

ORIGINAL RESEARCH

Decoding diversity in a coral reef fish species complex with restricted range using metagenomic sequencing of gut contents

Beverly J. French¹  | Yan Wei Lim² | Brian J. Zgliczynski¹ | Robert A. Edwards^{2,3} | Forest Rohwer² | Stuart A. Sandin¹ 

¹Center for Marine Biodiversity and Conservation, Scripps Institution of Oceanography, University of California, San Diego, CA, USA

²Department of Biology, San Diego State University, San Diego, CA, USA

³Department of Computer Science, San Diego State University, San Diego, CA, USA

Correspondence

Beverly J. French, Scripps Institution of Oceanography, Center for Marine Biodiversity and Conservation, University of California, San Diego, 9500 Gilman Dr. MC 0208, La Jolla, CA 9209300208, USA.
Email: bfrench@ucsd.edu

Funding information

Scripps Institution of Oceanography Director's Office; Private donors; San Diego State University; Division of Computer and Network Systems, Grant/Award Number: CNS-1305112; Moore Family Foundation; Gordon and Betty Moore Foundation, Grant/Award Number: GMBF-3781

Abstract

Aim: Identification of the processes that generate and maintain species diversity within the same region can provide insight into biogeographic patterns at broader spatiotemporal scales. Hawkfishes in the genus *Paracirrhites* are a unique taxon to explore with respect to niche differentiation, exhibiting diagnostic differences in coloration, and an apparent center of distribution outside of the Indo–Malay–Philippine (IMP) biodiversity hotspot for coral reef fishes. Our aim is to use next-generation sequencing methods to leverage samples of a taxon at their center of maximum diversity to explore phylogenetic relationships and a possible mechanism of coexistence.

Location: Flint Island, Southern Line Islands, Republic of Kiribati.

Methods: A comprehensive review of museum records, the primary literature, and unpublished field survey records was undertaken to determine ranges for four “arc-eye” hawkfish species in the *Paracirrhites* species complex and a potential hybrid. Fish from four *Paracirrhites* species were collected from Flint Island in the Southern Line Islands, Republic of Kiribati. Hindgut contents were sequenced, and subsequent metagenomic analyses were used to assess the phylogenetic relatedness of the host fish, the microbiome community structure, and prey remains for each species.

Results: Phylogenetic analyses conducted with recovered mitochondrial genomes revealed clustering of *P. bicolor* with *P. arcatus* and *P. xanthus* with *P. nesus*, which were unexpected on the basis of previous morphological work in this species complex. Differences in taxonomic composition of gut microbial communities and presumed prey remains indicate likely separation of foraging niches.

Main Conclusions: Our findings point toward previously unidentified relationships in this cryptic species complex at its proposed center of distribution. The three species endemic to the Polynesian province (*P. nesus*, *P. xanthus*, and *P. bicolor*) cluster separately from the more broadly distributed *P. arcatus* on the basis of relative abundance of metazoan sequences in the gut (presumed prey remains). Discordance between gut microbial communities and phylogeny of the host fish further reinforce the hypothesis of niche separation.

This is an open access article under the terms of the Creative Commons Attribution License, which permits use, distribution and reproduction in any medium, provided the original work is properly cited.

© 2020 The Authors. *Ecology and Evolution* published by John Wiley & Sons Ltd.

KEYWORDS

biodiversity, coexistence, hawkfish, metagenomics, rare species, trophic ecology

1 | INTRODUCTION

Although multiple ecological processes for species coexistence have been proposed (Chesson, 2000), a debate remains over what processes maintain morphologically and genetically similar species in a single habitat (Siepielski & McPeck, 2010; Velland, 2010). By the principle of competitive exclusion, complete competitors should not coexist (Hardin, 1960); yet, closely related and superficially ecologically equivalent species do co-occur in nature (Cothran, Noyes, & Relyea, 2015; Darwin, 1859; Gotelli, 1997). One mechanism contributing to initial species' divergence and continued persistence is trophic partitioning. Trophic shifts, especially feeding niche partitioning and associated morphological changes, have frequently been invoked to explain divergence among coral reef fishes on evolutionary timescales (Bellwood, Hoey, Bellwood, & Goatley, 2014; Bellwood & Wainwright, 2006), yet current trophic designations of reef fish are coarse and do not reflect likely fine-scale differences in diet (Sandin & Zgliczynski, 2015).

Dietary assessments are often conducted using stable isotope and gut content analysis, but these approaches have limitations. For gut content analysis, an important limitation is linked to difficulties with taxonomic identification of partially digested prey remains. Furthermore, even those identifiable remains may represent limited bouts of feeding (Baker, Buckland, & Sheaves, 2013; Dam, Peterson, & Okubo, 1991), especially for carnivorous fish species with rapid digestion processes. The accuracy of the results from stable isotopes is heavily dependent on whether sources have been accurately identified, despite advances in models incorporating uncertainty (Ward, Semmens, & Schindler, 2010).

High-throughput sequencing techniques, including metagenomic sequencing, of coral reef fish intestines and gut homogenates provide an opportunity to close current gaps in knowledge for investigations of high-resolution trophic partitioning among coral reef fishes. Results from metabarcoding studies have already shown promise for fishes (Casey et al., 2019; Leray, Boehm, Mills, & Meyer, 2012; Leray, Agudelo, Mills, & Meyer, 2013; Leray, Meyer, & Mills, 2015) and other species (Andriollo, Gillet, Michaux, & Ruedi, 2019; Deagle et al., 2018). However, examination of gut contents, both by traditional taxonomy and metabarcoding, provides only a snapshot of short-term feeding behavior, making extrapolation to general ecological behaviors difficult. Metagenomic sequencing, in contrast, provides information on the microbiome in addition to possible prey sequences (Srivathsan, Sha, Vogler, & Meier, 2014). Although prey sequences suffer from the same shortcomings of both morphological gut content analysis and metabarcoding in that they are still a snapshot of last meals, and should be interpreted as such, information from the gut microbiome holds more promise. One of the most important environmental factors that influences gut microbiota is host diet (Bolnick et al., 2014; Muegge et al., 2011; Smriga,

Sandin, & Azam, 2010; Sullam et al., 2012; Wong & Rawls, 2012), and in studies which have parsed the effects of diet and phylogeny, gut microbial communities cluster strongly by diet and only weakly by host species (Hale et al., 2018). In organisms including primates and fish, distinct gut microbiomes are associated with different diets (Hale et al., 2018; Uchii et al., 2006; Scott, Gratz, Sheridan, Flint, & Duncan, 2013), providing strong evidence that the gut microbiome can allow us to distinguish feeding habits within closely related species. For example, fish of the same species with different feeding habits (*Lepomis macrochirus*) have been shown to have distinct intestinal microbiota associated with different feeding modes, for example specialization on benthic invertebrates, aquatic plants, and zooplankton (Uchii et al., 2006). Thus, metagenomics integrates information on commonly consumed prey species and the intestinal microbiota. Furthermore, metazoan prey remnants may be more readily identifiable via sequencing than taxonomic identification of highly digested remains (Leray et al., 2012, 2013, 2015).

Exploring the utility and feasibility of this approach, we considered hawkfish in a *Paracirrhites* species complex from the remote island of Flint (Republic of Kiribati), in the likely center of the distribution for this genus in the Polynesian province of the Pacific (Donaldson, 1988). Hawkfish, including those in the genus *Paracirrhites*, are harvested for the aquarium trade (Donaldson, 2002), and rare species and colorations are highly prized. The arc-eye *Paracirrhites* group has been identified as a particularly interesting species complex for studies of biogeography (Donaldson, 1986). Donaldson (1986) noted the existence of two centers of distribution for Cirrhitidae, based on high levels of endemism. The center of maximum diversity for *Paracirrhites* in the arc-eye complex conflicts with a traditional view of Indo-West Pacific reef fish zoogeography in which there is a single center of highest diversity, the Indo-Malay-Philippine (IMP) biodiversity hotspot (Ekman, 1953), and instead points toward a hypothesis (Woodland, 1983, Gaither & Rocha, 2013) in which high levels of species richness in the IMP are due to pooling of species numbers from two distinct centers of distribution in confluence (center of overlap). In support of this, of the six Indo-West Pacific species of *Paracirrhites*, all six occur in the Polynesian area; three (*P. nissus*, *P. xanthus*, and *P. bicolor*) are endemic (Donaldson, 1988).

