Reevaluating claims of ecological speciation in *Halichoeres bivittatus*

Dan L. Warren¹ | Ron I. Eytan² | Alex Dornburg³ | Teresa L. Iglesias⁴ | Matthew C. Brandley⁵,⁶ | Peter C. Wainwright⁷

Abstract

Allopatry has traditionally been viewed as the primary driver of speciation in marine taxa, but the geography of the marine environment and the larval dispersal capabilities of many marine organisms render this view somewhat questionable. In marine fishes, one of the earliest and most highly cited empirical examples of ecological speciation with gene flow is the slippery dick wrasse, *Halichoeres bivittatus*. Evidence for this cryptic or incipient speciation event was primarily in the form of a deep divergence in a single mitochondrial locus between the northern and southern Gulf of Mexico, combined with a finding that these two haplotypes were associated with different habitat types (“tropical” vs. “subtropical”) in the Florida Keys and Bermuda, where they overlap. Here, we examine habitat assortment in the Florida Keys using a broader sampling of populations and habitat types than were available for the original study. We find no evidence to support the claim that haplotype frequencies differ between habitat types, and little evidence to support any differences between populations in the Keys. These results undermine claims of ecological speciation with gene flow in *Halichoeres bivittatus*. Future claims of this type should be supported by multiple lines of evidence that illuminate potential mechanisms and allow researchers to rule out alternative explanations for spatial patterns of genetic differences.

1 | INTRODUCTION

Marine environments are host to a great deal of the world’s biodiversity, yet the relative rarity of obvious barriers to dispersal coupled with long pelagic larval durations offers few opportunities for long-term genetic isolation or population structuring that are generally thought to be prerequisites of the speciation process (Palumbi, 1992, 1994). Not surprisingly, understanding the processes leading to speciation and the accumulation of biodiversity in marine environments has been an area of intense research in recent decades (Bowen et al., 2013; Cowman et al., 2017; Faria et al., 2021; Gaboriau et al., 2018; Hodge & Bellwood, 2016). Some have argued that speciation in marine systems is primarily allopatric—as in terrestrial systems (Mayr, 1954)—with barriers to gene flow being more cryptic in the ocean than on land (Goetze, 2005, 2011; Taylor & Hellberg, 2006). In contrast, a number of recent case studies have suggested that ecological processes likely play a major role in promoting speciation in marine systems, often with gene flow (Rocha et al., 2005; Prada & Hellberg, 2020; Taylor & Hellberg, 2005; Whitney et al., 2018; Momigliano et al., 2017; Teske et al., 2019; Faria et al., 2021; Rüber et al., 2003).
In a pioneering study, Rocha et al. (2005) presented evidence supporting the possibility of ecological speciation in coral reef fishes, presenting two possible cases of parapatric speciation in Atlantic Halichoeres. One of these case studies focused on Halichoeres bivittatus, in which they demonstrate a deep (3.6%) divergence in cytochrome B (cytb) sequences between a northern "subtropical" lineage (spanning the northern Gulf of Mexico, peninsular Florida, and the eastern coast of the United States) and a southern "tropical" lineage (spanning the Yucatan peninsula, Cuba, the eastern Bahamas, and all points south including the southern Caribbean and coastal Brazil).

Finding a deep divergence at a locus with geographic structure is not in itself evidence of speciation. For example, such a divergence can be expected even under neutral processes (Irwin, 2002), in particular with respect to mitochondrial loci such as cytochrome B (Irwin, 2002; Neigel & Avise, 1993; Taylor & Hellberg, 2006). However, Rocha et al. (2005) presented evidence that the two haplotypes were preferentially associated with different types of habitat in the Florida Keys and Bermuda, with individuals of the northern lineage being found in inshore areas that experience colder minimum temperatures, while the southern lineage dominated in populations that experienced warmer and more stable temperature regimes. They also pointed to the long pelagic larval stage of H. bivittatus and apparent connectivity between populations separated by large geographic distances to suggest that there was significant potential for gene flow between northern and southern H. bivittatus populations. In light of this, they argued that the genetic divergence seen between lineages represented by these major haplotype groups was unlikely to be explained by geographic distance. The finding of genetic divergence in the face of gene flow, combined with habitat partitioning in the contact zone between the two haplotypes, led the authors to conclude that ecological processes either had driven or were in the process of driving parapatric speciation in this system. If true, this represents a departure from the more common pattern of speciation in this clade (Wainwright et al., 2018) and other Caribbean fishes, in which new species seem to primarily arise from vicariance or long-distance dispersal events (Choat et al., 2012; Robertson et al., 2006).

