2026

Ecological and Evolutionary Divergence of a Shared Sympatric Color Polymorphism in Paracirrhites Hawkfish

Samuel Greaves
University of Central Florida, sa939468@ucf.edu

ECOLOGICAL AND EVOLUTIONARY DIVERGENCE OF A SHARED SYMPATRIC COLOR POLYMORPHISM IN *PARACIRRHITES* HAWKFISH

by

SAMUEL GREAVES
B.A. Biology, Carleton College, 2016

A dissertation submitted in partial fulfillment of the requirements
for the degree of Doctor of Philosophy
in the Department of Biology
in the College of Sciences
at the University of Central Florida
Orlando, Florida

Spring Term
2026

Major Professor: Michelle Gaither

ABSTRACT

Color polymorphism provides a tractable window into how selection, gene flow, and ecology interact to maintain phenotypic diversity and initiate divergence. I investigate the ecological and genomic basis of sympatric light and melanistic morphs in two reef hawkfishes, *Paracirrhites arcatus* and *P. hemistictus*, across Guam and Hawai‘i. This study aims to determine whether the color morphs of these hawkfishes are a stable polymorphism or have begun the process of ecological speciation. For *P. arcatus* revealed substantially lower density on Guam (0.086 ind/m²) than reported for Hawai‘i (0.292 ind/m²) and a strong skew away from the melanistic morph on Guam (7.0%) relative to Hawai‘i (39.6%), indicating geography-dependent morph structure and habitat context. Redundancy analysis showed that a simplified model explained slightly more variance than the model developed for the Hawaiian population, while multidimensional niche-overlap tests detected significant differentiation among morphs despite weak univariate habitat preferences. Whole-genome resequencing (148 genomes) recovered modest population differentiation (Guam-Hawai‘i mean F_{ST} = 0.0047) and slightly higher between-morph genome-wide F_{ST} on Guam (0.0015) than Hawai‘i (0.0010). Sliding-window scans and SNP outlier tests identified numerous regions of elevated differentiation, including pigmentation genes (e.g., *gpnmb*, *sox9*, *sfxn1*), supporting “islands of divergence” under gene flow. In *P. hemistictus* (69 genomes), morph divergence was markedly greater (mean F_{ST} = 0.0750), with spatial sorting across Orote Peninsula and selective signatures (e.g., *Chd1*) pointing to more advanced separation without single diagnostic loci. Together, these results show that similar light/melanistic polymorphisms can follow distinct trajectories driven by ecological conditions and including competitive environments. Sympatric incipient divergence is supported in *P. arcatus*, particularly in Hawai‘i, while there is deeper genome-wide separation in the morphs of *P. hemistictus*.

ACKNOWLEDGMENTS

I am immensely grateful to everyone who has helped me along this long journey. Thank you to my committee members for your guidance and giving me the tools and knowledge to conduct by study in the first place. Dr. Eric Hoffman taught me about population genetics and had valuable insight about color polymorphisms since he studied it for his dissertation. Dr. Robert Fitak taught me a lot about genomics and was our strongest supporter when Tiffany and I tried to help our department come back together after COVID though Tea Time. Dr Terry Donaldson provided his singular knowledge about hawkfish and immense logistical support by hosting me for my field season in Guam. Dr. Jonathan Whitney, who provided the basis for this study, samples, and has helped me make sense of the analyses. Finally, Dr. Michelle Gaither, has been more than I could have ever asked for in an advisor. You have taught me, encouraged me, and set me on a path for success. You have given me the opportunity and resources to pursue what I wanted to do, and been there for me on too many occasions when things were down. I am so grateful to have joined this lab and hope to keep working together for a long time.

I would also like to thank those who helped me with my project. First of all, the UoG team - Ka'ohinani Kawahigashi, who was my amazingly competent dive buddy for my entire field season; Nanny, our captain who got us exactly where we needed to be and kept us safe; Andrew McInnis, who collected the specimens for me while Guam was still on lockdown, and Kelsie Ebling-Whited, who carefully catalogued them and got them to us! Additionally, Jewel Rehpkof, Isabella Reyes, and Elijah Burgos pulled off the gargantuan task of classifying almost 1000 substrate photos. I would also like to thank my labmates, Pavi, Ashley, Emily, Kathryn, JK, Michelle (Shaffer!), Taylor, and Ella, as well as our postdocs Girish, Mel, and Mary, for all of their assistance, feedback, companionship, and shared trust issues with the Qubit over the years.

I am also grateful to my community over the past 8 years. My friends back home and the 4th Burtonites have always been there for me, mostly over Zoom now, but have been integral to my life. BGSA was helping me feel welcome even before I started and has provided much social and academic support. Our Wednesday Trivia group – TJ, Erin, Rachel, and Matt, has kept me sane by giving me a little break while finishing up my dissertation. I am happy I had such a good cohort in Davide, Katie, Pavi, and Tiffany. I am particularly grateful to have had such an amazing friend and roommate in Tiffany. I am happy to have been your dive and fruit picking buddy, and am glad you were there to discuss science at midnight, pick me up from the airport, and survive a pandemic with. My family, supporting me even though I'm on the other coast doing something that makes no sense.

Finally, I miss those who have departed while I was working on this. Dad, I wish you were still here. You probably started this whole mess taking me to aquariums and encouraging me when I was little. Loki – even though you clearly liked Tiffany more, you were there when I needed a distraction, consolation, or some funny (and surprisingly quiet) husky antics.

This work was made possible with funding from the University of Central Florida and the NSF Guam EPSCoR program under Grant No. OIA-1946352. I have been supported by the University of Central Florida Trustee's Fellowship and the National Science Foundation Graduate Research Fellowship Program under Grant No. DGE-2035702.

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

LIST OF ABBREVIATIONS

INT: Intermediate morph

MEL: Melanistic morph

PWS: Pink-white-stripe morph

YWS: Yellow-white-stripe morph

CHAPTER ONE: INTRODUCTION

Coloration is an important trait that serves a wide variety of biological functions. This includes communication and signaling, camouflage, photoprotection, and thermoregulation (Hubbard et al., 2010; Protas & Patel, 2008). Encompassing pigmentation, physiology, behavior, and development, coloration is a complex phenotype with potentially major impacts on an organism's fitness. Color polymorphism is the presence of two or more stable and heritable color phenotypes within a single population which occur at a frequency higher than expected under genetic drift (Ford, 1945; White & Kemp, 2016). Color polymorphisms are particularly useful to study because they are easily classified with non-invasive methods and have a high potential to affect both an organism's fitness as well as its mate choice. Of particular interest is the role of these polymorphisms in sympatric speciation, as they represent an opportunity for genetic and morphological differences to accumulate within a single population (Gray & McKinnon, 2007; Sergey, 2003).

My dissertation research takes advantage of the presence of sympatric color morphs across the ranges of the reef-associated fishes *Paracirrhites arcatus* and *P. hemistictus* to gain insight into how variation in both genotype and phenotype behaves between the morphs and species (Randall, 1963). Previous work on *P. arcatus* found that the color morphs there were partially sorting by microhabitat (DeMartini & Donaldson, 1996; Whitney, Donahue, et al., 2018), and demonstrated some limited genetic divergence in microsatellite loci as well as a single nucleotide polymorphism in the melanocortin 1 receptor gene, that is responsible for the expression of the pigment melanin (Whitney, Bowen, et al., 2018). This study aims to gain a broader genomic and ecological view to determine whether the color morphs of these hawkfishes are being maintained as a stable polymorphism or are in the early stages of ecological speciation, and in doing so, broaden our understanding of the evolutionary dynamics of variation.

This project consists of three parts: First, I analyze whole genome data from *P. arcatus* from Guam and Hawai'i to compare the genomic signatures of each morph, including separate populations. Next, I conducted *in situ* surveys to establish microhabitat occupancy and behavioral association between the morphs. Finally, I expanded the analysis by conducting additional surveys and whole genome resequencing of *P. hemistictus* on Guam. For purposes of publication, the chapters are organized as follows. Chapter 2 bundles the two reference genomes as a resource. I combined the first two aims for Chapter 3, focusing on an ecological and genomic comparison of *P. arcatus* by morph between Guam and Hawai'i. For Chapter 4, I take a similar look at the ecological and genomic differences between morphs of *P. hemistictus* and examine how the polymorphism compares between the two species along those factors. By approaching the light/melanistic dimorphism from ecology and genetics, across two populations of *P. arcatus*, and between two species within the genus *Paracirrhites*, this study is a major expansion to our knowledge of hawkfish color morphs, which were previously only examined in Hawai'i with *P. arcatus* (DeMartini & Donaldson, 1996; Whitney, Bowen, et al., 2018; Whitney, Donahue, et al., 2018) and the Red Sea with *P. forsteri* (Coker et al., 2017; Gaither et al., 2020). The replication across space and taxa will either reveal a common and thus more broadly applicable background and mechanism between locations and/or species, or that they are operating independently despite their shared phenotype.

CHAPTER TWO: GENOME ASSEMBLY AND ANNOTATION OF THE HAWKFISHES *PARACIRRHITES ARCATUS* AND *PARACIRRHITES HEMISTICTUS*

Introduction

The arc-eye hawkfish *Paracirrhites arcatus* (Cuvier & Valenciennes, 1829) and the black-spotted hawkfish *P. hemistictus* (Günther, 1874) are tropical reef-associated fishes of the family Cirrhitidae (Figure 1A &1B). These species are broadly distributed, with *P. arcatus* recorded across the tropical Indo-Pacific from the Red Sea to Hawai'i, while *P. hemistictus* has a smaller overlapping range, from Christmas Island in the Indian Ocean to French Polynesia, and is notably absent from Hawai'i (Randall, 2005). Hawkfish are mesopredators, which perch primarily on corals and prey on passing small fish and invertebrates (Kane et al., 2009). There are no available genomes within the family Cirrhitidae, with the closest being a recently generated assembly for the blackbarred morwong, *Cheilodactylus zonatus* (Chang et al., 2025) representing Cheilodactylidae within the suborder Cirrhitoidei.

Color polymorphisms within the genus *Paracirrhites* have been under study as there is a mix of similar and divergent polymorphisms across the species. There are six accepted species within the genus *Paracirrhites,* with *P. arcatus, P. forsteri,* and *P. hemistictus* having documented color polymorphisms that are sympatric across their geographic ranges (FishBase, 2026; Randall, 2005). While *P. forsteri* has more than four color morphs (Coker et al., 2017; Gaither et al., 2020), *P. arcatus* and *P. hemistictus* exhibit a simple and similar light/melanistic color dimorphism that is not linked with sex or ontogeny. Within these species, color morphs are sympatric and present on the same reefs (DeMartini & Donaldson, 1996; Gaither et al., 2020; Whitney, Donahue, et al., 2018). Existing work has shown some limited differentiation in habitat use between morphs in *P. arcatus* and *P. forsteri* (Coker et al., 2017; Whitney, Donahue, et al., 2018). Additionally, outlier loci have been found to distinguish the morphs with microsatellite loci for *P. arcatus* (Whitney, Bowen, et al., 2018) and *de novo* RADseq loci for *P. forsteri*

(Gaither et al., 2020). There have been no genetic studies on the other species within the genus, including *P. hemistictus*. However, *P. hemistictus* was originally described as two separate species based on color morphology (*Cirrhites hemistictus,* light, *Cirrhites polystictus*, melanistic) before being synonymized by Marshall (1950).

The presence of a shared color polymorphism within the genus *Paracirrhites* presents an opportunity to study the ecological and evolutionary factors that drive the creation and maintenance of color polymorphisms in closely related taxa. However, impeding this work has been a lack of genomic resources. Here, we report chromosome-level annotated genome assemblies for two species within the family Cirrhitidae: *P. arcatus* and *P. hemistictus.* PacBio HiFi long read, Hi-C proximity ligation, and RNA-seq libraries were generated from single individuals of each species collected from Guam. We leveraged these data to create chromosome-level high quality scaffolds and an expressed sequence tag-based annotation. These resources will facilitate genomic comparisons across color morphs both within and between species and help us to learn more about the evolution of the color polymorphism.

Results and Discussion

A total of 1.8 million HiFi reads and 137 million pairs of Hi-C Illumina reads were sequenced for *P. arcatus,* and 2.2 million HiFi reads and 156 million Hi-C read pairs for *P. hemistictus.* For Phase 0 (See Table 1 for Phase 1), the genome assemblies total 696 Mb for *P. arcatus* and 686 Mb for *P. hemistictus*, each consisting of 24 anchored scaffolds totaling 688 Mb and 667 Mb, respectively, or >97% of each assembly. These represent the same number of chromosome-level scaffolds resolved in the blackbarred morwong, *Cheilodactylus zonatus* (Chang et al., 2025). However, our two *Paracirrhites* genomes are almost 100 Mb larger than the *C. zonatus* assembly indicating differences in genome size within the suborder Cirrhitoidei. For *P. arcatus,* scaffold sizes range from 16-40 Mb while for *P. hemistictus* they range from 13-36 Mb.

Both genomes are highly complete, achieving BUSCO scores of greater than 99%. The annotation was consistent, with 18856 genes in *P. arcatus* and 18675 genes in *P. hemistictus.* RepeatMasker identified 21.68% of the *P. arcatus* and 22.10% of the *P. hemistictus* genomes as repetitive. While similar to each other, this is lower than the 33.72% repetitive sequence found in *C. zonatus*, which is primarily due to a lower amount of identified transposable elements, at about 4.3% for the two *Paracirrhites,* but 13.1% for *C. zonatus.* Synteny between *P. arcatus* and *P. hemistictus* shows the same number of scaffolds with no major translocations between putative chromosomes. There are a large number of translocations or inversions across most of the scaffolds (Figure 1c). For species within the same genus, this is relatively high and deserves some confirmation and follow-up study using additional long read sequencing.

Conclusion

Prior to this study, there were no species within the family Cirrhitidae with reference genomes. The closest relative is the recently published blackbarred morwong, *Cheilodactylus zonatus* (Genebank: GCA_050575295.2). These genomes will represent the second and third within the suborder Cirrhitoidei. Together, these genomic resources will allow comparative genomics within hawkfishes and particularly learn more about how color morphology evolves within this clade.

Materials and Methods

Sample Collection and Sequencing

We collected single individuals of the light morph of *P. arcatus* and *P. hemistictus* from Guam in 2021. These individuals were euthanized and immediately dissected. Muscle, skin, gills, eyes, brain, liver, and heart tissues were collected and flash frozen in liquid nitrogen, then stored at −80 °C. Hi-C libraries were prepared from muscle tissue using the Proximo Hi-C 4.0 Kit (Phase Genomics). DNA within intact cells was crosslinked using formaldehyde, then digested using the Sau3AI restriction enzyme. The ends were

filled in with biotinylated nucleotides, and then ligated to create chimeric fragments whose frequency reflected their proximity. Hi-C libraries were then sequenced on an Illumina NovaSeq 6000 S4 flowcell to generate 2x150 reads. To generate circular consensus sequencing (CCS) libraries we used the PacBio HiFi Prep Kit to prepare brain tissue into SMRTbell libraries. Libraries were then sequenced on an 8M HiFi flowcell on a Sequel 2 sequencer (Pacific Biosciences, Menlo Park, CA).

Assembly Pipeline

The initial assembly was generated from the CCS sequences using IPA v1.0.4 (Sovic et al., 2020) with default settings. Hi-C short reads were then aligned to the IPA assembly using BWA-MEM 0.7.17 (Li, 2013) with the −5SP and -t 8 options specified, and all other options as default settings. We used the program SAMBLASTER (Faust & Hall, 2014) to flag PCR duplicates, then alignments were then filtered with samtools (Li et al., 2009) to remove non-primary and secondary alignments and the PCR duplicates previously flagged. FALCON-Phase v2 (Kronenberg et al., 2021) was used to correct phase switching errors in the primary contigs and alternate haplotigs from IPA, forming assemblies for each phase. Next the Proximo Genome Scaffolding Platform (Phase Genomics) was run with 20,000 iterations to process the HI-C proximity ligation data to orient and order contigs into chromosome-scale scaffolds using the contact frequency matrix created by aligning the Hi-C reads to the phase-0 assembly (Bickhart et al., 2017; Burton et al., 2013). To correct scaffolding errors, Juicebox (Durand et al., 2016; Rao et al., 2014) was run, followed by a second run of FALCON-Phase v2 to correct phasing using scaffold data and generate assembly files for both phases with identical contig order and orientation. These data were used to generate the final diploid assembly via a custom script by Phase Genomics (https://github.com/phasegenomics/juicebox_scripts). Completeness of the final genome was assessed using Benchmarking Universal Single-Copy Orthologs (BUSCO) v6.0.0 (Simão et al., 2015) with the Actinopterygii_obd10 (2024-01-08) dataset.