Collection of species in their center of maximum diversity and endemism in the Polynesia province, along with the use of next-generation sequencing, enabled us to simultaneously re-examine phylogenetic relationships on the basis of new genetic information and assess the potential role for trophic partitioning in maintaining coexistence among closely related and morphologically similar species. For this study, we hypothesized that species of the arc-eye complex would exhibit distinct intestinal microbiota and differences in prey remnants obtained from guts, as a partial explanation for potential coexistence of this enigmatic species complex at the center of these species' ranges. We further

hypothesized that the three species endemic to the Polynesian province (*P. bicolor*, *P. xanthus*, and *P. nesus*) would show distinct gut microbiota and prey profiles from the more broadly distributed *P. arcatus*, thus exhibiting a relationship between dietary patterns and biogeographic distribution.

2 | METHODS

2.1 | Study system

Paracirrhites arcatus species complex: Hawkfish in the genus *Paracirrhites* are small (maximum published length 29 cm) coral reef fish in the family Cirrhitidae found across the Indo-Pacific (Randall, 1963; Randall, Allen, & Steene, 1990). These fish are sit-and-wait mesopredators lacking a swim bladder, generally found perched on branching coral species, such as *Pocillopora* (DeMartini, 1996). Morphological analysis of prey remains in gut contents from *P. arcatus* indicate a general diet of small fishes, shrimps, crabs, and other crustaceans (Randall et al., 1990), and a recent DNA metabarcoding approach revealed unexpected prey items such as *Paracalanus parvus*, a pelagic copepod species, in addition to coral crabs (genus *Trapezia*), species in the genus *Galathea* (squat lobster), and traces of DNA of *Pocillopora* coral (Leray et al., 2013, 2015). Although there are six recognized species in the genus *Paracirrhites*, there are only four recognized species of *Paracirrhites* with a characteristic and colorful oblique U-shaped mark behind the eye; throughout the rest of this manuscript, we will refer to this group as the “arc-eye complex.” Of the four species in the arc-eye complex, one (*P. bicolor*) is listed as data deficient by the IUCN, and remaining three as least concern (*P. arcatus*, *P. nesus*, and *P. xanthus*). *P. arcatus* is one of the more abundant and widespread of the cirrhitids, known from east Africa to Polynesia. This species occurs in two different color morphs: a light grayish brown to reddish body with a pale pink to white band over the region of the lateral line, and a variety that is olive to dark brown and lacks the pale band on the body. These color morphs appear to be behaviorally unmodifiable, and are unrelated to ontogeny, sex, body size, or maturation (DeMartini & Donaldson, 1996; Donaldson, 1990; Myers, 1989; Randall, 1963; Sadovy & Donaldson, 1995), but differences in microhabitat have been described (DeMartini & Donaldson, 1996; Whitney, Donahue, & Karl, 2018), and they have been considered a potential instance of incipient speciation (Whitney, Bowen, & Karl, 2018). Across the Pacific where the species spans a depth range of 1–27 m, the light color morph was found to be more abundant at depths greater than 10 m (DeMartini & Donaldson, 1996). A more recent examination of microhabitat associations in Hawaii found that the light color morph was observed more frequently in deeper, subsurge zones, the dark morph was found more frequently in shallow, steep surge zones, and phenotypic intermediates were restricted to intermediate habitats (Whitney, Donahue, et al., 2018). The distribution and ecology of the endemic species (*P. xanthus*, *P. nesus*, and *P. bicolor*) are understudied and less understood than for *P. arcatus*. Representative photographs

of each species and color morph, as well as distributions from occurrence records, are shown in Figure 1.

2.2 | Range determination and implications for biogeography

Occurrences of the four species were tabulated based on review of the literature, museum records, and observations from belt-transect surveys (see Table S3 and S4). Most surveys did not distinguish between color morphs of *P. arcatus*, so we have shown the range as one group in Figure 1. Some reports included relatively broad ranges (island groups; for example, Gambier, Society Islands, Tuamotus) and so we also included shaded ranges encompassing the entire region (eg., Siu et al., 2017), as detailed species lists were not always available by individual island in these areas. Observations of a presumed hybrid (e.g., *P. xanthus* by *P. nesus*) in transect surveys are shown on a map in Figure S2, in the context of observations of *P. xanthus*, *P. nesus*, and *P. bicolor*. Observations of *P. nesus* via visual census have also been recorded in Maratua Island, Indonesia (Madduppa, 2012). It is possible that these observations of *P. nesus* are misidentifications of *P. arcatus*, as *P. nesus* is considered endemic to the eastern central Pacific, and there are no voucher specimens or photographs to provide a definitive identification.

2.3 | Data collection

Samples were collected from Flint island (11.43°S, 149.82°W) in the Southern Line Islands (central Pacific) in the Republic of Kiribati. Sampling included five adult individuals of each of the five groups: *P. arcatus* (light morph), *P. arcatus* (dark morph), *P. bicolor*, *P. nesus*, and *P. xanthus*. Sampling for species was conducted from October 17th–21st, 2013 and restricted to forereef habitat at an average depth of 10 m to minimize variation due to sampling. (We will refer to these five groups operationally as “species” for brevity. Although the two color morphs of *P. arcatus* are not presumed to be taxonomically distinct, Whitney, Bowen, et al., 2018 found evidence for partial reproductive isolation and potential incipient speciation.) Individual fish were collected with 3-prong spears (Hawaiian slings) and fish anesthetic (clove oil). Upon collection, fish were stored on ice and brought back to the research ship for initial processing. Each fish was assigned a unique identification tag, and basic morphometric information was collected. Intestines were dissected, and hindgut homogenates from the five individuals per species were combined for DNA extraction using the MO-BIO Powersoil® DNA Isolation Kit, following the manufacturer's instructions. Sequence libraries were prepared using the Ion Xpress™ Plus Fragment Library Kit (Life Technologies, NY, USA) with slight protocol modification, and each library was barcoded using the Ion Xpress™ Barcode Adaptors 1–16 Kit. SPRI beads-based size selection was performed according to the published New England Bioscience (NEB) E6270 protocol (<https://www.neb.com/protocols/1/01/01/size-selection-e6270>)