At present, the support for parapatric ecological speciation in H. bivittatus hinges almost entirely on the demonstration of habitat segregation in areas where the two lineages overlap. However, support for habitat segregation in the Florida Keys was based on samples of only two populations and included larval samples, which may demonstrate patterns of microhabitat segregation due to reasons unrelated to speciation. Here, we present the results of an attempt to further explore patterns of habitat segregation for H. bivittatus in the Florida Keys by sampling additional populations of adults and conducting more extensive statistical analyses.

2 METHODS

To test habitat partitioning among Halichoeres bivittatus haplotypes, we analyzed the same mitochondrial cytochrome B fragment as Rocha et al. (2005) for thirteen additional populations/collection sites. We sampled eight populations in the Florida Keys including four populations on the edge of the continental shelf (Sombrero Light, 11 Foot Mound, XMuta, and Tennessee Reef), two populations on patch reefs in the inshore channel (East Washerman and East Turtle Shoal), and two grass beds located directly offshore in water <2m in depth (near mile marker 62 on Long Key and behind Keys Marine Lab (KML) on Vaca Key). For a broader geographic context, we also sampled fishes from two sites further north on the Gulf Coast of Florida, two sites in the Bahamas, and one site from Belize. Florida and Bahamas specimens were collected in 2005 and 2006, and Belize specimens were collected in 2006. In addition to comparing fore reef and inshore patch reef, we included the grass bed habitat as it experiences even greater seasonal and diurnal fluctuations in temperature than the inshore patch reef and as such provides an additional test of the proposed habitat segregation.

All animal handling procedures were approved by the University of California, Davis Institutional Animal Care and Use Committee. Fish were caught using a combination of hand nets, barrier nets, and otter trawls. Specimens were euthanized using MS-222 dissolved in seawater, and samples were taken from muscle tissue and preserved in 95% ethanol. We extracted DNA using DNeasy™ (Qiagen) columns and PCR-amplified a 723-base pair fragment of the mitochondrial cytochrome B gene using the L14768 and H15496 primers from Rocha et al. (2005). PCR products were cleaned using ExoSap-IT (USB Corp.). Purified templates were dye-labeled using BigDye (ABI) and sequenced on an ABI 3107 automated DNA Sanger sequencer.

Of the 225 individuals of H. bivittatus sampled for Rocha et al. (2005), only 12 sequences have been made available on GenBank (Benson et al., 2013), accession numbers AY823558.1 to AY823569. All of these are included in our analyses. We aligned sequences using ClustalW (Thompson et al., 2002) and inferred a population phylogeny using BEAST v.2.6.3 (Bouckaert et al., 2014). All cytb sequences were imported into BEAUTI and partitioned by codon position. All partitions had trees and clocks linked, while site models were allowed to vary. We used ModelTest with "transition-TransversionSplit" (Bouckaert & Drummond, 2017) to infer site models. BEAST analyses were run twice, with 50,000,000 steps of the Markov chain, sampling every 1,000 generations. After removing 10% of trees for burn-in and combining the two runs in LogCombiner, a maximum clade credibility tree was generated in TreeAnnotator with median node heights. For consistency with Rocha et al. (2005), we also conducted a separate analysis using the TN93 model. All analytical results from the trees inferred with this model were functionally identical to those from the full Bayesian procedure, however, and will not be presented here. We implemented a strict molecular clock and a constant coalescent tree model, as is appropriate for population genetic data when not inferring population size changes (Drummond & Rambaut, 2007). We constructed a strict consensus tree using the "contree" function in the APE R package, and used it to assign individuals to either "northern" or "southern" haplotypes for visualization and further analysis.