Gene Annotation

To create the annotation, we prepared cDNA libraries from the frozen muscle, skin, gills, eyes, liver, and heart tissues using the NEBNext Ultra II Directional RNA Library Prep Kit for Illumina (New England Biolabs), and sequence them on . We first used Trinity v2.12.0 (Grabherr et al., 2011) to assemble mRNA reads into a transcriptome. Trinity output was used as expressed sequence tag evidence, and the uniprot_sprot database (The UniProt Consortium, 2018) (downloaded January 2020) as protein homology evidence for the MAKER annotation pipeline v2.31.10 (Holt & Yandell, 2011). Repetitive elements were identified and classified using RepeatModeler v1.0.11 and RepeatMasker v4.0.9 (Smit & Hubley, 2017; Smit et al., 2019) relative to Repbase-derived RepeatMasker libraries v20181026 (Bao et al., 2015) and then used for soft repeat-masking of the assembly. We generated gene prediction models from the initial MAKER output using Augustus (Stanke et al., 2006), then used MAKER again to integrate it into the final assembly. To assign functional names to the annotated genes we used BLASTX (Altschul et al., 1990; Camacho et al., 2009) with the uniprot_sprot database release 2021_01 (Boutet et al., 2007). Finally, we ran a synteny analysis between the genomes using the program ntSynt v1.0.3 (Coombe et al., 2024) and a divergence preset of 4. The synteny plot was generated with ntSynt-viz v1.0.3 (Coombe et al., 2025).

Tables and Figures

Table 2-1. Genome assembly and annotation statistics for *P. arcatus* and *P. hemistictus*.

	P. arcatus		*P. hemistictus*	
Assembly	**Phase 0**	**Phase 1**	**Phase 0**	**Phase 1**
Assembly Length (Mb)	695.8	695.6	686.2	692.4
No. of scaffolds	24	24	24	24
Scaffold N50 (Mb)	30.6	30.2	28.4	28.5
No. of contigs	334	334	264	264
Contig N50	11.5	11.5	10.5	10.5
No. contigs in scaffolds	268 (80.2%)	268 (80.2%)	222 (84.1%)	222 (84.1%)
Length in scaffolds (Mbp)	688.2 (98.9%)	688.0 (98.9%)	667.4 (97.3%)	673.6 (97.28%)
GC (%)	43.74%	43.74%	43.77%	43.73%
Hi-C Read Pairs	137163193		155765762	
Hi-C Gb	41.1		46.7	
HiFi CCS Reads	1538349		2202433	
HiFi CCS Gb	26.9		33.0	
BUSCO Complete	3626 (99.6%)		3619 (99.4%)	
Annotation				
RNA-seq Read Pairs	31222364		68441818	
RNA-seq Gb	9.4		20.5	
No. of genes	18856		18675	
No. expressed transcripts	392985		397730	
Mean gene size (bp)	10888		10875	
Repetitive elements	21.68%		22.10%	
Retroelements	6.35%		5.76%	
DNA Transposons	4.27%		4.33%	

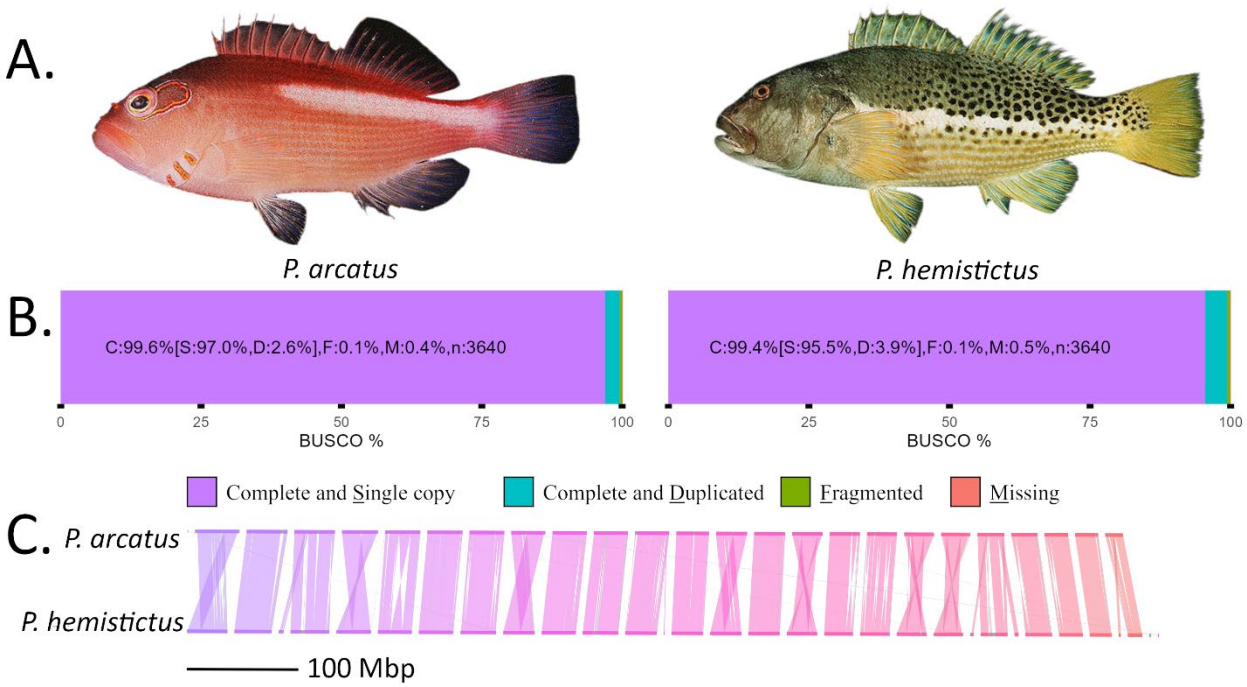


Figure 2-1 A) Left: Light morph of *Paracirrhites arcatus.* Right: Light morph of *P. hemistictus.* Images modified from Randall (1997) B) BUSCO results for each genome (Left: *P.* arcatus, right: *P. hemistictus*) using the Actinopterygii_obd10; 2024-01-08 reference c) Synteny Plot comparing *P. hemistictus,* and *P. arcatus*. Synteny was calculated with ntSynt (Coombe et al., 2024) then visualized with ntSynt-viz v1.0.3 (Coombe et al., 2025).

CHAPTER THREE: ECOLOGICAL AND GENOMIC DIVERGENCE IN THE SYMPATRIC COLOR MORPHS OF THE ARC-EYE HAWKFISH: A CASE STUDY ACROSS GUAM AND HAWAI'I

Introduction

Coloration in animals is an important trait that serves a number of key biological functions. Color itself can provide physiological benefits, such as photoprotection and thermoregulation, while in some cases pigment molecules have structural roles or function as antioxidants (Hubbard et al., 2010; McNamara et al., 2021). In animals, color is frequently involved in communication and signaling or to provide camouflage from predators, prey, and conspecifics (Hubbard et al., 2010; Protas & Patel, 2008). Color morphology is often used to distinguish and evaluate potential mates and is one of the most frequently implicated traits in sexual selection (Lande, 1980; McMillan et al., 1999; Price & Birch, 1996; Wagner et al., 2012). In some species, color variants are linked to ontogenetic changes with juveniles displaying markedly different phenotypes than adults, as is common in reef fishes. Some taxa are capable of temporary color changes that can convey important information to conspecifics, such as mate quality, and in some cases, individuals can change color rapidly to match backgrounds for camouflage (Lorenz, 1962).

Under specific circumstances, selection can promote the maintenance of more than a single discrete color phenotype within a species. The presence of multiple heritable color phenotypes within a population is known as a color polymorphism. This does not include temporary changes but instead the discrete phenotypes are stable and occur at a higher frequency than expected under genetic drift (Ford, 1945; Gray & McKinnon, 2007). A common form of color polymorphism in animals is sexually dimorphic coloration that often also includes ornamentation (Wellenreuther et al., 2014). These sex-linked color

polymorphisms have been the subject of extensive study and underpin much of our understanding of sexual selection (Wellenreuther et al., 2014). Alternatively, non-sex-linked polymorphisms vary widely and occur when multiple phenotypes have a selective advantage, with their maintenance generally attributed to balancing selection. For instance, the apostatic coloration of prey species can be affected by frequency-dependent selection as successful patterns increase in frequency, and then become recognized by pattern-searching of predators (Allen & Clarke, 1968; Endler, 1988). This has been shown to occur in Trinidadian guppies (Olendorf et al., 2006). There have been numerous studies linking non-sex-linked color polymorphisms to cryptic background matching (Harris et al., 2020) such as the classic examples of industrial melanism in the *Biston* moths during the Industrial Revolution in Great Britan (Kettlewell 1955; Majerus 1998) or the light and dark morphs of the desert pocket mouse (*Chaetodipus intermedius*) that occupy neighboring granite and lava habitats (Hoekstra et al., 2004) (Gray & McKinnon, 2007; Jamie & Meier, 2020). This form of background matching has also been observed in populations that occupy geographically distinct habitats yet that maintain high levels of gene flow (reviewed in Harris et al. (2020)). Color phenotype has also been shown to signal reproductive strategy (Kr-strategies). For instance, in the common wall lizard (*Podarcis muralis*), yellow-throated females lay many small eggs (r), white-throated females lay fewer large eggs (K), and red females represent the transition phase between the two strategies, and clutches had reduced viability when their mate was of a different morph (Galeotti et al., 2013).

As a trait, color can be under particularly strong selection and has been linked with speciation. Among the Midas cichlids of Lake Xiloá in Nicaragua, sexual selection is thought to have led to diversification in this group, which includes at least 13 species (Elmer et al., 2009). Color has also been shown to simultaneously affect individual survival and serves as an important cue for mate choice and thus acts as a self-reinforcing “magic trait” (Servedio et al., 2011) that drives differentiation between color morphs. For instance, Lake Victoria has a highly complex set of habitats within a single lake, and

the hundreds of species of African cichlids found there are thought to have diverged as they adapted to the environmental parameters of these habitats, such as different light environments. Under this scenario the co-evolution of signaling and sensory systems, known as sensory drive, is thought to have led to the diversification of this group (Carleton et al., 2005; Seehausen et al., 2008; Wright et al., 2020). In the Caribbean hamlets (genus *Hypoplectrus*), selection for aggressive mimicry and strong assortative mating among color variants are theorized to have led to that species radiation, which includes 17 described species (Lobel, 2011; Picq et al., 2016; Puebla et al., 2008; Puebla et al., 2007; Puebla et al., 2014; Puebla et al., 2022).

Hawkfishes (family Cirrhitidae) of the genus *Paracirrhites* are coral reef-associated fishes with several species that maintain sympatrically distributed color morphs (DeMartini & Donaldson, 1996; Gaither & Randall, 2012). Two species, *P. arcatus* and *P. hemistictus*, display light and melanistic color morphs, while a third species, *P. forsteri,* has a much more complex color polymorphism with at least four distinct morphs (Gaither et al., 2020). While there is some regional variation, the color polymorphisms extend throughout the geographic range of each species, which in the case of *P. arcatus* and *P. forsteri* extends from the east coast of Africa to the Hawaiian Archipelago. In the Red Sea, the color morphs of *P. forsteri* differ in estimated growth rate and depth preference (Coker et al., 2017; Gaither et al., 2020) with genetic differentiation between morphs detected at only a small subset of outlier loci (Gaither et al., 2020). Meanwhile for *P. arcatus* in Hawai'i, a moderate degree of fine scale niche partitioning has been identified, with melanistic individuals more abundant in the high energy, shallow portions of the reef while the light morph is more frequently associated with *Porites lobata* in calmer, often deeper water (DeMartini & Donaldson, 1996; Whitney, Donahue, et al., 2018). Genetic analyses show that the two color morphs of *P. arcatus* are distinguished by a single SNP within the *Mc1r* (melanocortin 1 receptor) gene (F_{ST} = 0.175) as well as shifts in allele frequencies in five microsatellite loci (Whitney, Bowen, et al., 2018). Together, these studies suggest that the color polymorphisms in

Paracirrhites hawkfishes may be due to adaptation to different environments within the same reef. However, geographic variation could play a role (McLean & Stuart-Fox, 2014), and it is unclear if the niche and genetic separations detected in these studies extend to other geographic locations across the species' range, and if genomic mechanisms of divergence are maintained.

Here, I work to resolve these questions and investigate the relationship between color polymorphism, habitat use, and genome level differentiation in the hawkfish *P. arcatus.* Extending the work of Whitney, Donahue, et al. (2018) in Hawai'i, I conduct field surveys in Guam to determine whether the color morphs (Figure 1A) demonstrate similar habitat partitioning in these two geographically distant locations. Furthermore, I generate a reference genome for *P. arcatus* and conduct whole genome resequencing of the original samples from Hawai'i (Whitney, Bowen, et al., 2018) as well as new collections from Guam. I employed genome scans to identify regions of differentiation that distinguish the two color morphs and analyze these in a geographical context, to ask whether the ecological and genomic level patterns exhibited by color morphs are similar across these two geographically distant populations (Toonen et al., 2016).

Materials and Methods

Ecological Field Surveys in Guam

Field surveys were conducted in June–July 2022 at five sites along the west coast of Guam (Figure 1C) following the methods of Whitney, Donahue, et al. (2018), with minor modifications. Transects were placed on the seaward slopes of each site, to obtain a depth gradient and because the species and their coral hosts are much more common there. All surveys were conducted by the same pair of scuba divers, in sets of 25 x 4 m transects. The transect tape was placed along depth contours running parallel to the reef crest. The shallowest transect was run in the surge zone or as shallow as the reef topology would allow. Subsequent transects were placed ~5-10 m downslope, descending in depth until the rarity of the

target species or logistical constraints (weather, site access) prohibited further transects. To determine substrate type, I randomly pre-selected 15 out of the possible 50 half-meter intervals along each transect using the *sample* function in R v3.6.0 (R Core Team, 2019) and took photos on alternating sides of the transect tape using a Hero 10 Black camera (GoPro, San Mateo, CA) mounted 47 cm above the center of a 50 x 70 cm PVC photoquadrat. In addition, the shallowest and deepest depths of every 5 x 4 m block of the transect were recorded as well as an estimate of wave action. Substrate photos were later analyzed to obtain estimates of percent cover of each substrate type. To do this I randomly generated 100 points per image using PhotoQuad v1.4 (Trygonis & Sini, 2012). Each photo was scored separately by two people who identified the substrate at each of the 100 random points as either sand, rock, algae, dead coral, or live coral identified to genus. Discrepancies in substrate scoring were resolved by a third evaluator. Finally, estimates of substrate cover were pooled per transect.

Along each transect, each individual *P. arcatus* was recorded including position along the transect, color phenotype, depth, and a visual estimate of total body length. The substrate upon which each individual rested was recorded as either sand, rock, algae, dead coral, or live coral identified to genus (DeMartini, 1996). It was noted when multiple *P. arcatus* individuals were observed on the same coral head or close to each other. One diver of the buddy pair recorded data while the other ensured that all individuals were recorded.

Statistical Analysis

Statistical analyses were conducted in R v4.3.1 (R Core Team, 2019). To determine if the color morphs differed in abundance across our transects (within sites), I calculated Levins' standardized spatial overlap index (Levins, 1968) by transect and by depth (2 m bins) using the *niche.overlap* function in the package spaa v0.5.3 (Levins, 1968; Zhang & Ma, 2013). This index is a measure of the overlap between pairs of taxa with values ranging from one (complete overlap) to 0 (no overlap or complete spatial sorting).

To determine if environmental variables could predict the abundance of color morphs, I ran redundancy analyses (RDA) in the package vegan v2,6-5 (Oksanen et al., 2023). I constructed the response matrix to include the abundance of each morph per transect, while the explanatory matrix included transect depth, slope, wave action, as well as percent cover of each of the substrate types described above. I tested for collinearity in the explanatory variables using variance inflation factor cutoff of 10. Two models were tested: one matching the best model identified by Whitney et al. (Whitney, Donahue, et al., 2018) in Hawai'i (hereafter the "Whitney model"), and the other chosen through forward selection (hereafter the "Guam model"). Model significance was tested using 100,000 permutations and the results were visualized using the first two axes. To determine if the color morphs demonstrate a preference for habitat types, I employed Manly's selectivity index (wi) type I (Manly et al., 2002) with the function 'widesI' in the R package adehabitatHS v0.3.16 (Calenge, 2006). An index of 1 indicates that the habitat is used according to availability, while an index of greater than 1 indicates habitat preference and an index of less than 1 indicates habitat avoidance. Habitat use was measured per transact as the percentage of fish per substrate type, while habitat availability was the percent cover per transect of each substrate type. Differences in Manly's selectivity indices between morphs were compared using a log-likelihood chi-square test in R. As a more nuanced measure of resource partitioning, I conducted a niche overlap analysis using the R scripts of Geange et al. (2011). This method creates a composite measure of the morphs utilization of habitat variables and compares it with null distribution of the niche space using 100,000 iterations of permutation tests to determine if two groups overlap, and how evenly.

