Marine Biology (2022) 169:32
https://doi.org/10.1007/s00227-022-04018-x

ORIGINAL PAPER

Maintenance of species boundaries within social aggregations of ecologically similar goby sister species

Jason D. Selwyn[1] · Elizabeth P. Hunt[1] · David S. Portnoy[1] · J. Derek Hogan[1]

Received: 28 June 2021 / Accepted: 4 January 2022
© The Author(s), under exclusive licence to Springer-Verlag GmbH Germany, part of Springer Nature 2022

Abstract

The maintenance of species boundaries when opportunities for admixture are abundant, is a poorly understood phenomenon for many taxa. While many mechanisms for maintaining species boundaries have been described their relative importance depends largely on the particulars of the system in question. Aggregating social behavior can be a means to keep sympatric sister species distinct if it leads to segregation during reproduction. The widespread Caribbean reef gobies *Coryphopterus personatus* and *C. hyalinus* are sympatric sister species with nearly identical morphology that spend their entire adult lives in shoals in which reproduction occurs. To date no studies have investigated whether shoals are species-specific, which would be expected if aggregating behavior helps to maintain species boundaries. To address this, the species of individual fishes collected from 16 shoals were identified using morphology, mitochondrial sequence data, and microsatellite allele frequencies. Levels of admixture between the species were also assessed. Shoals were generally composed of both species in similar proportions to their relative abundances on the reef, where the shoals were found, indicating that the species are not behaviorally segregating. For most specimens, morphological, mitochondrial, and nuclear data were congruent with a single species, but 18 individuals showed disagreements with microsatellite genotypes of 16 suggesting some level of historic/contemporary admixture. Of these, two were identified as likely first- or second-generation hybrids or backcrosses. Despite co-occurrence and evidence of some gene flow, the two species show little admixture overall suggesting that microscale differences in breeding site selection, allochrony, and/or cryptic mate choice may play an important role in the maintenance of species boundaries despite cooccurrence well within the range typically thought of as sympatry.

Keywords Social aggregations · Hybridization · Species boundaries · Sympatry · Isolating mechanisms

Introduction

Understanding processes that maintain species boundaries in sympatry is a major focus of evolutionary ecology. Many ecological and evolutionary mechanisms have been identified in the maintenance of species boundaries, with their relative importance depending on the specifics of the study system (Harrison et al. 2017). This is particularly true amongst closely related species living in sympatry, which can lead to frequent opportunities for hybridization. Research has found that allochronic, ecological, and behavioral isolation, as well as gametic incompatibility, are frequently important (Coyne and Orr 2004; Harrison et al. 2017). Allochronic isolation occurs when gamete release and production are offset in time, leading to a reduced opportunity for interaction between gametes of sympatrically distributed species (Levitan et al. 2011; Bouwmeester et al. 2021) and can occur on scales from hours to years (Knowlton et al. 1997; Rosser 2015; Tarpey et al. 2017). Ecological isolation occurs when species utilize different ecological niches which subsequently minimizes opportunities for mating interactions (Bovbjerg 1970). Behavioral isolation occurs when differences in behavior develop that impact the likelihood of heterospecific mating and can include differences in mating behavior (Parchman et al. 2013) and/or formation of spatially segregated social groups (Gerhardt 1974; Diabaté et al. 2009). Gametic incompatibility occurs when a viable zygote is not formed during fertilization (Rawson et al. 2003).

Responsible Editor: O. Puebla.

✉ Jason D. Selwyn
jdselwyn@gmail.com

[1] Department of Life Sciences, Texas A&M University - Corpus Christi, Corpus Christi, TX, USA

Published online: 22 January 2022

Sister species need not be isolated completely, and a continuum of states exists between the homogenization of once distinct gene pools to complete reproductive isolation, as levels of gene flow between species decrease and larger portions of genomes become isolated (Kopp and Frank 2005; Harrison and Larson 2014). Furthermore, as incipient species form, recurrent gene flow can lead to further diversification (through reinforcement or differential gene flow) or homogenization (Mallet 2005; Abbott et al. 2013). If species potentially interact across large geographic areas, there may be differences in levels of contemporary gene flow related to local conditions (Muhlfeld et al. 2009; Gagnaire et al. 2013). For example, in colonial nesting waterbirds, such as gulls, reproductive barriers are incomplete in areas of overlapping breeding habitat, and species complexes are known to form (Liebers et al. 2004). Alternatively, other isolating mechanisms (e.g., gametic isolation, mate recognition) may maintain species boundaries despite close association. For example, sub-social colonial spiders in the genus *Chikunia* form mixed-species colonies and have been observed to indiscriminately provide care for heterospecific broods and yet appear to maintain evolutionary independence (Grinsted et al. 2012; Smith et al. 2019). Because multiple mechanisms likely operate simultaneously, careful consideration of individual processes and their relative contribution to the cessation/interruption of gene flow is required to gain a holistic view of species boundaries (Coyne and Orr 2004).

For marine species with external fertilization gamete incompatibility and temporal/spatial offsets in gamete release, are thought to be particularly important (Levitan et al. 2004; Ohki et al. 2015), because gametes are released into a dispersive environment (Babcock et al. 1994) and may remain viable for hours to days (Williams and Bentley 2002). Social behavior seen in many mobile marine species including formation of spawning aggregations, schooling, and/or monogamous pairing (Domeier and Colin 1997; Pavlov and Kasumyan 2000; Whiteman and Côté 2004) can further decrease the opportunity for interaction between heterospecific gametes. However, many social units feature heterospecifics and the observation of incomplete isolation of interacting marine species is becoming increasingly common (Miranda et al. 2010; Montanari et al. 2012).

Gobies (Gobioidei) are small bodied short-lived fishes and comprise one of the most diverse sub-orders of vertebrates (Nelson et al. 2016). Habitat differentiation appears to be a primary driving force behind diversification and speciation (Thacker 2009) within the gobies, and sister taxa often segregate based on microhabitat (Brandl et al. 2018). Strong natural selection associated with habitat specialization can maintain species boundaries (Rice 1987; Teske et al. 2019; Öhlund et al. 2020) and is hypothesized to be important within the taxon (Brandl et al. 2018). For example, mudskipper diversification appears to have proceeded by differentiation into ecological guilds characterized by differences in salinity and water quality tolerance, as well as degree of terrestriality (Polgar et al. 2010). On coral reefs, microhabitat utilization is thought to be a contributing factor to the degree of diversification seen in the genus *Elacatinus*, which initially differentiated into sponge and coral-dwelling groups, with later diversification based on other ecological characteristics (Taylor and Hellberg 2005; Colin 2010). In a similar manner, diversification in the genus *Eviota* seems to be related to the degree of association with specific microhabitats including coral, rubble, or sand (Tornabene et al. 2013).

Most species of Caribbean gobies are solitary, however, in the genus *Coryphopterus* two sister species, *C. personatus* and *C. hyalinus* occupy social aggregations throughout juvenile and adult life stages (Allsop and West 2004). The genus *Coryphopterus* is a relatively recent radiation of fourteen species, one from the eastern Pacific with the remaining thirteen in the western Atlantic, arising within the last 30 million years (Baldwin and Robertson 2015; Thacker 2015). The two species, *C. personatus* and *C. hyalinus*, were initially split based on counts and positions of anterior head pores of the cephalic lateralis sensory system, with *C. hyalinus* having laterally paired pores and *C. personatus* having a single median pore (Böhlke and Robins 1962). Phylogenetic analysis consistently resolves *C. personatus* and *C. hyalinus* as sister species with a relatively recent common ancestor (Baldwin et al. 2009). In addition, their status as species is supported by 7.14% (6.79–7.65%) sequence divergence between the species at the mitochondrially encoded COI gene, as compared to 0.06% and 0.14% sequence divergence within each species, respectively (Baldwin et al. 2009). Despite these genetic differences the ecology of these two species is similar and their geographic distribution is nearly completely overlapping throughout the entirety of the Greater Caribbean (Robertson and Van Tassell 2019), leading many researchers to lump them together in ecological studies (e.g.,Serna Rodríguez et al. 2016; Chagaris et al. 2017). While *C. hyalinus* is generally collected from slightly deeper depths, the overall depth range of the two species shows near complete overlap. At Turneffe Atoll both species are frequently observed between 0 and 27 m with *C. hyalinus* occurring at depths greater than 30 m (Greenfield and Johnson 1999). Caribbean wide depth ranges for each are between 1 and 52 m, with *C. personatus* being observed to 70 m depths (Baldwin and Robertson 2015).

In addition, the species are reproductively similar, as both are protogynous hermaphrodites that lay and fertilize eggs within the reef structure and form large shoals of up to tens to thousands of individuals, with shoals in the current study area typically composed of fewer than 100 individuals (Böhlke and Robins 1962; Robertson and Justines 1982; Cole and Robertson 1988; Selwyn et al. 2021). These shoals

are spatially discrete, temporally stable aggregations on the reef, which reform rapidly when disturbed and exist in the same location for multiple days (J. Selwyn pers. obs). Shoals serve a number of purposes, including reproduction, with the ratio of males to females within a shoal influenced by shoal density and the ability of males to monopolize mates (Allsop and West 2004). In both species mating occurs between pairs of individuals within reef crevices, with males guarding and aerating the fertilized nest (Thresher 1984; Gardner 2000). Unfortunately, little else is known about the reproductive behaviors of these species in nature. The reproductive life span is short (~ 100 days; Beeken et al. 2021) and spawning is frequent (every 7–10 days; Gardner 2000). Because opportunity for adult dispersal is limited as adults are small and not highly mobile; individuals likely spend their entire reproductive life within a single shoal (Selwyn et al. 2016).

The main objective of this study was to identify whether species-specific aggregation was the mechanism maintaining species boundaries between the sympatric gobies, *C. personatus* and *C. hyalinus*. We hypothesized that *C. personatus* and *C. hyalinus* would form spatially segregated shoals composed of a single species or both species but in ratios significantly different than the background. If shoals contain both species, there is the potential for admixture which can be assessed using genetic techniques. Therefore, samples of individuals were taken from multiple, spatially explicit shoals spread across a single reef system, and analyzed using molecular and morphological characters to determine individual species identity, shoal composition, and degree of admixture.

