997); and BayArea-like model (Landis et al. 2013). Details of the phylogenomic dating and biogeographic analyses are given in the Supplementary Material.

RESULTS

Phylogenomic Analyses: Concatenation vs Coalescence

Maximum-likelihood analysis of the concatenated 75%-occupancy UCE + *COI* matrix yielded trees with well-supported nodes along the backbone of *Cirrhilabrus*, with monophyly of all species complexes (Fig. 1a). Except for taxa *incertae sedis*, the inferred relationships at deeper nodes were consistent with those previously inferred on the basis of mitochondrial DNA (Fig. 1b; Tea et al. 2020). This tree topology was congruent to the maximum-likelihood and Bayesian trees inferred from the 75%-occupancy UCE matrix (Supplementary Fig. S1a and S1d), but with the additional *C. condei* complex placed as the sister lineage to the *C. filamentosus* complex. Maximum-likelihood and Bayesian trees inferred using the 95%-occupancy UCE matrix and the variable-loci UCE matrix were less well resolved, with the *C. filamentosus*, *C. rubriventralis*, *C. cyanopleura*, and *C. temminckii* complexes not recovered as monophyletic (Supplementary Fig. S1b, S1c, S1e, and S1f). Analyses of concatenated UCE loci were able to resolve the relationships among species within all respective complexes, an outcome that was not achievable using *COI* sequences alone (Fig. 1c).

In contrast with the results of our analyses of concatenated UCE loci, phylogenomic analyses using the summary-coalescent approach in ASTRAL yielded poorly resolved trees. Across these analyses, monophyly was only supported for the *C. exquisitus*, *C. scottorum*, *C. lunatus*, *C. bathyphilus*, and *C. lubbocki* complexes (Supplementary Fig. S2a–f). The

Downloaded from https://academic.oup.com/sysbio/advance-article/doi/10.1093/sysbio/syab012/6146943 by guest on 01 March 2021

relationships among polyphyletic complexes were not improved with contraction of branches with low bootstrap support in the gene trees (Supplementary Fig. S2b–2d), when only a subset of most variable loci were used (Supplementary Fig. S2e), or with the inclusion of the *COI* gene tree (Supplementary Fig. S2f).

While the number of parsimony-informative sites for each UCE locus ranged from zero to 54, most loci contained 20 or fewer informative sites (Fig 2a). Only five UCE loci had more than 40 informative sites, and only four loci had the minimum number of informative sites needed to resolve relationships of the species tree given that the fully resolved tree has 43 internal branches. We found no evidence of model violations among the UCE loci. Trees inferred using the summary-coalescent approach had topologies that were generally dissimilar to each other and to the trees inferred from concatenated data (Fig. 2b). This is likely to be due to most UCE gene trees having low average bootstrap values (~45–55%; Fig. 2c), with no gene tree having an average bootstrap value exceeding 70%, and with nearly all having topologies discordant with the underlying species tree (Supplementary Fig. S3a). In contrast, the topologies of the trees inferred from concatenated data were more similar to each other, even between those inferred using maximum likelihood and Bayesian inference, and between trees from which we pruned species represented by only *COI* sequences.

Our GGI analysis of the five most problematic species complexes showed that topological hypothesis H_0 had the highest probability (185 tests in favor) and statistical support ($P < 0.05$). This result corroborates the relationships inferred using the concatenated UCE data set. All of the alternative hypotheses (H_1–H_{14}) received negligible support (Fig. 2d).

Downloaded from https://academic.oup.com/sysbio/advance-article/doi/10.1093/sysbio/syab012/6146943 by guest on 01 March 2021

Phylogenetic Relationships

Our maximum-likelihood and Bayesian analyses of the concatenated 75%-occupancy UCE + *COI* matrix were able to resolve relationships among the fairy wrasses (Fig. 1a and 3a). All species complexes were recovered as monophyletic and formed seven major lineages, which we designate as Lineages A–G. In all analyses, *Conniella* was inferred to be deeply nested within *Cirrhilabrus*, in particular within the *C. jordani* complex. The successive sister genera of *Cirrhilabrus* are *Paracheilinus*, *Pseudocheilinus*, and *Pseudocheilinops*. *Pteragogus* was found to be the sister lineage to all remaining pseudocheilines.

The maximum-likelihood and Bayesian trees were congruent across almost all nodes, except for relationships in Lineage G, where placement of the *C. condei* complex was not resolved with confidence (Fig. 3b). Relationships of the previously *incertae sedis C. claire*, *C. katoi*, and *C. laboutei* were resolved with high node support. Both *C. claire* and *C. katoi* were placed within the *C. lineatus* complex, whereas *C. laboutei* was found to occupy a monospecific sister lineage to a group comprising Lineages D–G.

We resolved the phylogenetic relationships of several species for which mitochondrial DNA showed little to no variation (Fig. 1b), in particular those from the *C. rubriventralis*, *C. filamentosus*, *C. cyanopleura*, and *C. lunatus* complexes. However, not all nodes within these complexes were well supported, in particular the positions of *C. hygroxerus* in the *C. rubriventralis* complex, *C. luteovittatus* + *C. ryukyuensis* in the *C. cyanopleura* complex, and *C. brunneus* + *C. squirei* in the *C. lunatus* complex. All analyses found strong support for deep divergences between populations within the *C. exquisitus* complex, which are separated into Indian Ocean and Pacific Ocean clades (Fig. 1a and 3a).

Downloaded from https://academic.oup.com/sysbio/advance-article/doi/10.1093/sysbio/syab012/6146943 by guest on 01 March 2021

Timescale of Diversification and Historical Biogeography

Our Bayesian relaxed-clock analysis places the crown age of the pseudocheilines at 44.2 Ma (95% credibility interval [CI] 33.4–51.4 Ma). The split between *Paracheilinus* and *Cirrhilabrus* was estimated at 12.2 Ma (95% CI 8.2–16.3 Ma), in the mid to late Miocene (Supplementary Table S1). The crown age of *Cirrhilabrus* was estimated at 9.1 Ma (95% CI 5.9–13 Ma), with several major divergences occurring between 5.7 and 7.4 Ma (Lineages B–E). Together, these lineages account for more than a third of all *Cirrhilabrus* species. Lineages F and G account for nearly all remaining species of *Cirrhilabrus*, diverging from each other approximately 3.6 Ma (2.1–5.6 Ma). These deeper divergences within *Cirrhilabrus* were succeeded by a set of cladogenetic events leading to several species complexes, occurring between 3.3 and 4.7 Ma around the Miocene-Pliocene boundary (Supplementary Table S1). The topological differences in Lineage G between the Bayesian and maximum-likelihood trees had a small effect on most of the divergence-time estimates across the tree, including the crown age of the pseudocheilines (mean age of 44.2 versus 43.8 Ma; Supplementary Fig. S3b–c).

In the BioGeoBEARS analysis, we found that the BayArea-like model had the highest support based on the corrected Akaike information criterion (BayArea-like: 370.7; DEC: 374.3; DIVA-like: 376.0). Under this model, the most recent common ancestor of *Cirrhilabrus* occurred in an area comprising the central Indo-Pacific and Melanesia (Fig. 3c). These areas were inferred for all nodes along the backbone of the tree, suggesting a common area of origin for many of the major lineages of *Cirrhilabrus*. We resolve biogeographic histories for several taxa occurring in peripheral regions, including the western Indian Ocean, Hawaii, and the southeast Pacific Ocean. Our reconstruction of ancestral locations supports at least five independent invasions of the Indian Ocean by the *C. rubriventralis*, *C. filamentosus*, *C. cyanopleura*, *C. jordani*, and *C. exquisitus* complexes, via the boundaries of

Downloaded from https://academic.oup.com/sysbio/advance-article/doi/10.1093/sysbio/syab012/6146943 by guest on 01 March 2021

the central Indo-Pacific and Indian Ocean. In the *C. rubriventralis* complex, this occurred around 3 Ma (95% credibility interval 1.8–4.5 Ma) and was followed by expansion into the western Indian Ocean and dispersal into the Red Sea (*C. rubriventralis*).