Genomic Assessments: Guam and Hawai'i

Sample Collection

I collected 74 individuals of *P. arcatus* from Guam using pole spears including 27 light [referred to as pink white-stripe (PWS) in Whitney et al., (2018) and hereafter], 18 intermediate (INT), and 29 melanistic

(MEL) morphs. Muscle and fin tissues were taken from each individual and preserved in 95% ethanol and stored at -20°C. In addition, I obtained tissues from 72 individuals collected from the island of Hawai'i and previously examined in Whitney, Bowen, et al. (2018) including 24 PWS, 24 INT, and 24 MEL morphs (Table S1). For whole genomic assembly and annotation, a separate single individual of the PWS morph was collected from Guam (see below).

Genome Resequencing

Genomic DNA was extracted from 74 individuals from Guam and 72 from Hawai'i using the E.Z.N.A. Tissue Kit (Omega Bio-tek, USA) following the manufacturer's protocol. DNA was eluted in two rounds of 75 µL of 10 mM Tris-HCl pH 8. DNA extracts were purified using Sera-Mag SpeedBead Carboxylate Modified (E3) Magnetic Beads (Cytiva Life Sciences, USA) at a 1.5x bead concentration and then quantified using Qubit High-Sensitivity 1X assay on a Qubit 4.0 (Thermo Fisher Scientific, USA). To prepare samples for sequencing I used the SparQ DNA Frag and Library Prep Kit (QuantaBio, USA) following its PCR-free workflow. I started with 500 ng of DNA and used a fragmentation time of 7 min to target a peak fragment size of 400 bp. I ligated sparQ UDI (Unique Dual Index) adapters (QuantaBio) (250 nM) to each sample, then cleaned and size selected each using 0.8x, 0.6x, and 0.75x bindings of sparQ PureMag Beads (QuantaBio) for a final size distribution of approximately 450-600 bp. I ran a random selection of 74 of the 155 libraries on an Agilent TapeStation 4200 using the D1000 DNA ScreenTape (Agilent, USA) to verify size distribution and obtain the average fragment size. The libraries were then quantified using the NEBNext Library Quantification Kit for Illumina (New England Biolabs, USA) on a Bio-Rad CFX384 qPCR machine (Hercules, CA). Libraries were pooled in equimolar amounts into two final pools (Table S1) which were assessed using the TapeStation and quantified with the NEBNext Quantification Kit. Pools were sequenced on a S4 flowcell (2x150) at the University of Florida ICBR Sequencing Core on an Illumina NovaSeq 6000. Each pool was run across two lanes shared with ddRAD libraries (approximately 500 bp) from a different project.

Following sequencing, samples with less than 10 million reads were omitted from the dataset. Initial quality filtering was done using fastp v0.23.2 (Chen et al., 2018). Reads were processed for polyG-tail and adapter sequence removal, paired-end overlap base correction, and trimmed from the 5' end using a 4 bp sliding window until the average window quality was Q25. Reads that were less than 25 bp long or had less than 75% of the bases below Q25 were removed. The remaining reads were aligned to the reference genome using BWA-MEM2 v 2.2.1 using default settings (Vasimuddin et al., 2019). Samtools v1.15 (Danecek et al., 2021; Li et al., 2009) was employed to sort reads, mark optical duplicates, and to output the mapped reads. Variants were called separately between the Guam and Hawaiian populations using the 'mpileup' and 'call' commands in bcftools v1.15 (Danecek et al., 2021). Finally, raw SNP calls were filtered using bcftools 'view' to remove loci with a sequencing depth of less than 10 per sample (approximately the expected coverage) and alleles with a minor allele frequency of less than 0.02. Finally, I retained only biallelic SNPs with less than 20% missing data.

Population Genetic Analysis and GWAS

To estimate the degree of genetic divergence between geographic locations and between color morphs, I calculated the genetic diversity statistics of overall average nucleotide diversity (π) and average heterozygosity within each population using vcftools v0.1.16 (Danecek et al., 2011). I then calculated F_{ST} estimates using 5 kb sliding windows with a 1 kb step using the '--weir-fst-pop function' in vcftools v0.1.16 (Danecek et al., 2011). Average F_{ST} was calculated from these values. F_{ST} outlier windows between color morphs were identified within each population by examining windows that exceeded the 99% bootstrapped confidence interval of the 99.9% quantile of F_{ST} estimates (calculated per scaffold). Contiguous outlier windows were combined to identify regions of high F_{ST}. In addition, F_{ST} outlier SNPs were identified per scaffold (due to computational limitations) using the R package OutFLANK (Whitlock & Lotterhos, 2015) with a q-value cutoff of 0.05. For both outlier analyses, overlapping genes were identified using the reference annotation. A PANTHER Overrepresentation analysis was run for the

outlier genes against the zebrafish reference list to identify Gene Ontology Biological Processes linked with the F_{ST} outliers, using PANTHER release 19.0 (Mi et al., 2021).

I ran a Genome-Wide Association Study on the light and melanistic morphs using the Genomic Association and Prediction Integrated Tool (GAPIT v3.3) (Wang & Zhang, 2021). A linkage-pruned dataset was made for this analysis using BCFtools v1.15, with a max R^2 of 0.25 in a 10 Mb window. The linkage-pruned dataset was converted to hapmap format using TASSEL v4 for this analysis (Bradbury et al., 2007). I used the Bayesian based BLINK algorithm (Huang et al., 2019) within GAPIT, which integrates the potentially confounding factors of linkage disequilibrium and population structure into its model. I identified candidate genes as those within 10 kb of SNPs deemed significant after FDR correction. I also tested for soft selective sweeps using the 'saltiLASSI' algorithm in lassip (DeGiorgio & Szpiech, 2022), with a window size of 50 SNPs.

Survey of Candidate Coloration Genes

A list of 128 candidate teleost coloration genes were identified from (Braasch et al., 2009). To ensure that no genes were missed due to the quality of our annotation, the sequences for all genes were downloaded from ENSEMBL release 113 (Dyer et al., 2024) using the biomaRt R package (Durinck et al., 2005; Durinck et al., 2009). Exon sequences were aligned to the reference genome using minimap2 v2.28 with the 'splice' preset (Li, 2018). Final gene coordinates were then expanded to the full gene or the expressed sequence tag that any exon overlapped with. Out of the 128 genes in the initial list, 106 had likely matches in the *P. arcatus* genome. Based on the position of these 106 genes, I exported all SNPs within the target sequences using bcftools using the filters Genotyping Quality >15, depth>10, and fraction missing <0.2. Next, I calculated F_{ST} for each SNP within the coloration genes using vcftools as above, but with the window size set to 1. Outliers were identified by scaling the F_{ST} within each gene using a Z-score to create ZF_{ST} estimates (Lai et al., 2016), adjusting them for multiple comparisons based

on a false discovery rate, and selecting those past a threshold of 0.000001. Genes containing outliers were assessed for potential biological significance by identifying those with a SNP with a maximum F_{ST} above 0.15, or an average F_{ST} of 0.1 within the gene. Candidate coloration genes were also assessed for selection by first LD pruning their SNPS using 'indep-pairwise' in plink v1.90 (1 Mb window size, R^2 cutoff of 0.1), then using BayeScan v2.1 (Foll & Gaggiotti, 2008) to identify SNPs under selection.

Results

Ecological Field Surveys in Guam

I conducted a total of 53 transects, each 25 m long, across five sites along the west coast of Guam (Figure 1C). These surveys covered depths from 0m to 14m. A total of 455 *P. arcatus* were observed across 5300 m^2 of reef habitat, resulting in a density of 0.086 individuals/m^2, in comparison to 0.292 individuals/m^2 observed in Hawai'i (Whitney, Donahue, et al., 2018). The PWS (light) morph was the most abundant constituting 58.7% (267) of the individuals observed, while 34.3% (156) were the INT (intermediate) morph, and 7.0% (32) were the MEL (melanistic) morph, while these ratios were 39.3%, 20.1%, and 39.6% respectively in Hawai'i. I recorded both PWS and INT individuals at all sites, while MEL individuals were absent from the surveys at the two sites on the unsheltered south side of Orote Peninsula (Turtle and Barracuda Rocks, Figure 1C). Based on Levin's standardized spatial overlap index (*O),* I recorded a moderate level of spatial sorting between the PWS and MEL morphs across our transects with *O* = 0.413 (CI = 0.197-0.657), while INT phenotypes had a higher level of spatial overlap with the PWS morphs (*O* = 0.771, CI = 0.587-0.960) than the melanistic morphs (*O* = 0.160, CI = 0.064-0.378). My data indicate that the morphs in Guam are not sorting by depth (Figure 1B), with all pairwise calculations of the Levin's index at or near 1 (PWS-MEL *O* = 0.988, CI = 0.767-1.052; PWS-INT *O* = 1.034, CI = 0.764-1.093; INT-MEL *O* = 0.865, CI = 0.725-1.041). Finally, based on the calculation of Manly's selectivity index (w_i), *P. arcatus* color morphs did not demonstrate a preference for a specific habitat type with insignificant values across all samples (p = 0.559), as well as by morph (PWS, p = 0.231; INT, p =

0.429; MEL, p = 0.527) driven by large variances in the w_i estimates. On a per-substrate basis, none were significant except for an avoidance of rock (w_i = 0.056, p < 0.001).

Redundancy analysis of the Guam data fit to the Whitney model (Whitney, Donahue, et al. (2018); *occurrence ~ depth + slope + wave_action + Pocillopora.meandrina + Porites.lobata + Rock*) was marginally significant (p = 0.040) and explained 27.5% of the variation, while the best fit model for the Guam data (Guam model) was *occurrence ~ Rock + Pocillopora + Porites* which was significant (p = 0.0001) and explained slightly more of the variance in the data at 29.5% (Table 1, Figure 2A). Niche overlap analysis identified evidence of significant differentiation between morphs within multidimensional niche space (all 3 morph pairs p < 0.0001, individual niche components in Table 2), indicating that morphs are responding to niches instead of individual environmental variables.

Genome Resequencing: Guam and Hawai'i

Illumina sequencing resulted in 9.1 billion reads, with a mean of 59.5 million reads and standard deviation of 48 million reads per sample. From the Illumina sequencing, 89.7% of the reads passed quality filtering and aligned to the newly assembled reference genome (Table S1). Five samples were omitted from the dataset due to low read counts (<10 million reads). The remaining 148 samples had an average of ~61.5 million reads (~13x coverage) including 67 from Guam (26 PWS, 12 INT, and 29 MEL) and 71 from Hawai'i (23 PWS, 25 INT, and 23 MEL morphs) (Table S1). After variant calling, I resolved 40 million raw and 19.4 million SNPs after filtering in the individuals from Hawai'i and 62.4 million raw and 18.2 million SNPs after filtering in the individuals from Guam.

Nucleotide diversity was higher in Guam than Hawai'i, but heterozygosity was lower (Table 3). The mean F_{ST} between the Guam and Hawai'i populations (without considering color morph) was 0.0047. When color morphology was considered, mean F_{ST} between morphs on Hawai'i was 0.0010, but 50% higher on Guam at 0.0015. Based on the 5 kb sliding window F_{ST} analysis, 459 genomic regions with elevated F_{ST}

estimates and ranging from 5-89 kb were identified including 223 in the Guam dataset and 246 in Hawai'i (Figure 3). Of these, 179 could be assigned to genes including 104 in Guam and 96 in Hawai'i (Table S2). Eight genes were identified as outliers in both locations (Table S2). Several are involved in ion transport (Ano, CALM2, slc25a1b), or cell structure (DSP, tubb2b). With 105 of the genes mapped into its database, PANTHER analysis based on the sliding window F_{ST} outliers identified several GO Biological Processes that were associated with color morph (Table 5), notably involving development. Our OUTFLANK analysis identified 6161 outlier SNPs between the color morphs in Guam and 118 in Hawai'i. These outlier SNPs could be assigned to 2269 genes, of which 16 were shared between the populations (Table S2). While there were fewer outliers detected in Hawai'i, the mean OUTFLANK outlier F_{ST} between morphs was 0.343, while it was 0.200 in Guam; about 40% lower.

After FDR correction, the GWAS did not identify any SNPs associated with color morphology in the overall dataset. However, when the data were subset by population, our GWAS analyses identified one SNP in Hawai'i and two in Guam as significantly associated with color morph. In Hawai'i, the single SNP was within 3 kb of the Inpp4a gene (p = 1.1e-12, Effect = 0.493). In Guam, one SNP was within the rabgef5 gene coding sequence (p = 1.0e-8, Effect = 0.470), and the other was within 1 kb of the Rnd3 gene (p = 1.34e-8, Effect = -0.456). When the effect is positive, the SNP is associated with the melanistic (MEL) morph, while a negative effect shows an association with the light (PWS) morph.

LASSI analysis returned 9,031 haplotypes in Hawai'i and 13,727 haplotypes in Guam as under selection. After merging overlapping windows and matching these regions to the annotation I identified 162 genes and 1,159 Expressed Sequence Tags (ESTs) in Hawai'i and 110 genes and 795 ESTs in Guam that are under selection (Figure S2). Out of these, 25 genes and 266 ESTs of unknown function were found in both populations (Table S2). While none of these genes are known to affect coloration, many are involved in cellular signaling. The shared gene with the highest LASSI likelihood scores are the C6 and

C7 immune genes, which are in close proximity on the same chromosome and code for proteins involved in pathogen binding.

Coloration Genes

While Bayescan did not identify any of the 106 candidate genes as being under selection, the SNP based F_{ST} outlier analyses detected highly differentiated SNPs between morphs, with the maximum reaching F_{ST} = 0.27 for the gpnmb gene in the Hawaiian population and F_{ST} = 0.31 for the sfxn1 gene between morphs in Guam; many times higher than the genome average of 0.0047. The number of outlier SNPs detected in each gene varied, with some genes exhibiting a small number or no outliers, while others had outliers evenly distributed throughout the gene (Figure S3). Guam had 961 outlier SNPs across 99 genes, while Hawai'i had 2,312 outlier SNPs across 97 of the coloration genes. I detected 16 genes with at least one outlier SNP with a F_{ST} over 0.15 that were shared between Guam and Hawai'i . Of these 16 genes, 15 are involved in melanophore (melanin-containing pigment cells) activity according to Braasch et al. (2009) (Table 6). Three genes (gpnmb, sox9, and rps20) had an average F_{ST} over 0.1 in both populations, all of which are linked to melanophores (Table 6).

Discussion

Studies on color polymorphism have been foundational to our understanding of how phenotypic variation persists in natural populations. Classic examples include the divergent morphs associated with crypsis or aggressive mimicry (e.g., Cook et al., 2012)), and sex-linked phenotypes maintained by sexual selection (Elmer et al., 2009). Although color polymorphism is an important topic in evolutionary biology, studies have typically focused on only one or a few populations, lacking a geographic context (McLean & Stuart-Fox, 2014), and only recently have advances in sequencing technologies allowed us to explore the links between phenotype and genotype in non-model systems (Casas et al., 2021; Corl et al., 2010a;

Orteu & Jiggins, 2020). Phylogeographic studies indicate a link between geographic variation in polymorphism and the formation of non-interbreeding populations (Iserbyt et al., 2010). Furthermore, color polymorphisms have been associated with accelerated rates of speciation in taxa of birds (Hugall & Stuart-Fox, 2012) and reef fishes (Streelman et al., 2002) that have color polymorphic phenotypes. These indicate that variation in the number, type, and frequency of phenotypes within and across populations can have a significant evolutionary impact.

Reef fishes of the genus *Paracirrhites* display distinct color morphs that co-occur on reefs across vast areas of the Pacific and Indian Oceans, making them are an ideal group for the study of geographic variation in color polymorphisms. Specifically, the arc-eye hawkfish, *P. arcatus,* displays two relatively easy to score color morphs, in which individuals found on the same reefs have either light (PWS) or melanistic (MEL) phenotypes, with intermediate (INT) phenotypes found at varying frequencies. Previous studies in Hawai'i show that the two morphs sort incompletely across microhabitats with low but significant genetic differentiation (F_{ST} = -0.004 to 0.007) between phenotypes at microsatellite loci. To determine if similar ecological and genetic patterns are maintained in geographically isolated populations, I replicated the field surveys of Whitney et al. (2018) and conduct new collections on the island of Guam in Micronesia, 6,400 km southwest of Hawai'i. The surveys I conducted indicate that *P. arcatus*, and in particular the MEL morph, is far more abundant in Hawai'i than in Guam. While I detected some niche partitioning between morphs in Guam, it is far less pronounced than in Hawai'i. The whole genome data show seemingly conflicting results. On average, the two morphs were genetically less differentiated in Hawai'i (pairwise FST = 0.0010) compared to Guam (pairwise FST = 0.0015) However, where genetic differences were present, the difference between morphs was much stronger in Hawai'i compared to Guam In the list of candidate coloration genes, I did not detect any SNPs that were diagnostic between the morphs, but I did identify several genes with elevated F_{ST} that

were associated with melanophores but not other pigmentation cell types, thus giving a potential genetic basis for a cumulative effect that was missed by most of the other analyses.