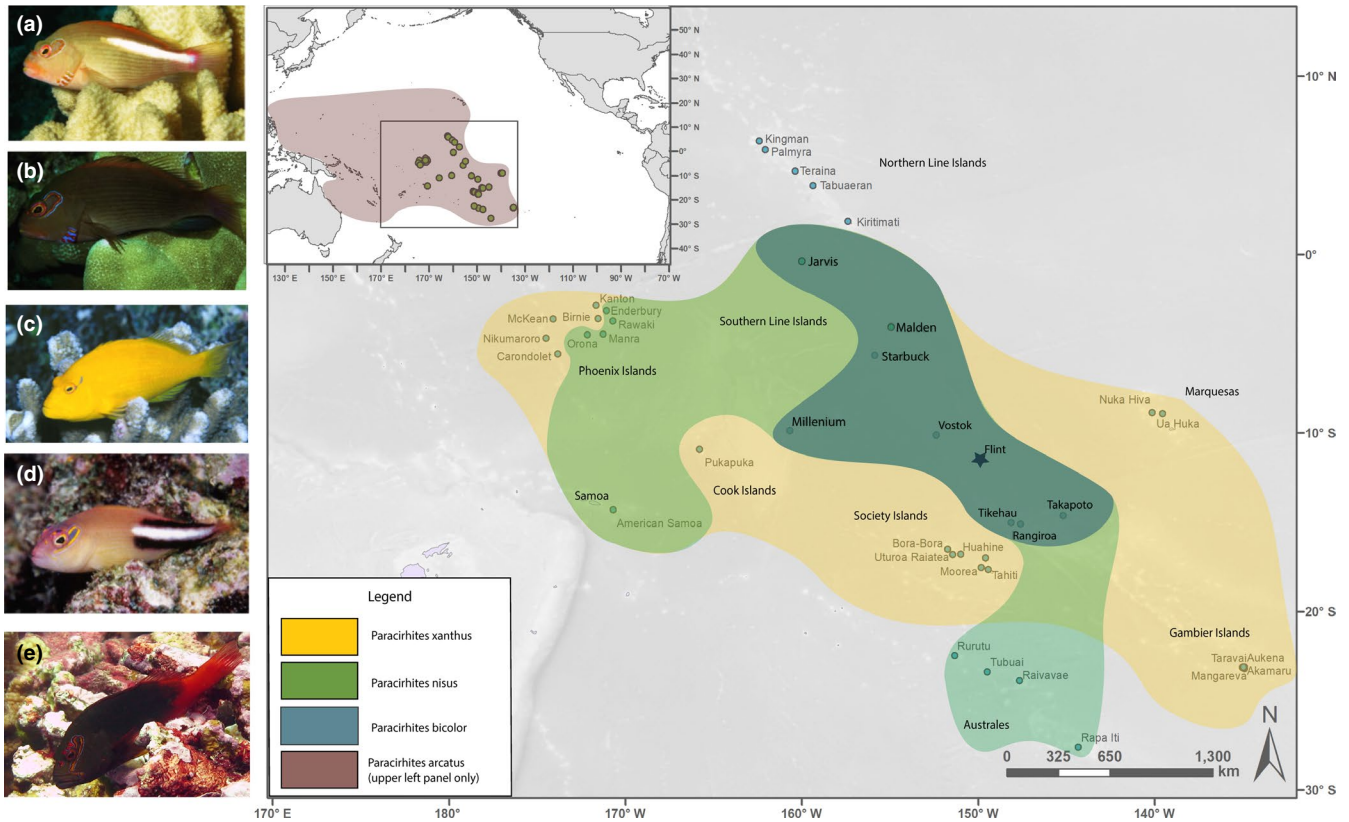


FIGURE 1 (Left panel) Illustration of color and morphological differentiation for species in the *arc-eye* species complex. Panel (a) corresponds to *Paracirrhites arcatus*; light morph; panel (b) corresponds to *Paracirrhites arcatus*, dark morph; panel (c) is *Paracirrhites xanthus*; panel (d) is *Paracirrhites nesus*; and panel (e) is *Paracirrhites bicolor*. (Right panel) Biogeographic range for species in the *arc-eye* complex. The range of *Paracirrhites arcatus* is shown in the upper left (smaller figure) only, as this range encompasses the ranges of the other species. The square box denotes the enlarged portion. The sampling site for collection of specimens for metagenomic analysis is labeled with a star (Flint Island, Republic of Kiribati). (Photographs: B.J. Zgliczynski and M.J. Adams)

for 200–300 bp fragment size selection after adaptor ligation. Emulsion PCR was performed on an 8-cycle amplified library using the OneTouch supplemented with Ion Torrent PGM Template OT2 200 Kit. Template libraries were sequenced on the Ion Torrent PGM Sequencing 200 Kit v2 and Ion 318TM Chip Kit v2 on the research vessel M/Y Hanse Explorer (Lim et al., 2014).

2.4 | Bioinformatic analyses

DNA isolated from gut homogenates was pooled for five individuals per species and sequenced to yield 5 metagenomic libraries (*P. arcatus*: light morph, *P. arcatus*: dark morph, *P. xanthus*, *P. nesus*, and *P. bicolor*). Metagenomic sequence reads were filtered for quality using the Preprocessing and Information of Sequences tool (PRINSEQ; Schmieder & Edwards, 2011). The sequenced reads were translated in silico into predicted protein sequences and compared with the SEED database to provide taxonomic identifications and information on organism abundance using Blastx (Atschul, Gish, Miller, Myers, & Lipman, 1990) with a maximum e-value cutoff of 1×10^{-5} , a minimum alignment cutoff of 15, and a

minimum percent identity of 80%. Metagenomes were deposited in the MG-RAST metagenomics analysis server (Meyer et al., 2008) and made publicly available using the unique identifiers listed in Table S1. Taxonomic assignment of bacterial species was confirmed using a blastn search against whole bacteria genomes deposited in GenBank (<ftp://ftp.ncbi.nlm.nih.gov/genomes/Bacteria/>). Bacterial taxa were categorized at the phylum level, with the exception of Proteobacteria and Firmicutes, which were categorized by class. Unclassified bacteria were categorized as “other bacteria.” While conducting analyses at this taxonomic level aids in comparability to results from other studies conducted on fish gut microbiome, there may be hidden diversity in the identity and function of the gut microbial community that warrants a deeper look. For this reason, the 16s rRNA, rpoB, and recA gene sequences were extracted from the unassembled reads of each genome and clustered into OTUs using the program GenomePeek (McNair & Edwards, 2015). Further analyses focused on the results for 16S rRNA exclusively, as the database for this marker has the most coverage, thereby making comparisons of relative sequence abundance more accurate. For the analysis of putative prey items, the metagenomes were subset into 20,000 sequences to speed processing time. We used Midori-UNIQUE, a curated reference dataset of metazoan mitochondrial

DNA sequences containing all unique haplotypes for each species for 16,697 unique species (Machida, Leray, Ho, & Knowlton, 2019) as our reference dataset. The Midori reference datasets are divided into multiple markers, including two ribosomal RNA and thirteen mitochondrial protein-coding genes found in most phyla of metazoans. The cytochrome oxidase subunit I (COI) gene had the largest number of sequences, so we conducted our analyses with this dataset. We performed an initial blastn search for each metagenome in CLC Genomics Workbench (Qiagen Bioinformatics). Retaining only those sequences with a maximum e-value cutoff of 1×10^{-4} , we then removed sequences with matches to *Paracirrhites*, as we assumed these were likely the result of host tissue contamination, rather than cannibalism, as customary for these analyses (Casey et al., 2019). We took all remaining annotations and calculated relative abundance within each metagenome at the phylum level. At lower taxonomic levels (eg. species), the best sequence hit sometimes matched to an organism that either (a) does not occur in the study location, according to existing censuses, or (b) does not occur in the marine environment (e.g., species in the orders Lepidoptera and Hemiptera within the phylum Arthropoda). To avoid any erroneous assumptions produced by using false species IDs or excluding important prey that could not be confidently assigned to lower taxonomic levels, we therefore conducted all analyses with metazoan sequences at the phylum level.

2.5 | Phylogenetic analyses

Sequence reads from the metagenome for each species and color morph were aligned against the reference genome for the *P. arcatus* mitochondrion (17,336 bp, NCBI RefSeq AP006012.1) using alignment tools in CLC Genomics Workbench 7 (Qiagen, CA, USA). Using *Cheilodactylus quadricornis* (NCBI RefSeq: KT357695.1), a species in the sister group to Cirrhitidae (Donaldson, 1997), and *Cirrhitichthys aprinus* (NCBI RefSeq: AP006011.1) as outgroups, the resultant FASTA mitochondrial DNA (mtDNA) sequences were aligned using MAFFT, a multiple alignment program for amino acid or nucleotide sequences. Maximum likelihood (ML), maximum parsimony (MP), and Bayesian inference (BI) analyses were conducted in order to assess congruence between the methods. An ML tree was constructed with RaxML-HPC2 (Randomized Axelerated Maximum Likelihood) using the GTRGAMMA model with 100 bootstrap replications (Stamatakis, 2014) run on the CIPRES portal (Miller, Pfeiffer, & Schwartz, 2010). MP was carried out using PAUP 4.0a166 (Swofford, 2003). Node support was assessed using jackknifing of sites with 100 replicates to produce a 50% majority-rule consensus tree (Figure S3). BI was conducted in MrBayes v.3.2.7a (Ronquist & Huelsenbeck, 2003) using two runs of four chains with ten million generations and a sampling frequency of 1,000 generations for the Markov chain Monte Carlo analysis. The first 25% of trees were discarded as burn-in and the posterior probabilities were estimated by combining the remaining trees from each run into a majority-rule consensus tree.

2.6 | Statistical analyses

In order to assess differences in bacterial community composition and metazoan prey composition between the four species and two color morphs, separate similarity profile routine (SIMPROF) tests were conducted on relative abundance of bacterial taxonomic groups and the relative abundance of metazoan sequences using the *clustsig* package (Whitaker & Christman, 2010) implemented in R (R Development Core Team, www.r-project.org). SIMPROF is a permutation procedure, which tests for the presence of sample groups in a priori unstructured sets of samples (Clarke, Somerfield, & Gorley, 2008). Analyses were based on 10,000 random permutations of the annotated metagenomic data. Following procedure which has been used in other analyses of bacterial structure in metagenomic samples, the two most abundant phyla (Firmicutes and Proteobacteria) were further divided into the lower taxonomic grouping of class to provide clearer resolution of the groups driving differences in microbiome composition between species (Kelly et al., 2014). Shannon–Weiner (H') diversity indices were calculated using the R package *vegan* (Oksanen, Blanchet, & Kindt, 2012). All statistics were conducted using R 3.1.2.

3 | RESULTS

3.1 | Phylogenetic analyses of hawkfish

The ML and MP phylogenetic analyses of the whole mitochondrial genome (mtDNA) for all four species, including the two color morphs analyzed separately, provided almost identical reconstructions with strong node support for the separation of two clades within the arcatus complex, with one clade composed of *P. arcatus* (both color morphs) and *P. bicolor*, and one clade composed of *P. xanthus* and *P. nesus*. The maximum likelihood tree showed clustering of *P. xanthus* with *P. nesus*, in addition to clustering of the two color morphs of *P. arcatus* with *P. bicolor*. Under maximum parsimony and Bayesian inference, the phylogenetic tree showed additional structure among species, with *P. arcatus* (dark morph) and *P. bicolor* forming a sister clade to *P. arcatus* (light morph) (Figure 3, right).