The widespread *C. exquisitus* is also represented by populations occurring in the western Indian Ocean, but the ancestral area of origin for this complex remains unresolved due to a lack of representation from the eastern Indian Ocean. For the same reason, the biogeographic history of the *C. jordani* complex remains unresolved. The ancestral location for the Hawaiian endemic *C. jordani* occurred in an area comprising several regions of the western and southwestern Pacific Ocean. From these regions, southern Japan was inferred as the most likely ancestral area leading to cladogenesis and subsequent occupancy of Hawaii. The ancestor of the southeastern Pacific *C. claire* was placed in Melanesia.

DISCUSSION

Phylogenomic Analyses

Our phylogenomic analyses of concatenated UCE loci have resolved major relationships among the fairy wrasses, revealing that the genus has an evolutionary history originating in the Pacific Ocean, but involving multiple invasions of the Indian Ocean. In contrast, we found that summary-coalescent analyses of our data set yielded poor phylogenetic resolution, with several species complexes found to be non-monophyletic. This result is consistent with those of several previous analyses of UCE data that have compared concatenation and summary-coalescent approaches, particularly in groups of organisms that underwent recent divergences (Meiklejohn et al. 2016; Longo et al. 2017; Ochoa et al. 2020).

In contrast with some previous findings (Longo et al. 2017), we found that contraction of poorly supported branches in the gene trees and targeted analyses focusing on informative

Downloaded from https://academic.oup.com/sysbio/advance-article/doi/10.1093/sysbio/syab012/6146943 by guest on 01 March 2021

loci did little to improve resolution in our species trees. This outcome is likely to be due to poor resolution in individual gene trees (Xi et al. 2015). Analysis of these data using a summary-coalescent approach led to trees that were topologically dissimilar to each other, even when gene trees with low information content were excluded. In contrast, trees inferred using a concatenation approach had relatively more consistent topologies, even between those obtained using reduced data sets and different methods of phylogenetic inference.

Coalescent theory predicts that lineages evolving under a combination of short internal branches and large effective population sizes have phylogenetic histories that are prone to incomplete lineage sorting (Degnan and Rosenberg, 2009; Arcila et al. 2017). For taxa that diverged rapidly or recently, these conditions can lead to gene trees that are discordant with the underlying species tree. To investigate this possibility, we performed a GGI focusing on the relationships among five of the most poorly resolved species complexes in *Cirrhilabrus*, inferred on the basis of summary-coalescent approaches. The GGI results support the topology that is congruent with the relationships inferred using the concatenated UCE data set, as well as those of previous studies (Tea et al., 2018b).

On an individual level, the topologies of gene trees inferred from UCE data do not match that of the concatenated species tree, with nearly all of the relationships found to be incongruent. However, our tests of model adequacy did not detect any model violations, which can lead to errors in estimates of gene trees and species trees (Jeffroy et al. 2006; Whelan et al. 2015). Our results suggest that gene-tree incongruence is driven by a lack of phylogenetic information in individual UCE loci, supporting previous assertions that a concatenation approach is preferred over summary-coalescent approaches when estimation error is the dominant cause of gene-tree incongruence (Gatesy and Springer 2014; Bryant and Hahn 2020; but see Jiang et al. 2020). Nonetheless, we cannot exclude the possibility that our inference based on concatenated data has been misled by the effects of incomplete lineage

Downloaded from https://academic.oup.com/sysbio/advance-article/doi/10.1093/sysbio/syab012/6146943 by guest on 01 March 2021

sorting and interspecific gene flow. This should be looked for between closely related species within species complexes, given that the monophyly of most *Cirrhilabrus* complexes are well supported on the basis of morphology, juvenile and terminal male coloration patterns (Kuiter, 2010), and mitochondrial DNA. While incongruence between species trees and gene trees in rapidly diversifying taxa is not uncommon due to short internal branches, these same evolutionary histories can complicate the inference of gene trees by not allowing enough time for nucleotide substitutions on these branches. For this reason, we urge caution in interpreting whether coalescent or concatenation approaches might be more biased as a result of these factors.

Phylogenetic Relationships Among the Fairy Wrasses

Our phylogenomic analyses of UCE loci have resolved some relationships that could not be determined on the basis of mitochondrial and/or morphological evidence. In particular, we find mutual monophyly of all species complexes and infer strongly supported relationships among species within the *C. rubriventralis*, *C. filamentosus*, *C. cyanopleura*, and *C. lunatus* complexes. We note, however, that taxonomic representation in these complexes was incomplete. For example, the putative sister species of *C. hygroxerus* from the *C. rubriventralis* complex, *C. humanni* and *C. morrisoni*, were absent in our data set. Similarly, *C. johnsoni* from the *C. lunatus* complex, and *C. randalli*, and *C. aurantidorsalis* from the *C. cyanopleura* complex were not sampled in our study.

Our maximum-likelihood and Bayesian phylogenomic analyses yielded disagreements in some inferred relationships within Lineage G, in particular the placement of the *C. condei* complex. One possible explanation is the limited representation of the *C. condei* complex, which includes three valid species, *C. condei*, *C. marinda*, and *C. walshi*, of which the last two are rare and obtainable only with great logistical difficulty. Only *C. condei*

Downloaded from https://academic.oup.com/sysbio/advance-article/doi/10.1093/sysbio/syab012/6146943 by guest on 01 March 2021

was included in our analyses, and only on the basis of mitochondrial *COI*; we justify its inclusion in order to have representation of all species complexes.

Members of Lineage G are distinctive in having one row of cheek scales versus the usual two in all other species of *Cirrhilabrus.* This apomorphy was one of two characters justifying the erection of the genus *Cirrhilabrichthys* (for *C. filamentosus*) by Klausewitz (1976), with the other character being prolongation of the three posteriormost dorsal-fin spines into a single filament. Dorsal-fin prolongation is unusual in *Cirrhilabrus*, occurring almost exclusively within Lineage G, specifically the *C. filamentosus* and *C. rubriventralis* complexes (also independently acquired in *C. nahackyi*).

Although dorsal-fin prolongation is not synapomorphic for either the *C. filamentosus* or *C. rubriventralis* complex, the occurrence of these characters at the complete exclusion of the *C. condei* complex suggests a sister relationship between the two on the basis of morphology (Tea et al. 2018a). Furthermore, the *C. rubriventralis* and *C. filamentosus* complexes have allopatric distributions in the Indian and Pacific Ocean basins, reflecting a common evolutionary history shared by many Indo-Pacific reef fishes (Randall 1988; Gaither and Rocha 2013; Liu et al. 2014; Ludt and Rocha 2014). Based on these additional lines of morphological and biogeographic evidence, we consider the relationships between these complexes inferred in our Bayesian analysis as having better support.

The placements of *C. claire*, *C. katoi*, and *C. laboutei* could not be previously resolved with confidence in analyses of mitochondrial *COI* (see Tea et al. 2017; 2018b; 2020). Here we place *C. claire* and *C. katoi* in the *C. lineatus* complex, but with coloration and morphological characters that depart from other members the group (see Tea et al. 2018b; 2020). The unusual *C. laboutei* occupies a monospecific lineage and is not found to be closely related to any of the extant species of *Cirrhilabrus*. On the basis of morphology, *C. laboutei* departs from all other species of *Cirrhilabrus* in possessing greatly elongate first and

Downloaded from https://academic.oup.com/sysbio/advance-article/doi/10.1093/sysbio/syab012/6146943 by guest on 01 March 2021

second anal-fin spines (Fig. 4a), a character that within the pseudocheilines is found only in the genus *Pseudocheilinus*.