Materials and methods

Sample collection

Individual *Coryphopterus hyalinus/personatus* were collected by divers on SCUBA using hand nets from a single reef (17° 16′ 40.55″ N, 87° 48′ 18.08″ W, Fig. 1) in Turneffe Atoll, Belize, in August 2014. Turneffe Atoll is

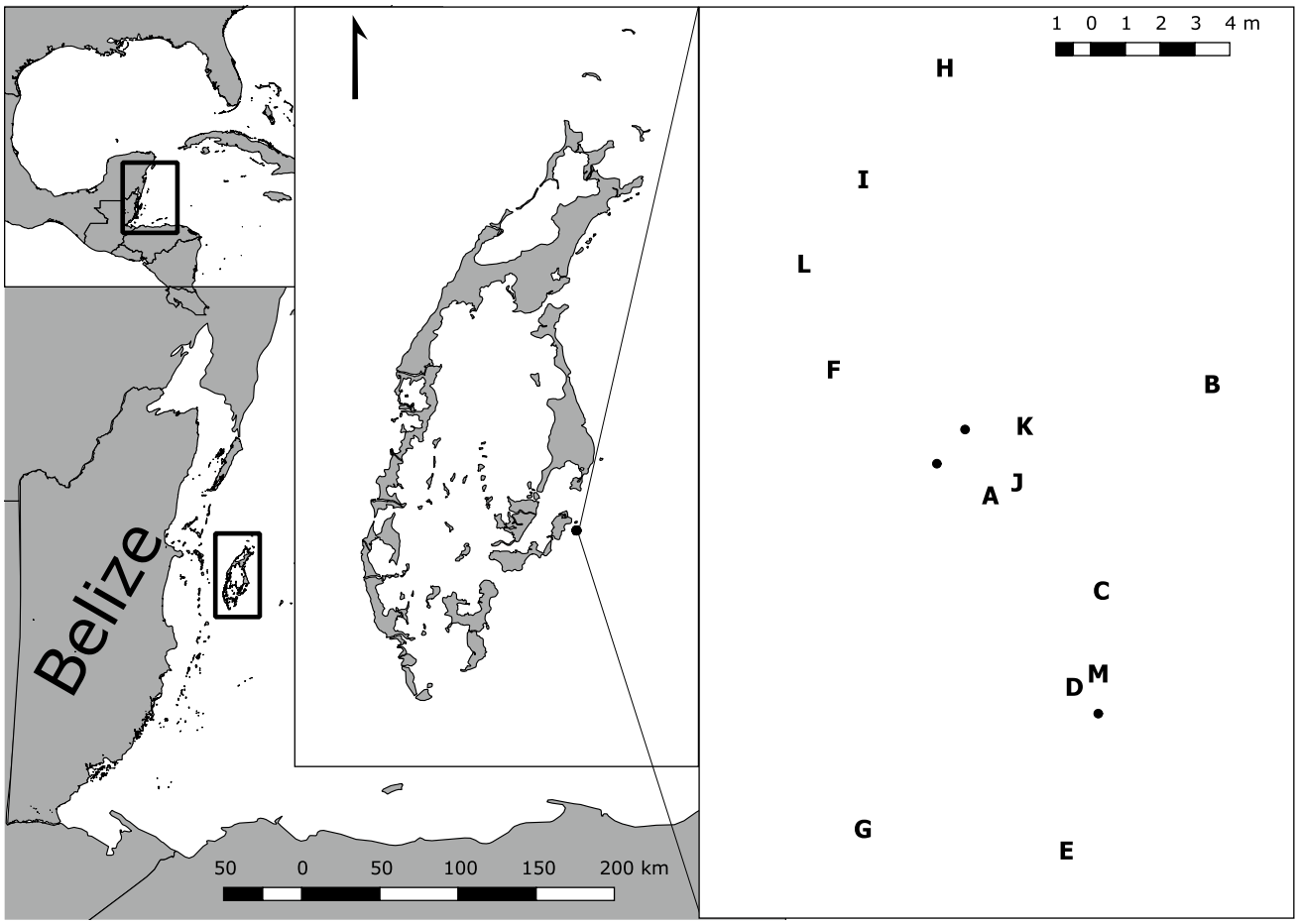

Fig. 1 Map of Turneffe Atoll, Belize including sampling site (17° 16′ 40.55 N, 87° 48′ 18.08 W) with inset showing spatial arrangement of sampled shoals on the reef. Letters indicate shoal ID and are consistent with Fig. 6. Points indicate shoals which were sampled but did not have individuals identified using both molecular techniques

composed of numerous mangrove islands approximately 9–23 km offshore from the main Belize Barrier Reef. This area has records of both *C. hyalinus* and *C. personatus* and is composed of suitable forereef habitat in depths, where both species are commonly found (Greenfield and Johnson 1999). The habitat composition is typical of shallow, windward forereef locations in Turneffe Atoll (for more detailed description see: Garcia and Holtermann 1998). This reef was selected for study of small spatial scale interactions of *C. personatus* and *C. hyalinus*, as the depth (16 m) of this reef is where the maximum number of both *C. hyalinus* and *C. personatus* have been observed at Turneffe Atoll previously (Greenfield and Johnson 1999), located between the shallow water coral heads preferred by *C. personatus*, and the deeper coral walls preferred by *C. hyalinus* (Victor 2019).

A total of 428 individual *Coryphopterus sp.* were collected from 16 shoals. Fish were humanely euthanized using buffered MS222 and stored in 95% non-denatured ethanol. Individual shoals were kept separate during and after collection. All collections were performed in accordance with the ethical guidelines of Texas A&M University—Corpus Christi (TAMUCC-AUP-05-14) and in compliance with standards outlined in the US National Research Council's Guide for the Care and Use of Laboratory Animals. Collections were made with the express permission of the government of Belize (Aquatic Scientific Research Permit 000044-13).

Morphology

To distinguish between *C. personatus* and *C. hyalinus* using morphology, anterior interorbital cephalic (AIC) pores were counted using a dissecting microscope; *Coryphopterus personatus* have one pore, while *C. hyalinus* have two (Böhlke and Robins 1962). Because these pores only develop in larger individuals (> 10 mm SL), morphological identification was attempted only with individuals larger than this threshold (Victor 2019) and a total of 134 (31.3%) specimens were examined.

Mitochondria

The mitochondrial gene cytochrome c oxidase subunit 1 (COI) was amplified from DNA extracted from fin and muscle tissue from the caudal end of each fish using either E.Z.N.A.® DNA extraction kit (Omega Bio-tek) or Chelex extraction using a multiplex reaction with four universal fish primers (de Lamballerie et al. 1992; Ward et al. 2005): FishF1 (5′-TCAACCAACCACAAAGAGATTGGCAC-3′), FishF2 (5′-TCGACTAATCATAAAGATATCGGCAC-3′) and FishR1 (5′-TAGACTTCTGGGTGGCCAAAGAATCA-3′), FishR2 (5′-ACTTCAGGGTGACCGAAGAATCAGAA-3′). Each 30 µl reaction contained 1 X buffer (pH 8.5), 1.5 mM $MgCl_2$, 0.20 mM dNTPs each, 0.04% Tween, 250 nM forward and reverse primers each (F1 and F2, R1 and R2), 1.0 U *Taq* polymerase and 1.0 µl of template. PCR amplification was run with initial denaturing at 95 °C for 2 min, followed by 35 cycles of denaturation at 95 °C for 1 min, annealing at 50–55 °C for 1 min, and elongation at 72 °C for 1 min. A final round of elongation was run at 72 °C for 10 min. Amplicons were cleaned using 0.7X Mag-Bind®Total Pure NGS beads (Omega Bio-Tek) and sequenced at the Genomics Core Lab at Texas A&M University—Corpus Christi or at Retrogen, Inc. (San Diego, CA) on 96-capillary ABI 3730xl Genetic Analyzer (Applied Biosystems Inc.). Chromatographs were edited by eye using SEQUENCHER v.5.4.6 (GeneCodes Corporation). All COI sequences from this study can be found on GenBank (accession numbers MT784949–MT785286).

To determine species identity from COI sequences, a haplotype network was created using individuals sequenced from this sampling along with 11 sequences of each species downloaded from GenBank, which had been morphologically identified by Baldwin et al (2009; Accession numbers: GQ367313—GQ367334). Sequences from this study and those from Baldwin et al. (2009) were then aligned using CLUSTAL W implemented in MSA (Thompson et al. 1994; Bodenhofer et al. 2015) and trimmed to retain a core region of 547 bp for all individuals. A median joining network was created using POPART (Bandelt et al. 1999; Leigh and Bryant 2015) and used to assign individuals to species. For individuals, where morphological and the genetic identity were incongruent, specimens were reexamined microscopically without prior knowledge of whether COI sequences were consistent with *C. hyalinus* or *C. personatus*. Net genetic divergence between species and average genetic distance within species was calculated using STRATAG (Nei and Kumar 2000; Archer et al. 2017).

Microsatellites

Nine microsatellite loci were amplified for 384 specimens (89.7%; Table S1; Hepburn et al. 2005; Hogan et al. 2010). Each 10 µl reaction contained 1X buffer (pH 8.5), 3.0–4.5 mM $MgCl_2$, 0.8 mM each dNTPs, 100–500 nM fluorescent labelled forward and unlabelled reverse primers, 0.5 U *Taq* polymerase, and 1.0 µl of template (Table S1). PCR amplification was run using a touchdown protocol with an initial denaturation of 94 °C for 3 min, followed by a 40–50 cycles (Table S1) of denaturation for 15 s at 94 °C, annealing for 45 s at 68 to 64 °C—58 to 52 °C (Table S1), and elongation for 30 s at 72 °C, followed by 5 min elongation at 72 °C. Amplicons were analysed at the Texas A&M University—Corpus Christi Genomics core lab on a 96-capillary ABI 3730xl Genetic Analyzer with the Liz 600® (Applied Biosystems Inc.) size standard. Size polymorphisms were

scored by eye using GeneMarker 2.6.4 software (SoftGenetics Inc.).

To determine species identity of individuals based on the microsatellite loci, Structure was run using the admixture model with correlated allele frequencies (Pritchard et al. 2000). Prior to running any models two loci (COPE10 and CPER52) were removed due to failure to amplify in more than 20% of individuals. The models were run for all values of K (number of distinct clusters) between 1 and 15, using 1,000,000 burn-in iterations followed by 10,000,000 sampling iterations with a thinning interval of 100 and 10 replicate runs. Mixing, proper exploration of parameter space, and chain convergence were confirmed by visually inspecting trace plots, which show parameter values across MCMC steps, and ensuring the $\widehat{R}$ value, a measure of chain convergence, equaled one (Fig. S1; Vehtari et al. 2021). The optimal value of K was then determined using the Evanno method (Evanno et al. 2005). After performing Structure analysis, species identity as determined from COI was matched with each individual's cluster assignments, and clusters were associated with one species or the other, based on the majority of individuals in a cluster having the same COI species identity.

Joint species identification

Because DNA quality of some specimens was poor, not all fish could be identified using both molecular methods (Fig. 2). Therefore, a final species identification was made only when both COI and microsatellite data were available ($n = 321$, number of shoals = 13). To visualize differences among specimens based on all three identification methods, a principal components analysis (PCA) combining morphological, mitochondrial, and microsatellite data was performed. Morphological identification was not required as the informative character does not develop until the fish are larger than 10 mm standard length (Victor 2019). The size at which AIC pores become informative characters is coincidentally similar to the minimum size observed for the transition from female to male (~ 13 mm TL, Cole and Robertson 1988, TL = − 1.5 + 1.3 × SL, 10 mm SL = 11.1 mm TL; 95% prediction interval = 10.1–12.2, unpublished data) though it should be noted there are many mature females larger than this size (Cole and Robertson 1988). Individuals for which microsatellite and mtDNA-based identification disagreed or one of the two marker types failed to amplify, were excluded from analysis of shoal composition ($n = 107$).