Niche partitioning between color morphs

Based on a comparable survey area (5,300 m^2 and 6,400 m^2 in Guam and Hawai'i, respectively), I found a much lower abundance of *P. arcatus* in Guam (0.086 individuals/m^2) compared to Hawai'i (0.292 individuals/m^2) (Whitney, Donahue, et al., 2018). Guam is located in the western Pacific and is the southernmost island in the Mariana Archipelago. The region sustains a rich marine biodiversity, including over 1,100 species of reef fishes (Myers & Donaldson, 2003). The Hawaiian Archipelago is geologically young with Kure in the northwest having emerged ~30 mya, and the island of Hawai'i, where our Hawaiian samples are from, having emerged ~ 0.5 mya , while Guam began emerging 43 mya (Clague & Dalrymple, 1987; Riegl et al., 2008). The geographic isolation of the islands coupled with their young geologic age contributes to the depauperate nature of the Hawaiian fauna (Bhattacharya et al., 2022). Approximately ~615 species of fishes (Randall, 2005) and 70 species of scleractinian corals (Veron et al., 2016) have been recorded in the Hawaiian Archipelago compared to >1,000 species of reef fishes and ~377 species of corals in Guam (Myers & Donaldson, 2003). Species diversity decreases with distance from the Indo-Malayan region, a well-documented phenomena that has also been observed within the family Cirrhitidae (Donaldson, 1986).

The morph distribution was also skewed away from the melanistic morphs in Guam, with 58.7% PWS, 34.3% INT, and 7.0% MEL, while these ratios were 39.3%, 20.1%, and 39.6% respectively in Hawai'i. These proportions align with data from DeMartini (1996), that reported that ~32% of the *P. arcatus* on the island of Oahu in Hawai'i were the MEL morph, compared with ~7% at over 100 sites examined across Oceania (DeMartini & Donaldson, 1996). With a simpler community structure in Hawai'i, *P. arcatus* morphs were potentially able to expand to new niches filled by other species of fish in reefs with

higher biodiversity. The 3-fold higher abundance of *P. arcatus* in Hawai'i compared to Guam could be due to decreased competition. Furthermore, my surveys of benthic habitat reflect the broader trend of higher diversity in Guam at the specific sites where we conducted them. In Hawai'i, the shallow reefs are *Pocillopora*-dominated, replaced by species of *Porites* (*P. lobata* and *P. compressa*) with increased depth/distance to shore. These corals are preferred by *P. arcatus* in Hawai'i with little evidence of usage of other species as perches (Kane et al., 2009; Whitney, Donahue, et al., 2018). Those surveys reported that 98% of the coral cover in their survey area was comprised of the species *Pocillopora meandrina, Porites compressa, and P. lobata*. However, thirteen coral groups, plus fire coral, were above 1% of the coral cover on our surveys in Guam. While *Pocillopora spp.* were still the dominant coral in shallow areas, deeper habitat usually did not have a high *Porites spp.* cover, and the predominant species of *Porites* shifted to *P. rus* which has a plate and pillar morphology and is infrequently used by *P. arcatus*, therefore microhabitat availability also likely influenced the different distributions between the islands.

In Guam, I detected a lower degree of niche partitioning between the PWS and MEL morphs compared to Hawai'i (Levin's spatial overlap index GU = 0.413, HI = 0.360; Niche Overlap GU = 0.437, HI = 0.640). Notably, where intermediate morphs had a high overlap with melanistic morphs in Hawai'i, they instead had high overlap with the light morphs in Guam. When projecting Guam's transect data onto the RDA of (Whitney, Donahue, et al., 2018) based on surveys in Hawai'i (Figure 2B), it becomes clear that while I recorded fewer melanistic individuals in Guam, both morphs and the intermediates co-occur in the shallow, high energy niche space that in Hawai'i was occupied largely by the melanistic morph (shallow reefs with high surge). On the other hand, the niche space that was dominated by the PWS morph in Hawai'i (deeper water, low surge, with primarily *Porities compressa* and *P. lobata*) was not present in Guam. Instead, those deeper transects were dominated by *Porites rus,* which in the survey area has a plate and pillar growth form instead of the lobe form of *P. compressa* and *P. lobata* preferred by *P arcatus* that is more conducive to their camouflage (Whitney, Donahue, et al., 2018). Interestingly,

our surveys in Guam show that the congenerics *P. forsteri* and *P. hemistictus*, the latter of which is absent from Hawaiian reefs, occupy the shallower habitat that the melanistic morphs prefer (Figure S1). These species are much larger than *P. arcatus* and occupy a similar position in the food web. All three species are mesopredators that perch on corals and defend territories (Randall, 2005). It is possible that they may compete with *P. arcatus* for the same resources and thus keep their populations at lower numbers, particularly in the shallow high energy habitats that are preferred by melanistic morphs in Hawai'i. In Guam, lower abundance of preferred habitat, and greater diversity of potential competitors, including congenerics, may create a more competitive seascape that lends to reduced ecological differentiation between the morphs of *P. arcatus*

Genomic divergence between color morphs

A review of phylogeographic studies show that Hawaiian populations have lower genetic diversity and are at least partially isolated from other population across the Pacific (Gaither et al. 2011). Thus, my finding of $F_{ST} \approx 0.005$ between Guam and Hawaiian populations of *P. arcatus* was surprisingly low, especially given the difference in the ecological analyses. Whitney, Donahue, et al. (2018) observed greater genetic divergence between color morphs on the same reefs than between the same morphs at different study sites on the island of Hawai'i, but I did not observe the same across the 6,400 km between Guam and Hawai'i (Table 4). Here, comparisons between the morphs within each archipelago are nearly the same as when the populations are combined, and less than the population-level differentiation. Interestingly the pairwise F_{ST} estimate between-morphs in Guam was 50% higher than the same comparison in Hawai'i (FST = 0.0015 vs 0.0010, respectively) (Table 4). I did find that the Hawaiian population has lower genetic diversity nucleotide diversity (Table 3). Despite relatively low average heterozygosity, neither population demonstrates high inbreeding, with the Guam population slightly higher (Table 3).

I detected hundreds of outlier regions across methods, but only a few of these overlapped between Guam and Hawaii. Also, there is little consistency in the gene lists between the outlier detection methods. Window-based methods, such as the sliding window F_{ST} analysis and LASSI do detect regions potentially under selection, but the magnitude of differentiation is relatively low. Along with a lack of fixed differences between morphs, this indicates that the outlier loci are not directly responsible for the shared color morph phenotypes. If phenotype is genetically determined, it is an epistatic trait, however GWAS results only identified genes associated with morph when subset by population, and the genes are not known to affect color. My sample size had limited power for GWAS, so only a very strong association would have been detected, so this is not precluded. Nevertheless, there are some flickers of differentiation present between morphs. When F_{ST} was calculated per-base within the candidate coloration genes, I found higher levels of differentiation over smaller (1bp – 89kb) regions of the genome. This pattern is expected when divergence is occurring under gene flow, known as "genomic islands of differentiation", in which gene flow rapidly erodes the width of a locus under selection.

GO analyses of the F_{ST} outlier SNPs and genes were primarily associated with development and gene expression. We know that the white stripe and body coloration are separate traits because there are intermediate individuals with dark coloration and a stripe, or light individuals without one, but they most commonly occur together to form the light morph. The white stripe lies along the lateral line and differences in its development may be responsible for that portion of the PWS morph's phenotype. The outlier genes gnb1, fgf3 (shared), and sh3pxd2a are associated with the "posterior lateral line neuromast primordium migration" GO term, providing a possible mechanism for the stripe. Pigmentation cells are derived from neural crest stem cells (Hou et al., 2000), so it is also possible that genes associated with neural GO terms influence the pigmentation paths without being tagged as such.

I paid special attention to genes known to be involved in pigmentation of fish, identifying the locations of 106 of those identified by Braasch et al. (2009). Direct analysis of coloration genes via

Bayescan indicated that all genes were neutral; however, the F_{ST} results indicate the presence of outlier SNPs in many coloration genes. The overlap of pigmentation pathways with other essential neural and developmental pathways provides a constraint on some genes, such as wnt1, a highly conserved signaling protein essential to development. Additionally, all but one of the 19 high F_{ST} coloration genes were associated with melanophores (Table 6). While no single SNP or gene can be identified as responsible for color morphology, this adds to the signals of differentiation initially found by Whitney, Bowen, et al. (2018).

The Hawaiian morphs exhibit a few large shifts in allele frequency consistent with genomic islands of differentiation. The average F_{ST} between morphs in the Hawaiian population was an order of magnitude lower than Guam. However, on a more granular level, the sliding window analysis found eight clusters where average $F_{ST} > 0.06$ in Hawai'i, but only one cluster met this criterion in Guam. Within the coloration genes, there were over two times more outlier SNPs between Hawaiian morphs, as well as a slightly higher number of genes with SNPs exceeding an F_{ST} of 0.15 – much higher than the background rate of $F_{ST} = 0.001$, and similar to inter-population differentiation in terrestrial species. This pattern of higher F_{ST} isolated to regions of differentiation against a background of lower differentiation could indicate the Hawaiian morphs are diverging despite either higher gene flow or a more recent migration event. A pattern of genomic islands of differentiation linked with coloration and speciation were identified within the hamlet (Puebla et al., 2014) and cichlid (Keller et al., 2013) systems.

All the available evidence indicates that color phenotype in *P. arcatus* is stable within an individual's lifetime (DeMartini & Donaldson, 1996; Whitney, Donahue, et al., 2018). Therefore, color phenotype is expected to be controlled by heritable genes or perhaps epigenetically. Here I found no specific locus with sufficient fixation by morph to be fully responsible for controlling this phenotype, nor did GWAS identify an association with groups of genes. Even so, there are signs of at least partial reproductive isolation with the accumulation of regions of elevated genetic differentiation. It has been

theorized that divergence can begin as the accumulation of epigenetic differences between morphs, with genetic differentiation primarily occurring once one morph is sufficiently more fit than the other in a particular niche (West-Eberhard, 1986). If the phenotype is under epigenetic control, the observed shifts in allele frequencies could be signs of genetic conversion, particularly in the Hawaiian population.

Conclusion

Combining the geographic replication of an ecological study with deeper genomic analysis has changed the context of Whitney, Bowen, et al. (2018) and Whitney, Donahue, et al. (2018) by indicating their findings may not be representative of the species as a whole, yet are still valid. The genomic data does support the finding that Hawaiian *P. arcatus* exhibit signals of divergence. A potential explanation is that the isolation of Hawai'i is contributing to an opportunity for divergence that is not present on Guam. Hawai'i represents the north-eastern edge of the range of *P. arcatus*, while Guam is much closer, and better connected by currents, to other populations in the Coral Triangle. In addition, Guam is an older island with a different coral cover regime and more complex species assemblage. This includes an additional species from the genus, *Paracirrhites hemistictus*, which demonstrates the same light/melanistic dimorphism and may compete for the same niche space and explain the lower abundance of the dark morph in the high energy, shallower portions of the reef.

There are additional species of *Paracirrhites*, specifically *P. nisus, P. xanthus,* and *P. bicolor* that are restricted to island groups in the south-central tropical Pacific and are sometimes considered part of a species complex with *P. arcatus*. These species may represent examples of peripatric speciation, and the Hawaiian *P. arcatus* population could also be heading down this evolutionary path as new environments and limited migration provide new opportunities for divergence. The wide range of hawkfishes in *Paracirrhites* combined with the presence of several modes of color polymorphism present

an opportunity to learn more about the evolution of color morphs. By looking at fish from Hawaii and Guam, this study suggests that the edge of the species' range in Hawaii potentially releases the morphs from balancing selection and their normal niches, allowing them to begin to diverge.

Tables and Figures

Table 3-1 Comparison of redundancy analysis between the Whitney model (optimized for Hawaii populations; Whitney, Donahue, et al. (2018)) and the Guam-optimized model. The Whitney model considered the environmental variables of depth, slope, wave action, and the percent cover of *Pocillopora meandrina*, *Porites lobata*, and rock. The Guam model is a simpler model including just the percent cover of *Pocillopora spp., Porites spp.,* and rock.

	Whitney Model	**Guam Model**
RDA1 Eigenvalue	0.0529	0.0603
RDA2 Eigenvalue	0.0080	0.0049
Constrained proportion	0.2757	0.2954
ANOVA p-value	0.0402	0.0001

Table 3-2 Niche Overlap Analysis estimates based on Geange et al. (2011). Shown are pairwise estimates for light (PWS), melanistic (MEL), and intermediate (INT) morphs of *P. arcatus* on Guam. Bolded values are significantly different based on a Bonferroni corrected cutoff of $p < 0.00625$ (sep.pvals and overall.pvals).

	PWS-MEL	**PWS-INT**	**INT-MEL**	Mean
depth	**0.275**	**0.747**	**0.423**	0.482
slope	**0.481**	**0.790**	**0.454**	0.575
wave action	**0.576**	**0.718**	0.772	0.689
perch	**0.422**	0.959	**0.399**	0.593
live coral cover	0.433	0.831	0.442	0.569
Composite niche overlap (mean ± SD)	**0.437** ± 0.109	**0.809** ± 0.094	**0.498** ± 0.154	0.581

Table 3-3 Genome statistics by geographic location. Number of samples and SNPs are given post-filtering. The observed and expected heterozygosity, as well as nucleotide diversity were calculated using vcftools (Danecek et al., 2011).

	# samples	# SNPs	pi	Ho	He	F
Guam	67	18260884	0.00347	0.112	0.122	0.0685
Hawai'i	71	19365504	0.00278	0.157	0.160	0.0180

Table 3-4 Pairwise F_{ST} between populations and color morphs. F_{ST} was calculated by vcftools (Danecek et al., 2011) and averaged across the genome.

Populations	**F_{ST}**
Guam vs Hawai'i (all)	0.0047
Guam vs Hawai'i (MEL)	0.0051
Guam vs Hawai'i (PWS)	0.0046
Color morphs	**F_{ST}**
MEL vs PWS	0.0082
MEL vs PWS (Guam)	0.0015
MEL vs PWS (Hawai'i)	0.0010

Table 3-5 Gene Ontology analysis of outlier genes identified using the sliding window F_{ST} analysis. Analysis was conducted using the PANTHER Overrepresentation test against the *Danio rerio* reference gene list and corrected using the false discovery rate.

GO Biological Process	GO Accession #	Fold Enrichment	FDR
animal organ development	0048513	2.50	0.015
anatomical structure development	0048856	1.96	0.017
developmental process	0032502	1.87	0.042
system development	0048731	2.43	0.004
multicellular organism development	0007275	2.15	0.011
multicellular organismal process	0032501	1.97	0.013
transport	0006810	2.24	0.042

Table 3-6 Outlier Coloration genes. Coloration genes with SNPs that have an F_{ST} estimate greater than 0.15, or mean F_{ST} greater than 0.1. All genes except skiv2l2 are associated with melanophores. The genes vps39 and sox9 are also involved in iridophore development.