3.2 | Fish gut microbiomes and metabolic functions

Investigation of the bacterial community structure shows divergence between species, but not color morphs (SIMPROF, $\alpha < .05$, see Figure 2). Firmicutes and Proteobacteria together made up 94%–96% (Firmicutes = 81%–91%; Proteobacteria = 5%–13%) of the bacterial community and thus were divided into class. SIMPROF analysis of bacterial community structure in the five composite metagenomes for each of the five members of the arc-eye complex revealed three significant ($\alpha = .05$) groupings: (1) *P. arcatus* (light), *P. arcatus* (dark), and *P. xanthus*; (2) *P. nesus*, and (3) *P. bicolor* (Figures 2c and 3). From the analyses of 16S rRNA OTUs, *Clostridium perfringens* was found in highest abundance of all bacterial taxa for all species, but comprises only 42.6% and 37.4% of the relative

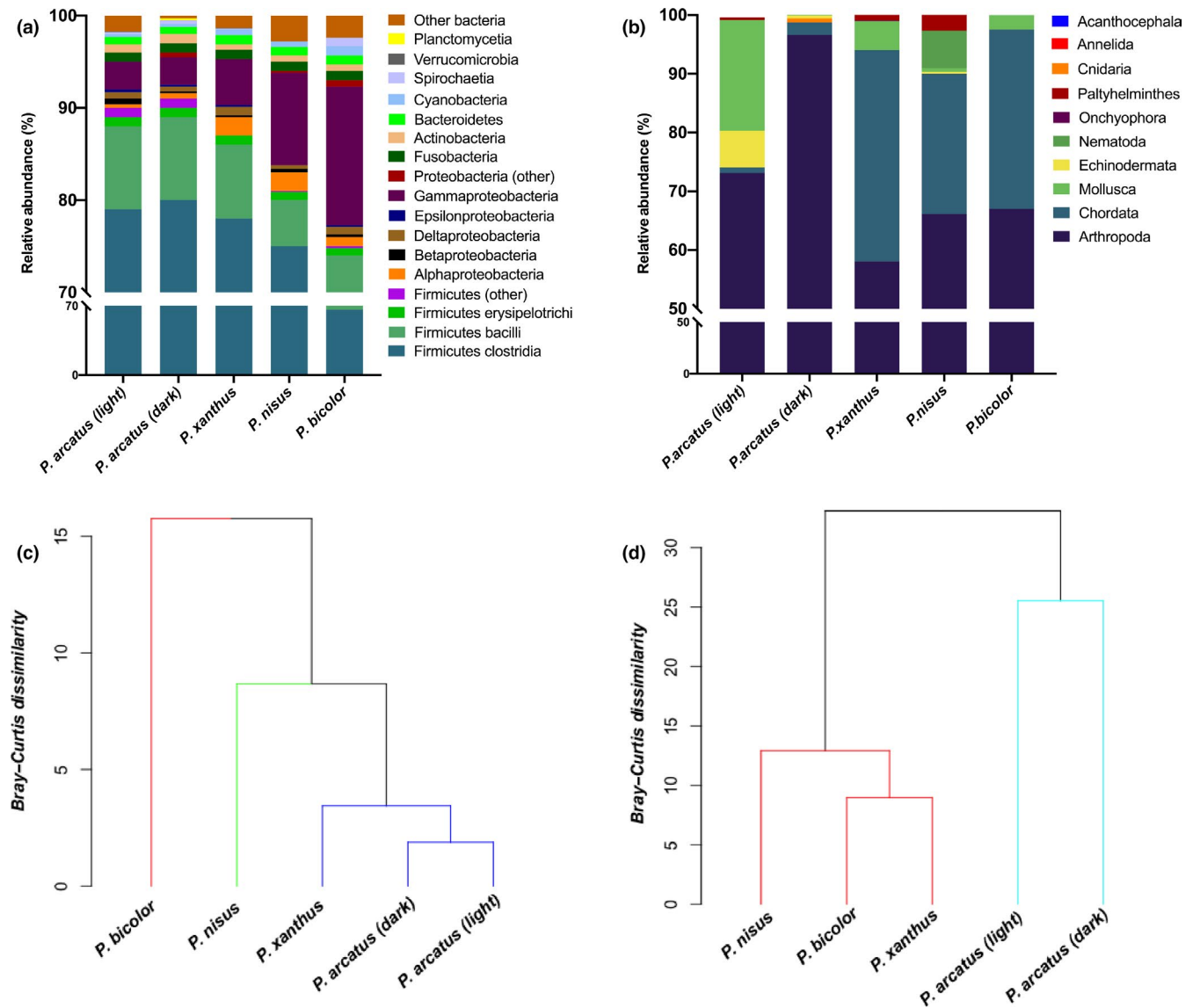


FIGURE 2 (a) Gut microbial community structure and metazoan composition of metagenomes. Stacked barplot illustrating relative abundance of bacterial phyla within total bacterial reads. The two most abundant phyla (Proteobacteria and Firmicutes) are broken down into class. (b) Putative prey sequences in hawkfish gut contents. Stacked barplot illustrating relative abundance of all High-scoring Segment Pairs (HSPs) with $e < 1 \times 10^{-4}$ from BLASTn search of the Midori-UNIQUE reference dataset for the cytochrome oxidase subunit I (COI) gene, after removal of HSPs where the accession was a match to *Paracirrhites*. (c) Dendrogram from Similarity Profile Routine (SIMPROF) test on hawkfish gut microbial community structure. Colors indicate significantly different clusters at $\alpha = .05$. (d) Dendrogram from Similarity Profile Routine (SIMPROF) test on metazoan prey sequence composition in hawkfish guts. Colors indicate significantly different clusters at $\alpha = .15$

abundance of all 16S contigs for *P. nesus* and *P. bicolor*, respectively, in comparison with 78%–94% of the bacterial community for *P. xanthus* and *P. arcatus* (both morphs). Accordingly, Shannon–Weiner diversity indices of the gut microbial communities using 16S data are highest for *P. bicolor* ($H' = 2.27$) and *P. nesus* ($H' = 1.51$), and consistent with the clustering identified at higher taxonomic levels (Figure 2c).

3.3 | Putative prey

Chordate sequences were found in higher relative abundance for the species with a more restricted range (*P. nesus*, *P. xanthus*, and *P.*

bicolor); that is, 24%–36% compared to 0.9%–2% for *P. arcatus* (dark) and *P. arcatus* (light), respectively (Figure 2b). SIMPROF analysis of the metazoan data ($\alpha = .15$) shows two groupings; grouping of the two color morphs of *P. arcatus*, and a separate grouping with the three endemic species (*P. bicolor*, *P. xanthus*, and *P. nesus*), see Figure 2d.

One limitation of this approach is the difficulty in distinguishing secondary predation (that is, prey remains in a prey item's gut) from targeted prey, or incidental consumption in the course of feeding. For instance, sequences with a best match to Cnidaria were found in the metagenome for the dark morph of *P. arcatus*. *P. arcatus* are known to feed on mutualistic coral crab species of the genus *Trapezia*