Conniella apterygia*: The Mutant Wrasse*

We show in all analyses that *Conniella* is deeply nested within *Cirrhilabrus*. This enigmatic genus, with only one species (*Conniella apterygia* Allen 1983), was erected on the basis of lacking pelvic fins and associated osteology (Fig. 4b). In all other aspects of gross morphology, the genus is inseparable from *Cirrhilabrus* (Allen 1983). On the basis of coloration patterns, *Conniella* shows a remarkable similarity to *C. earlei* (Fig. 4c). Both species possess scales, along with several osseous elements, that retain and/or develop purple pigmentation even in preservation. This potentially plesiomorphic character is restricted to several species in the *C. jordani* complex (Tea et al. 2018b, 2019). Although these two species were not grouped together in any of our inferred trees, we note that some relationships within the complex were not confidently resolved, and that several species were not sampled for UCEs (including *C. earlei*).

The loss of pelvic fins and associated osteology is rare in percoid fishes (Yamanoue et al. 2010) and, within the Labridae, is reported in only one other species, *Siphonognathus argyrophanes* Cuvier 1829. Our phylogenomic analyses suggests that the loss of pelvic elements in *Conniella* is apomorphic. A parsimonious explanation is that character was lost through genetic drift; in extreme endemics like *Conniella* with small effective population sizes, naturally occurring deleterious mutations can become fixed by chance (Price & Hadfield, 2014; Sonsthagen et al. 2017). Indeed, fin loss has been reported in other species of *Cirrhilabrus* (Figs 4d & 4e), indicating that loss of fins can occur naturally, albeit as a rare congenital defect. In order to preserve the monophyly of *Cirrhilabrus*, we recommend

Downloaded from https://academic.oup.com/sysbio/advance-article/doi/10.1093/sysbio/syab012/6146943 by guest on 01 March 2021

Conniella Allen 1983 be recognized as a junior synonym of *Cirrhilabrus* Temminck & Schlegel 1845.

Evolutionary Timescale and Biogeographic History

Our dating analyses place the crown age of the pseudocheilines at 44.2 Ma (95% CI 33.4–51.4 Ma), an estimate that is concordant with those of several labrid studies using independent data sets (44.1 Ma [95% CI 36.9–51.4 Ma] in Cowman and Bellwood 2011; also see Alfaro et al. 2009). We note, however, that no study has previously included all component pseudocheiline genera, and that the different relationships inferred for the pseudocheilines by Cowman and Bellwood (2011) are likely to have been a result of incomplete sampling and missing data.

The fairy wrasses underwent a period of diversification that involved multiple invasions of the Indian Ocean. The genus has attained a geographic distribution spanning the entire Indo-Pacific, but with a disproportionate number of species occurring within the Coral Triangle. In contrast, species diversity decreases peripherally, with only eight species in the western Indian Ocean (including two species in the Red Sea), one species in Hawaii, and three species in the southeast Pacific Ocean. Our divergence-time estimates suggest the occurrence of two major periods of diversification: divergences among major *Cirrhilabrus* lineages during the late Miocene-Pliocene, followed by cladogenesis of species complexes during the Pliocene-Pleistocene. The former coincides with patterns of vicariance in other coral reef groups (including the Labridae) in the Indo-Pacific (4.1–7 Ma; Cowman and Bellwood 2013). For most of the *Cirrhilabrus* complexes, diversification occurred around the Pliocene-Pleistocene boundary, which marked the start of extensive glacial cycles (Ludt and Rocha 2014).

Downloaded from https://academic.oup.com/sysbio/advance-article/doi/10.1093/sysbio/syab012/6146943 by guest on 01 March 2021

The combination of glacial cycles, the widespread dispersal capabilities of marine fishes (Quenouille et al. 2011; Gaither et al. 2015), and the permeability of ephemeral barriers on either side of the Indo-Australian Archipelago has led to shared faunal affinities, provincial endemism, and widespread distributions of species exhibited today, particularly for species in the central Indo-Pacific and Coral Triangle (Bellwood and Wainwright 2002; Connolly et al. 2003; Briggs, 2005; Hobbs & Allen, 2014; DiBattista et al. 2016, 2018). Our biogeographic analyses suggest that *Cirrhilabrus* and its major lineages originated in this region, in particular the central Indo-Pacific and Melanesia.

The Indian Ocean was colonized independently by five *Cirrhilabrus* complexes. Of these, only two are represented in peripheral regions of the western Indian Ocean (i.e., the Red Sea). Although the Red Sea began forming in the late Oligocene, environments conducive to coral reefs emerged only within the last 5 Myr (Siddall et al. 2003; Bosworth et al. 2005). Our temporal and biogeographic estimates for divergences of the Red Sea endemic *C. rubriventralis* (680 ka; 95% CI 20 ka – 1.6 Ma) and *C. blatteus* (840 ka; 95% CI 11 ka – 1.6 Ma) from their respective sister taxa are consistent with these dates, suggesting the possibility of dispersal from the western Indian Ocean.

The process of dispersal-mediated range expansion could also explain the regional endemism of the Hawaiian endemic *C. jordani*. Our biogeographic reconstruction placed the ancestor of *C. jordani* within several regions of the western Pacific Ocean, with southern Japan being the most likely ancestral area. The divergence of *C. jordani* from its sister lineage was estimated at 2.3 Ma (95% CI 1.1–3 Ma), which coincides with substantial lowering of sea levels in the Pleistocene. During this period, contemporary atolls would have been raised into table islands and exposed reefs, providing stepping-stones for the widespread dispersal of pelagic organisms (Kobayashi, 2006; Eble et al. 2009). Colonization of Hawaii from southern Japan could have occurred via the numerous archipelagos connecting both

Downloaded from https://academic.oup.com/sysbio/advance-article/doi/10.1093/sysbio/syab012/6146943 by guest on 01 March 2021

regions, notably Midway Atoll (Randall et al. 1993). This region hosts a mixture of endemic Hawaiian and Japanese fauna, with recent studies suggesting that periodic connectivity occurred through to the last glacial cycle (see Tea et al. 2019). The alternative explanation for the occurrence of endemic species in peripheral regions is relictualism through extinction of intermediate populations. However, given the uneven distribution of *Cirrhilabrus* throughout its range and the recent evolutionary divergences for many species, isolation by dispersal appears to be a more parsimonious explanation.

The colonization history of the Indian Ocean is less clear for species in the *C. exquisitus* and *C. jordani* complexes. The former includes a single, highly polymorphic and widespread species spanning the entire range of the genus (but is curiously absent from the Red Sea and Hawaii; Supplementary Fig. S4). Our results suggest that divergence between the Indian and Pacific Ocean lineages of *C. exquisitus* occurred around 3.1 Ma (95% CI 1.6–5 Ma), during a period when fluctuating sea levels are likely to have cut off exchange between the eastern Indian Ocean and the Indo-Australian Archipelago (Ludt and Rocha 2014). This distinction between the Indian and Pacific Ocean lineages of *C. exquisitus* is supported by both UCE data and mitochondrial *COI* (Supplementary Fig. S4). However, due to the lack of samples from the eastern Indian Ocean and other regions of the Pacific Ocean, we were unable to confidently determine the ancestral area of origin for this species complex, nor infer the processes leading to diversification within the two major lineages in each ocean basin.

Identifying the biogeographic drivers of speciation for other *Cirrhilabrus* complexes remains problematic, particularly in the *C. cyanopleura* complex, where contemporary distribution patterns include extensive sympatry between species. While evidence from both empirical and theoretical studies indicate the possibility of sympatric speciation in coral reefs, the process remains contentious (Rocha & Bowen, 2008). Our results explain the tight

Downloaded from https://academic.oup.com/sysbio/advance-article/doi/10.1093/sysbio/syab012/6146943 by guest on 01 March 2021

clustering of *Cirrhilabrus* species within the Coral Triangle through late Miocene diversification of major lineages, followed by extensive cladogenesis of species complexes during the late Pliocene to Pleistocene. While this period was characterized by glacial cycles that promoted diversification in coral reef fishes, the formation of barriers was often ephemeral across short spatial and temporal scales (particularly around the Indo-Australian Archipelago; Rocha & Bowen, 2008), with few opportunities for prolonged and continuous dispersal into wider, more peripheral regions. In short, the repeated fluctuations of sea levels during the late Pliocene-Pleistocene might have acted as a species pump, leading to diversification on a localized scale through episodic barrier formation, but with few opportunities for more widespread colonization.