Gene Name	Gene Function	Guam Max F_{ST}	Hawai'i Max F_{ST}
adam17	melanophore development	0.167	0.202
egfr	melanophore development	0.156	0.177
itgb1	melanophore development	0.182	0.238
sfxn1	melanophore development	0.308	0.152
gpnmb	components of melanosomes	0.176	0.269
ap3d1	melanosome construction	0.164	0.157
fig4	melanosome construction	0.170	0.163
hps3	melanosome construction	0.177	0.178
lyst	melanosome construction	0.164	0.158
vps39	melanosome construction	0.152	0.205
myo5a	melanosome transport	0.157	0.177
myo7a	melanosome transport	0.185	0.160
atrn	regulation of melanogenesis	0.171	0.195
drd2	regulation of melanogenesis	0.164	0.153
mycbp2	pteridine synthesis	0.164	0.153
skiv2l2	uncategorized function	0.203	0.163

Gene Name	Function	Guam Mean F_{ST}	Hawai'i Mean F_{ST}
sox9	melanophore development	0.139	0.114
gpnmb	components of melanosomes	0.104	0.124
rps20	systemic effects	0.122	0.103

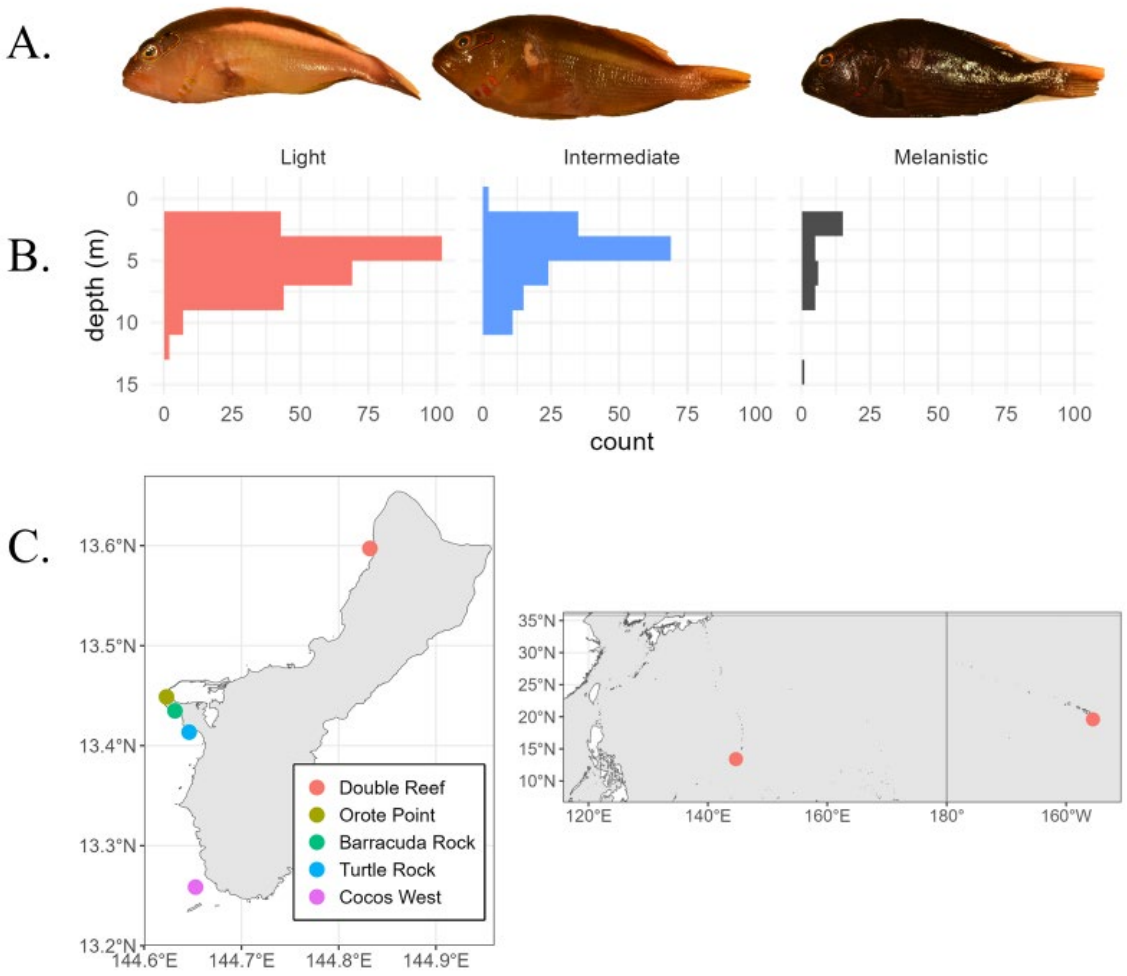
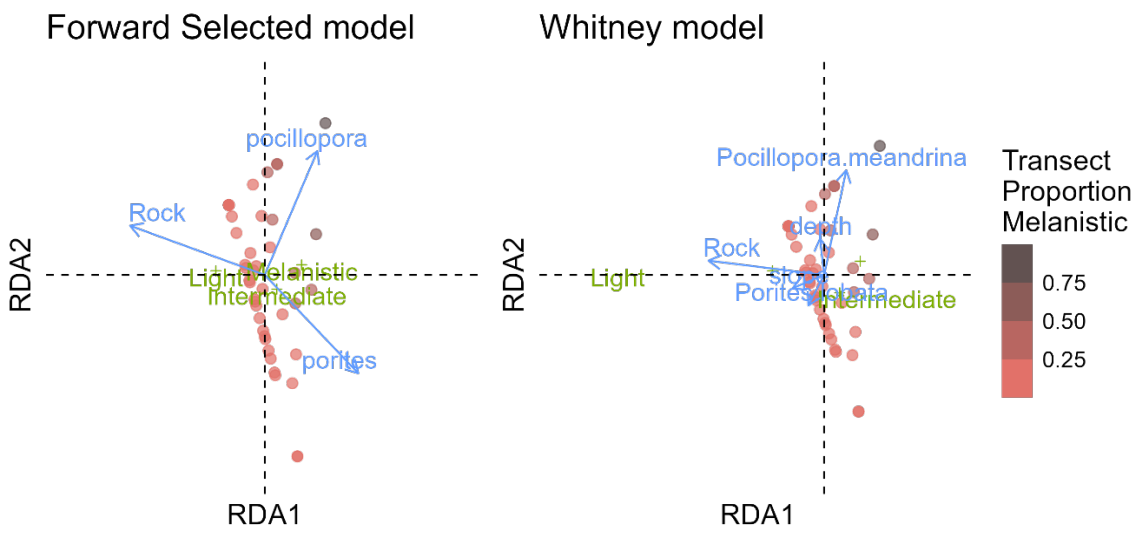


Figure 3-1 A) The light morph (pink-white stripe: PWS) is characterized by a pink body color and a white stripe along the lateral line while the melanistic morph (MEL) is melanated across their entire body. Intermediate morphs (INT) have a melanated body

A.



B

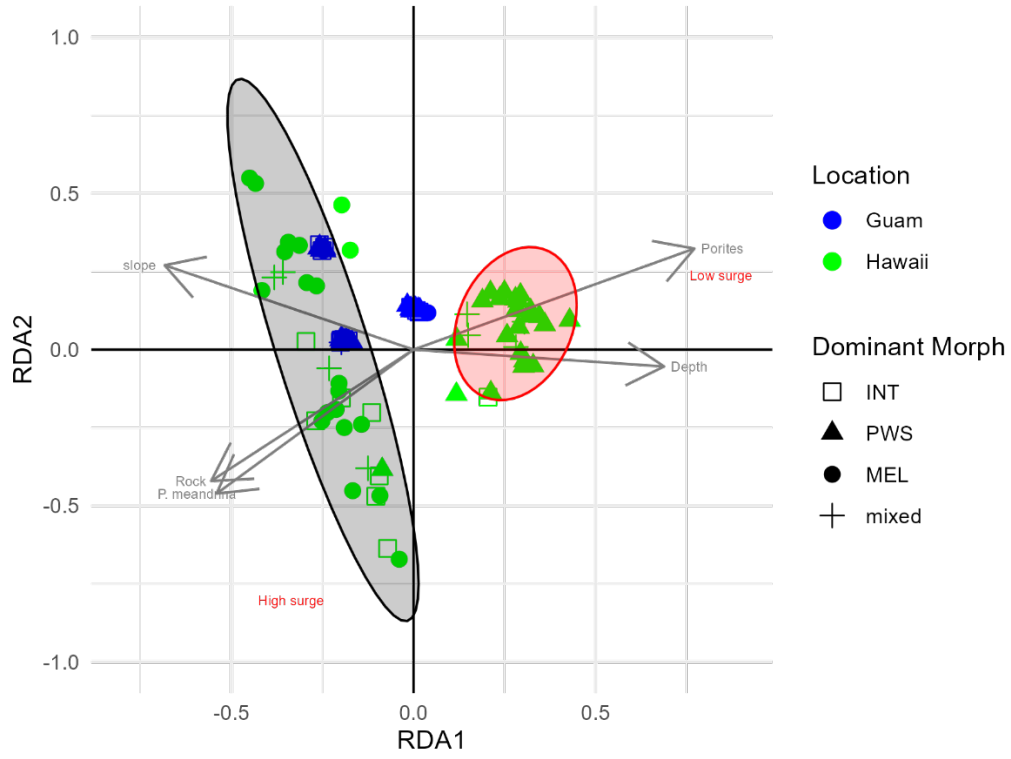


Figure 3-2 A) Redundancy Analysis. Redundancy analysis triplots showing a comparison of the RDA model developed for this study with that from Whitney, Donahue, et al. (2018). The forward-selected

model from this study predicted the relative abundance of the *P. arcatus* color morphs on Rock + *Pocillopora* + *Porites*, while the Whitney model used depth + slope + wave action + *Pocillopora meandrina* + *Porites lobata* + Rock. Note that only five of the transects had >=50% melanistic morphs. B) RDA projection of Guam transect data (blue) onto the Hawaiian ecospace from Figure 3a of (Whitney, Donahue, et al., 2018). There is no clustering of transects by dominant morph in the Guam data, and transect environmental data is more centered (less distinct) than those in Hawai'i.

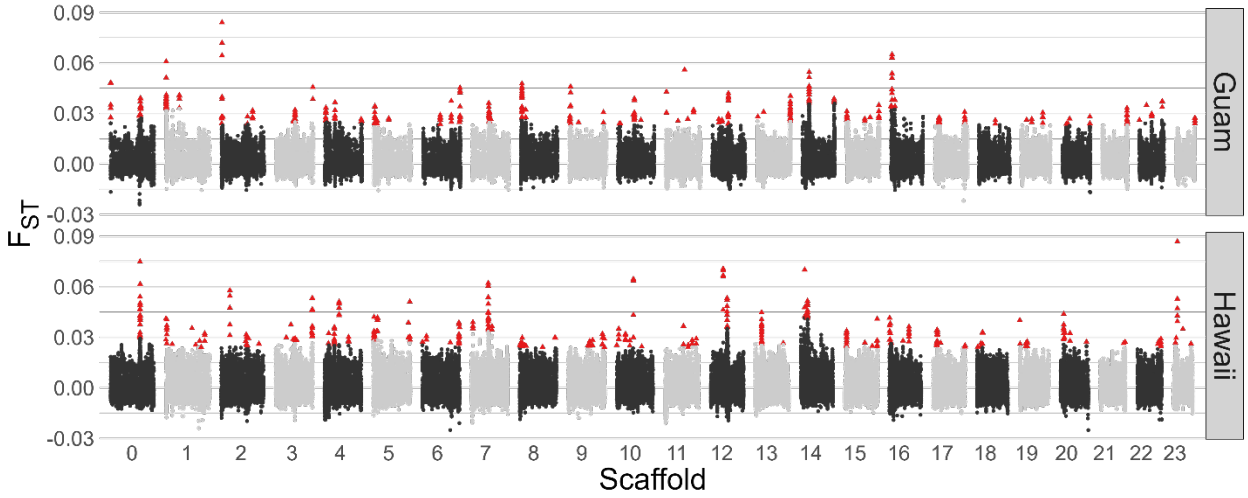


Figure 3-3 F_{ST} outliers. F_{ST} outliers for *Paracirrhites arcatus* are calculated using 5 kb sliding windows with 1 kb step size. Outliers are calculated per-scaffold using windows that exceeded the 99% bootstrapped confidence interval of the 99.9% quantile.

CHAPTER FOUR: COLOR MORPH DIVERGENCE IN *PARACIRRHITES HEMISTICTUS*

Introduction

The study of polymorphisms has long been essential to understanding evolutionary and ecological processes. Polymorphisms are defined as the presence of two or more stable variants within a single population at a frequency greater than expected due to genetic drift (Ford, 1945). Color polymorphisms, in which multiple color morphs co-exist, are not uncommon in animals and have been documented to persist across speciation events and to occur in closely related species. Examples can be found among cichlids and raptors where related species exhibit similar polymorphisms (Elmer et al., 2009; Galeotti & Rubolini, 2004). Damselflies of the genus *Ischnura* present an ancestral female-limited color polymorphism common to over 30 species (Sánchez-Guillén et al., 2020), but with new species continuing to evolve distinct phenotypes while retaining the ancestral polymorphism. Similarly, ancestral color polymorphisms are found in arctic bumblebees (*Alpinobombus*) (Williams et al., 2015) and Hawaiian spiders (*Tetragnatha*) (Oxford & Gillespie, 2001). In other groups, such as wall lizards, the ancestral polymorphism has been lost in new lineages (Uller et al., 2026). Ancestral color polymorphisms specifically have been linked with an increased rate of speciation in lizard, bird, and fish taxa, often accompanied by loss of a polymorphism (Corl et al., 2010b; Hugall & Stuart-Fox, 2012; Puebla et al., 2007). Studies on the genetic mechanisms responsible for trans-species polymorphisms often focus on pigmentation genes, such as *MC1R* (melanocortin receptor 1) or *ASIP* (agouti signaling protein) (Mundy & Kelly, 2006; Rosenblum et al., 2004; Switonski et al., 2013). These have shown that even in closely related taxa, similar morph phenotypes are coded by different variant positions within the same gene, indicating that even seemingly shared ancestral phenotypes may be the result of convergent evolution (Manceau et al., 2010). This has been observed at the intraspecific level with the cavefish *Astyanax mexicanus* (Gross et al., 2009), as well as across the poison dart frog genus *Ranitomeya* (Twomey et al.,

2023). Evidence of identical trans-species polymorphisms linked with color morph is rare, although an intronic transposon was found to be responsible for the gold morph within the Midas cichlid group (Kratochwil et al., 2022).

Among the hawkfishes of the genus *Paracirrhites* (family Cirrhitidae), multiple species exhibit intraspecific color polymorphisms (DeMartini & Donaldson, 1996; Gaither & Randall, 2012). *Paracirrhites forsteri* has a complex color polymorphism with at least four distinct morphs, (Gaither et al., 2020) that are distinguished by shifts in alleles frequency at a small number of genetic outliers as well as differences in growth rate and depth preference. On the other hand, *P. hemistictus* and *P. arcatus* share a similar light/melanistic dimorphism (Figure 1). The light morphs of *P. hemistictus* (YWS; Yellow white stripe) have yellow bodies with black spots on the dorsal surface and a white longitudinal stripe, while the melanistic morphs (MEL; melanistic, formerly recognized as *Cirrhites polystictus* in Günther (1874), but synonymized and both moved to the genus *Paracirrhites* by Randall (1963)) have dark bodies and a single white spot instead of the stripe. Light morphs of *P*. *arcatus* are red/pink with a lighter ventral surface and a white longitudinal stripe, while the melanistic morph is completely dark except for its fins and the iridescent markings behind its eyes and on its operculum. While the primary color of the light morphs differs, they share the white longitudinal stripe and ventral lightening, while the dark morphs include the loss of the stripe. These patterns are present across the species geographic ranges, with some minor variation. There have been three published studies on *P. arcatus*. DeMartini and Donaldson (1996) conducted a broad survey across Oceania and first noted that the light morph tended to be found deeper than the dark morph. This was followed by Whitney, Donahue, et al. (2018) and Whitney, Bowen, et al. (2018)who closely examined the Hawaiian population, and constructed an ecological niche model for the morphs. They broadly concluded that the morphs were undergoing the very beginning stages of speciation, potentially driven by improved crypsis in different portions of the reef. Finally, in Chapter 3, I replicated the Hawaiian study on Guam to compare, and found a weaker, but still present signal of niche

partitioning, and that the two morphs are distinguished by differentiation across outlier genes, particularly key genes associated with developmental processes and melanosomes.

While some attention has been given to the color polymorphism in *P. arcatus,* there have been no comparable ecological or genetic studies yet conducted on *P. hemistictus.* This species overlaps with *P. arcatus* across much its geographic range, but it is not present in Hawai'i (Randall, 2005). To compare it with *P. arcatus,* I conducted field transects and whole genome resequencing of *P. hemistictus* in Guam and then compare the results with data for *P. arcatus* from Chapter 3 to determine whether the ecological and genomic patterns that distinguish the color morphs are similar across these two species.

Materials and Methods

Ecological Surveys

I conducted 53 25 x 4 m belt transects along descending depth gradients at five sites along the west coast of Guam (See Chapter 3). The scale of this survey was optimized for *P. arcatus* and did not yield a sufficient sample size. So, I conducted a GPS-tow survey (Colin, 2012) along Orote Peninsula (Figure 2). Surveys were conducted simultaneously by two individuals: a snorkeler at the surface towing the GPS unit and locating shallow individuals, and a diver at approximately 3-4 meters. The diver recorded all data including the color phenotype, microhabitat, depth, estimated total length, and timestamp (to obtain location) of every *P. hemistictus* observed. On the outer portion of the peninsula (approximately 5 km), the substrate is very steep, enabling the observation of large areas of potential habitat in one pass, while inside the harbor, the divers employed a zigzag pattern to cover the available habitat (approximately 2.5 km).