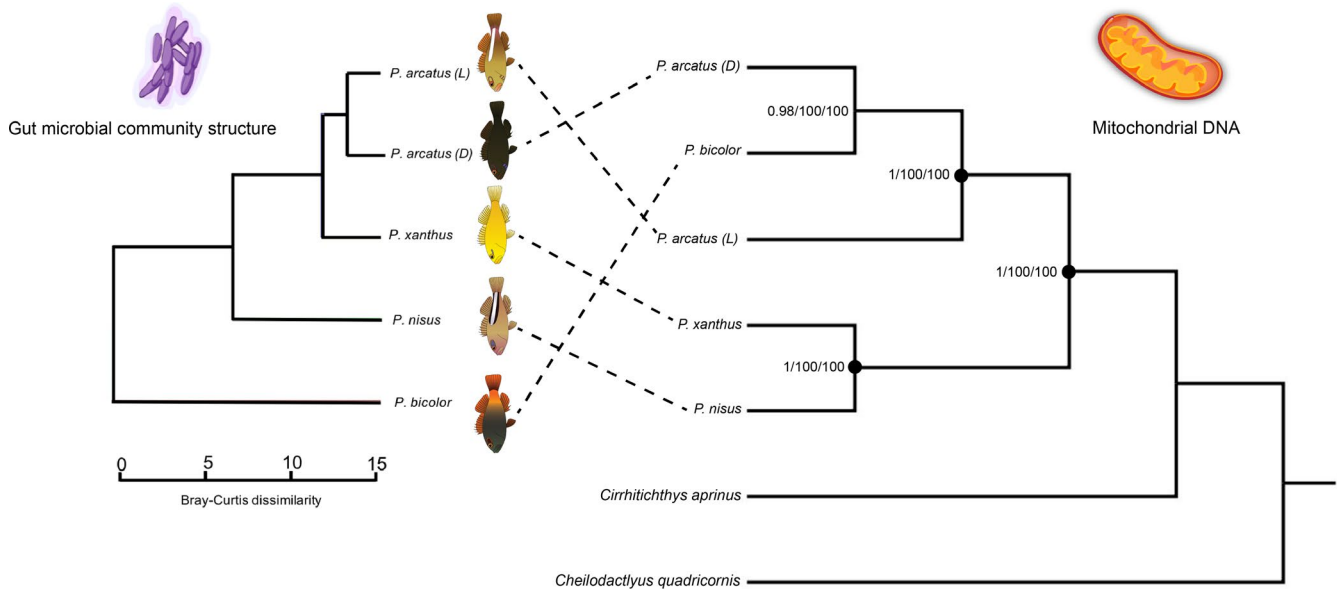


FIGURE 3 Dendrogram from SIMPROF analysis on hawkfish gut microbial communities (left). *Paracirrhites bicolor* and *Paracirrhites nesus* show significant ($p < .05$) separation from *Paracirrhites arcatus* (both color morphs) and *Paracirrhites xanthus*. The microbiomes of the two color morphs and *Paracirrhites xanthus* are not significantly different from one another. Phylogenetic analysis of hawkfish species using genome sequences for the mitochondrion (right). Depicted is a phylogenetic tree for *Paracirrhites* species with *Cheilodactylus quadricornis* (KT357695.1) and *Cirrhichthys aprinus* (AP006011.1) as outgroups. Support values are indicated at the nodes in the order Bayesian inference/maximum likelihood/maximum parsimony. Maximum support values (1/100/100) are indicated at the nodes by black circles

(Leray et al., 2015), which themselves feed on coral mucus (Knudsen, 1967). Given the low relative abundance of sequences matching to Cnidaria (>1%), it is certainly possible these sequences are from secondary prey.

4 | DISCUSSION

The importance of the expansion of feeding forms over evolutionary time for the diversification of reef fishes has been well-supported, with niche partitioning and morphological innovation identified as playing key roles in the development of herbivory and detritivory (Bellwood et al., 2014; Bellwood & Wainwright, 2006). Behavioral innovations may be equally as important as morphological innovations but have not received the same attention likely due to difficulties in inference from the fossil record. Diversification by this mode, however, may have left traces in the current biogeography of reef fishes, and may be more readily studied by examining potential ecological differentiation in present-day genera. Local adaptation may reduce the likelihood of successful colonization of novel habitats after dispersal, and dietary specialization has been linked to reduced species durations (Balisi, Casey, & Van Valkenburgh, 2018), as many of the traits that increase the chance of speciation also increase the chances of extinction (Jablonski, 2008). Compressed ranges may therefore be the result of either (a) limitations in dispersal or successful colonization, or (b) the result of past extinction events. Simultaneous examination of phylogenetic relationships, potential ecological differentiation, and current ranges within a genus—especially one for which incipient speciation has been proposed (Whitney, Bowen,

et al., 2018)—may provide clues to the mechanisms involved in creating and maintaining the observed species' distributions. To that end, our findings point toward previously unidentified relationships in a cryptic species complex at its proposed center of distribution, and discordance between gut microbial communities and phylogeny of the host fish which support the hypothesis of niche separation.

The findings here further demonstrate that metagenomic techniques can provide data comprising three important lines of information for a variety of uses: (1) host genetics, (2) microbial communities in the gut, and (3) the identity of putative prey items. When these tools were applied to understudied and closely related fish from a remote island, the value of the multidimensional data are revealed well. Species-level designations for hawkfish included in this study were originally determined on the basis of morphometric differences. In the initial description by Randall (1963), it was suggested that, of the five species with postocular marks in *Paracirrhites*, *P. bicolor* might be most closely related to *P. nesus*. Our phylogenetic analysis of mtDNA, in contrast, suggests that *P. nesus* and *P. xanthus* are more closely related than *P. bicolor* and *P. nesus*. This finding further matches observations of *P. nesus* by *P. xanthus* hybrids observed in the wild (see Figure S1, Tables S3 and S4), as species that are more closely related are more likely to hybridize. The phylogenetic analyses conducted here further raise the possibility that *P. bicolor* is a color morph variation of *P. arcatus*, given that mtDNA sequences for the *P. arcatus* dark morph and *P. bicolor* were nearly identical (>99% similarity). As mtDNA is inherited solely from the female parent, it is also possible that *P. bicolor* is a hybrid arising from a spawning event in which *P. arcatus* was the female parent with an unknown male species. In support

of this possibility is the extreme rarity of *P. bicolor*. Rarity of at least one parental species has been shown to play an important role in hybrid formation for some reef fishes, with hybridization resulting from mating of closely related species due to scarcity of conspecific partners (Hobbs, Frisch, Allen, & Herwerden, 2008). Incorporating nuclear markers with the mtDNA results, as well as microsatellite markers, would further clarify phylogenetic relationships between species in this study. These possibilities are being investigated with additional molecular research.

A unique gut microbiome and differential relative abundance of likely dietary items suggests that there may be some ecological differences among species, despite high genetic similarity, that warrant further investigation. Previous studies have demonstrated an unexpected negative relationship between diet diversity and diversity of gut microbiota for two fish species; that is, dietary generalists had a less diverse microbiome (Bolnick et al., 2014). Indeed, we found that *P. bicolor* and *P. nissus* have the most diverse microbiota, perhaps indicating dietary specialization. Additionally, although the two color morphs of *P. arcatus* cluster in analyses of gut microbial communities and metazoan sequences, differences in relative abundance of sequences for prey items found in the guts warrant further study of potential fine-scale differences. As noted previously, it has been suggested that *P. arcatus* may be an example of incipient speciation in a sympatric species (Whitney, Bowen, et al., 2018). Differences in composition of diet between color morphs could be driven by these differences in microhabitat associations.

Analyses of putative prey remains revealed broad convergence with what was generally known regarding *Paracirrhites* diets, including a high relative percentage of sequences belonging to the phyla Arthropoda and Chordata. This is consistent with a reported diet consisting predominantly of crustaceans and other fishes, but also revealed some interesting differences. In particular, the three species with the most restricted ranges had a higher relative abundance of sequences matching Chordata. An analysis of gut contents by morphologic identification of prey remains for 199 individuals of *Paracirrhites arcatus* (both color morphs) show a range of 0.62%–9.4% relative abundance for other fish species (Cordner, 2013, see Table S2), suggesting that other fish species may make up a higher proportion of the diet for those species with more restricted ranges (*P. bicolor*, *P. nissus*, and *P. xanthus*). One potential caveat is that all sequences with a match to the host organism were removed to avoid conflation of host tissue with estimates of prey consumption. Any instances of intraspecific predation would therefore be impossible to detect. As this processing was applied equally across all species, if intraspecific predation is more frequent among *P. arcatus*, for instance, the estimates of relative abundance of prey fish would be underestimated in comparison.

There are some considerations that must be noted for the interpretation of our results, and in particular what the relative abundance of metazoan sequences indicates for hawkfish diets over long timescales. The number of High-scoring Segment Pairs (HSPs) is not a direct correlate to the number of organisms in the gut; a single large organism could result in many HSPs, and gene copy number

differs among eukaryotes (Prokopowich, Gregory, & Crease, 2003). Furthermore, carnivorous fish have relatively short gut passage times (~8 hr, Clements, Angert, Montgomery, & Choat, 2014), so the identification of prey remains, even with combined samples to homogenize variability, only reflects a limited selection of what these fish are eating over time. Gut microbiomes may therefore be a more stable predictor of integrated food habits over time. Dietary interventions in well-studied human gut microbiomes have shown community changes on timescales of months (Ley, Turnbaugh, Klein, & Gordon, 2014) to years (Faith et al., 2013), and others showing return to original structure of gut microbiota only 2 days after the end of a dietary intervention (David et al., 2014). Studies that have examined this for fish have had similarly varying results. A study on captive clownfish found that certain bacterial groups such as *Clostridia perfringens* did not vary significantly with time postfeeding, whereas other groups such as *Bacilli* (Firmicutes) and *Photobacterium sp.* increased after feeding (Parris, Morgan, & Stewart, 2019). Other studies found no or minimal changes in diversity, but observed changes in relative abundance of bacterial phyla 3–24 hr after feeding (Zhang et al., 2018) and 3–12 days post-fasting (Xia et al., 2014). Taken together, there does appear to be a core microbiome for fish that is sensitive to diet on longer timescales. We further posit that the general concordance between the metazoan and microbial data that we show here provides strong evidence for real ecological differences between species.