CONCLUSIONS

Our study demonstrates both the capacity and limitations of genome-wide UCEs in resolving phylogenetic relationships in a recently diverged group characterized by few robust morphological characters. We show that a summary-coalescent approach is unable to resolve the evolutionary relationships of *Cirrhilabrus*, probably because of poorly resolved gene trees. In contrast, analysis of concatenated sequence data appears to resolve the relationships among species even across shallow timescales, but we cannot rule out errors due to introgression or incomplete lineage sorting. Importantly, our investigation highlights the importance of scrutinizing the information content of loci used in phylogenomic analyses.

Our phylogenomic analyses of the fairy wrasses have provided novel insights into the evolutionary history, timing of divergences, and historical biogeography of one of the largest and most charismatic genera of labrid fishes. Although the emergence of the fairy wrasses coincided with patterns of Miocene vicariance typical of other groups of coral reef fishes, much of their diversification took place in the Pliocene-Pleistocene. Further investigation

Downloaded from https://academic.oup.com/sysbio/advance-article/doi/10.1093/sysbio/syab012/6146943 by guest on 01 March 2021

with an expanded data set will provide additional insights into the extent of incomplete lineage sorting and introgression among the fairy wrasses, leading to a more confident resolution of their evolutionary history.

Author Contributions

YKT, PFC, and SYWH devised the study, with input from NL. JDD and PFC provided tissue samples. YKT and XX conducted the analyses. YKT wrote the manuscript. All authors approved the final version of the manuscript and agree to be held accountable for the content. All authors declare no conflict of interest.

Supplementary Material

Detailed methods, results, and figures for all taxonomic sampling and phylogenomic analyses reported in this study are available in the Supplementary Material. Data sets, alignment files, tree files, custom scripts, and output files are provided in the Dryad repository (https://doi.org/10.5061/dryad.4tmpg4f6q).

Conflict Of Interest

All authors declare no conflict of interest.

Funding

YKT was funded by a Research Training Program Scholarship from the Australian Government and by an Australian Museum Research Institute Postgraduate Award. NL and SYWH were funded by the Australian Research Council (FT160100463 and FT160100167).

Downloaded from https://academic.oup.com/sysbio/advance-article/doi/10.1093/sysbio/syab012/6146943 by guest on 01 March 2021

PFC was funded by an ARC DECRA Fellowship (DE170100516) and the ARC Centre of Excellence Program (CE140100020).

ACKNOWLEDGEMENTS

We thank Amanda Hay, Sally Reader, and Kerryn Parkinson of the Australian Museum for curatorial support. Gerald R. Allen (Western Australian Museum), David R. Bellwood (James Cook University), Michael P. Hammer (Museum and Art Gallery of the Northern Territory), Luiz A. Rocha (California Academy of Sciences), Kelvin Lim (Lee Kong Chian Natural History Museum), Benjamin C. Victor, and Fenton Walsh provided valuable tissue samples. All genetic materials obtained from museums were procured in accordance with guidelines and approval from the respective institutions. Tissue curation and DNA extractions were performed by Martina de Freitas Prazeres, Hillary Smith, and Julia Yun-Hsuan Hung. Benjamin W. Frable and Joseph Rowlett provided valuable discourse on the subject matter. Charles P. Foster and Kevin Love provided custom scripts for data handling. We dedicate this study to the late John Ernest Randall, who was a fierce and pioneering force in the modern age of ichthyology, and to Hiroyuki Tanaka, who has contributed significantly towards our understanding of the genus. The paper benefited from comments by Bryan Carstens, Rayna Bell, Prosanta Chakrabarty, and four anonymous reviewers.

Downloaded from https://academic.oup.com/sysbio/advance-article/doi/10.1093/sysbio/syab012/6146943 by guest on 01 March 2021

REFERENCES

Aberer A.J., Kobert K., Stamatakis A. 2014. ExaBayes: massively parallel Bayesian tree inference for the whole-genome era. Mol. Biol. Evol. 31: 2553–2556.

Alfaro M.E., Brock C.D., Banbury B.L., Wainwright P.C. 2009. Does evolutionary innovation in pharyngeal jaws lead to rapid lineage diversification in labrid fishes? BMC Evol. Biol. 9: 255.

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

Allen G.R. 1983. A new genus and species of wrasse (Pisces: Labridae) from Rowley Shoals, Western Australia. Rev. Fr. d'Aquariol. 10: 43–46.

Allen G.R., Erdmann M.V., Dailami, M. 2015. *Cirrhilabrus marinda*, a new species of wrasse (Pisces: Labridae) from eastern Indonesia, Papua New Guinea, and Vanuatu. J. Ocean Sci. Foun. 15: 1–15.

Barber P.H., Bellwood, D.R. 2005. Biodiversity hotspots: evolutionary origins of biodiversity in wrasses (*Halichoeres*: Labridae) in the Indo-Pacific and new world tropics. Mol. Phylogenet. Evol. 35: 235–253.

Bellwood, D.R. 1994. A phylogenetic study of the parrotfish family Scaridae (Pisces: Labroidea), with a revision of genera. Rec. Aus. Mus. Supp. 20: 1–86.

Bellwood, D.R., Wainwright, P.C. 2002. The history and biogeography of fishes on coral reefs. In: P.S. Sale (ed.) Coral reef fishes: dynamics and diversity in a complex ecosystem. Elsevier, San Diego, USA. Pp. 5–32.

Bosworth, W., Huchon, P., McClay, K. 2005. The Red Sea and Gulf of Aden basins. J. Afr. Earth Sci. 43: 334–378.

Downloaded from https://academic.oup.com/sysbio/advance-article/doi/10.1093/sysbio/syab012/6146943 by guest on 01 March 2021

Briggs, J.C. 2005. The marine East Indies: Diversity and speciation. J. Biogeogr. 32: 1517–1522.

Bryant D., Hahn M. 2020. The concatenation question. In: Scornavacca C., Delsuc F., Galtier N. (eds) Phylogenetics in the genomic era. No commercial publishers | Authors' open access book.

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

Cowman, P.F., Bellwood, D.R. 2011. Coral reefs as drivers of cladogenesis: expanding coral reefs, cryptic extinction events, and the development of biodiversity hotspots. J. Evol. Biol. 23: 2543–2562.

Cowman, P.F., Bellwood, D.R. 2013. Vicariance across major marine biogeographic barriers: temporal concordance and the relative intensity of hard versus soft barriers. Proc. R. Soc. B. 280: 20131514.

Cowman, P.F., Bellwood, D.R., van Herwerden, L. 2009. Dating the evolutionary origins of wrasse lineages (Labridae) and the rise of trophic novelty on coral reefs. Mol. Phylogenet. Evol. 52: 621–631.

Cowman, P.F., Quattrini, A.M., Bridge, T.C., Watkins-Colwell, G.J., Grimblat, M., Roberts, T.E., McFadden, C., Miller, D.J., Baird, A.H. 2020. An enhanced target-enrichment bait set for Hexacorallia provides phylogenomic resolution of the staghorn corals (Acroporidae) and close relatives. Mol. Phylogenet. Evol. 153: 106944.

DiBattista, J.D., Alfaro, M.E., Sorenson, L., Choat, J.H., Hobbs, J.P.A., Sinclair-Taylor, T.H., Rocha, L.A., Chang, J., Luiz, O.K., Cowman, P.F., Friedman, M., Berumen, M.L. 2018. Ice ages and butterflyfishes: phylogenomics elucidates the ecological and evolutionary history of reef fishes in an endemism hotspot. Ecol. Evol. 8: 10989–11008.

Downloaded from https://academic.oup.com/sysbio/advance-article/doi/10.1093/sysbio/syab012/6146943 by guest on 01 March 2021

DiBattista, J.D., Whitney, J., Craig, M.T., Hobbs, J.P.A., Rocha, L.A, Feldheim, K.A., Berumen, M.L., Bowen, B.W. 2016. Surgeons and suture zones: hybridization among four surgeonfish species in the Indo-Pacific with variable evolutionary outcomes. Mol. Phylogenet. Evol. 101: 203–215.