We conducted population genetic analyses using the R packages adegenet and hierfstat (Goudet, 2005; Jombart, 2008). Because
some sites were represented by only a few individuals, we pooled sites by habitat types: "offshore reef," "inshore reef," or "inshore grass bed." To assess whether haplotypes were segregating between different populations, we measured pairwise Fst (Nei, 1987) between all pairs of habitat types in the Florida Keys. To evaluate the statistical significance of these patterns, we compared the observed genetic distance between habitat types with that expected if the assortment of haplotypes was random. The expected patterns under this null hypothesis were estimated using a permutation test in which sequences were randomly assigned to habitat types, keeping sample sizes consistent with those from the empirical data. In order to test whether the results were robust to our assignment of populations to habitat types, we repeated the analyses without pooling sites. Further details of the analysis and all code are provided in the supplemental materials.

Additionally, we used a single-threshold generalized mixed Yule-coalescent (GMYC) model (Pons et al., 2006) for single-locus species delimitation analyses. The GMYC model uses an ultrametric tree to infer a shift between Yule speciation and coalescent processes, using this shift to delimit species. The GMYC method was implemented using the "splits" package in R and the consensus tree inferred using BEAST. We then used a likelihood-ratio test to test the hypothesis that more than one species was present in our dataset.

3 | RESULTS

Phylogenetic and broad-scale biogeographic patterns were concordant with those seen in Rocha et al. (2005), showing a deep (~5%) divergence between a broadly northern and a broadly southern lineage (Figures 1 and 2). We found an approximately equal mix of the two haplotypes in the Florida Keys. In contrast, the Bahamas were dominated by the southern haplotype, with only one individual out of the forty having the northern haplotype. We note that this individual was the only Bahamas specimen obtained from GenBank and that no fine-scale locality information was available for it. Given that the authors providing the original data (Rocha et al., 2005) report the Bahamas as being home to the southern lineage of *Halichoeres bivittatus*, however, it is possible that this sequence was misidentified when it was posted to GenBank. Similarly, we find that of the two examples from the Virgin Islands in the original study that were available on GenBank, one was from the southern lineage and one was from the northern lineage.

Our GMYC analysis showed no evidence for more than one species in our dataset (LRT: $p = .313$). Moreover, as shown in Figure 3, the finer-scale analyses of haplotypes in the Florida Keys do not support the hypothesis of habitat segregation presented in Rocha et al. (2005). Under that hypothesis, we would expect to see a significant association between haplotype group and collection site, with inshore patch reef and grass bed populations primarily represented by the southern "subtropical" haplotype and continental shelf populations dominated by the southern "tropical" haplotype. With higher power to detect differences as a consequence of sampling more individuals in this region (78 specimens vs. 36 specimens), more populations (8 vs. 2), and more diverse habitats (i.e., with the inclusion of fore reef, inshore patch reef, and shallow grass bed populations), we find no strong evidence for the hypothesis that there are differences in allele frequencies between habitat types or individual populations. Comparison of sites grouped by habitat type showed no significant differences (Figure S1.1). For the site-level analysis, the only statistically significant difference between any pairs of populations was between the fore reef site XMutA and the single individual from the KML grass bed site. This result is likely an artifact of permutation tests conducted with a small sample size (4 samples from one population and 1 from the other; see Figure S1.2). Moreover, in this sole exception, the direction of the difference was opposite to that expected: The specimens sampled from the fore reef were of the northern haplotype, while the lone individual from the shallow grass bed was of the southern haplotype (Figure 3).

4 | DISCUSSION

In the current study, we attempted to replicate a widely cited study of parapatric ecological speciation in marine fishes. Our results do not support the hypothesis that the northern and southern lineages of *Halichoeres bivittatus* represent a product of either cryptic or incipient ecological speciation. We find no evidence that these two lineages represent different species. Further, we find no evidence for habitat partitioning between inshore patch reefs, grass beds, and reefs on the edge of the continental shelf in the Florida Keys. On the contrary, our study finds that northern and southern lineages are randomly distributed among habitat types and populations in this region. The only site-by-site comparison in the Florida Keys that was significantly different from random assortment was in the opposite direction to that predicted.