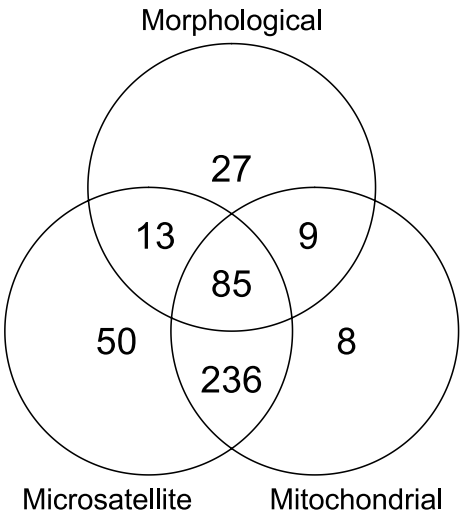

Fig. 2 Venn diagram showing the methods of identification used for the samples. The number is the total number of fish identified using that combination of methods

Characterizing genetic diversity

After making final species identifications, microsatellites were used to characterize variation within and between *C. hyalinus* and *C. personatus*, excluding all samples with uncertain species identity. Within species diversity measures included, per locus rarified allelic richness (A_R, Hurlbert 1971), number of private alleles, corrected expected heterozygosity (H_e, Nei and Chesser 1983), and the inbreeding coefficient (F_{IS}, Nei and Chesser 1983). Each locus within each species was tested for Hardy–Weinberg Equilibrium with the p values corrected for familywise error using the sequential Bonferroni correction (Holm 1979). In addition, both overall and per locus fixation indices (F_{ST}, Nei 1973) were calculated with significance tested using 10,000 permutations. The maximum F_{ST} given the genetic diversity of these loci was also calculated (Hedrick 2005). These metrics were calculated using in R v3.5.1 using the packages adegenet, pegas, poppr, and hierfstat (Jombart 2008; Paradis 2010; Kamvar et al. 2014; Goudet 2005; R Core Team 2018).

Hybridization

To discriminate between historic and contemporary gene flow a Bayesian analysis of hybridization, based upon the microsatellite data, was performed using NewHybrids (Anderson and Thompson 2002). This analysis was done to identify admixed individuals, including first- and second-generation (the offspring of two first-generation hybrids) hybrids and first-generation backcrosses. For the analysis, 5 independent MCMC chains, with 100,000 burnin iterations and a subsequent 1,000,000 sampling iterations, were run. Uniform priors were used for both mixing proportions and allele frequencies to minimize the influence of rare alleles on the classifications. Proper mixing, exploration of parameter space, and convergence were confirmed

by visually inspecting trace plots and confirming that the $\widehat{R}$ value equaled one (Fig. S2; Vehtari et al. 2021). Due to skew in sample sizes between the species, a general lack of fixed alleles, and the relatively small number of loci, combined with evidence of strong genetic differentiation between individuals confidently assigned to one species or the other, only specimens with mismatching species identifications and/or low STRUCTURE assignment probabilities were assessed as potential hybrids, to minimize the rate of Type I error (False Positives, Table 1). To further mitigate misclassification errors, all hybrid categories were merged into a single category (hereafter admixed).

Shoal composition

Because the individuals collected represent a random sample taken from each shoal that are composed of less than ~ 100 individuals, a 95% credible interval around the observed proportion of *C. personatus*/*hyalinus* present in each shoal, and

Table 1 Identification of each of the 18 specimens with disagreements between the three identification methods and the 16 specimens with less than 0.9 STRUCTURE assignment probability

ID	Shoal	Morphological	Mitochondrial	Microsatellite	STRUCTURE		NEWHYBRIDS		
					C. hyalinus	*C. personatus*	*C. hyalinus*	Admixed	*C. personatus*
0002	A	*C. personatus*	*C. hyalinus*	*C. hyalinus*	0.977	0.023	0.995	0.005	0
0032	J	*C. personatus*	*C. hyalinus*	*C. hyalinus*	0.990	0.010	0.993	0.007	0
0090	F	*C. personatus*	*C. hyalinus*	*C. hyalinus*	0.993	0.007	0.997	0.003	0
0103	F	*C. personatus*	*C. hyalinus*	*C. hyalinus*	0.988	0.012	0.992	0.008	0
0104	F	*C. personatus*	*C. hyalinus*	*C. hyalinus*	0.991	0.009	0.996	0.004	0
0119	L	*C. personatus*	*C. hyalinus*	*C. hyalinus*	0.990	0.010	0.994	0.006	0
0130	L	*C. personatus*	*C. hyalinus*	*C. hyalinus*	0.991	0.009	0.995	0.005	0
0151	B	*C. personatus*	*C. hyalinus*	*C. hyalinus*	0.993	0.007	0.994	0.006	0
0158	B	*C. personatus*	–	*C. hyalinus*	0.993	0.007	0.997	0.003	0
0382	N	*C. personatus*	–	*C. hyalinus*	0.995	0.005	0.998	0.002	0
1284	O	*C. hyalinus*	*C. personatus*	–	–	–	–	–	–
1313	O	*C. personatus*	*C. hyalinus*	–	–	–	–	–	–
1314	O	*C. personatus*	*C. hyalinus*	–	–	–	–	–	–
0039	J	–	*C. hyalinus*	*C. personatus*	0.200	0.800	0.693	0.200	0.108
0084	**K**	**–**	***C. hyalinus***	***C. personatus***	**0.272**	**0.728**	**0.281**	**0.636**	**0.083**
0094	F	*C. hyalinus*	*C. hyalinus*	*C. personatus*	0.432	0.568	0.927	0.061	0.012
0216	M	–	*C. hyalinus*	*C. personatus*	0.098	0.902	0.508	0.301	0.191
0302	I	–	*C. hyalinus*	*C. personatus*	0.185	0.815	0.553	0.197	0.249
0025	A	*C. hyalinus*	*C. hyalinus*	*C. hyalinus*	0.899	0.101	0.907	0.093	0.001
0040	J	–	*C. hyalinus*	*C. hyalinus*	0.889	0.111	0.922	0.076	0.002
0051	K	–	*C. hyalinus*	*C. hyalinus*	0.859	0.141	0.787	0.213	0
0081	**K**	**–**	***C. personatus***	***C. personatus***	**0.186**	**0.814**	**0.102**	**0.671**	**0.227**
0145	L	–	*C. personatus*	*C. personatus*	0.138	0.862	0	0.050	0.95
0152	B	–	*C. hyalinus*	*C. hyalinus*	0.848	0.152	0.780	0.212	0.008
0157	B	–	*C. personatus*	*C. personatus*	0.196	0.804	0.035	0.365	0.600
0252	C	–	*C. hyalinus*	*C. hyalinus*	0.897	0.103	0.986	0.013	0
0262	C	–	*C. hyalinus*	*C. hyalinus*	0.858	0.142	0.976	0.024	0
0287	I	*C. hyalinus*	*C. hyalinus*	*C. hyalinus*	0.638	0.362	0.971	0.028	0.001
0305	I	–	*C. hyalinus*	*C. hyalinus*	0.897	0.103	0.964	0.036	0
0383	N	–	–	*C. personatus*	0.202	0.798	0.870	0.110	0.019

Specimens in the first section showed disagreement between the morphological and molecular identification methods. Specimens in the middle section have disagreements between the nuclear and mitochondrial identification methods. Specimens in the final section have assignment probabilities < 0.9. STRUCTURE assignment probability shows the probability of assignment to species indicated in the Microsatellite column from the STRUCTURE analysis. The STRUCTURE assignment probabilities for both species cluster is shown for each specimen with microsatellite data. The posterior probability of each specimen being a pure *C. hyalinus*/*personatus* or a first- or second-generation hybrid based on the NEWHYBRIDS analysis is also shown for each specimen with microsatellite data. Individuals in bold are those with > 50% probability of being a first- or second-generation hybrid

overall on the reef, was estimated using an algebraic solution of the binomial distribution using the beta distribution as the conjugate prior (Gelman et al. 2013), for all models a uninformative conjugate prior was used ($\beta(1, 1)$). To determine if species proportions in shoals differed from species proportions at the collection site, an algebraic solution to the difference in two proportions was calculated in R v3.5.1 (Pham-Gia et al. 1993; R Core Team 2018). If zero was contained within the 95% credible interval of the difference between the site and shoal proportions then the shoal was determined to be composed of a random mixture of species. If zero was not contained within the 95% credible interval, then the shoal was determined to contain a biased mixture with either more *C. personatus* (greater than zero) or more *C. hyalinus* (less than zero) than would be expected from a random sample taken across shoals at the site.

Results

Morphology

The collected specimens ranged from 5 to 22 mm standard length (SL). This distribution of lengths was not normal and skewed to smaller individuals with ~ 78% of the specimens likely being new recruits (Beeken et al. 2021). Thirty (22.4%) individuals had a single anterior interorbital cephalic (AIC) pore, consistent with *C. personatus* and 104 (77.6%) individuals had two AIC pores, consistent with *C. hyalinus*. Specimens which were able to be identified morphologically were significantly larger than those which could not be identified morphologically (Kruskal–Wallis's $\chi^2_{(1)} = 176.8$, $p < 0.0001$). Moreover, specimens which were misidentified based on morphology tended to be larger than those which were correctly identified (Kruskal–Wallis's $\chi^2_{(1)} = 6.04$, $p = 0.014$). There was a significant difference in SL between *C. hyalinus* and *C. personatus* based on morphological identification (Kruskal–Wallis's $\chi^2_{(1)} = 9.38$, $p = 0.0022$) with no differences observed when using all other methods of species identification (see below; COI: Kruskal–Wallis's $\chi^2_{(1)} = 1.23$, $p = 0.269$; Microsatellite: Kruskal–Wallis's $\chi^2_{(1)} = 0.81$, $p = 0.397$; Joint Method: Kruskal–Wallis's $\chi^2_{(1)} = 0.72$, $p = 0.370$).

Mitochondria

The COI locus was successfully amplified for 338 (79.0%) individuals. After trimming sequences to contain only a shared core of 547 bp, there were 34 alternately fixed sites in the samples from Baldwin et al. (2009). The net sequence divergence between the two species in this sampling was 6.0%, while mean within species divergence was 0.12% in *C. hyalinus* and 0.13% in *C. personatus*. Consistent with Baldwin et al. (2009), there were two distinct groups of haplotypes separated by 27 mutations (Fig. 3). Based on COI alone 62 (18.3%) individuals were identified as *C. personatus* and 276 (81.7%) individuals as *C. hyalinus*.

Microsatellites

Structure identified two distinct genetic clusters following the Evanno method (Fig. S3). Most individuals fully assigned (> 90% assignment probability) to either one cluster or the other, with only 16 (4.2%) individuals showing evidence of more than 10% admixture (Fig. 4, Table 1). All individuals assigned to cluster 1 were identified as *C. hyalinus* using COI data, and cluster 1 was, therefore, assumed to represent *C. hyalinus*. Ninety-one percent of individuals in cluster 2 were identified as *C. personatus* using the COI data, and cluster 2 was, therefore, assumed to represent *C. personatus*. The rate of disagreement between nuclear and mitochondrial markers was 1.6%. Based on Structure analysis alone, 55 (14.3%) individuals were identified as *C. personatus* and 329 (85.7%) as *C. hyalinus*.