Duchêne, D., Duchêne, S., Ho, S.Y.W. 2018. PhyloMAd: efficient assessment of phylogenomic model adequacy. Bioinformatics 34: 2300–2301.

Eble, J.A., Toonen, R.J., Bowen, B.W. 2009. Endemism and dispersal: comparative phylogeography of three surgeonfishes across the Hawaiian Archipelago. Mar. Biol. 156: 689–698.

Evans, K.M., Williams, K.L., Westneat, M.W. 2019. Do coral reefs promote morphological diversification? Exploration of habitat effects on labrid pharyngeal jaw evolution in the era of big data. Integr. Comp. Biol. 59: 696–704.

Faircloth, B.C. 2016. PHYLUCE is a software package for the analysis of conserved genomic loci. Bioinformatics 32: 786–788.

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

Faircloth, B.C., Sorensen, L., Santini, F., Alfaro, M.E. 2013. Phylogenomic perspective on the radiation of ray-finned fishes based upon target sequencing of ultraconserved elements (UCEs). PLOS ONE 8: e65923.

Fricke, R., Eschmeyer, W.N. van der Laan, R. 2020. Eschmeyer's catalog of fishes: Genera, species, references. (http://researcharchive.calacademy.org/research/ichthyology/catalog/fishcatmain.asp). Electronic version accessed 17 April 2020.

Downloaded from https://academic.oup.com/sysbio/advance-article/doi/10.1093/sysbio/syab012/6146943 by guest on 01 March 2021

Gaither, M.R., Bowen, B.W., Rocha, L.A., Briggs, J.C. 2015. Fishes that rule the world: circumtropical distirbutions revisited. Fish Fish. 17: 664–679.

Gaither, M.R., Rocha, L.A. 2013. Origins of species richness in the Indo-Malay-Philippine biodiversity hotspot: evidence for the centre of overlap hypothesis. J. Biogeogr. 40: 1638–1648.

Gatesy, J., Springer, M.S. 2014. Phylogenetic analysis at deep timescales: unreliable gene trees, bypassed hidden support, and the coalescence/concatalescence conundrum. Mol. Phylogenet. Evol. 80: 231–266.

Harris, R.B., Alström, P., Ödeen, A., Leaché, A.D. 2018. Discordance between genomic divergence and phenotypic variation in a rapidly evolving avian genus (*Motacilla*). Mol. Phylogenet. Evol. 120: 183–195.

Hench, K., Vargas, M., Höppner, M.P., McMillan, W.O., Puebla, O. 2019. Inter-chromosomal coupling between vision and pigmentation genes during genomic divergence. Nat. Ecol. Evol. 3: 657–667.

Hobbs, J.P.A., Allen, G.R. 2014. Hybridisation among coral reef fishes at Christmas Island and the Cocos (Keeling) Islands. Raffles B. Zool. 30: 220–226.

Huertas, V., Bellwood, D.R. 2020. Trophic separation in planktivorous reef fishes: a new role for mucus? Oecologia 192: 813–822.

Jeffroy, O, Brinkmann, H., Delsuc, F., Philippe, H. 2006. Phylogenomics: the beginning of incongruence? Trends Genetc. 22: 225–231.

Jiang, X., Edwards, S.V., Liu, L. 2020. The multispecies coalescent model outperforms concatenation across diverse phylogenomic data sets. Syst. Biol. 69: 795–812.

Kaufman, L.S., Liem, K.F. 1982. Fishes of the suborder Labroidea (Pisces: Perciformes): phylogeny, ecology, and evolutionary significance. Breviora 472: 1–19.

Downloaded from https://academic.oup.com/sysbio/advance-article/doi/10.1093/sysbio/syab012/6146943 by guest on 01 March 2021

Klausewitz, W. 1976. *Cirrhilabrichthys filamentosus* n. gen., n. sp., aus der Javasee (Pisces: Labridae). Senck. biol. 57: 11–14.

Kobayashi, D.R. 2006. Colonization of the Hawaiian Archipelago via Johnston Atoll: a characterization of oceanographic transport corridors for pelagic larvae using computer simulation. Coral Reefs 25: 407–417.

Kuiter, R. H. 2010. Labridae fishes: Wrasses. 1st ed. Aquatic Photographics, Seaford, Australia.

Kulbicki, M., Parravinci, V., Bellwood, D.R., Arias-Gonzàlez, E, Chabanet, P., Floeter, S.R., Friedlander, A., McPherson, J., Myers, R.E., Bigliola, L., Mouillot, D. 2013. Global biogeography of reef fishes: a hierarchical quantitative delineation of regions. PLOS ONE 8: e81847.

Landis, M.J., Matzke, N.J., Moore, R.B., Huelsenbeck, J.P. 2013. Bayesian analysis of biogeography when the number of areas is large. Syst. Biol. 62: 789–804.

Liu, S.Y., Chang, F.T., Borsa, P., Chen, W.J., Dai, C.F. 2014. Phylogeography of the humbug damselfish, *Dascyllus aruanus* (Linnaeus, 1758): evidence of Indo-Pacific vicariance and genetic differentiation of peripheral populations. Biol. J. Linn. Soc. 113: 931–942.

Longo, S.J., Faircloth, B.C., Meyer, A., Westneat, M.W., Alfaro, M.E., Wainwright, P.C. 2017. Phylogenomic analysis of a rapid radiation of misfit fishes (Sygnathiformes) using ultraconserved elements. Mol. Phylogenet. Evol. 113: 33–48.

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.

Ludt, W.B., Rocha L.A. 2014. Shifting seas: the impacts of Pleistocene sea-level fluctuations on the evolution of tropical marine taxa. J. Biogeogr. 42: 25–38.

Downloaded from https://academic.oup.com/sysbio/advance-article/doi/10.1093/sysbio/syab012/6146943 by guest on 01 March 2021

Matzke, N. 2013. BioGeoBEARS: BioGeography with Bayesian (and likelihood) evolutionary analysis in R scripts. R package, version 0.2.1, published July 27, 2013.

Mayr, E. 1942. Systematics and the Origin of Species. Columbia University Press, New York, USA.

Meiklejohn, K.A., Faircloth, B.C., Glenn, T.C., Kimball, R.T., Braun, E.L. 2016. Analysis of a rapid evolutionary radiation using ultraconserved elements: Evidence for a bias in some multispecies coalescent methods. Syst. Biol. 65: 612–627.

Nguyen, L.T., Schmidt, H.A., von Haeseler, A., Bui, Q.M. 2015. IQ-TREE: a fast and effective stochastic algorithm for estimating maximum-likelihood phylogenies. Mol. Biol. Evol. 32: 268–274.

Parenti P., Randall, J.E. 2000. An annotated checklist of the species of the labroid fish family Labridae and Scaridae. Ichth. Bull. JLB Smith Inst. Ichth. 68: 1–97.

Price, M.R., Hadfield, M. 2014. Population genetics and the effects of a severe bottleneck in an ex situ population of critically endangered Hawaiian tree snails. PLOS ONE 9:e114377.

Quattrini, A.M., Faircloth, B.C., Dueñas, L.F., Bridge, T.C.L., Brugler, M.R., Calixto-Botía, I.F., DeLeo, D.M., Forêt, S., Herrera, S., Lee, S.M.Y., Miller, D.J., Prada, C., Rádis-Baptista, G., Ramírez-Portilla, C., Sánchez, J.A., Rodríguez, E., McFadden, C.S. 2018. Universal target-enrichment baits for anthozoan (Cnidaria) phylogenomics: New approaches to long-standing problems. Mol. Ecol. Res. 18: 251–295.

Quenouille, B., Hubert, N., Bermingham, E., Planes, S. 2011. Speciation in tropical seas: allopatry followed by range change. Mol. Phylogenet. Evol. 58: 546–552.

Randall, JE. 1988. Zoogeography of shore fishes of the Indo-Pacific region. Zool. Stud. 37: 227–268.