Sample collection and sequencing

I collected 69 *P. hemistictus* from Guam using pole spears. Muscle and fin tissues were taken from each individual and preserved in 95% ethanol and stored at -20°C. A separate single individual of the YWS

morph was collected from Guam for genomic assembly and annotation following the methods outlined in Chapter 3.For this step sequences were generated using Hi-C libraries and PacBio HiFi CCS reads, and assembled using Phase Genomics' proprietary methods. The genome annotation used a MAKER-based pipeline informed by RNA-seq of muscle, skin, gills, eyes, brain, liver, and heart tissues from RNA libraries prepared from the reference individual. Synteny for the reference genome was compared with the *P. arcatus* genome previously assembled In Chapter 2. I used the program ntSynt v1.0.3 (Coombe et al., 2024) to calculate synteny using a divergence preset of 4, then generated a plot with ntSynt-viz v1.0.0 (Coombe et al., 2025).

Whole genome resequencing was conducted for 69 individuals of *P. hemistictus* following the methods of Chapter 3. Briefly, genomic DNA was extracted following the manufacturer's protocol with the E.Z.N.A. Tissue Kit (Omega Bio-tek, USA). PCR-free whole genome libraries were prepared with SparQ DNA Frag and Library Prep Kit (QuantaBio, USA) using 500 ng input DNA. Equimolar pools of the libraries were sequenced along with other projects across two lanes of an Illumina NovaSeq 6000 S4 flowcell on 2x150 mode at the University of Florida ICBR Sequencing Core. Raw sequence data was filtered using fastp v0.23.2 (Chen et al., 2018). Reads were processed for polyG tail and adapter sequence removal, paired-end overlap base correction, and trimmed from the 5' end using a 4 bp sliding window until the average window quality was Q25. Reads that were less than 25 bp long or had less than 75% of the bases below Q25 were removed. The remaining reads were aligned to the reference genome (Chapter 2) using BWA-MEM2 v 2.2.1 using default settings (Vasimuddin et al., 2019). Samtools v1.15 (Danecek et al., 2021; Li et al., 2009) was employed to sort reads, mark optical duplicates, and to output the mapped reads. Variants were called using the 'mpileup' and 'call' commands in bcftools v1.15 (Danecek et al., 2021). Finally, raw SNP calls were filtered using bcftools 'view' to remove loci with a sequencing depth of less than 10 per sample and alleles with a minor allele frequency of less than 0.02, then extracted biallelic SNPs with less than 20% missing data.

Bioinformatics

I calculated the F_{ST} estimates between the two color morphs over 5 kb sliding windows with a 1 kb step using the '--weir-fst-pop function' in vcftools v0.1.16 (Danecek et al., 2011). F_{ST} outlier windows between the light and melanistic color morphs were identified by examining windows that exceeded the 99% bootstrapped confidence interval of the 99.9% quantile of F_{ST} estimates (calculated per scaffold). Adjacent outlier windows were combined to form contiguous regions of high F_{ST}. Genes overlapping outlier windows were identified and tested for statistical overrepresentation of Gene Ontology Biological Process terms against a zebrafish reference using the PANTHER 19.0 Overrepresentation Test (Mi et al., 2021; Mi et al., 2019). I also tested for soft selective sweeps using the 'saltiLASSI' algorithm in lassip (DeGiorgio & Szpiech, 2022) and a window size of 50 SNPs.

Candidate Coloration Genes

Following the methods in Chapter 3, I isolated 128 candidate teleost coloration genes identified from (Braasch et al., 2009). To ensure that no genes were missed due to the quality of the annotation, the sequences for all genes were downloaded from ENSEMBL release 113 (Dyer et al., 2024) using their biomaRt R package (Durinck et al., 2005; Durinck et al., 2009). Exon sequences were aligned to the reference genome using minimap2 v2.28 with the 'splice' preset (Li, 2018), then the alignment was expanded to the ends of the gene or expressed sequence tag annotation they overlapped with. Out of the 128 genes in the initial list, 108 had likely matches in the *P. hemistictus* genome. Based on the position of these 108 genes, I exported all SNPs within the target sequences using bcftools with the filters Genotyping Quality >15, depth>10, and fraction missing <0.2. Next, I calculated F_{ST} estimates for each SNP using vcftools as above, but with the window size set to 1. Outliers were identified by scaling the F_{ST} within each gene using a Z-score to create ZF_{ST} estimates (Lai et al., 2016), adjusting them for multiple comparisons based on a false discovery rate, and selecting those past a threshold of 0.01.

Genes containing outliers were assessed for potential biological significance by identifying those with a maximum F_{ST} of 0.15, or an average of 0.1 within the gene. Coloration genes were also assessed for selection using BayeScan v2.1 (Foll & Gaggiotti, 2008) after first LD pruning the SNPS using 'indep-pairwise' with a window size of 1Mb and an R^2 cutoff of 0.1 using PLINK.

Results

Survey

In total, the belt transects originally conducted in Chapter 3 covered 1.3 km, while the diver-towed GPS distance transect included an additional 7.5 km. Along the 1.3 km of belt transects, I observed only 7 YWS and 14 MEL for a total of 21 fish, while along the longer diver-towed transect I counted 100 individuals of *P. hemistictus*. including 52 YWS, 47 MEL, and a single intermediate (YWS coloration with a dorsal dot instead of the white lateral stripe). *Paracirrhites hemistictus* was found at higher density within the outer harbor, with 46 observations on a 2.5 km long transect, or 18.9 observations per kilometer, compared with 10.4 observations per kilometer outside the harbor. Belt transect density was 16.2 observations per kilometer (similar to sheltered reef habitat to the harbor), indicating consistency between the survey methods. Individuals of the two morphs recorded along the diver-towed transect were unevenly distributed. Inside the harbor, 10 of the 46 (22%) individuals recorded there were the melanistic morph, while outside the harbor 37 of the 54 (68%) were melanistic (Figure 2). While the outer reef was much steeper and deeper (deepest observation was at 11.9 m), the median depths were almost identical (harbor 1.75 m, outer 1.8 m) and not significantly different (t-test $p = 0.15$). Furthermore, the substrate on which *P. hemistictus* were perched on differed between the habitats. Individuals recorded on the outer reef were found on fire coral (primarily *Millepora platyphylla*) the majority of the time (31 fire coral, 8 *Pocillopora spp.*, 15 other), while individuals within the harbor were almost exclusively found associated with the coral *Pocillopora* (0 fire coral, 45 *Pocillopora spp.*, 1 other). Fire coral and *Pocillopora* were present in both habitats, but I did not conduct an empirical survey of

substrate availability on this transect, therefore I cannot assess whether resource use was proportionally biased.

Resequencing Analysis

I conducted whole genome resequencing for all 69 individuals of *P. hemistictus*, including 32 YWS (yellow white-stripe) and 37 melanistic (MEL) morphs. No intermediate morphs were observed during collections. I obtained 4.6 billion reads from 69 samples, averaging 66.3 million reads (~14x coverage) per sample. Of these reads 98.5% passed quality filtering and aligned to the newly assembled reference genome (Supplementary Table S4). After variant calling, I resolved 22.9 million SNPs. Mean F_{ST} between the color morphs was 0.0750. Analysis of F_{ST} outlier windows revealed 204 outlier regions (Figure 3), of which 53 overlapped genes within the annotation (Table S2). The GO overrepresentation analysis did not identify any overrepresentation in the set of outlier genes. With only 26 of the 53 gene identifiers mapping to the PANTHER database, there was little statistical power. For the 128 candidate coloration genes (Braasch et al., 2009), neither F_{ST} (Figure 3) nor Bayescan analysis revealed any outliers. However, the saltiLASSI analysis identified 101 outlier genes that reached the top 0.5 % likelihood (lambda) cutoff of 790.53 (Figure 4, Supplementary Table S2), with the *Chd1* locus (Chromodomain-Helicase DNA-binding protein) identified as under selection in both *P. hemistictus* and *P. arcatus*.

Discussion

Studies of color variation have typically focused either on polymorphisms within a single species or on species flocks demonstrating a large variety of coloration, but do not represent a polymorphism at the species level. By comparing the sister taxa *Paracirrhites hemistictus* and *P. arcatus,* this study provides insight into the evolutionary history of a shared color polymorphism. Since the color morphs of these

species have similar, but not identical phenotypes, the polymorphisms may pre-date the speciation events within the genus, or there may be a predisposition to parallel or convergent evolution of coloration. If the color morphs are an ancestral relic that is no longer maintained by balancing selection, each color morph could represent a diverging lineage at different points of the process.

Color polymorphism provides a tool for a single species to occupy multiple niches (Bond & Kamil, 2006), which can be valuable when there is high competition, which often is associated with more fine niche partitioning (Finke & Snyder, 2008). The Mariana Islands are more speciose than Hawai'i, with over 1000 species of reef fishes and nine species of hawkfishes (Myers & Donaldson, 2003), compared with about 600 reef fishes and six hawkfishes in Hawai'i (Randall, 2005). The range of *P. hemistictus* and *P. arcatus* overlap in the center of their ranges around west Pacific, but *P. hemistictus* is not present in Hawai'i or most of the Indian Ocean range of *P. arcatus* (Donaldson, 1986; Randall, 1963). *P. hemistictus* is a large bodied hawkfish (up to 30 cm) that is largely restricted to shallow reef habitat (<20 m), while *P. arcatus* is smaller (up to 7 cm) and can be found at a depth of up to 90 m (Randall, 2005). Competition for coral perches, territory, and prey with congenerics and other reef predators could limit *P. arcatus* on Guam more strongly than on Hawai'i.

In Guam, I found *P. hemistictus* at density of approximately 15 individuals per linear kilometer which is significantly lower than the 350 *P. arcatus* per kilometer reported along the same belt transects Chapter 3. Additionally, the morphs of *P. hemistictus* were roughly even in abundance (47% MEL) but showed a greater (although still incomplete) physical separation (Figure 2) than *P. arcatus*, which was 7% MEL, and partitions habitat on a finer scale. Surveys revealed a distribution of 68% melanistic morphs outside the harbor compared to 22% within it – a large shift. Depths were comparable between locations, but reef preferences varied, with individuals on the outer reef predominantly found on fire coral, in contrast with those found within the harbor which were largely associated with *Pocillopora*. It is notable on Guam there is a much lower overall density of *P. arcatus* compared with Hawai'I, where *P.*

hemistictus is absent, and in Guam there were vastly fewer melanistic *P. arcatus* (Whitney, Donahue, et al., 2018). It has been shown that the absence of closely related competitors allows for niche expansion in fish (Robinson & Wilson, 1994). While the two species are unlikely directly competing on prey (Leray et al., 2012), hawkfish can be aggressively territorial and share a preference for perching on *Pocillopora spp.* coral heads. Possibly, the dark morphs of *P. arcatus* are being outcompeted by larger species of hawkfish, where they are present. My surveys show that *P. hemistictus* occupies the same shallow, high energy niche of the reef that Whitney, Donahue, et al. (2018) showed was preferred by the melanistic morph of *P. arcatus* in Hawai'i. Release from competition as *P. arcatus* colonized Hawai'i may have allowed the melanistic morph to expand into this niche.

The morphs of *P. hemistictus* demonstrate a moderate level of genome-wide divergence, with a mean F_{ST} = 0.0750 indicating that there is incomplete reproductive isolation and a high amount of neutral variation. This is much higher than the differentiation between the color morphs of *P. arcatus* in both Guam (F_{ST}=0.0015) and Hawai'i (F_{ST}=0.001) (Chapter 3). While I recorded some degree of habitat sorting as evidenced by the morph proportions changing between inside the harbor (22% MEL) and outside (68% MEL), the surveys show that the morphs still overlap, indicating that genetic differentiation is likely the result of mate choice. Outlier analyses indicated 204 F_{ST} outlier regions, but no coloration gene was more differentiated than the background variation. In contrast with *P. arcatus*, there were much fewer outlier loci, including no outlier SNPs within the candidate coloration genes. The standing variation between the morphs is of sufficient magnitude that the whole genome is differentiated, instead of the islands of differentiation identified in *P. arcatus.*

I found a single gene that was under selection in both species according to the saltiLASSI analysis, as well as five shared F_{ST} outliers (Table S3). The *Chd1* locus identified by saltiLASSI as under selection is responsible for a highly-conserved chromatin-remodeling protein, Chromodomain-Helicase DNA-binding 1, which while not directly responsible for coloration, may be involved in epigenetic control

of some aspect of the morphs. Additionally, there were five F_{ST} outlier genes shared between the species: gnb1 (g-protein subunit), NAA25 (acetyltransferase subunit), and Pi4ka (PIP_2 synthesis) are involved in signaling; tubb4bl (Tubulin) and COL6A3 (collagen subunit) are structural. These genes have no known link to coloration and may be related to some other trait of the morphs. However, color expression is a complex trait, and these signaling genes might contribute to the phenotype by affecting the regulation of melanin expression or the migration of melanophores during development. Within *P. hemistictus,* any gene directly involved in coloration is likely captured within the background differentiation between the morphs, without any signals of direct selection detected here.

Color polymorphism as a vehicle for divergence

I detected a moderate level of genetic differentiation between color morphs in *P. hemistictus*, with F_{ST}=0.075. This estimate is higher than the differentiation between the majority of the described species of Caribbean hamlets (Hench et al., 2022). My surveys showed that they are present within close proximity of each other, with the morphs exhibiting different frequencies in different habitat rather than fully separating. *Paracirrhites hemistictus* is much more differentiated than the morphs of *P. arcatus* (Guam F_{ST}=0.0015, Hawai'i F_{ST}=0.0010; Chapter 3), and the spatial sorting, while still incomplete, is much more severe. While I did not observe a significant depth difference, Donaldson (1990) did find that the melanistic morph of *P. hemistictus* prefers shallower habitat. However, if the preference is instead for a higher energy habitat, the melanistic preference for the outer reef is similar to and often correlated with depth and matches the findings for *P. arcatus* in Hawai'i by Whitney, Donahue, et al. (2018). Overall, these results suggest that the morphs of *P. hemistictus* are acting as separate populations, with some gene flow. To establish whether they should be regarded again as separate species or subspecies, there should be additional observation of mating behavior to determine whether there is mating segregation, as well as examination of morphology for differences other than color.

The relationship between color phenotype and population divergence in fish has been studied in a number of systems. The Midas cichlids (*Amphilophus*) of Nicaraguan crater lakes and the African rift lake cichlids were among the first to be studied (Barlow, 1973; McKaye & Barlow, 1976; Seehausen et al., 1999). These groups underwent rapid radiation within confined lake systems, with coloration in some species evolving in new photic environments while assisting in mate recognition (Allender et al., 2003; Torres-Dowdall et al., 2017; Wright et al., 2020). Sexual selection on ornamental coloration leading to divergence has been demonstrated in stickleback (*Gasterosteus aculeatus*) (Bakker, 1993; Wright et al., 2015), and recently in the radiation of Caribbean hamlets (*Hypoplectrus*) (Helmkampf et al., 2025; Puebla et al., 2007). These systems have provided significant support for hypotheses of ecological or sympatric speciation, often with selection on color driving or reinforcing the differentiation (Matsubayashi & Yamaguchi, 2022). However, these examples represent rapid radiations, where the majority of the species emerged within 10,000 generations (Fang et al., 2018; Hench et al., 2022; Kautt et al., 2020; Sturmbauer et al., 2001), although some African cichlid radiations are potentially up to 100,000 generations (Sturmbauer et al., 2001). Because divergence was rapid, overall genetic differentiation between nascent species can be low, ranging from F_{ST}=0.02-0.46 in Nicaraguan cichlids (Recknagel et al., 2013), African cichlids F_{ST}=0.01-0.25 (Bezault et al., 2011), stickleback morphs F_{ST}≈0.05 (Jenck et al., 2022), stickleback populations F_{ST}=0.031-0.383 (Jones et al., 2012), hamlets F_{ST}=0.003-0.1 (Hench et al., 2022), with most specific species pairs towards the lower ends of these ranges. The color morphs within *Paracirrhites* provide a system to study polymorphism that is more typical of an average reef fish. This genus has not undergone an explosive radiation, while having a much larger range than the cichlids and more differentiation between species than the stickleback and hamlets.

Conclusion

The results of this study show that there is substantial divergence between the color morphs of *P. hemistictus*, despite sometimes perching on corals within meters of each other. Genetic data suggests

that reproductive isolation is incomplete, but contact is rare. When compared with a similar analysis of *P. arcatus,* there is little shared signal of differentiation between the species which could indicate the polymorphism is not ancestral. In fact, there is little sign of a direct genetic mechanism at all, and instead both within-species and shared outliers point towards a regulatory background for color morphs, with the outlier genes and . Finally, the "Flickers of Speciation" initially identified in Hawaiian *P. arcatus* by Whitney, Bowen, et al. (2018) may have advanced to effectively separate populations in *P. hemistictus.* Further study of both species in different archipelagos can help us learn more about whether populations are determined more by morph or by geography as I look across their ranges.