Joint species identification

In general, species identification methods agreed but there were several potentially interesting examples of disagreements between the various methods (Table 1). A total of 13 specimens showed disagreements between morphological and molecular species identification, an error rate of 12.1%. Twelve of these samples were morphologically identified as *C. personatus* (one AIC pore) and one as *C. hyalinus* (two AIC pores). All 13 individuals were reexamined for morphology, blind to the original species identification, and the original morphological identification was confirmed in all cases. Five individuals with a COI species identification of *C. hyalinus* had a high Structure assignment probability < 0.9 to cluster 2 (*C. personatus*, Table 1).

Characterizing genetic diversity

A total of 107 (25.0%) individuals were only identified using one method or using only morphology and one molecular method and were excluded as a result. Of the remaining 321 specimens, 308 (96.0%) showed agreement among all three methods, or both molecular techniques. Of these, 15.6% were identified as *C. personatus* and 84.4% as *C. hyalinus*. Both species tend to exhibit homozygote excess across the same loci and overwhelmingly had elevated inbreeding coefficients. The high level of inbreeding and ubiquity of homozygote excess across loci likely results from the reproductive strategy and relatively short distance of larval dispersal leading to a heightened frequency of inbreeding rather than genotyping artifacts (Waples 2015; Selwyn

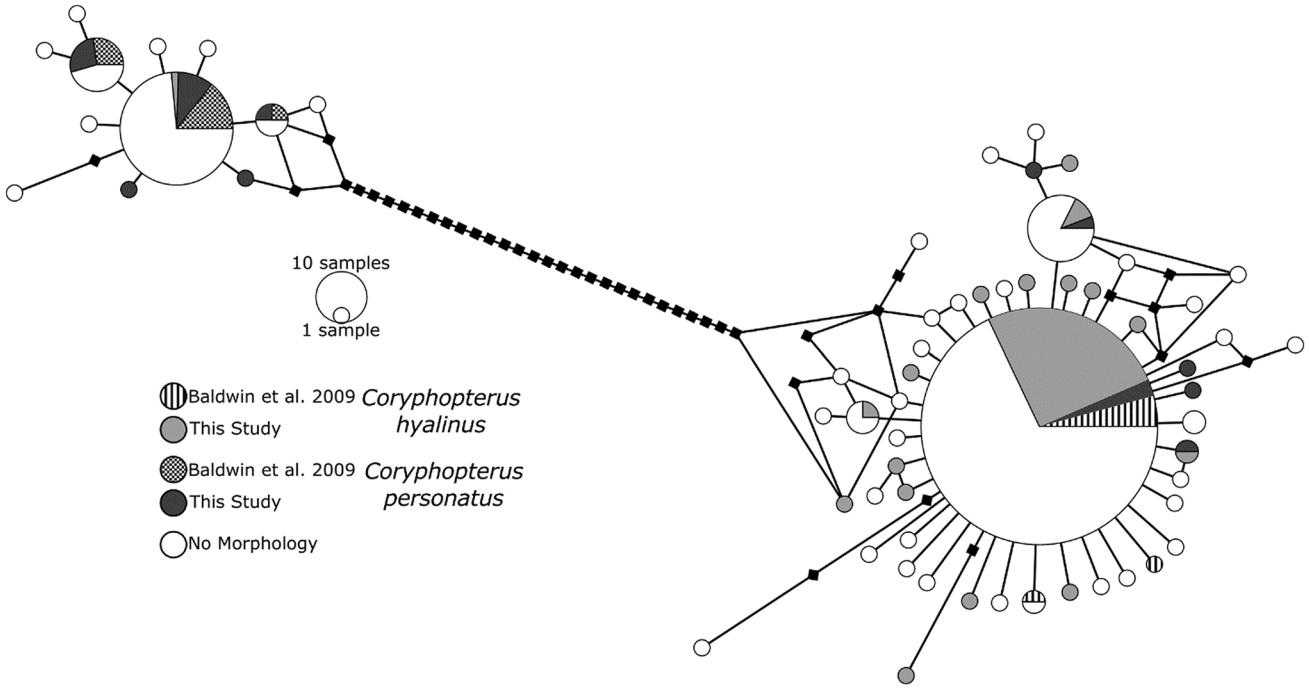

Fig. 3 Haplotype network showing COI haplotypes with the size of each circle representing the number of individuals observed with that haplotype. The shade and hatching of the circle shows the species identification based on either morphological identification or from Baldwin et al. (2009). Small squares represent inferred haplotypes

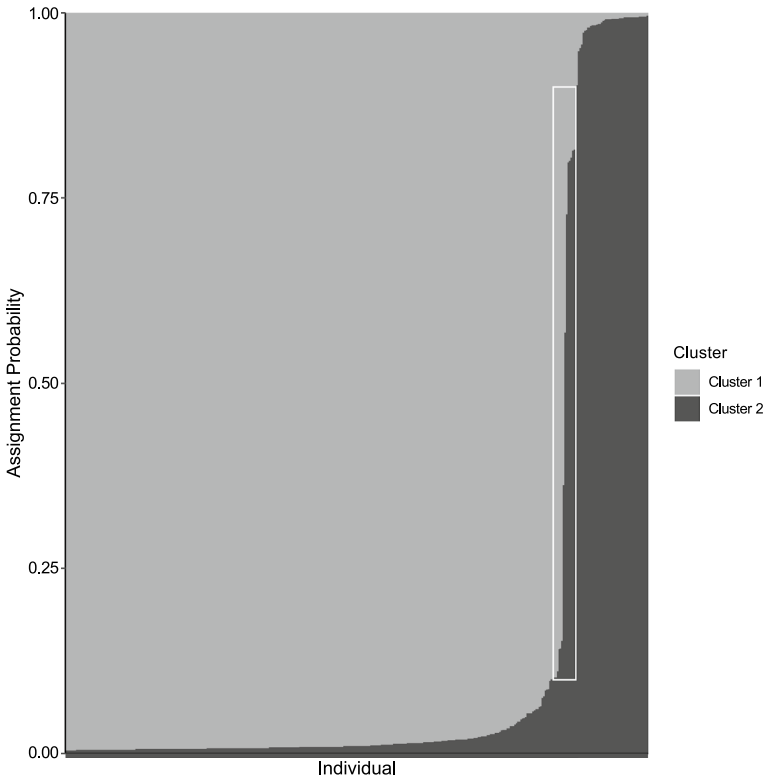

Fig. 4 STRUCTURE plot showing cluster assignments of each individual. Individuals are along the *x*-axis and assignment probability on the *y*-axis to each cluster (bar color). The white box indicates the 16 individuals with assignment probabilities between 10 and 90% suggesting possible mixed heritage

Table 2 Per locus population summary statistics for each species (value in parenthesis after species name shows number of individuals)

	Coryphopterus hyalinus (260)					*Coryphopterus personatus* (48)					Fixation Index	
Locus	A	A_P	A_R	H_e	F_{IS}	A	A_P	A_R	H_e	F_{IS}	F_{ST}	p
COPE5	69	57	32.36	*0.98*	0.21	14	4	12.49	0.60	0.18	0.20	0.0007
COPE9	10	2	7.75	*0.73*	0.21	9	1	8.80	*0.77*	0.43	0.01	0.0068
CPER26	10	4	5.87	*0.59*	0.15	8	2	7.21	0.79	0.26	0.29	0.0007
CPER92	12	2	6.01	*0.50*	0.41	9	0	10.00	*0.70*	0.45	0.05	0.0226
CPER99	4	0	2.48	*0.11*	− 0.05	5	0	4.78	0.63	− 0.07	0.77	0.0007
CPER119	20	2	13.61	*0.90*	0.46	19	2	17.98	*0.95*	0.56	0.01	0.0008
CPER188	12	7	6.68	0.51	0.03	7	3	5.72	0.32	− 0.06	0.03	0.0008

Summary statistics included are number of alleles (A), number of private alleles (A_P), rarified allelic richness (A_R), expected (H_e) heterozygosity, and the inbreeding coefficient (F_{IS}). The p value (p) indicates if the F_{ST} is significantly different from zero based on 10,000 permutations and has been corrected for family-wise error using sequential Bonferroni (Holm 1979). Italicized H_e indicates significant homozygote excess (all $p < 0.0028$). Finally, locus specific fixation indices (F_{ST}) are included with italicized F_{ST} values indicating significant differentiation

2015; Selwyn et al. 2016). The two species were significantly differentiated from each other ($F_{ST} = 0.19$, $p < 0.0001$, $F_{STmax} = 0.21$, Table 2).

Individuals definitively classified as either *C. personatus* or *C. hyalinus* fell into two well-formed clusters separating along PC1 (Fig. 5). Specimens which did not assign clearly to one species or the other in Structure, and/or showed disagreements among identification techniques generally fell between the two clusters. The variables most associated with PC1 were alleles of microsatellite loci that had the highest frequency differences between the two species (Fig. S4, Table 2). While PC2 is strongly associated with specimen morphology, matching with the observation that 12 out of the 13 morphological/molecular disagreements involved a specimen observed with a single AIC pore (Fig. S4, Table 1). Specimens where disagreement was observed between morphology and molecular identification were significantly shifted to the right on PC1 and higher on PC2 (MANOVA $\eta_p^2 = 0.32$, Pillai's trace statistic = 0.64, $F_{(4,54)} = 6.4$, $p = 0.00027$, Fig. 5), relative to specimens with low assignment probabilities and specimens with mitochondrial–nuclear discordance. The observed placement of specimens with molecular–morphological discordance on PC1 seems to indicate that genetically these specimens are *C. hyalinus*. Meanwhile, specimens with either nuclear–mitochondrial discordance or low Structure assignment probabilities were generally found between the two main species clusters on PC1 (Fig. 5), and closer to zero on PC2.

Hybridization

All specimens which showed disagreement between morphological and molecular identification methods were confidently classified as pure *C. hyalinus* (minimum posterior probability *C. hyalinus* assignment = 0.992). For all other specimens there was greater uncertainty in the delineation between admixed and pure species, with only two specimens identified as admixed with > 50% posterior probability (0081 & 0084; Table 1). Based on body lengths (SL = 6 & 7 mm, respectively) both individuals were likely new recruits (Beeken et al. 2021). Specimen 0084 was positioned centrally between the two species clusters in the PCA, while specimen 0081 was located near the *C. personatus* cluster. This is likely a manifestation of the agreement between nuclear and mitochondrial markers in 0081 and the disagreement between these markers in 0084 (Table 1). For both specimens the second highest posterior probability was associated with species identified using mtDNA.