Downloaded from https://academic.oup.com/sysbio/advance-article/doi/10.1093/sysbio/syab012/6146943 by guest on 01 March 2021

Randall, J.E., Allen, G.R., Steene, R.C. 1990. Fishes of the Great Barrier Reef and the Coral Sea. Crawford House Press, Bathurst, NSW, Australia.

Randall, J.E., Earle, J.L., Pyle, R.L., Parrish, J.D., Hayes, Y. 1993. Annotated checklist of the fishes of Midway Atoll, Northwestern Hawaiian Islands. Pac. Sci. 47: 356–400.

Ree, R.H., Smith, S.A. 2008. Maximum likelihood inference of geographic range evolution by dispersal, location extinction, and cladogenesis. Syst. Biol. 57: 4–14.

Rocha, L.A., Bowen, B.W. 2008. Speciation in coral-reefs fishes. J. Fish Biol. 72: 1101–1121.

Ronquist, F. 1997. Dispersal-vicariance analysis: a new approach to the quantification of historical biogeography. Syst. Biol. 46: 195–203.

Schliep, K.P. 2011. phangorn: phylogenetic analysis in R. Bioinformatics 27: 592–593.

Schultz, O. 1978. Neue und fehlinterpretierte Fischformen aus dem Miozän des Wiener Beckens. Ann. Naturhist. Mus. Wien. 81: 203–219.

Siddall, M., Rohling, E.J., Almogi-Labin, A., Hemleben, C., Meischner, D., Schmelzer, I., Smeed, D.A. 2003. Sea-level fluctuations during the last glacial cycle. Nature 423: 853–858.

Smith, S.A, Moore, M. J., Brown, J.W., Yang, Y. 2015. Analysis of phylogenomic datasets reveals conflict, concordance, and gene duplications with examples from animals and plants. BMC Evol. Biol. 15:150.

Sonsthagen, S.A., Wilson, R.E., Underwood, J.G. 2017. Genetic implications of bottleneck effects of differing severities on genetic diversity in naturally recovering populations: an example from Hawaiian coot and Hawaiian gallinule. Ecol. Evol. 7: 9925–9934.

Springer, V.G., Randall, J.E. 1974. Two new species of the labrid fish genus *Cirrhilabrus* from the Red Sea. Isr. J. Zool. 23: 45–54.

Downloaded from https://academic.oup.com/sysbio/advance-article/doi/10.1093/sysbio/syab012/6146943 by guest on 01 March 2021

Tea, Y.K, Gill. A.C. 2017. *Cirrhilabrus shutmani*, a new species of fairy wrasse from the Babuyan Islands, northern Philippines. Zootaxa 4341: 77–88.

Tea, Y.K., Frable, B.W., Gill, A.C. 2018a. *Cirrhilabrus cyanogularis*, a new species of fairy wrasse from the Philippines and Indonesia (Teleostei: Labridae). Zootaxa 4418: 577–587.

Tea, Y.K., Frable, B.W., Van Der Wal, C. 2018b. Redescription and phylogenetic placement of *Cirrhilabrus sanguineus* Cornic (Teleostei: Labridae), with first documentation of the female form. Zootaxa 4526: 358–372.

Tea, Y.K., Pinheiro, H.T., Shepherd, B., Rocha, L.A. 2019. *Cirrhilabrus wakanda*, a new species of fairy wrasse from mesophotic ecosystems of Zanzibar, Tanzania, Africa (Teleostei, Labridae). ZooKeys 863: 85–96.

Tea, Y.K., Pyle, R.L., Rocha, L.A. 2020. A new species of fairy wrasse (Teleostei: Labridae: *Cirrhilabrus*) from mesophotic ecosystems of the Verde Island Passage, Philippines. Copeia 108: 91–102.

Tea, Y.K., Senou, H., Greene, B.D. 2016. *Cirrhilabrus isosceles*, a new species of wrasse (Teleostei: Labridae) from the Ryukyu Archipelago and the Philippines, with notes on the *C. lunatus* complex. J. Ocean Sci. Found. 21: 18–37.

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

Victor, B.C., Randall, J.E. 2014. *Pseudojuloides edwardi*, n. sp. (Perciformes: Labridae): an example of evolution of male-display phenotype outpacing divergence in mitochondrial genotype. J. Ocean Sci. Found. 11: 1–12.

Downloaded from https://academic.oup.com/sysbio/advance-article/doi/10.1093/sysbio/syab012/6146943 by guest on 01 March 2021

Warner, R.R. 1984. Mating behavior and hermaphroditism in coral reef fishes: the diverse forms of sexuality found among tropical marine fishes can be viewed as adaptations to their equally diverse mating systems. Am. Sci. 72: 128–136.

Westneat, M.W. 1993. Phylogenetic relationships of the tribe Cheilinini (Labridae: Perciformes). Bull. Mar. Sci. 52: 351–394.

Westneat, M.W., Alfaro, M.E. 2005. Phylogenetic relationships and evolutionary history of the reef fish family Labridae. Mol. Phylogenet. Evol. 36: 370–390.

Whelan, N.V., Kocot, K.M., Moroz, L.L., Halanych, K.M. 2015. Error, signal, and the placement of Ctenophora sister to all other animals. Proc. Natl Acad. Sci. U.S.A. 112: 5773–5778.

Xi, Z., Liu, L., Davis, C.C. 2015. Genes with minimal phylogenetic information are problematic for coalescent analyses when gene tree estimation is based. Mol. Phylogenet. Evol. 92:63–71.

Yamanoue, Y., Setiamarga, D.H.E., Matsuura, K. 2010. Pelvic fins in teleosts: structure, function and evolution. J. Fish Biol. 77: 1173–1208.

Yang, Z. 2007. PAML 4: phylogenetic analysis by maximum likelihood. Mol. Biol. Evol. 24: 1596–1591.

Yang Z., Rannala B. 2006. Bayesian estimation of species divergence times under a molecular clock using multiple fossil calibrations with soft bounds. Mol. Biol. Evol. 23:212–226.

Zhang, C, Rabiee, M., Sayyari, E., Mirarab, S. 2018. ASTRAL-III: polynomial time species tree reconstruction from partially resolved gene trees. BMC Bioinformatics 19: 153.

Zhang, Y.M., Williams, J.L., Lucky, A. 2019. Understanding UCEs: A comprehensive primer on using ultraconserved elements for arthropod phylogenomics. Insect Syst. Diver. 3: 3.

Downloaded from https://academic.oup.com/sysbio/advance-article/doi/10.1093/sysbio/syab012/6146943 by guest on 01 March 2021

Figure Legends

Figure 1. (a) Maximum-likelihood phylogenetic tree of Cirrhilabrus, inferred using the 75%-occupancy UCE + mitochondrial COI matrix. Asterisks indicate species that are represented by only COI. Open star indicates nesting of the monotypic Conniella. The outgroup genera Pseudocheilinus, Pseudocheilinops, and Pteragogus have been pruned from the tree. Circles at nodes indicate levels of SH-aLRT and UFBS support. SH-aLRT: Shimodaira-Hasegawa-like approximate likelihood ratio test. UFBS: ultrafast bootstrapping. This phylogenomic approach has resolved previously intractable relationships among several species and the placements of those that were uncertain in analyses of mitochondrial COI ([b] Tree of Cirrhilabrus relationships, inferred from mitochondrial COI, adapted from Tea et al. 2018b, 2020). Blue, yellow, and orange arrows indicate placement of taxa incertae sedis in both trees. (c) Mitochondrial COI gene tree demonstrating sharing of haplotypes for several members of the Cirrhilabrus cyanopleura complex. Photographs by G. R. Allen, E. Fleishauer, M. P. Hammer, R. H. Kuiter, J. E. Randall, Y. K. Tea, and B. C. Victor.