Some of the limitations associated with this method can help guide additional research in this field and contribute to the discussion on best practices in using next-generation sequencing of gut contents as a technique for assessing dietary habits within animals. As guts were pooled to produce composite metagenomes, information on individual variability between microbiomes and diet within species was impossible to obtain for this study, although we had a priori evidence that intraspecific variability in this genus would be low. In analyses excluding all empty stomachs (~10%), stomach content analysis of 199 *P. arcatus* individuals was indistinguishable across islands (Zgliczynski et al. 2019). In instances where there are constraints to the number of samples that can be analyzed, whether financial or logistical, specimen pooling is often promoted as an option with very little loss of statistical power (Weinberg & Umbach, 1999). Shisterman and Vexler (2008) show that estimations based on pooled data increase efficiency over the use of individual measurements when an assay has a detection threshold, as pooling samples minimize the amount of information lost below the detection threshold. This technique is used extensively in the literature for population-based epidemiological studies (Keeler & Berwick, 1976; Shisterman & Vexler, 2008) and other situations with high-dimensional biological data, such as that generated via high-throughput sequencing techniques (Clarke, Ransom, et al., 2008).

Sample sizes for this study were also limited. *P. bicolor* in particular is so rare that a photograph of one was only published recently (Bacchet, Zysman, & Lefevre, 2017), and this species is known from only four specimens from two localities (Randall, 1963). Future

studies that incorporate analyses of intraspecific variability with higher sample sizes for species where additional collection is both logistically possible and ethically defensible are warranted to verify and extend the findings here.

Notwithstanding the limitations, a major strength of this approach includes the ability to look at the host genetics, microbiome, and the presumed prey fragments simultaneously. This approach further allows the investigation of the results using multiple genes, rather than one preselected marker that must be chosen a priori (as in the use of metabarcoding). In most situations, and particularly for highly complex ecosystems, reference barcodes are not available for all potential prey species. Combined with the knowledge that no single marker is ideal for resolving all taxonomic groups (Deagle, Jarman, Coissac, Pompanon, & Taberlet, 2014; Hebert, Ratnasingham, & de Waard, 2003), the use of this technique represents a significant advancement in the available tools for investigating realistic patterns of prey consumption in wild animals. As databases further improve through international efforts, our ability to address complex ecological questions with genetic tools will continue to grow. It is our hope that this study will encourage more community ecologists and conservation scientists to consider metagenomics as a powerful tool that can be used to leverage multidimensional and ecologically informative data from rare and understudied species.

ACKNOWLEDGMENTS

This work was carried out under research permits from the Environment and Conservation Division of the Republic of Kiribati. For logistical support, we thank the Captain Martin Graser, officers and crew of the M/Y Hanse Explorer. Dive support was provided by Christian McDonald (SIO). We thank Life Technologies for providing library preparation and sequencing reagents for the study. The SDSU Vice President of Research, Director's Office of Scripps Institution of Oceanography, Moore Family Foundation, and several private donors provided cruise support. Computing resources and support were provided by NSF grant CNS-1305112 to Robert A. Edwards. We further thank Adi Khen and Kiefer Forsch for the artwork, Gregory Rouse, Emily McLaughlin, and Josefine Stiller for conversations and recommendations regarding phylogenetic analyses, Katelyn McNair and Genivaldo Gueiros Z. Silva for bioinformatics assistance and advice, and Gilles Siu for providing detailed information on hawkfish observations in the Tuamotus. The funders had no role in study design, data collection, and analysis, decision to publish, or preparation of the manuscript. This manuscript is dedicated to Rachel A. Morrison, whose work laid the foundations for the investigation conducted here.

CONFLICT OF INTEREST

None Declared.

AUTHOR CONTRIBUTIONS

BZ and SS initiated the project and collected the samples. YL conducted the sequencing on the expedition. BF conducted the analyses and wrote the paper. RE, YL, and FR contributed to the analyses

and design thereof. All authors contributed to the final version of the manuscript.

DATA AVAILABILITY STATEMENT

Metagenomes were deposited in the MG-RAST metagenomics analysis server (Meyer et al., 2008) and made publicly available using the following unique identifiers: *P. arcatus*; light morph: mgm4632779.3, *P. arcatus*; dark morph: mgm4856250.3, *P. xanthus*: mgm4632786.3, *P. nesus*: mgm4631833.3, and *P. bicolor*: mgm4856249.3/mgm4856251.3.

ORCID

Beverly J. French  <https://orcid.org/0000-0001-8848-0093>

Stuart A. Sandin  <https://orcid.org/0000-0003-1714-4492>

REFERENCES

- Andriollo, T., Gillet, F., Michaux, J. R., & Ruedi, M. (2019). The menu varies with metabarcoding practices: A case study with the bat *Plecotus auratus*. *PLoS ONE*, *14*, e0219135.
- Atschul, S. F., Gish, W., Miller, W., Myers, E. W., & Lipman, D. J. (1990). Basic local alignment search tool. *Journal of Molecular Biology*, *27*, 863–864.
- Bacchet, P., Zysman, T., & Lefevre, Y. (2017). *Guide des poissons de Tahiti et ses îles. Quatrième édition revue et augmentée*. Tahiti (Polynésie Française): Éditions Au Vent des Îles, 342.
- Baker, R., Buckland, A., & Sheaves, M. (2013). Fish gut content analysis: Robust measures of diet composition. *Fish and Fisheries*, *15*, 170–177. <https://doi.org/10.1111/faf.12026>
- Balasi, M., Casey, C., & Van Valkenburgh, B. (2018). Dietary specialization is linked to reduced species durations in North American fossil canids. *Royal Society Open Science*, *5*, 171861. <https://doi.org/10.1098/rsos.171861>
- Bellwood, D. R., Hoey, A. S., Bellwood, O., & Goatley, C. H. R. (2014). Evolution of long-toothed fishes and the changing nature of fish-benthos interactions on coral reefs. *Nature Communications*, *5*, 3144. <https://doi.org/10.1038/ncomms4144>
- Bellwood, D. R., & Wainwright, P. C. (2002). The History and Biogeography of Fishes on Coral Reefs. In P. F. Sale (Ed.), *Coral reef fishes: dynamics and diversity in a complex ecosystem*. (pp. 5–32). San Diego: Academic Press.
- Bolnick, D. I., Hirsch, P. E., Lauber, C. L., Knight, R., Caporaso, J. G., & Svanback, R. (2014). Individuals' diet diversity influences gut microbial diversity in a freshwater fish (threespine stickleback and Eurasian perch). *Ecology Letters*, *17*, 979–987.
- Casey, J. M., Meyer, C. P., Morat, F., Brand, S. J., Planes, S., & Parravicini, V. (2019). Reconstructing hyperdiverse food webs: Gut content metabarcoding as a tool to disentangle trophic interactions on coral reefs. *Methods in Ecology and Evolution*, *10*, 1157–1170. <https://doi.org/10.1111/2041-210X.13206>
- Chesson, P. (2000). Mechanisms of maintenance of species diversity. *Annual Review of Ecology and Systematics*, *31*, 343–366. <https://doi.org/10.1146/annurev.ecolsys.31.1.343>
- Clarke, K. R., Somerfield, P. J., & Gorley, R. N. (2008). Testing of null hypotheses in exploratory community analyses: Similarity profiles and biota-environment linkages. *Journal of Experimental Marine Biology and Ecology*, *366*, 56–69. <https://doi.org/10.1016/j.jembe.2008.07.009>
- Clarke, R., Ransom, H. W., Wang, A., Xuan, J., Liu, M. C., Gehan, E. A., & Wang, Y. (2008). The properties of high-dimensional data spaces: Implications for exploring gene and protein expression data. *Nature Reviews Cancer*, *8*, 37–49. <https://doi.org/10.1038/nrc2294>
- Clements, K., Angert, E., Montgomery, W. L., & Choat, J. R. (2014). Intestinal microbiota in fishes: What's known and what's not.