Shoal composition

Overall, the reef was estimated to be composed of 12.0–20.1% (95% CI) *C. personatus* and 79.9–88.0% (95% CI) *C. hyalinus*. Only two shoals (out of 13) differed significantly in the proportion of the two species present compared to the site-level proportion (Fig. 6). One of these shoals had significantly more *C. personatus* than expected (45.5%, 95% CI 18.7–73.8%; Difference from site 95% CI 0.019–0.581), while the other had significantly fewer *C. personatus* (3.7%, 95% CI 0.1–13.2%; Difference from site 95% CI − 0.202 to − 0.0402).

Discussion

Coryphopterus personatus and *C. hyalinus* are sympatric sister-taxa that live in social aggregations throughout their reproductive lifespans. The general agreement between nuclear and mitochondrial identification (98.4%) and

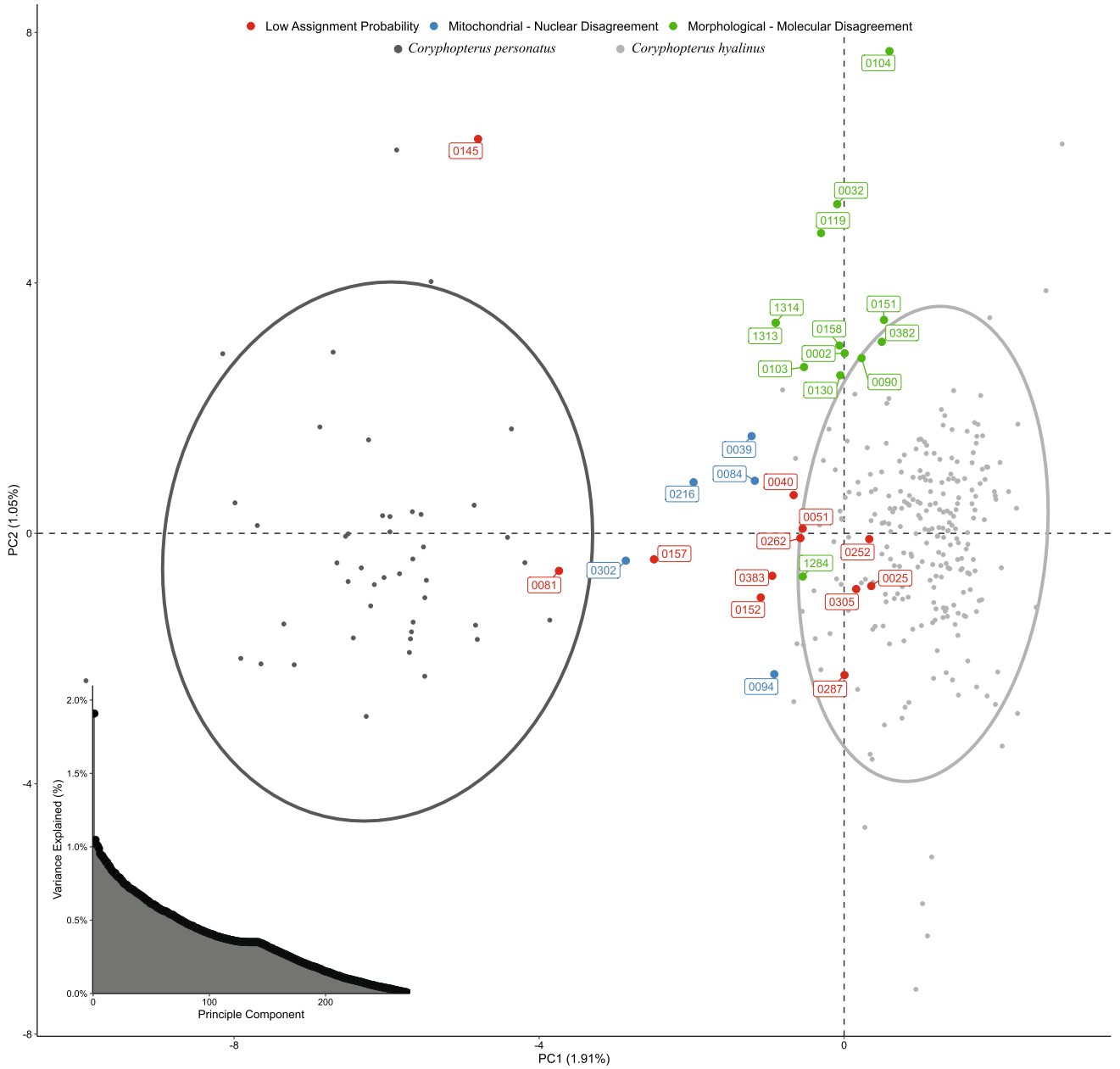

Fig. 5 Principal component analysis of morphological, mitochondrial, and microsatellite data from specimens (points) of Coryphopterus hyalinus (light gray) and Coryphopterus personatus (dark gray). Colored and labelled points show specimens with disagreements between either: **A** morphological and molecular identification methods (green), **B** mitochondrial and nuclear markers (blue), or **C** low STRUCTURE assignment probabilities (red). These individuals are the same as found in Table 1. The inset plot shows the variance explained by each principal component

the paucity of admixed individuals (4.2%) suggests that the boundary between the two species is generally well maintained. However, gene flow has not ceased entirely as there was evidence for ongoing hybridization (two putative hybrids were observed) and historic introgression (admixture and disagreement between mitochondrial and nuclear loci). Contrary to the hypothesis that species boundaries are maintained by the formation of species-specific shoals, all shoals were composed of a mixture of *C. personatus* and *C. hyalinus*. Furthermore, the proportions of the species in each shoal generally conformed to the proportion of these species on the entire reef, indicating that the species are not segregating due to social behavior. In addition, there was disagreement between molecular and morphological methods in 12.1% of individuals and between the two molecular methods in 1.6% of the individuals,

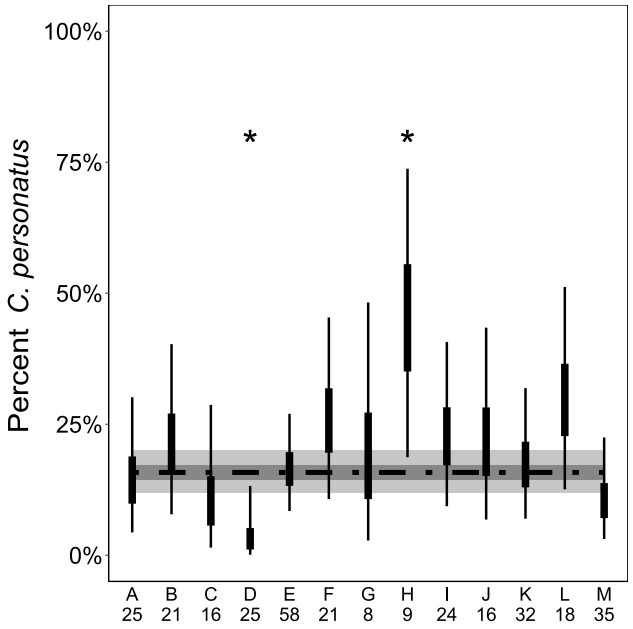

Fig. 6 Percentage of each shoal (A–M) which was identified as *Coryphopterus personatus*. The horizontal dashed line shows the percent *C. personatus* across shoals at Turneffe Atoll, with 50% and 95% credible intervals shaded. The vertical intervals show the 50% and 95% credible intervals of the percentage of *C. personatus* in each shoal. Stars indicate a shoal that is significantly different in species composition relative to Turneffe Atoll. Numbers under letters along the *x*-axis show the total number of individuals analyzed from each shoal

suggesting that gene flow between the species is not completely interrupted.

Mismatches in species identification between methods occurred but only for a small percentage of samples. Disagreements between morphological and molecular methods are likely attributable to one or two sources of error and showed a bias of morphologically misidentifying larger individuals as *C. personatus*. First morphological characters may be difficult to accurately assess following preservation (Kristoffersen and Salvanes 1998; Martinez et al. 2013). In *Coryphopterus*, anterior interorbital cephalic pores are difficult to see prior to ethanol preservation, which causes the pores to dilate, making them more readily visible (Baldwin et al. 2009). The preservation process could act asymmetrically, causing one pore to become more easily visible, appear to be more centrally located, and/or tear the tissue dividing pores, forming what appears to be a single pore and lead to misidentification. Alternatively, there may be natural variation in the character state within species (one or two interorbital cephalic pores) leading to misidentification. Overlap in the distribution of meristic characters is a common problem in ichthyology (Hubbs 1922; Tåning 1952; McKay and Miller 1997) and variation in the number and arrangement of sensory pores within species is a common phenomenon (Ahnelt et al. 2004; Vanderpham et al. 2013; Ito et al. 2017). Furthermore, because this character develops in larger individuals (> 10 mm SL), it may not be fully developed in some smaller individuals (Victor 2019). While morphological identification is possible and, in this study, seemed to be reasonably accurate (~ 87.9%), species identification using molecular methods is likely more reliable. By contrast, disagreements between mitochondrial and nuclear markers are likely the result of historic introgression or contemporary gene flow, including hybridization (Toews and Brelsford 2012). The presence of recently admixed individuals suggests that barriers to reproduction may be incomplete. However, the loci available, while capable of distinguishing between species (Table 2), do not offer sufficient resolution to distinguish between hybrid categories.

Despite indications of low levels of recurrent gene flow and regular co-occurrence in social aggregations, the two species remain genetically distinct with an estimated pairwise F_{ST} (0.19) at ~ 90% of its maximum value (0.21). The benefits of aggregating in this system seem clear, since larger aggregations make predators less efficient and decrease the probability of any individual being depredated (Hamilton 1971; Landeau and Terborgh 1986). Small reef fishes are constantly under high predation risk and both *C. personatus* and *C. hyalinus* spend the entirety of their life with elevated predation risk due to their small maximum body sizes, likely heightened by their behavior of hovering above the reef structure (Goatley and Bellwood 2016). In heterospecific aggregations the individuals of the less numerous species may benefit from allying with the more numerous species, if co-aggregation increases aggregation sizes beyond that which the minor species could achieve on its own (Parrish 1989; Gibson et al. 2002; Wood and Ackland 2007). The idea that these heterospecific shoals form in part to offer protection is further supported by the presence of a third phylogenetically distant species of similar size within *Coryphopterus* shoals, the arrow blenny (*Lucayablennius zingaro*; Greenfield 1972).

Spending a significant proportion of the reproductive life span interacting with congeners comes with the risk of a reduction in fitness caused by production of inviable or less viable hybrids (Dobzhansky 1940; Coyne 1974; Friberg et al. 2013). Hybrid inviability can be induced through epistatic gene interactions (Dobzhansky 1936; Goodnight 2000), often resulting in viable, fertile F1 hybrids which are unable to produce viable F2 hybrids, but can backcross with parental lineages (Stelkens et al. 2015). As species diverge the number of incompatibilities tends to increase, further reinforcing isolation (Bolnick and Near 2005). Hybrid inviability in turn can reinforce pre-mating reproductive isolation and prevent the formation of hybrid swarms (Liou and Price 1994; Sadedin and Littlejohn 2003). For example, species boundaries are maintained between sympatric darters

(family: Percidae) because of epistatic incompatibilities leading to elevated mortality in backcrossed individuals (Moran et al. 2019). Reproduction is energetically costly and as such the production of inviable hybrids represents a disproportionately large energetic cost (Dobzhansky 1940; Wootton 1985) that would be borne more heavily in species with short reproductive lifespans, such as *C. personatus* and *C. hyalinus*. Selection might quickly cause pre-mating isolation to develop when the probability of wasted energy via hybridization is high (Ortiz-Barrientos et al. 2004). Consistent with this idea, contemporary geneflow between the two species of *Coryphopterus* appears to be minimal, despite opportunity for frequent hybridization, suggesting that another mechanism may be maintaining species boundaries.