Figure 2. (a) Plot of numbers of parsimony-informative sites for all UCE loci. The dashed line represents the minimum number of informative sites needed to resolve a species tree with 43 internal branches. (b) Heat map of normalized Robinson-Foulds topological distances between trees inferred from different data sets and using different methods of analysis. Topological distances increase with color intensity. Asterisks indicate trees from which taxa represented only by COI sequences have been pruned. (c) Plot of average bootstrap support across individual UCE gene trees. (d) Gene genealogy interrogation applied to UCE loci. Colored lines represent each of the 15 possible unrooted topologies (H0–H14) based on the five sampled species complexes (rubriventralis, filamentosus, cyanopleura, punctatus, and temminckii complexes). Values on x-axis represent cumulative number of UCE gene trees supporting each hypothesis with highest probability and their associated P-values (y-axis). Values above dashed line indicate hypotheses that are statistically significant (P < 0.05). H0 indicates the highest-ranking hypotheses (represented by the inset tree topology).

Downloaded from https://academic.oup.com/sysbio/advance-article/doi/10.1093/sysbio/syab012/6146943 by guest on 01 March 2021

Figure 3. (a) Time-calibrated phylogeny of Cirrhilabrus, inferred using a Bayesian relaxed-clock analysis of the 75%-occupancy UCE + mitochondrial COI matrix. All outgroup taxa are shown. Node bars represent 95% credibility intervals for node ages. Red arrow indicates point of topological difference from the maximum-likelihood tree. Abbreviations are as follows: Great Barrier Reef (GBR); Coral Sea (CS); East Timor (ET); Maldives (Md); Kenya (Ky). (b) Cladogram for component species complexes of Lineage G, along with selected character states. (c) Most probable ancestral ranges in the genus Cirrhilabrus, inferred using a BayArea-like model in BioGeoBEARS. Outgroup taxa have been pruned from the tree. Squares at nodes represent the most probable ancestral areas, or a combination of most probable areas. Unless specified (with an asterisk), the relative probability of areas or area combinations for all nodes were ≥35%. Squares at bases of terminal branches leading to tips represent most probable ancestral ranges immediately after a cladogenetic event. Squares at tips represent present-day distributions, with colors corresponding to biogeographic regions. Photographs by B. D. Greene, R. H. Kuiter, and B. C. Victor.

Figure 4. Selected species of fairy wrasses. (a): Cirrhilabrus laboutei, USNM FIN 29549. Note the elongate anal-fin spines (open arrowhead). (b): Conniella apterygia, USNM FIN 29616. Note the absence of pelvic fins (white arrow). (c): Cirrhilabrus earlei. Note the general similarities in coloration and external morphology to Conniella. (d) & (e): Cirrhilabrus pylei. Note absence of pelvic fins (white arrowhead) in (d), a rare congenital defect in this otherwise fin-bearing species. Photographs by J. E. Randall, B. D. Greene, and Y. K. Tea.

Downloaded from https://academic.oup.com/sysbio/advance-article/doi/10.1093/sysbio/syab012/6146943 by guest on 01 March 2021

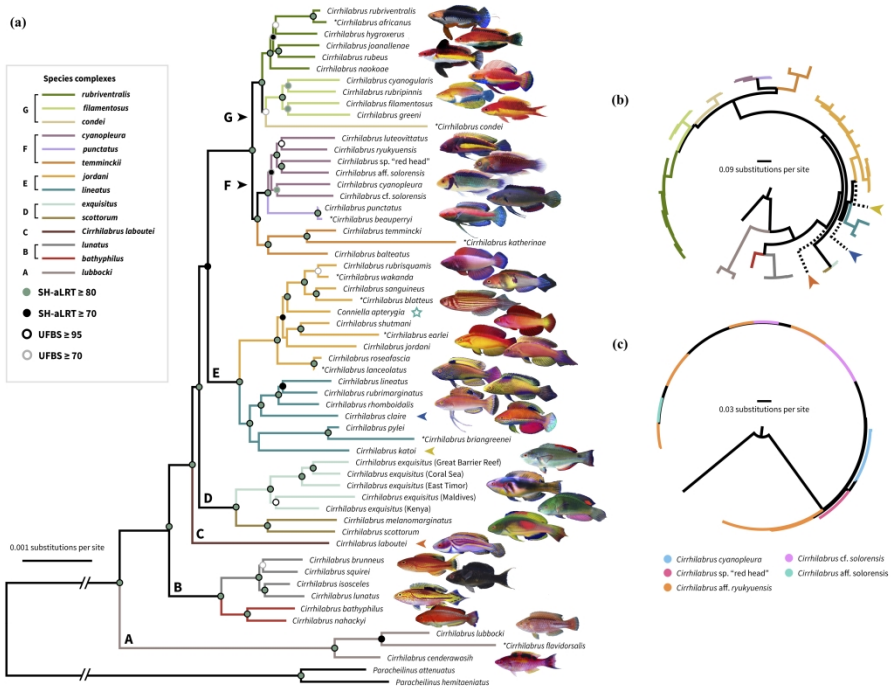


Figure 1. (a) Maximum-likelihood phylogenetic tree of *Cirrhilabrus*, inferred using the 75%-occupancy UCE + mitochondrial *COI* matrix. Asterisks indicate species that are represented by only *COI*. Open star indicates nesting of the monotypic *Conniella*. The outgroup genera *Pseudocheilinus*, *Pseudocheilinops*, and *Pteragogus* have been pruned from the tree. Circles at nodes indicate levels of SH-aLRT and UFBS support. SH-aLRT: Shimodaira-Hasegawa-like approximate likelihood ratio test. UFBS: ultrafast bootstrapping. This phylogenomic approach has resolved previously intractable relationships among several species and the placements of those that were uncertain in analyses of mitochondrial *COI* ([b] Tree of *Cirrhilabrus* relationships, inferred from mitochondrial *COI*, adapted from Tea et al. 2018b, 2020). Blue, yellow, and orange arrows indicate placement of taxa *incertae sedis* in both trees. (c) Mitochondrial *COI* gene tree demonstrating sharing of haplotypes for several members of the *Cirrhilabrus cyanopleura* complex. Photographs by G. R. Allen, E. Fleishauer, M. P. Hammer, R. H. Kuiter, J. E. Randall, Y. K. Tea, and B. C. Victor.

501x360mm (300 x 300 DPI)

Downloaded from https://academic.oup.com/sysbio/advance-article/doi/10.1093/sysbio/syab012/6146943 by guest on 01 March 2021