Figures

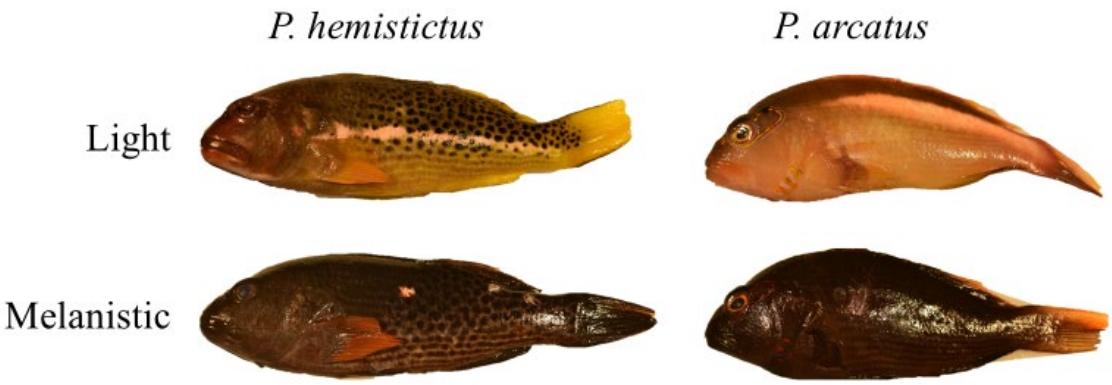


Figure 4-1 Color morphs of *Paracirrhites hemistictus* and *P. arcatus.*

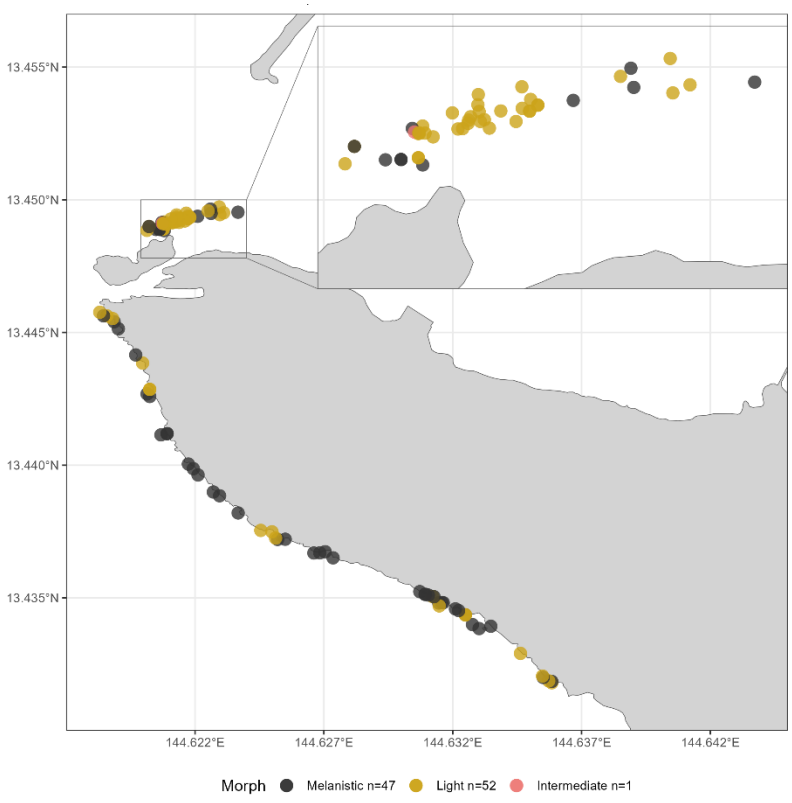


Figure 4-2 Distribution of *Paracirrhites hemistictus* at Orote Point, Guam

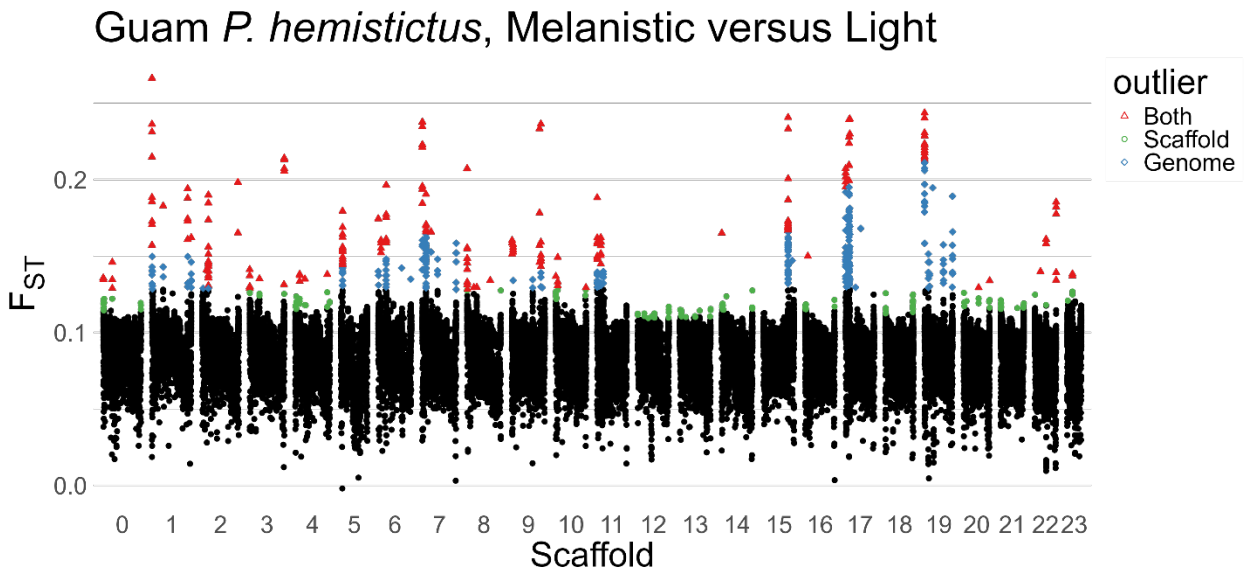


Figure 4-3 F_{ST} outliers for *Paracirrhites hemistictus* are calculated using 5 kb sliding windows with 1 kb step size. Outliers are calculated per-scaffold using windows that exceeded the 99% bootstrapped confidence interval of the 99.9% quantile.

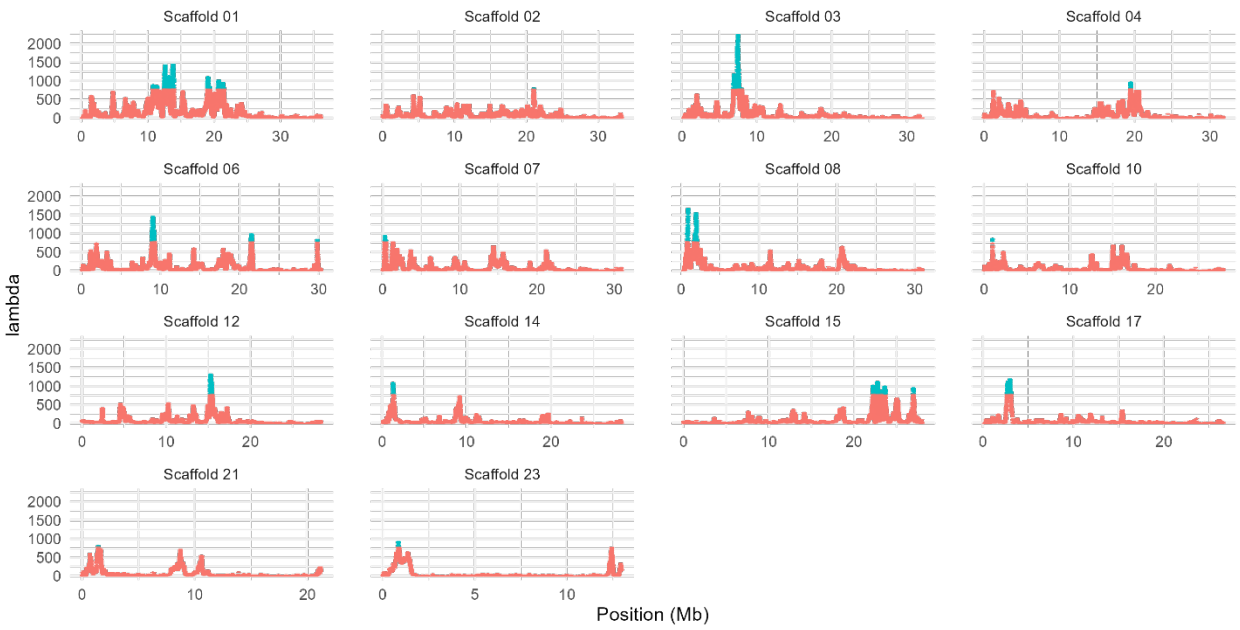


Figure 4-4 Loci under selection calculated with LASSI (**L**ikelihood-based **A**pproach for **S**elective **S**weep **I**nference) (DeGiorgio & Szpiech, 2022). Only scaffolds containing windows with the top 0.1% of lambda values are shown. The shared locus under selection, *Chd1*, is present on Scaffold 17.

CHAPTER FIVE: CONCLUSION

My dissertation set out to understand how a shared sympatric color dimorphism in *Paracirrhites arcatus* and *P. hemistictus* can be maintained or diverge. After establishing chromosome-level genomic resources for these two species (Chapter 2), I replicated ecological and population-genomic analyses across geography and taxa (Chapters 3-4) and show that similar phenotypes can reflect different evolutionary states. The Hawaiian and Guam populations of *P. arcatus,* and *P. hemistictus* follow distinct trajectories depending on the balance among selection, ecology, and gene flow.

I produced high-quality, chromosome-level reference genomes for *P. arcatus* and *P. hemistictus*. These assemblies fill a major resource gap for Cirrhitidae and enabled the genome-wide divergence scans and candidate-gene mapping in the later chapters. In addition to supporting this dissertation, these resources provide a foundation for future work on reef fish adaptation, the evolution of coloration, and the genomic basis of divergence in widely dispersing marine taxa.

For *P. arcatus*, I replicated ecological surveys done by Whitney, Donahue, et al. (2018) in Hawai'i at my study site in Guam, and by comparing the two revealed geographic differences in both the demographics of the color morphs and the opportunity for niche segregation. Guam populations occurred at substantially lower density than previously reported for Hawai‘i, and Guam reefs showed a pronounced skew away from the melanistic morph. While multivariate niche analyses detected statistically significant differentiation among morphs, the ecological separation in Guam was weak relative to Hawai‘i and did not reproduce the same clear habitat associations as Whitney, Donahue, et al. (2018).

These ecological patterns were mirrored by population genomics. Genome-wide divergence between morphs was low in both archipelagos, and no diagnostic loci emerged that could cleanly explain the phenotype across the range. However, *P. arcatus* exhibited evidence of genomic islands of

differentiation, particularly in Hawai'i. Differentiation was detectable in developmental, gene regulatory, and melanosome-associated genes. This pattern is consistent with divergence occurring under gene flow, where selection can generate narrow genomic regions of differentiation while gene flow homogenizes the remainder of the genome. Guam and Hawai‘i share the same phenotypes for their color morphs, but the underlying evolutionary dynamics appear different. The Guam morphs may receive sufficient gene flow, potentially from western Pacific populations, to prevent them from diverging more. In contrast, the Hawaiian *P. arcatus* showed fewer but stronger peaks of differentiation indicating the initial stages of the formation of barriers to gene flow around select genomic regions.

In comparison to *P. arcatus, P. hemistictus* has a much larger degree of genome-wide divergence between its morphs, even within Guam. This is especially striking because the morphs can still occur within meters of each other, underscoring that geographic distance is not required for strong genetic structure when reproductive isolation mechanisms are emerging. Unlike *P. arcatus*, where there were peaks within a background of little differentiation, *P. hemistictus* exhibited broad genomic separation without any diagnostic loci. A buildup of neutral differences indicates that the *P. hemistictus* color morphs are behaving as partially isolated populations.

Together, these findings support the idea that the previously hypothesized “flickers of speciation” emerge where isolation and local ecological opportunity allow differentiation to persist long enough to accumulate (Whitney, Bowen, et al., 2018). While the average genetic differentiation is higher in *P. arcatus* from Guam than Hawai'i, this can also be expected due to founder effects. Hawaiian genomic islands of differentiation are more substantial, along with increased ecological differentiation. At the very edge of the range of *P. arcatus,* Hawai‘i may provide that opportunity for *P. arcatus* in a way that Guam does not. Gene flow to Guam may repeatedly reset divergence, maintaining the trait as a polymorphism, while increased competition keeps the morphs from specializing into different niches. Meanwhile, *P. hemistictus* appears further along the continuum of divergence, although further work is

needed to establish how different the morphs are on a morphological and ecological basis. Establishing the phylogenetic relationships of the species within *Paracirrhites*, ecological surveys focused on *P. hemistictus*, and an analysis of expression and epigenetic data between the morphs would greatly improve our understanding of the system.

This dissertation shows that color polymorphism can be both a stable evolutionary outcome and a potential vehicle for divergence, but only when ecological conditions align. By combining ecological and genomic approaches across space and taxa, I show that sympatric divergence is occurring within this system, but the magnitude of gene flow and the opportunity to occupy different niches are important factors. Furthermore, similar to the Caribbean hamlets (Puebla et al., 2014), this differentiation likely starts out as small genomic islands while gene flow is strong, before smoothing out as divergence progresses. Future work that integrates behavioral observations of mate choice, broader sampling across these species' ranges, and a similar ecological assessment of *P. hemistictus* as *P. arcatus* will be key to determining how pervasive these patterns are across more populations and better establish the link between genetic variants and phenotype. Potentially, those "flickers of speciation" previously hypothesized for *P. arcatus* in Hawai'i is just one outcome of the genetic architecture and population dynamics of these hawkfish.

APPENDIX A: SUPPLEMENTAL FIGURES AND TABLES FOR CHAPTER 2

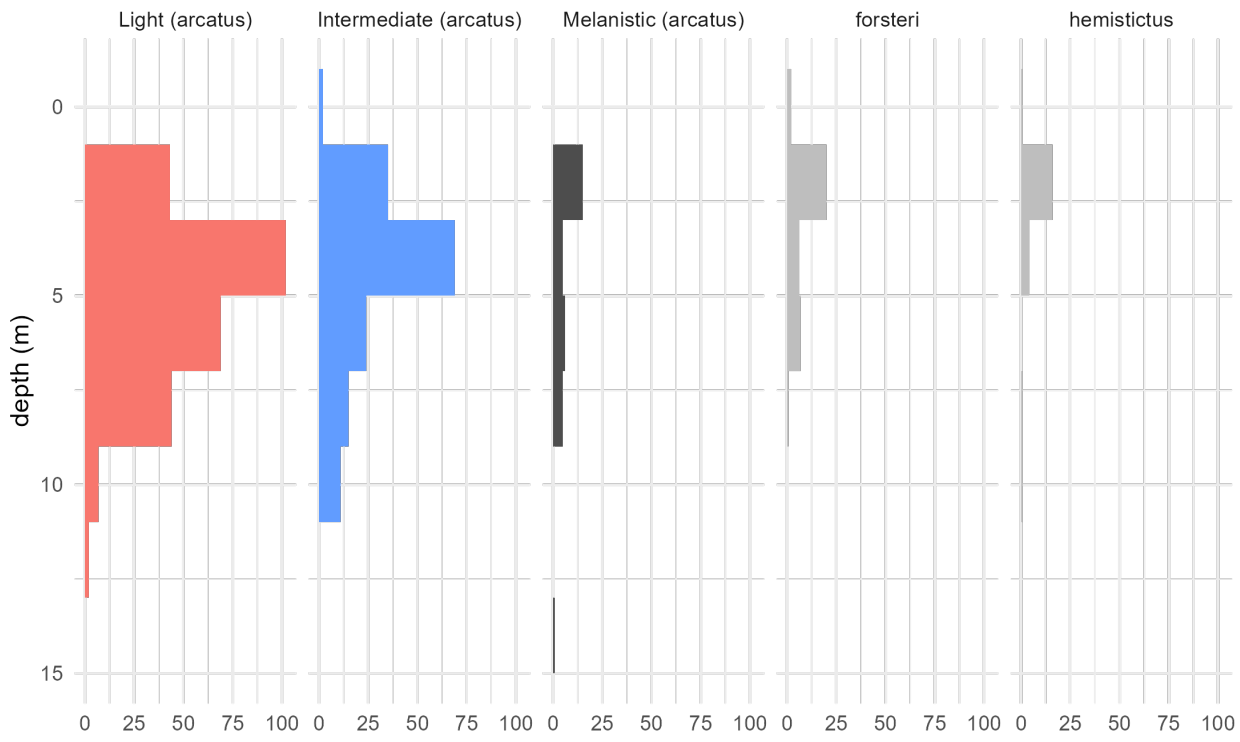
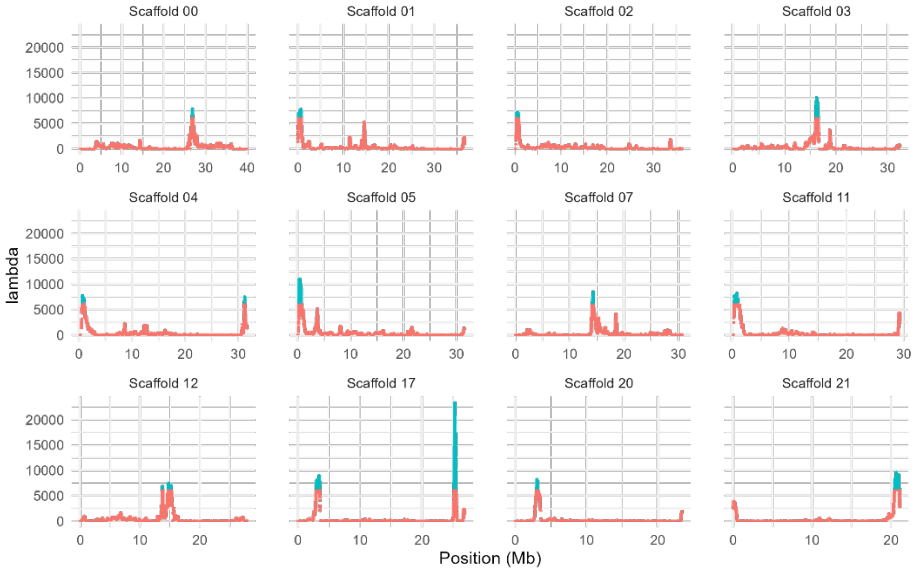


Figure S1 Depth distribution of *P. arcatus, P. hemistictus,* and *P. forsteri* recorded during field surveys in Guam.