Unidirectional hybridization is a commonly observed pattern and occurs for a variety of reasons (Wirtz 1999). In general, when females are the choosy sex, hybridization occurs between females of the rare species and males of the common species (Wirtz 1999). However, in this case when there was mito-nuclear discordance within individuals the maternal lineage was the more common species, *C. hyalinus* (Table 1). This observation could be explained by several different mechanisms. First, females of the less common species, *C. personatus* may be more discriminatory in choosing mates than female *C. hyalinus*. It could be selectively advantageous for the less common species to be more discerning due to the increased probability of heterospecific mating (Cooley 2007), especially if genetic incompatibilities have developed between the species leading to high fitness costs (i.e., less viable hybrid offspring) that outweigh potential costs associated with mate discrimination (Milinski and Bakker 1992; Wong and Jennions 2003). Under this mechanism locations, where the relative abundances of the species are reversed, should result in female *C. hyalinus* being more discriminatory than *C. personatus*. The observed pattern could also result from female-biased sex ratios within shoals. Both *C. personatus* and *C. hyalinus* are protogynous (i.e., change sex from female to male) and research has demonstrated associated female skew in sex ratios (Cole and Robertson 1988; Cole and Shapiro 1990; Allsop and West 2004). This could result in larger dominant males that defend nesting sites being the choosy sex, rather than females, with the less common males (*C. personatus*) involved in more interspecies matings due to the relative lack of intraspecies females (Thresher 1984; Kramer et al. 2009).

While mate choice may be an isolating mechanism these data do not preclude the possibility of ultra-fine scale spatial segregation or asynchronous reproduction as alternative plausible mechanisms. Sympatric species of triplefin blennies (Family Tripterygiidae) in New Zealand, utilize distinct nesting micro-habitats leading to reduced reproductive encounters between species (Wellenreuther and Clements 2007). However, triplefin nests are spread over the expanse of rocky reefs (Feary and Clements 2006). *Coryphopterus personatus/hyalinus* are thought to nest in reef crevices in the immediate vicinity of the shoal, suggesting that any differentiation in nest site preference between the species is occurring on at a much finer spatial scale. Asynchronous reproduction, a common mechanism of reproductive isolation seen across taxa (Aspinwall 1974; Palumbi 1994), does not seem to be a likely for *C. personatus* and *C. hyalinus*, because they only live ~ 100 days post-settlement (Beeken et al. 2021) and congeners reproduce continuously through the lunar cycle (Kramer et al. 2009).

Despite ample opportunity and evidence of ongoing/past hybridization, *C. hyalinus* and *C. personatus* remain distinct but the exact mechanism(s) keeping them distinct remain unclear. Whatever mechanism(s) are at work, they likely occur at within-shoal scales and are mechanisms not often explored as first principles when seeking to explain the maintenance of species boundaries. Breeding site selection within shoals and/or cryptic mate recognition may play a large role in reducing gene flow between the species, as may be the case in other social species that co-aggregate with closely related taxa.

Supplementary Information The online version contains supplementary material available at https://doi.org/10.1007/s00227-022-04018-x.

Acknowledgements This research would not have been possible without the field assistance of Emily Anderson and Alan Downey-Wall. Field logistics were provided by the University of Belize Calabash Caye Field Station. We would also like to thank Sharon Magnuson and Christopher Bird for their helpful comments and suggestions on the manuscript. We would like to thank funding sources: the Carl R. Beaver Memorial Scholarship, the Karen Koester Dodson Memorial Fund Grant, and Grants-in-Aid of Graduate Student Research Award by Texas Sea Grant College Program awarded to JDS, the Texas Research Development Fund award to JDH and funds provided by TAMUCC to JDH.

Author contributions All authors contributed to study conception and design. Data collection and analysis were performed by JDS and EPH. The first draft of the manuscript was written by JDS with all authors commenting on subsequent versions. All authors read and approved the final manuscript.

Funding The Carl R. Beaver Memorial Scholarship, the Karen Koester Dodson Memorial Fund Grant, and Grants-in-Aid of Graduate Student Research Award by Texas Sea Grant College Program awarded to JDS and the Texas Research Development Fund award to JDH and funds provided by TAMUCC to JDH.

Availability of data and materials All data generated or analyzed during this study are included in this published article [and its supplementary information files], CO1 Sequence data available at NCBI Genbank Accession Numbers: MT784949—MT785286.

Code availability https://github.com/jdselwyn/Mixed_Shoal.

Declarations

Conflicts of interest On behalf of all authors, the corresponding author states that there is no conflict of interest.

Ethics approval All collections were performed in accordance with the ethical guidelines of Texas A&M University—Corpus Christi (TAMUCC-AUP-05-14) and in compliance with standards outlined in the US National Research Council's Guide for the Care and Use of Laboratory Animals. Collections were made with the express permission of the government of Belize (Aquatic Scientific Research Permit 000044-13).

References

Abbott R, Albach D, Ansell S, Arntzen JW, Baird SJE, Bierne N, Boughman J, Brelsford A, Buerkle CA, Buggs R, Butlin RK, Dieckmann U, Eroukhmanoff F, Grill A, Cahan SH, Hermansen JS, Hewitt G, Hudson AG, Jiggins C, Jones J, Keller B, Marczewski T, Mallet J, Martinez-Rodriguez P, Möst M, Mullen S, Nichols R, Nolte AW, Parisod C, Pfennig K, Rice AM, Ritchie MG, Seifert B, Smadja CM, Stelkens R, Szymura JM, Väinölä R, Wolf JBW, Zinner D (2013) Hybridization and speciation. J Evol Biol 26:229–246. https://doi.org/10.1111/j.1420-9101.2012.02599.x

Ahnelt H, Dawson MN, Jacobs DK (2004) Geographical variation in the cephalic lateral line canals of *Eucyclogobius newberryi* (Teleostei, Gobiidae) and its comparison with molecular phylogeography. Folia Zool 53:358–398

Allsop DJ, West SA (2004) Sex allocation in the sex-changing marine goby, *Coryphopterus personatus*, on atoll-fringing reefs. Evol Ecol Res 6:843–855

Anderson EC, Thompson EA (2002) A model-based method for identifying species hybrids using multilocus genetic data. Genetics 160:1217–1229

Archer FI, Adams PE, Schneiders BB (2017) stratag: An r package for manipulating, summarizing and analysing population genetic data. Mol Ecol Resour 17:5–11. https://doi.org/10.1111/1755-0998.12559

Aspinwall N (1974) Genetic Analysis of North American populations of the pink salmon, *Oncorhynchus gorbuscha*, possible evidence for the neutral mutation-random drift hypothesis. Evolution 28:295–305. https://doi.org/10.2307/2407331

Babcock RC, Mundy CN, Whitehead D (1994) Sperm diffusion models and in situ confirmation of long-distance fertilization in the free-spawning asteroid *Acanthaster planci*. Biol Bull 186:17–28. https://doi.org/10.2307/1542033

Baldwin CC, Robertson DR (2015) A new, mesophotic *Coryphopterus* goby (Teleostei, Gobiidae) from the southern Caribbean, with comments on relationships and depth distributions within the genus. ZooKeys 513:123–142. https://doi.org/10.3897/zookeys.513.9998

Baldwin CC, Weigt LA, Smith DG, Mounts JH (2009) Reconciling genetic lineages with species in western Atlantic *Coryphopterus* (Teleostei: Gobiidae). Smithson Contrib Mar Sci 38:111–138

Bandelt H-J, Forster P, Röhl A (1999) Median-joining networks for inferring intraspecific phylogenies. Mol Biol Evol 16:37–48

Beeken NS, Selwyn JD, Hogan JD (2021) Determining the life history strategy of the cryptobenthic reef gobies *Coryphopterus hyalinus* and *C. personatus*. Mar Ecol Prog Ser 659:161–173. https://doi.org/10.3354/meps13573

Bodenhofer U, Bonatesta E, Horejš-Kainrath C, Hochreiter S (2015) msa: an R package for multiple sequence alignment. Bioinformatics 31:3997–3999. https://doi.org/10.1093/bioinformatics/btv494

Böhlke JE, Robins CR (1962) The Taxonomic position of the West Atlantic Goby, *Eviota personata*, with descriptions of two new related species. Proc Acad Nat Sci Phila 114:175–189

Bolnick DI, Near TJ (2005) Tempo of hybrid inviability in centrarchid fishes (Teleostei: Centrarchidae). Evolution 59:1754–1767. https://doi.org/10.1111/j.0014-3820.2005.tb01824.x

Bouwmeester J, Edwards AJ, Guest JR, Bauman AG, Berumen ML, Baird AH (2021) Latitudinal variation in monthly-scale reproductive synchrony among *Acropora* coral assemblages in the Indo-Pacific. Coral Reefs. https://doi.org/10.1007/s00338-021-02129-3

Bovbjerg RV (1970) Ecological isolation and competitive exclusion in two crayfish (Orconectes Virilis and Orconectes Immunis). Ecology 51:225–236. https://doi.org/10.2307/1933658

Brandl SJ, Goatley CHR, Bellwood DR, Tornabene L (2018) The hidden half: ecology and evolution of cryptobenthic fishes on coral reefs. Biol Rev 93:1846–1873. https://doi.org/10.1111/brv.12423

Chagaris D, Binion-Rock S, Bogdanoff A, Dahl K, Granneman J, Harris H, Mohan J, Rudd MB, Swenarton MK, Ahrens R, Patterson WF III, Mohan JA Jr, Allen M (2017) An ecosystem-based approach to evaluating impacts and management of invasive lionfish. Fisheries 42:421–431. https://doi.org/10.1080/03632415.2017.1340273

Cole KS, Robertson DR (1988) Protogyny in the Caribbean reef goby, *Coryphopterus personatus*: gonad ontogeny and social influences on sex-change. Bull Mar Sci 42:317–333

Cole KS, Shapiro DY (1990) Gonad Structure and Hermaphroditism in the Gobiid Genus *Coryphopterus* (Teleostei: Gobiidae). Copeia 1990:996–1003. https://doi.org/10.2307/1446482

Colin PL (2010) Fishes as living tracers of connectivity in the tropical western North Atlantic: I. Distribution of the neon gobies, genus *Elacatinus* (Pisces: Gobiidae). Zootaxa 2370:36–52. https://doi.org/10.11646/zootaxa.2370.1.2

Cooley JR (2007) Decoding asymmetries in reproductive character displacement. Proc Acad Nat Sci Phila 156:89–96. https://doi.org/10.1635/0097-3157(2007)156[89:DAIRCD]2.0.CO;2