- Molecular Ecology*, 23, 1891–1898. <https://doi.org/10.1111/mec.12699>
- Cordner, E. G. (2013). *Diet Shifts in coral reef fishes in response to changes in food availability and predator density*. San Diego, CA: University of California, ProQuest Dissertations Publishing, 1540928.
- Cottrah, R. D., Noyes, P., & Relyea, R. A. (2015). An empirical test of stable species coexistence in an amphipod species complex. *Oecologia*, 178, 819–831. <https://doi.org/10.1007/s00442-015-3262-1>
- Dam, H. G., Peterson, W. T., & Okubo, A. (1991). A simple mathematical analysis of the limitations to inferring feeding behavior of zooplankton from gut content. *Marine Ecology Progress Series*, 69, 41–45.
- Darwin, C. (1859). *On the origin of species*. London: John Murray.
- David, L. A., Maurice, C. F., Carmody, R. N., Gootenberg, D. B., Button, J. E., Wolfe, B. E., ... Turnbaugh, P. J. (2014). Diet rapidly and reproducibly alters the human gut microbiota. *Nature*, 505, 559–563.
- Deagle, B. D., Jarman, S. N., Coissac, E., Pompanon, F., & Taberlet, P. (2014). DNA metabarcoding and cytochrome c oxidase subunit I marker: Not a perfect match. *Biology Letters*, 10, 20140562.
- Deagle, B. E., Thomas, A. C., McInnes, J. C., Clarke, L. J., Vesterinen, E. J., Clare, E. L., ... Eveson, J. P. (2018). Counting with DNA in metabarcoding studies: How should we convert sequence reads to dietary data? *Molecular Ecology*, 28, 391–406.
- DeMartini, E. E. (1996). Sheltering and foraging substrate uses of the arc-eye hawkfish *Paracirrhites arcatus* (Pisces: Cirrhitidae). *Bulletin of Marine Science*, 58, 826–837.
- DeMartini, E. E., & Donaldson, T. J. (1996). Color Morph-habitat relations in the Arc-Eye Hawkfish *Paracirrhites arcatus* (Pisces: Cirrhitidae). *Copeia*, 1996(2), 362–371. <https://doi.org/10.2307/1446852>
- Donaldson, T. J. (1986). Distribution and species richness patterns of Indo-West Pacific Cirrhitidae: support for Woodland's hypothesis. In T. Uyeno, R. Arai, T. Taniuchi, & K. Matsuura (Eds.), *Proceedings of the Second International Conference on Indo-Pacific fishes* (pp. 623–628). Tokyo: Ichthyological Society of Japan.
- Donaldson, T. J. (1988). *Social Organization, Reproductive Behavior, and Zoogeography of Some Indo-Pacific Hawkfishes (Cirrhitidae)*. LSU Historical Dissertations and Theses, 4633.
- Donaldson, T. J. (1990). Reproductive behavior and social organization of some Pacific hawkfishes (Cirrhitidae). *Japanese Journal of Ichthyology*, 36, 439–458. <https://doi.org/10.1007/BF02905464>
- Donaldson, T. J. (1997). Assessing phylogeny, historical ecology, and the mating systems of hawkfishes (Cirrhitidae). *Proc. 5th Indo-Pacific Fish Conference*, Noumea.
- Donaldson, T. J. (2002). Phylogeny, reef fish conservation biology, and the Live Reef Fish Trade. *Fisheries science*, 68(1), 143–147.
- Ekman, S. (1953). *Zoogeography of the sea*. London: Sidgwick and Jackson.
- Faith, J. J., Guruge, J. L., Charbonneau, M., Subramanian, S., Seedorf, H., Goodman, A. L., ... Gordon, J. I. (2013). The long-term stability of the human gut microbiota. *Science*, 341, 6141. <https://doi.org/10.1126/science.1237439>
- 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. *Journal of Biogeography*, 40, 1638–1648. <https://doi.org/10.1111/jbi.12126>
- Gotelli, N. (1997). Competition and coexistence of larval ant lions. *Ecology*, 78, 1761–1773. [https://doi.org/10.1890/0012-9658\(1997\)078\[1761:CACOLA\]2.0.CO;2](https://doi.org/10.1890/0012-9658(1997)078[1761:CACOLA]2.0.CO;2)
- Hale, V. L., Tan, C. L., Niu, K., Yang, Y., Knight, R., Zhang, Q., ... Amatao, K. R. (2018). Diet versus phylogeny: A comparison of gut microbiota in captive colobine monkey species. *Microbial Ecology*, 75, 515–527. <https://doi.org/10.1007/s00248-017-1041-8>
- Hardin, G. (1960). The competitive exclusion principle. *Science, New Series*, 131(3409), 1292–1297.
- Hebert, P. D. N., Ratnasingham, S., & deWaard, J. R. (2003). Barcoding animal life: cytochrome c oxidase subunit 1 divergences among closely related species. *Proceedings of the Royal Society B Biological Sciences*, 270, S96–S99.
- Hobbs, J. P., Frisch, A. J., Allen, G. R., & Van Herwerden, L. (2008). Marine hybrid hotspot at Indo-Pacific biogeographic border. *Biology Letters*, 5, 258–261. <https://doi.org/10.1098/rsbl.2008.0561>
- Jablonski, D. (2008). Extinction and the spatial dynamics of biodiversity. *Proceedings of the National Academy of Sciences*, 105(Suppl 1), 11528–11535.
- Keeler, E., & Berwick, D. (1976). Effects of pooled samples. *Health Laboratory Science*, 13(2), 121–128.
- Kelly, L. W., Williams, G. J., Barott, K. L., Carlson, C. A., Dinsdale, E. A., Edwards, R. A., ... Rohwer, F. (2014). Local genomic adaptation of coral reef-associated microbiomes to gradients of natural variability and anthropogenic stressors. *Proceedings of the National Academy of Sciences*, 111, 10227–10232. <https://doi.org/10.1073/pnas.1403319111>
- Knudsen, J. W. (1967). Trapezia and Tetralia (Decapoda, Brachyura, Xanthidae) as Obligate Ectoparasites of Pocilloporid and Acroporid Corals. *Pacific Science*, 21, 51–57.
- Leray, M., Agudelo, N., Mills, S. C., & Meyer, C. P. (2013). Effectiveness of annealing blocking primers versus restriction enzymes for characterization of generalist diets: Unexpected prey revealed in the gut contents of two coral reef fish species. *PLoS ONE*, 8(4), e58076. <https://doi.org/10.1371/journal.pone.0058076>
- Leray, M., Boehm, J. T., Mills, S. C., & Meyer, C. P. (2012). Moorea BIOCDE barcode library as a tool for understanding predator prey interactions – insights into the diet of common predatory reef fishes. *Coral Reefs*, 31, 383–388.
- Leray, M., Meyer, C. P., & Mills, S. C. (2015). Metabarcoding dietary analysis of coral dwelling predatory reef demonstrates the minor contribution of coral mutualists to their highly partitioned, generalist diet. *PeerJ*, 3, e1047.
- Ley, R. E., Turnbaugh, P. J., Klein, S., & Gordon, J. I. (2006). Human gut microbes associated with obesity. *Nature*, 444(7122), 1022–1023.
- Lim, Y. W., Cuevas, D. A., Silva, G. G. Z., Aguinaldo, K., Dinsdale, E. A., Haas, A. F., ... Edwards, R. A. (2014). Sequencing at sea: challenges and experiences in Ion Torrent PGM sequencing during the 2013 Southern Line Islands Research Expedition. *PeerJ*, 2, e520. <https://doi.org/10.7717/peerj.520>
- Machida, R. J., Leray, M., Ho, S. L., & Knowlton, N. (2017). Metazoan mitochondrial gene sequence reference datasets for taxonomic assignment of environmental samples. *Scientific data*, 4(1), 1–7.
- Madduppa, H. H. (2012). Fish biodiversity in coral reefs and lagoon at the Maratua Island, East Kalimantan. *Biodiversitas, Journal of Biological Diversity*, 13(3), 145–150. <https://doi.org/10.13057/biodiv/d130308>
- McNair, K., & Edwards, R. A. (2015). GenomePeek – an online tool for prokaryotic genome and metagenome analysis. *PeerJ*, 3, e1025. <https://doi.org/10.7717/peerj.1025>
- Meyer, F., Paarmann, D., D'Souza, M., Olson, R., Glass, E. M., Kubal, M., ... Edwards, R. A. (2008). The metagenomics RAST server – a public resource for the automatic phylogenetic and functional analysis of metagenomes. *BMC Bioinformatics*, 9, 386. <https://doi.org/10.1186/1471-2105-9-386>
- Miller, M. A., Pfeiffer, W., & Schwartz, T. (2010). Creating the CIPRES Science Gateway for inference of large phylogenetic trees. In *Proceedings of the Gateway Computing Environments Workshop (GCE)*. (pp. 1–8). New Orleans, LA: IEEE.
- Muegge, B. D., Kuczynski, J., Knights, D., Clemente, J. C., Gonzalez, A., Fontana, L., ... Gordon, J. I. (2011). Diet drives convergence in gut microbiome functions across mammalian phylogeny and within humans. *Science*, 332, 970–974. <https://doi.org/10.1126/science.1198719>
- Myers, R. F. (1989). *Micronesian reef fishes: A practical guide to the identification of the coral reef fishes of the tropical central and western Pacific*. Barrigada: Coral graphics.
- Oksanen, J. F., Blanchet, F. G., & Kindt, R. et al (2012) *vegan: Community ecology package*. R package version 2.0-4. Retrieved from <http://cran.r-project.org/index.html>