Demonstrating speciation with gene flow is notoriously difficult. For these purposes, we find lists of criteria such as those presented by Potkamp and Fransen (2019) to be of particular value; they allow us to quickly quantify the strength of evidence for a given process and adjust our level of confidence accordingly. They suggested six criteria that needed to be addressed:

1. Are populations reproductively isolated?
2. Is there (potential for) disruptive selection?
3. Do populations mate assortatively in sympathy?
4. Is the selected trait linked to the assortment trait?
5. Is there evidence for gene flow between populations at the time of divergence?
6. Do geographic ranges of populations overlap?

Similarly, Nosil (2012) established criteria for considering a case of speciation to be "ecological speciation," which are effectively the same as criteria 1, 2, and 4 above. The case for parapatric speciation and ecological speciation in *Halichoeres bivittatus* so far only consists of direct support for criterion 6: The presence of both major
FIGURE 1  Haplotype networks showing similarity between cytb sequences at a broad geographic scale (top), between different sampling locations in the Florida Keys (middle), and between different habitat types in the Florida Keys (bottom). Circle sizes represent the count of each haplotype, while colors represent locality (top and middle) and habitat type (bottom).
cytochrome B haplotype groups in some locations strongly supports the presence of geographically overlapping populations. However, this pattern could also come about if the divergence seen in cytochrome B were entirely due to allopatric divergence followed by secondary contact, or due to neutral processes (Irwin, 2002), and as such is not sufficient to support any mechanism of speciation. Consideration of barriers to north/south dispersal that may contribute to allopatric speciation may be particularly relevant, as the broad geographic distribution of haplotypes presented in Rocha et al. (2005, figure 1) and the current study very closely matches one of the most significant faunal breaks in the region for fishes (Robertson & Cramer, 2014, figure 2) and other marine groups (summarized in Robertson & Cramer, 2014, figure 1), suggesting that geography and current regimes in the region may create barriers to gene flow that are relatively consistent across a broad range of taxa. Examining the other criteria, we find that questions 1 and 3 have not been addressed in any study, while the remainder are supported only by verbal arguments based on the dispersal capability of the group (criterion 5) and the previous finding of habitat assortment in the Keys and Bermuda (criteria 2 and 4). We note that in parapatric ecological speciation, substantial differentiation in mitochondrial haplotypes would only be expected to arise well after strong selection linking ecological divergence and assortative mating, and as such, we should expect to see fairly strong evidence for criteria 1–4 in this system.

As we could not replicate the sampling of Rocha et al. (2005) in Bermuda or Key Largo, it is still possible that habitat partitioning is occurring in those localities. In light of our findings and the lack of any demonstration of morphological differentiation or assortative mating between northern and southern lineages, however, we find it difficult to see how such highly localized habitat partitioning could be considered evidence for either ecological speciation or speciation with gene flow in the rest of the Caribbean and Gulf of Mexico. Instead, we caution that the apparent differences in haplotype frequencies in these populations demonstrated by Rocha et al. (2005) could be driven by a number of processes that are not necessarily associated with speciation, including lottery recruitment and postrecruitment selection related to local conditions (Bernardi et al., 2012; Grorud-Colvert & Sponaugle, 2011; Searcy & Sponaugle, 2001; Selkoe et al., 2006). The use of sequences from larvae in Rocha et al. (2005) in some populations makes it particularly difficult to eliminate these processes as alternative explanations for patterns seen in *H. bivittatus*. As such, we would suggest that even those localized results should be viewed with extreme caution until they have been replicated with adult fish over a longer timescale.