Coyne JA (1974) The evolutionary origin of hybrid inviability. Evolution 28:505–506. https://doi.org/10.1111/j.1558-5646.1974.tb00781.x

Coyne JA, Orr HA (2004) Speciation, 1st edn. Sinauer Associates, Sunderland

de Lamballerie X, Zandotti C, Vignoli C, Bollet C, de Micco P (1992) A one-step microbial DNA extraction method using "Chelex 100" suitable for gene amplification. Res Microbiol 143:785–790. https://doi.org/10.1016/0923-2508(92)90107-Y

Diabaté A, Dao A, Yaro AS, Adamou A, Gonzalez R, Manoukis NC, Traoré SF, Gwadz RW, Lehmann T (2009) Spatial swarm segregation and reproductive isolation between the molecular forms of *Anopheles gambiae*. Proc R Soc B Biol Sci 276:4215–4222. https://doi.org/10.1098/rspb.2009.1167

Dobzhansky TG (1936) Studies on Hybrid sterility. II. Localization of sterility factors in drosophila pseudoobscura hybrids. Genetics 21:113–135

Dobzhansky TG (1940) Speciation as a stage in evolutionary divergence. Am Nat 74:312–321. https://doi.org/10.1086/280899

Domeier ML, Colin PL (1997) Tropical reef fish spawning aggregations: defined and reviewed. Bull Mar Sci 60:698–726

Evanno G, Regnaut S, Goudet J (2005) Detecting the number of clusters of individuals using the software STRUCTURE: a simulation study. Mol Ecol 14:2611–2620. https://doi.org/10.1111/j.1365-294X.2005.02553.x

Feary DA, Clements KD (2006) Habitat use by triplefin species (*Tripterygiidae*) on rocky reefs in New Zealand. J Fish Biol 69:1031–1046. https://doi.org/10.1111/j.1095-8649.2006.01178.x

Friberg M, Leimar O, Wiklund C (2013) Heterospecific courtship, minority effects and niche separation between cryptic butterfly species. J Evol Biol 26:971–979. https://doi.org/10.1111/jeb.12106

Gagnaire P-A, Pavey SA, Normandeau E, Bernatchez L (2013) The Genetic Architecture of Reproductive Isolation During Speciation-with-Gene-Flow in Lake Whitefish Species Pairs Assessed by Rad Sequencing. Evolution 67:2483–2497. https://doi.org/10.1111/evo.12075

Garcia E, Holtermann K (1998) Calabash Caye, Turneffe Islands Atoll, Belize. In: Kjerfve B (ed) Caricomp: Caribbean coral reef, seagrass and mangrove sites. Coastal region and small islands unit. UNESCO, Paris, France, pp 1–345

Gardner T (2000) Breeding the marine goby *Coryphopterus personatus* in Aquaria. http://gobiidae.com/breeding_c_pers.htm. Accessed 29 Mar 2017

Gelman A, Carlin JB, Stern HS, Dunson DB, Vehtari A, Rubin DB (2013) Bayesian data analysis, 3rd edn. Chapman and Hall/CRC, Boca Raton

Gerhardt HC (1974) Behavioral Isolation of the Tree Frogs, *Hyla cinerea* and *Hyla andersonii*. Am Midl Nat 91:424–433. https://doi.org/10.2307/2424332

Gibson RM, Aspbury AS, McDaniel LL (2002) Active formation of mixed–species grouse leks: a role for predation in lek evolution? Proc R Soc Lond B Biol Sci 269:2503–2507

Goatley CHR, Bellwood DR (2016) Body size and mortality rates in coral reef fishes: a three-phase relationship. Proc R Soc B Biol Sci 283:20161858. https://doi.org/10.1098/rspb.2016.1858

Goodnight CJ (2000) Modeling gene interaction in structured populations. In: Wolf JB, Brodie ED, Wade MJ (eds) Epistasis and the evolutionary process. Oxford University Press, Oxford, pp 129–145

Goudet J (2005) HIERFSTAT, a package for R to compute and test hierarchical *F*-statistics. Molec Ecol Notes 5:184–186. https://doi.org/10.1111/j.1471-8286.2004.00828.x

Greenfield DW (1972) Notes on the biology of the arrow blenny, *Lucayablennius zingaro* (Böhlke) from British Honduras. Copeia 1972:590–592. https://doi.org/10.2307/1442939

Greenfield DW, Johnson RK (1999) Assemblage structure and habitat associations of western Caribbean gobies (Teleostei: Gobiidae). Copeia 1999:251–266. https://doi.org/10.2307/1447470

Grinsted L, Agnarsson I, Bilde T (2012) Subsocial behaviour and brood adoption in mixed-species colonies of two theridiid spiders. Naturwissenschaften 99:1021–1030. https://doi.org/10.1007/s00114-012-0983-4

Hamilton WD (1971) Geometry for the selfish herd. J Theor Biol 31:295–311

Harrison RG, Larson EL (2014) Hybridization, Introgression, And The Nature Of Species Boundaries. J Hered 105:795–809. https://doi.org/10.1093/jhered/esu033

Harrison HB, Berumen ML, Saenz-Agudelo P, Salas E, Williamson DH, Jones GP (2017) Widespread hybridization and bidirectional introgression in sympatric species of coral reef fish. Mol Ecol 26:5692–5704. https://doi.org/10.1111/mec.14279

Hedrick PW (2005) A standardized genetic differentiation measure. Evol Int J Org Evol 59:1633–1638

Hepburn RI, Mottillo EP, Bentzen P, Heath DD (2005) Polymorphic microsatellite loci for the masked goby, *Coryphopterus personatus* (Gobiidae). Conserv Genet 6:1059–1062. https://doi.org/10.1007/s10592-005-9088-3

Hogan JD, Galarza JA, Walter RP, Heath DD (2010) Bayesian analysis of population structure and the characterization of nine novel microsatellite markers for the study of a Caribbean coral reef Gobiid (*Coryphopterus personatus*) and related taxa. Mol Ecol Resour 10:1–11. https://doi.org/10.1111/j.1755-0998.2009.02827.x

Holm S (1979) A simple sequentially rejective multiple test procedure. Scand J Stat 6:65–70

Hubbs CL (1922) Variations in the number of vertebrae and other meristic characters of fishes correlated with the temperature of water during development. Am Nat 56:360–372. https://doi.org/10.1086/279875

Hurlbert SH (1971) The nonconcept of species diversity: a critique and alternative parameters. Ecology 52:577–586. https://doi.org/10.2307/1934145

Ito T, Fukuda T, Morimune T, Hosoya K (2017) Evolution of the connection patterns of the cephalic lateral line canal system and its use to diagnose opsariichthyin cyprinid fishes (Teleostei, Cyprinidae). ZooKeys. https://doi.org/10.3897/zookeys.718.13574

Jombart T (2008) adegenet: a R package for the multivariate analysis of genetic markers. Bioinformatics 24:1403–1405

Kamvar ZN, Tabima JF, Grünwald NJ (2014) Poppr: an R package for genetic analysis of populations with clonal, partially clonal, and/or sexual reproduction. PeerJ 2:e281

Knowlton N, Maté JL, Guzmán HM, Rowan R, Jara J (1997) Direct evidence for reproductive isolation among the three species of the *Montastraea annularis* complex in Central America (Panamá and Honduras). Mar Biol 127:705–711. https://doi.org/10.1007/s002270050061

Kopp A, Frank AK (2005) Speciation in progress? A continuum of reproductive isolation in drosophila bipectinata. Genetica 125:55–68. https://doi.org/10.1007/s10709-005-4787-8

Kramer AL, Van Tassell JL, Patzner RA (2009) Aspects of spawning behaviour in five gobiids of the genus *Coryphopterus* (Pisces: Gobiidae) in the Caribbean Sea. Open Fish Sci J 2:50–54

Kristoffersen JB, Salvanes AGV (1998) Effects of formaldehyde and ethanol preservation on body and otoliths of *Maurolicus muelleri* and *Benthosema glaciale*. Sarsia 83:95–102. https://doi.org/10.1080/00364827.1998.10413675

Landeau L, Terborgh J (1986) Oddity and the "confusion effect" in predation. Anim Behav 34:1372–1380

Leigh JW, Bryant D (2015) popart: full-feature software for haplotype network construction. Methods Ecol Evol 6:1110–1116

Levitan DR, Fukami H, Jara J, Kline D, McGovern TM, McGhee KE, Swanson CA, Knowlton N (2004) Mechanisms of reproductive isolation among sympatric broadcast-spawning corals of the *Montastraea annularis* species complex. Evolution 58:308–323. https://doi.org/10.1111/j.0014-3820.2004.tb01647.x

Levitan DR, Fogarty ND, Jara J, Lotterhos KE, Knowlton N (2011) Genetic, spatial, and temporal components of precise spawning synchrony in reef building corals of the *Montastraea annularis* species complex. Evolution 65:1254–1270. https://doi.org/10.1111/j.1558-5646.2011.01235.x

Liebers D, de Knijff P, Helbig AJ (2004) The herring gull complex is not a ring species. Proc R Soc Lond B Biol Sci 271:893–901. https://doi.org/10.1098/rspb.2004.2679

Liou LW, Price TD (1994) Speciation by reinforcement of premating isolation. Evolution 48:1451–1459. https://doi.org/10.1111/j.1558-5646.1994.tb02187.x

Mallet J (2005) Hybridization as an invasion of the genome. Trends Ecol Evol 20:229–237. https://doi.org/10.1016/j.tree.2005.02.010

Martinez PA, Berbel-Filho WM, Jacobina UP (2013) Is formalin fixation and ethanol preservation able to influence in geometric morphometric analysis? Fishes as a case study. Zoomorphology 132:87–93. https://doi.org/10.1007/s00435-012-0176-x

McKay SI, Miller PJ (1997) The affinities of European sand gobies (Teleostei: Gobiidae). J Nat Hist 31:1457–1482. https://doi.org/10.1080/00222939700770791

Milinski M, Bakker TCM (1992) Costs influences sequential mate choice in sticklebacks, *Gasterosteus aculeatus*. Proc R Soc Lond B Biol Sci 250:229–233. https://doi.org/10.1098/rspb.1992.0153

Miranda MBB, Innes DJ, Thompson RJ (2010) Incomplete Reproductive Isolation in the Blue Mussel (*Mytilus edulis* and *M. trossulus*) Hybrid Zone in the Northwest Atlantic: Role of Gamete Interactions and Larval Viability. Biol Bull 218:266–281. https://doi.org/10.1086/BBLv218n3p266

Montanari SR, van Herwerden L, Pratchett MS, Hobbs J-PA, Fugedi A (2012) Reef fish hybridization: lessons learnt from butterflyfishes (genus *Chaetodon*). Ecol Evol 2:310–328. https://doi.org/10.1002/ece3.83

Moran RL, Catchen JM, Fuller RC (2019) Genomic resources for darters (Percidae: Etheostominae) provide insight into postzygotic barriers implicated in speciation. Mol Biol Evol. https://doi.org/10.1093/molbev/msz260