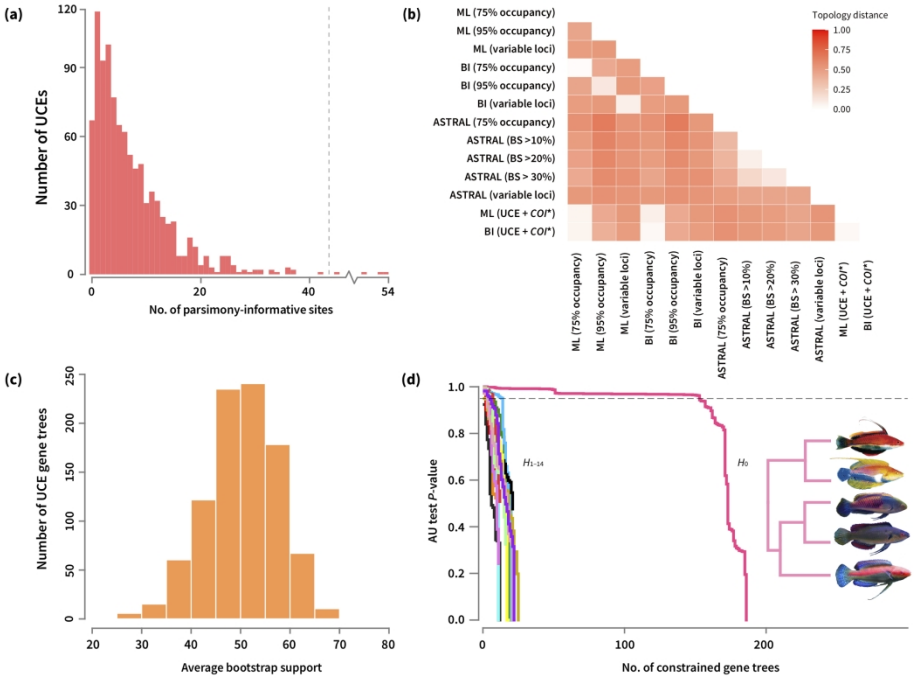


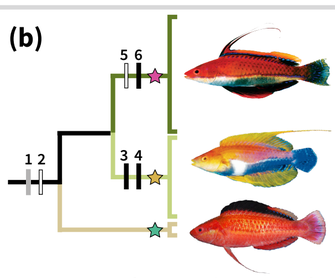
Figure 2. (a) Plot of numbers of parsimony-informative sites for all UCE loci. The dashed line represents the minimum number of informative sites needed to resolve a species tree with 43 internal branches. (b) Heat map of normalized Robinson-Foulds topological distances between trees inferred from different data sets and using different methods of analysis. Topological distances increase with color intensity. Asterisks indicate trees from which taxa represented only by *COI* sequences have been pruned. (c) Plot of average bootstrap support across individual UCE gene trees. (d) Gene genealogy interrogation applied to UCE loci. Colored lines represent each of the 15 possible unrooted topologies (*H0–H14*) based on the five sampled species complexes (*rubriventralis*, *filamentosus*, *cyanopleura*, *punctatus*, and *temminckii* complexes). Values on x-axis represent cumulative number of UCE gene trees supporting each hypothesis with highest probability and their associated *P*-values (y-axis). Values above dashed line indicate hypotheses that are statistically significant ($P < 0.05$). *H0* indicates the highest-ranking hypotheses (represented by the inset tree topology).

249x184mm (300 x 300 DPI)

Downloaded from https://academic.oup.com/sysbio/advance-article/doi/10.1093/sysbio/syab012/6146943 by guest on 01 March 2021

(a)

(b)



Apomorphy
Homoplasy
W. P. Ocean
Synapomorphy
Indian Ocean
S. P. Ocean

1 One row of cheek scales
2 Caudal scintilla
3 Throat and isthmus blue
4 Posterior dorsal filaments
5 Anterior dorsal filaments
6 Pendulous pelvic fins

Species complexes

G	*rubriventralis*
	filamentosus
	condei
F	*cyanopleura*
	punctatus
	temminckii
E	*jordani*
	lineatus
D	*exquisitus*
	scottorum
C	*Cirrhilabrus laboutei*
B	*lunatus*
	bathyphilus
A	*lubbocki*

Pp ≥ 0.95 ≤ 0.99
Pp ≥ 0.80 ≤ 0.94
Pp ≥ 0.50 ≤ 0.79

0.05 substitutions/site/my

C. africanus
C. rubriventralis
C. hygroxerus
C. rubeus
C. joanallenae
C. naokoae
C. greeni
C. filamentosus
C. cyanogularis
C. rubripinnis
C. condei
C. ryukyuensis
C. luteovittatus
C. aff. *solorensis*
C. sp. "red head"
C. cyanopleura
C. cf. *solorensis*
C. punctatus
C. beauperryi
C. katherinae
C. temmincki
C. balteatus
C. sanguineus
C. blatteus
C. wakanda
C. rubrisquamis
Co. apterygia
C. earlei
C. shutmani
C. jordani
C. lanceolatus
C. roseafascia
C. rubrimarginatus
C. lineatus
C. rhomboidalis
C. claire
C. briangreenei
C. pylei
C. katoi
C. exquisitus (GBR)
C. exquisitus (CS)
C. exquisitus (ET)
C. exquisitus (Md)
C. exquisitus (Ky)
C. melanomarginatus
C. scottorum
C. laboutei
C. isosceles
C. lunatus
C. brunneus
C. squirei
C. bathyphilus
C. nahackyi
C. flavidorsalis
C. lubbocki
C. cenderawasih
P. hemitaeniatus
P. attenuatus
Ps. evanidus
Ps. ocellatus
Ps. octotaenia
Ps. hexataenia
Ps. tetrataenia
Psops. ataenia
Pt. amboinensis
Pt. cryptus

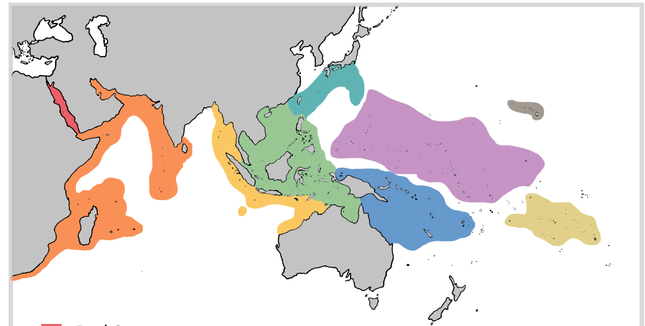
A
B
C
D
E
F
G

Mio. | Pli. | Ple.

50 40 20 10 0

Time (million years ago)

(c)



Red Sea
Western Indian Ocean
Eastern Indian Ocean
Central Indo-Pacific
Southern Japan
Melanesia
Central Pacific Ocean
Southeast Pacific Ocean
Hawaii

Ple. | Pli. | Mio.

0 2 4 6 8 10 12

Time (million years ago)

Downloaded from https://academic.oup.com/sysbio/advance-article/doi/10.1093/sysbio/syab012/6146943 by guest on 01 March 2021

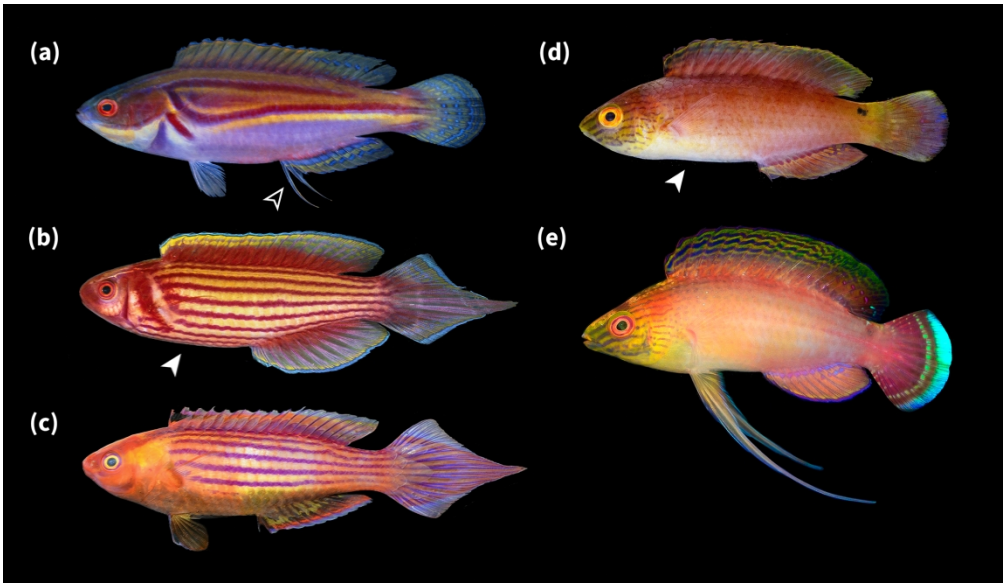


Figure 4. Selected species of fairy wrasses. (a): *Cirrhilabrus laboutei*, USNM FIN 29549. Note the elongate anal-fin spines (open arrowhead). (b): *Conniella apterygia*, USNM FIN 29616. Note the absence of pelvic fins (white arrow). (c): *Cirrhilabrus earlei*. Note the general similarities in coloration and external morphology to *Conniella*. (d) & (e): *Cirrhilabrus pylei*. Note absence of pelvic fins (white arrowhead) in (d), a rare congenital defect in this otherwise fin-bearing species. Photographs by J. E. Randall, B. D. Greene, and Y. K. Tea.

Downloaded from https://academic.oup.com/sysbio/advance-article/doi/10.1093/sysbio/syab012/6146943 by guest on 01 March 2021