There is a growing body of evidence that ecological factors play an important role in structuring the genetic diversity of marine populations and promoting speciation (Prada & Hellberg, 2020; Taylor & Hellberg, 2005; Whitney et al., 2018; Momigliano et al., 2017; Teske et al., 2019; Holt et al., 2020; Bird et al., 2011; Choat et al., 2012; Potkamp & Fransen, 2019; Faria et al., 2021; Rüber et al., 2003), and failure to replicate one study is not sufficient cause to question the growing consensus that ecological speciation and speciation with gene flow play an important role in generating marine biodiversity. Likewise, there is an abundance of evidence that allopatry has also promoted speciation in marine settings (Chenuil et al., 2018;
Ekimova et al., 2019; Holt et al., 2020; Laakkonen et al., 2021; Wainwright et al., 2018). We should not be surprised that in such a species-rich and unique environment, there is evidence for a variety of speciation mechanisms. The question is when should we conclude that the weight of evidence supports a given scenario. While many in the field might suggest that our default position should be one of assuming allopatric speciation until proven otherwise, we are less confident that this is the appropriate stance to take for marine environments. Rather, we suggest that when we do not know the answer to four of the six criteria for demonstrating speciation with gene flow, or to any of the three criteria for demonstrating ecological speciation, the most appropriate position is to simply acknowledge that we have insufficient evidence to argue for any mechanism of speciation in this system. “We don’t know” is a deeply unsatisfying answer, but it is the only one that accurately reflects the currently available evidence.

**ACKNOWLEDGMENTS**

The authors would like to thank the generous assistance of the Keys Marine Lab, FSU Coastal and Marine Laboratory, Dolphin Encounters, and Gulf Specimen Marine Laboratory. We would also like to thank Lisa Tipsword, Lonnie Anderson, and Colleen Young for assistance in the field, as well as the help of Kyle Miller from the Florida Fish and Wildlife Conservation Commission and Roland Albury in the Bahamas Department of Marine Resources (Fisheries). The authors are also grateful to Luiz Rocha, who provided helpful feedback on an earlier version of the manuscript.

**CONFLICT OF INTEREST**

None declared.

**AUTHOR CONTRIBUTIONS**

Dan L. Warren: Conceptualization (equal); data curation (lead); formal analysis (lead); funding acquisition (supporting); investigation (lead); methodology (lead); project administration (lead); resources (supporting); visualization (lead); writing-original draft (lead); writing-review & editing (lead). Ron I. Eytan: Conceptualization (supporting); data curation (supporting); formal analysis (equal); investigation (supporting); methodology (equal); writing-original draft (equal); writing-review & editing (equal). Alex Dornburg: Conceptualization (supporting); formal analysis (supporting); investigation (supporting); methodology (supporting); writing-original draft (equal);
writing-review & editing (equal). **Teresa L. Iglesias**: Conceptualization (supporting); data curation (supporting); funding acquisition (supporting); investigation (equal); methodology (supporting); project administration (supporting); visualization (equal); writing-original draft (equal); writing-review & editing (equal). **Matthew C. Brandley**: Conceptualization (supporting); Data curation (supporting); formal analysis (supporting); funding acquisition (supporting); investigation (supporting); methodology (supporting); resources (supporting); writing-original draft (equal); writing-review & editing (equal). **Peter C. Wainwright**: Conceptualization (equal); formal analysis (supporting); funding acquisition (equal); investigation (equal); methodology (equal); project administration (supporting); resources (equal); supervision (equal); writing-original draft (equal); writing-review & editing (equal).

**DATA AVAILABILITY STATEMENT**

The DNA alignment used here is available on Dryad, https://doi.org/10.5061/dryad.fj6q573vf. Cytochrome B sequences for this study are available from GenBank, accession numbers MZ516972–MZ517136. All R code for analyses and figures is included in the online supplementary materials.

**ORCID**

Dan L. Warren [https://orcid.org/0000-0002-8747-2451](https://orcid.org/0000-0002-8747-2451)

Ron I. Eytan [https://orcid.org/0000-0002-4625-4589](https://orcid.org/0000-0002-4625-4589)

Alex Dornburg [https://orcid.org/0000-0003-0863-2283](https://orcid.org/0000-0003-0863-2283)

Teresa L. Iglesias [https://orcid.org/0000-0002-0237-8539](https://orcid.org/0000-0002-0237-8539)

Matthew C. Brandley [https://orcid.org/0000-0002-3734-913X](https://orcid.org/0000-0002-3734-913X)

Peter C. Wainwright [https://orcid.org/0000-0003-0498-4759](https://orcid.org/0000-0003-0498-4759)

**REFERENCES**


**SUPPORTING INFORMATION**

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