Muhlfeld CC, McMahon TE, Boyer MC, Gresswell RE (2009) Local habitat, watershed, and biotic factors influencing the spread of hybridization between Native Westslope Cutthroat Trout and Introduced Rainbow Trout. Trans Am Fish Soc 138:1036–1051. https://doi.org/10.1577/T08-235.1

Nei M (1973) Analysis of gene diversity in subdivided populations. Proc Natl Acad Sci 70:3321–3323

Nei M, Chesser RK (1983) Estimation of fixation indices and gene diversities. Ann Hum Genet 47:253–259. https://doi.org/10.1111/j.1469-1809.1983.tb00993.x

Nei M, Kumar S (2000) molecular evolution and phylogenetics, 1st edn. Oxford University Press, Oxford

Nelson JS, Grande TC, Wilson MVH (2016) Fishes of the world. JWiley

Ohki S, Kowalski RK, Kitanobo S, Morita M (2015) Changes in spawning time led to the speciation of the broadcast spawning corals *Acropora digitifera* and the cryptic species *Acropora* sp. 1 with similar gamete recognition systems. Coral Reefs 34:1189–1198. https://doi.org/10.1007/s00338-015-1337-4

Öhlund G, Bodin M, Nilsson KA, Öhlund S-O, Mobley KB, Hudson AG, Peedu M, Brännström Å, Bartels P, Præbel K, Hein CL, Johansson P, Englund G (2020) Ecological speciation in European whitefish is driven by a large-gaped predator. Evol Lett 4:243–256. https://doi.org/10.1002/evl3.167

Ortiz-Barrientos D, Counterman BA, Noor MAF (2004) The genetics of speciation by reinforcement. PLoS Biol 2:e416. https://doi.org/10.1371/journal.pbio.0020416

Palumbi SR (1994) Genetic Divergence, Reproductive Isolation, and Marine Speciation. Annu Rev Ecol Syst 25:547–572. https://doi.org/10.1146/annurev.es.25.110194.002555

Paradis E (2010) pegas: an R package for population genetics with an integrated–modular approach. Bioinformatics 26:419–420

Parchman TL, Gompert Z, Braun MJ, Brumfield RT, McDonald DB, Uy JAC, Zhang G, Jarvis ED, Schlinger BA, Buerkle CA (2013) The genomic consequences of adaptive divergence and reproductive isolation between species of manakins. Mol Ecol 22:3304–3317. https://doi.org/10.1111/mec.12201

Parrish JK (1989) Layering with depth in a heterospecific fish aggregation. Environ Biol Fishes 26:79–85. https://doi.org/10.1007/BF00001024

Pavlov DS, Kasumyan AO (2000) Patterns and mechanisms of schooling behavior in fish: a review. J Ichthyol 40:S163–S231

Pham-Gia T, Turkkan N, Eng P (1993) Bayesian analysis of the difference of two proportions. Commun Stat Theory Methods 22:1755–1771. https://doi.org/10.1080/03610929308831114

Polgar G, Sacchetti A, Galli P (2010) Differentiation and adaptive radiation of amphibious gobies (Gobiidae: Oxudercinae) in semi-terrestrial habitats. J Fish Biol 77:1645–1664. https://doi.org/10.1111/j.1095-8649.2010.02807.x

Pritchard JK, Stephens M, Donnelly P (2000) Inference of population structure using multilocus genotype data. Genetics 155:945–959

R Core Team (2018) R: a language and environment for statistical computing. R Foundation for Statistical Computing, Vienna

Rawson PD, Slaughter C, Yund PO (2003) Patterns of gamete incompatibility between the blue mussels *Mytilus edulis* and *M. trossulus*. Mar Biol 143:317–325. https://doi.org/10.1007/s00227-003-1084-x

Rice WR (1987) Speciation via habitat specialization: the evolution of reproductive isolation as a correlated character. Evol Ecol 1:301–314. https://doi.org/10.1007/BF02071555

Robertson DR, Justines G (1982) Protogynous hermaphroditism and gonochorism in four Caribbean reef gobies. Environ Biol Fishes 7:137–142

Robertson DR, Van Tassell J (2019) Shorefishes of the Greater Caribbean: online information system. Version 2.0, Smithsonian Tropical Research Institute, Balboa, Panamá

Rosser NL (2015) Asynchronous spawning in sympatric populations of a hard coral reveals cryptic species and ancient genetic lineages. Mol Ecol 24:5006–5019. https://doi.org/10.1111/mec.13372

Sadedin S, Littlejohn MJ (2003) A spatially explicit individual-based model of reinforcement in hybrid zones. Evolution 57:962–970. https://doi.org/10.1111/j.0014-3820.2003.tb00308.x

Selwyn JD (2015) How does dispersal affect chaos: viscous populations and families in a common Caribbean goby (*Coryphopterus personatus*). Texas A&M University, Corpus Christi

Selwyn JD, Hogan JD, Downey-Wall AM, Gurski LM, Portnoy DS, Heath DD (2016) Kin-aggregations explain chaotic genetic patchiness, a commonly observed genetic pattern, in a marine fish. PLoS ONE 11:e0153381. https://doi.org/10.1371/journal.pone.0153381

Selwyn JD, Usseglio P, Hogan JD (2021) Spatial distribution and habitat usage of *Coryphopterus personatus* and *C hyalinus* in Turneffe Atoll, Belize. Zenodo. https://doi.org/10.5281/zenodo.5348484

Serna Rodríguez KM, Zapata FA, Mejía-Ladino LM (2016) Diversity and distribution of fishes along the depth gradient of a coral reef wall at San Andrés Island, Colombian Caribbean. Bol Investig Mar Costeras-INVEMAR 45:15–39

Smith C, Cotter A, Grinsted L, Bowolaksono A, Watiniasih NL, Agnarsson I (2019) In a relationship: sister species in mixed colonies, with a description of new *Chikunia* species (Theridiidae). Zool J Linn Soc 186:337–352. https://doi.org/10.1093/zoollinnean/zly083

Stelkens RB, Schmid C, Seehausen O (2015) Hybrid breakdown in cichlid fish. PLoS ONE 10:e0127207. https://doi.org/10.1371/journal.pone.0127207

Tåning ÅV (1952) experimental study of meristic characters in fishes. Biol Rev 27:169–193. https://doi.org/10.1111/j.1469-185X.1952.tb01392.x

Tarpey CM, Seeb JE, McKinney GJ, Templin WD, Bugaev A, Sato S, Seeb LW (2017) Single-nucleotide polymorphism data describe contemporary population structure and diversity in allochronic lineages of pink salmon (*Oncorhynchus gorbuscha*). Can J Fish Aquat Sci. https://doi.org/10.1139/cjfas-2017-0023

Taylor MS, Hellberg ME (2005) Marine raiations at small geographic scales: speciation in neotropical reef Gobuies (*Elacatinus*). Evolution 59:374–385. https://doi.org/10.1111/j.0014-3820.2005.tb00996.x

Teske PR, Sandoval-Castillo J, Golla TR, Emami-Khoyi A, Tine M, von der Heyden S, Beheregaray LB (2019) Thermal selection as a driver of marine ecological speciation. Proc R Soc B Biol Sci 286:20182023. https://doi.org/10.1098/rspb.2018.2023

Thacker CE (2009) Phylogeny of Gobioidei and placement within Acanthomorpha, with a new classification and investigation of diversification and character evolution. Copeia 2009:93–104

Thacker CE (2015) Biogeography of goby lineages (Gobiiformes: Gobioidei): origin, invasions and extinction throughout the Cenozoic. J Biogeogr 42:1615–1625. https://doi.org/10.1111/jbi.12545

Thompson JD, Higgins DG, Gibson TJ (1994) CLUSTAL W: improving the sensitivity of progressive multiple sequence alignment through sequence weighting, position-specific gap penalties and weight matrix choice. Nucleic Acids Res 22:4673–4680. https://doi.org/10.1093/nar/22.22.4673

Thresher RE (1984) Reproduction in reef fishes. T.F.H. Publications, Neptune City

Toews DPL, Brelsford A (2012) The biogeography of mitochondrial and nuclear discordance in animals. Mol Ecol 21:3907–3930. https://doi.org/10.1111/j.1365-294X.2012.05664.x

Tornabene L, Ahmadia GN, Berumen ML, Smith DJ, Jompa J, Pezold F (2013) Evolution of microhabitat association and morphology in a diverse group of cryptobenthic coral reef fishes (Teleostei: Gobiidae: *Eviota*). Mol Phylogenet Evol 66:391–400. https://doi.org/10.1016/j.ympev.2012.10.014

Vanderpham JP, Nakagawa S, Closs GP (2013) Habitat-related patterns in phenotypic variation in a New Zealand freshwater generalist fish, and comparisons with a closely related specialist. Freshw Biol 58:396–408. https://doi.org/10.1111/fwb.12067

Vehtari A, Gelman A, Simpson D, Carpenter B, Bürkner P-C (2021) Rank-normalization, folding, and localization: an improved $\widehat{R}$ for assessing convergence of MCMC (with discussion). Bayesian analysis. https://doi.org/10.1214/20-ba1221

Victor BC (2019) *Coryphopterus* and *Lythrypnus* gobies of the Western Atlantic. In: Photogr. Guide Larvae Coral Reef Fishes. http://www.coralreeffish.com/gobiidae2adult.html. Accessed 22 Sep 2019

Waples RS (2015) Testing for Hardy-Weinberg proportions: have we lost the plot? J Hered 106:1–19. https://doi.org/10.1093/jhered/esu062

Ward RD, Zemlak TS, Innes BH, Last PR, Hebert PDN (2005) DNA barcoding Australia's fish species. Philos Trans R Soc B Biol Sci 360:1847–1857. https://doi.org/10.1098/rstb.2005.1716

Wellenreuther M, Clements KD (2007) Reproductive isolation in temperate reef fishes. Mar Biol 152:619–630. https://doi.org/10.1007/s00227-007-0713-1

Whiteman EA, Côté IM (2004) Monogamy in marine fishes. Biol Rev 79:351–375. https://doi.org/10.1017/S1464793103006304

Williams ME, Bentley MG (2002) Fertilization success in marine invertebrates: the influence of gamete age. Biol Bull 202:34–42. https://doi.org/10.2307/1543220

Wirtz P (1999) Mother species–father species: unidirectional hybridization in animals with female choice. Anim Behav 58:1–12. https://doi.org/10.1006/anbe.1999.1144

Wong BBM, Jennions MD (2003) Costs influence male mate choice in a freshwater fish. Proc R Soc Lond B Biol Sci. https://doi.org/10.1098/rsbl.2003.0003

Wood AJ, Ackland GJ (2007) Evolving the selfish herd: emergence of distinct aggregating strategies in an individual-based model. Proc R Soc B Biol Sci 274:1637–1642. https://doi.org/10.1098/rspb.2007.0306

Wootton RJ (1985) Energetics of reproduction. In: Tytler P, Calow P (eds) Fish energetics: new perspectives. Springer Netherlands, Dordrecht, pp 231–254

Publisher's Note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.