- Parris, J. D., Morgan, M. M., & Stewart, F. J. (2019). Feeding rapidly alters microbiome composition and gene transcription in the clownfish gut. *Applied and Environmental Microbiology*, 85, e02479–e2518.
- Prokopowich, C. D., Gregory, T. R., & Crease, T. J. (2003). The correlation between rDNA copy number and genome size in eukaryotes. *Genome*, 46, 48–50. <https://doi.org/10.1139/g02-103>
- Randall, J. E. (1963). Review of the Hawkfishes (Family Cirrhitidae). *Proceedings of the United States National Museum*, 114(3472), 389–451. <https://doi.org/10.5479/si.00963801.114-3472.389>
- Randall, J. E., Allen, G. R., & Steene, R. C. (1990). *Fishes of the great barrier reef and coral sea* (p. 506). Honolulu: University of Hawaii Press.
- Ronquist, F., & Huelsenbeck, J. P. (2003). MrBayes 3: Bayesian phylogenetic inference under mixed models. *Bioinformatics*, 19, 1572–1574. <https://doi.org/10.1093/bioinformatics/btg180>
- Sadovy, Y., & Donaldson, T. J. (1995). Sexual pattern of *Neocirrhites armatus* (Cirrhitidae) with notes on other hawkfish species. *Environmental Biology of Fishes*, 42, 143–150.
- Sandin, S. A., & Zgliczynski, B. J. (2015). Inverted trophic pyramids. In C. Mora (Ed.), *Ecology of Fishes on Coral Reefs* (pp. 247–251). Cambridge: Cambridge University Press.
- Schmieder, R., & Edwards, R. (2011). Quality control and preprocessing of metagenomic datasets. *Bioinformatics*, 27, 863–864. <https://doi.org/10.1093/bioinformatics/btr026>
- Scott, K. P., Gratz, S. W., Sheridan, P. O., Flint, H. J., & Duncan, S. H. (2013). The influence of diet on the gut microbiota. *Pharmacological Research*, 69(1), 52–60.
- Shisterman, E. F., & Vexler, A. (2008). To pool or not to pool, from whether to when: Applications of pooling to biospecimens subject to a limit of detection. *Paediatric and Perinatal Epidemiology*, 22(5), 486–496. <https://doi.org/10.1111/j.1365-3016.2008.00956.x>
- Siepielski, A. M., & McPeck, M. A. (2010). On the evidence for species coexistence: A critique of the coexistence program. *Ecology*, 91, 3153–3164. <https://doi.org/10.1890/10-0154.1>
- Siu, G., Bacchet, P., Bernardi, G., Brooks, A. J., Carlot, J., Causse, R., ... Galzin, R. (2017). Shore fishes of French Polynesia. *Cybius*, 41, 245–278.
- Smriga, S., Sandin, S. A., & Azam, F. (2010). Abundance, diversity, and activity of microbial assemblages associated with coral reef fish guts and feces. *FEMS Microbiology Ecology*, 73, 31–42. <https://doi.org/10.1111/j.1574-6941.2010.00879.x>
- Srivathsan, A., Sha, J. C. M., Vogler, A. P., & Meier, R. (2014). Comparing the effectiveness of metagenomics and metabarcoding for diet analysis of a leaf-feeding monkey (*Pygathrix nemaeus*). *Molecular Ecology Resources*, 15, 250–261.
- Stamatakis, A. (2014). RAXML version 8: A tool for phylogenetic analysis and post-analysis of large phylogenies. *Bioinformatics*, 30, 1312–1313. <https://doi.org/10.1093/bioinformatics/btu033>
- Sullam, K. E., Essinger, S. D., Lozupone, C. A., O'Connor, M. P., Rosen, G. L., Knight, R., ... Russell, J. A. (2012). Environmental and ecological factors that shape the gut bacterial communities of fish: A meta-analysis. *Molecular Ecology*, 21, 3363–3378. <https://doi.org/10.1111/j.1365-294X.2012.05552.x>
- Swofford, D. L. (2003). *PAUP*. Phylogenetic Analysis Using Parsimony (*and Other Methods)*. Version 4. Sunderland, MA: Sinauer Associates.
- Uchii, K., Matsui, K., Yonekura, R., Tani, K., Kenzaka, T., Nasu, M., & Kawabata, Z. (2006). Genetic and physiological characterization of the intestinal bacterial Microbiota of Bluegill (*Lepomis macrochirus*) with three different feeding habits. *Microbial Ecology*, 51, 277–284.
- Velland, M. (2010). Conceptual synthesis in community ecology. *The Quarterly Review of Biology*, 85, 183–206. <https://doi.org/10.1086/652373>
- Ward, E. J., Semmens, B. X., & Schindler, D. E. (2010). Including source uncertainty and prior information in the analysis of stable isotope mixing models. *Environmental Science & Technology*, 44, 4645–4650. <https://doi.org/10.1021/es100053v>
- Weinberg, C. R., & Umbach, D. M. (1999). Using pooled exposure assessment to improve efficiency in case-control studies. *Biometrics*, 55(3), 718–726. <https://doi.org/10.1111/j.0006-341X.1999.00718.x>
- Whitaker, D., & Christman, M. (2010). *clustsig: Significant cluster analysis*. R package Version 1.0. Retrieved from www.r-project.org
- Whitney, J. L., Bowen, B. W., & Karl, S. A. (2018). Flickers of speciation: Sympatric colour morphs of the arc-eye hawkfish, *Paracirrhites arcatus*, reveal key elements of divergence with gene flow. *Molecular Ecology*, 27, 1479–1493.
- Whitney, J. L., Donahue, M. J., & Karl, S. A. (2018). Niche divergence along a fine-scale ecological gradient in sympatric color morphs of a coral reef fish. *Ecosphere*, 9, e02015. <https://doi.org/10.1002/ec2.2015>
- Wong, S., & Rawls, J. (2012). Intestinal microbiota composition in fishes is influenced by host ecology and environment. *Molecular Ecology*, 21, 3100–3102. <https://doi.org/10.1111/j.1365-294X.2012.05646.x>
- Woodland, D. J. (1983). Zoogeography of the Siganidae (Pisces): an interpretation of distribution and richness patterns. *Bulletin of marine science*, 33(3), 713–717.
- Xia, J., Lin, G., Fu, G. H., Wan, Z. Y., Lee, M., Wang, L., ... Yue, G. H. (2014). The intestinal microbiome of fish under starvation. *BMC Genomics*, 5, 266. <https://doi.org/10.1186/1471-2164-15-266>
- Zgliczynski, B. J., Williams, G. J., Hamilton, S. L., Cordner, E. G., Fox, M. D., Eynaud, Y., ... & Sandin, S. A. (2019). Foraging consistency of coral reef fishes across environmental gradients in the central Pacific. *Oecologia*, 191(2), 433–445.
- Zhang, Z., Li, D., Refaey, M. M., Xu, W., Tang, R., & Li, L. (2018). Host age affects the development of Southern Catfish Gut bacterial community divergent from that in the food and rearing water. *Frontiers in Microbiology*, 9, 495. <https://doi.org/10.3389/fmicb.2018.00495>

SUPPORTING INFORMATION

Additional supporting information may be found online in the Supporting Information section.

How to cite this article: French BJ, Lim YW, Zgliczynski BJ, Edwards RA, Rohwer F, Sandin SA. Decoding diversity in a coral reef fish species complex with restricted range using metagenomic sequencing of gut contents. *Ecol Evol*. 2020;10:3413–3423. <https://doi.org/10.1002/ece3.6138>