S2A



S2B

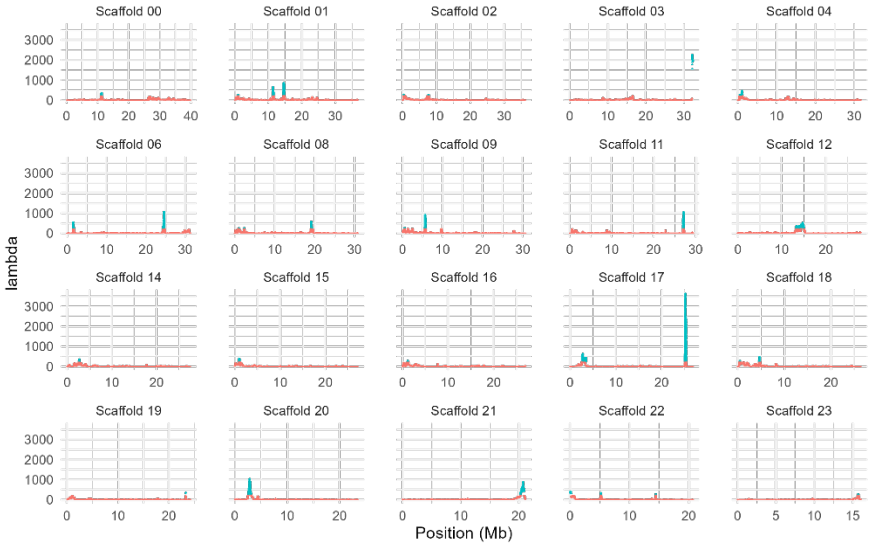


Figure S2 LASSI Manhattan plot, showing the composite likelihood statistic lambda plotted against position within a scaffold (DeGiorgio et al., 2014). Analysis was done using a window size of 50 haplotypes. Outliers (blue) are the top 0.1% of the composite likelihood statistic lambda. A) LASSI outliers from the Guam population B) LASSI outliers from the Hawai'i population.

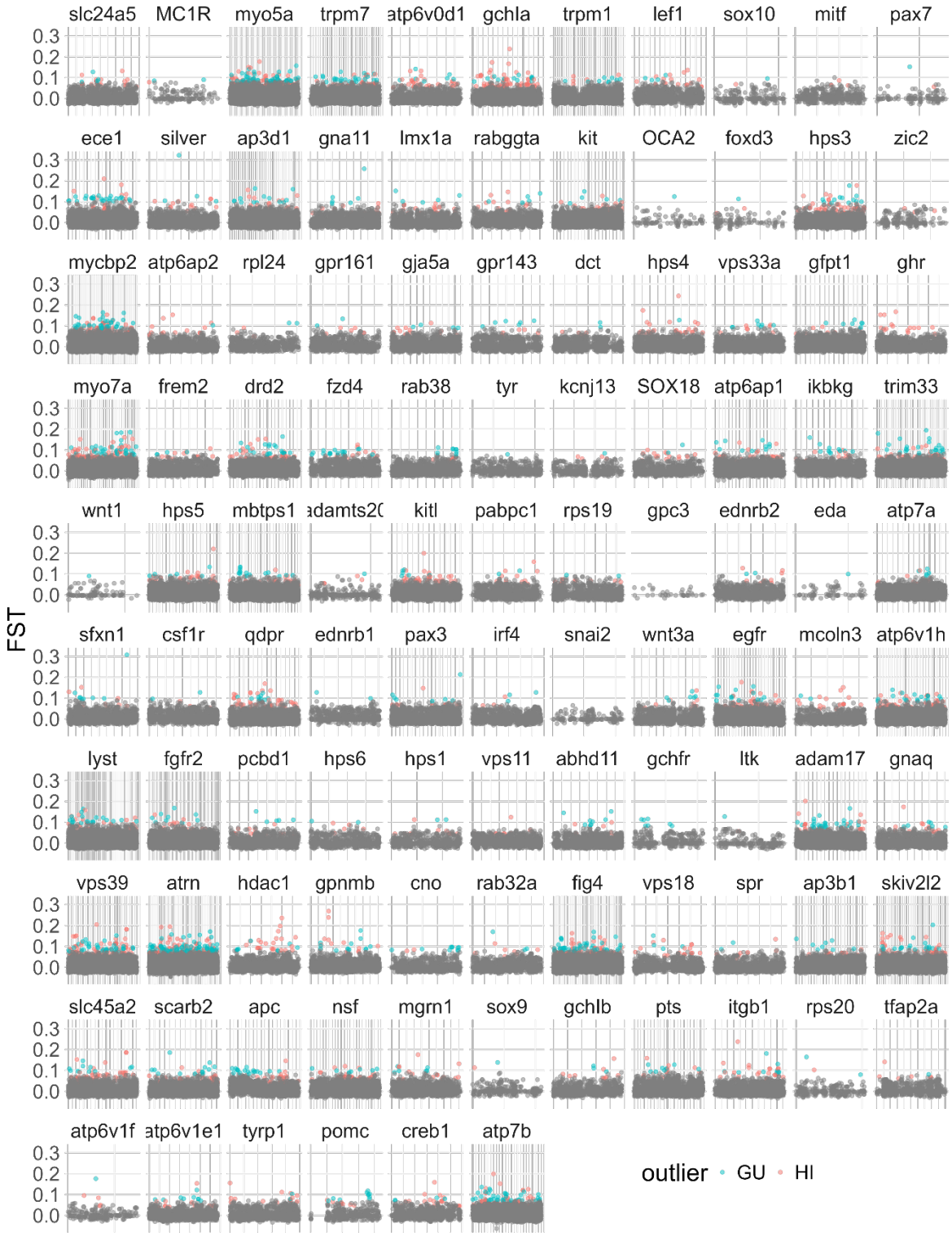


Figure S3 F_{ST} estimates calculated per-SNP for coloration genes using vcftools (Danecek et al., 2011). Outliers are identified using an FDR < 0.000001 on Z-adjusted F_{ST} estimates.

Table S1 Sample information and sequencing statistics for *P. arcatus*.

Sample ID	Island	Morph	Year	adapter	Mapped reads	Mean quality	Mean depth	coverage	notes
PARAR_0056	Guam	MEL	2019	QDIY-028	32162508	35.3	6.96063	0.97586	
PARAR_0057	Guam	MEL	2019	QDIY-051	98975262	35.2	21.3361	0.98676	
PARAR_0058	Guam	MEL	2019	QDIY-023	23521602	35.3	5.18382	0.96115	
PARAR_0059	Guam	IFS	2019	QDIY-001	665	35.2	1.23493	0.00011	QC failed
PARAR_0060	Guam	PWS	2019	QDIY-055	81932247	35.3	17.7455	0.98533	
PARAR_0061	Guam	IWS	2019	QDIY-005					library QC failed
PARAR_0062	Guam	PWS	2019	QDIY-053	93542324	35.3	20.1944	0.98639	
PARAR_0063	Guam	MEL	2019	QDIY-021	36486111	35.3	7.94795	0.97812	
PARAR_0067	Guam	PWS	2014	QDIY-054	110292859	35.2	23.8028	0.98741	
PARAR_0068	Guam	IWS	2014	QDIY-004	229698428	35.2	49.0637	0.99028	
PARAR_0069	Guam	IWS	2018	QDIY-006	23155765	35.3	5.07452	0.96174	
PARAR_0070	Guam	PWS	2018	QDIY-052	90901635	35.4	19.0835	0.98664	
PARAR_0073	Guam	PWS	2017	QDIY-056	123124771	35.1	26.5399	0.98757	
PARAR_0076	Guam	PWS	2017	QDIY-057	133519829	35.3	28.5762	0.98815	
PARAR_0077	Guam	IWS	2017	QDIY-017	35000518	35.4	7.46093	0.97634	
PARAR_0078	Guam	PWS	2017	QDIY-067	75651068	35.4	15.6105	0.98587	
PARAR_0079	Guam	PWS	2017	QDIY-062	125731702	35.3	26.8834	0.98796	
PARAR_0080	Guam	PWS	2017	QDIY-079	171679961	35.3	36.7511	0.98881	
PARAR_0081	Guam	IWS	2017	QDIY-018	33489863	35.3	7.2265	0.97694	
PARAR_0082	Guam	IWS	2017	QDIY-007	27373247	35.3	5.95893	0.97082	
PARAR_0083	Guam	PWS	2017	QDIY-072	138701387	35.3	29.3221	0.98791	
PARAR_0084	Guam	PWS	2019	QDIY-058	194705311	35.3	41.4898	0.98892	
PARAR_0085	Guam	PWS	2019	QDIY-063	100951360	35.3	21.7283	0.98705	
PARAR_0086	Guam	PWS	2020	Ref					Reference
PARAR_0088	Hawaii	MEL	2010	QDIY-061	40739688	34.8	8.74651	0.97985	
PARAR_0089	Hawaii	MEL	2010	QDIY-001	38886654	34.9	8.35606	0.97973	
PARAR_0090	Hawaii	MEL	2010	QDIY-037	40565539	35	8.7002	0.97910	
PARAR_0091	Hawaii	MEL	2010	QDIY-002	31988646	34.9	6.88677	0.97543	
PARAR_0092	Hawaii	MEL	2010	QDIY-038	36441391	34.9	7.88345	0.97814	
PARAR_0093	Hawaii	MEL	2010	QDIY-019	26044177	34.9	5.64558	0.96811	
PARAR_0094	Hawaii	MEL	2010	QDIY-039	37932816	34.8	7.97717	0.97718	
PARAR_0095	Hawaii	MEL	2010	QDIY-040	45136109	35	9.54111	0.98054	
PARAR_0096	Hawaii	MEL	2010	QDIY-020	33676652	34.7	7.27237	0.97714	
PARAR_0097	Hawaii	MEL	2010	QDIY-004	22837740	34.7	5.03421	0.96306	
PARAR_0098	Hawaii	MEL	2010	QDIY-041	40509052	35.1	8.34892	0.97690	
PARAR_0099	Hawaii	MEL	2010	QDIY-005	40675800	34.9	8.55184	0.97874	
PARAR_0100	Hawaii	MEL	2010	QDIY-042	23854943	34.9	5.2289	0.96427	
PARAR_0101	Hawaii	MEL	2010	QDIY-006	28710316	34.9	6.09293	0.96993	
PARAR_0102	Hawaii	MEL	2010	QDIY-007	36662063	34.9	7.91349	0.97854	

Sample ID	Island	Morph	Year	adapter	Mapped reads	Mean quality	Mean depth	coverage	notes
PARAR_0103	Hawaii	MEL	2010	QDIY-008	38053969	34.9	8.05152	0.97745	
PARAR_0104	Hawaii	MEL	2010	QDIY-062	40626845	34.9	8.68429	0.97981	
PARAR_0105	Hawaii	MEL	2010	QDIY-009	39814049	34.9	8.5598	0.97987	
PARAR_0106	Hawaii	MEL	2010	QDIY-010	33033656	34.9	6.99998	0.97505	
PARAR_0107	Hawaii	MEL	2010	QDIY-011	32511245	34.8	7.03179	0.97695	
PARAR_0108	Hawaii	MEL	2010	QDIY-021	35938446	34.8	7.77672	0.97895	
PARAR_0109	Hawaii	MEL	2010	QDIY-043	31587389	34.8	6.79326	0.97349	
PARAR_0110	Hawaii	MEL	2010	QDIY-044	38786635	34.9	8.33406	0.97909	
PARAR_0111	Hawaii	PWS	2010	QDIY-023	38883138	34.9	8.30578	0.97886	
PARAR_0112	Hawaii	PWS	2010	QDIY-045	37226884	35	8.03799	0.97592	
PARAR_0113	Hawaii	PWS	2010	QDIY-046	42477384	34.9	9.04348	0.97988	
PARAR_0114	Hawaii	PWS	2010	QDIY-013	27039086	34.8	5.87419	0.96998	
PARAR_0115	Hawaii	PWS	2010	QDIY-047	49188491	34.9	10.3012	0.98202	
PARAR_0116	Hawaii	PWS	2010	QDIY-063					QC failed
PARAR_0117	Hawaii	PWS	2010	QDIY-014	21617762	34.8	4.74822	0.95701	
PARAR_0118	Hawaii	PWS	2010	QDIY-048	35328378	35	7.20923	0.97302	
PARAR_0119	Hawaii	PWS	2010	QDIY-049	27360935	34.8	5.95989	0.97108	
PARAR_0120	Hawaii	PWS	2010	QDIY-064	39431159	35	8.37222	0.97959	
PARAR_0121	Hawaii	PWS	2010	QDIY-065	63192193	35	13.1549	0.98347	
PARAR_0122	Hawaii	PWS	2010	QDIY-024	41556296	34.8	8.6436	0.97948	
PARAR_0123	Hawaii	PWS	2010	QDIY-025	39750917	34.8	8.54182	0.97967	
PARAR_0124	Hawaii	PWS	2010	QDIY-050	37319559	35	7.86472	0.97658	
PARAR_0125	Hawaii	PWS	2010	QDIY-051	45382734	35	9.69669	0.98018	
PARAR_0126	Hawaii	PWS	2010	QDIY-066	27202759	34.9	5.8695	0.96866	
PARAR_0127	Hawaii	PWS	2010	QDIY-052	33112710	35	7.10175	0.97484	
PARAR_0128	Hawaii	PWS	2010	QDIY-026	31340461	34.7	6.75448	0.97511	
PARAR_0129	Hawaii	PWS	2010	QDIY-015	34570709	34.8	7.43029	0.97560	
PARAR_0130	Hawaii	PWS	2010	QDIY-016	27105486	34.8	5.904	0.97117	
PARAR_0131	Hawaii	PWS	2010	QDIY-067	26912120	34.9	5.86168	0.96941	
PARAR_0132	Hawaii	PWS	2010	QDIY-068	40227658	34.9	8.44336	0.97786	
PARAR_0133	Hawaii	PWS	2010	QDIY-027	33984090	34.8	7.32917	0.97717	
PARAR_0134	Hawaii	PWS	2010	QDIY-028	38430426	34.8	8.27434	0.97860	
PARAR_0135	Hawaii	IWS	2010	QDIY-070	53518253	34.9	11.4061	0.98299	
PARAR_0136	Hawaii	IWS	2010	QDIY-029	16992530	34.8	3.82383	0.93293	
PARAR_0137	Hawaii	IWS	2010	QDIY-031	44794387	35	9.48331	0.97995	
PARAR_0138	Hawaii	INS	2010	QDIY-053	40460939	35	8.64711	0.98000	
PARAR_0139	Hawaii	IWS	2010	QDIY-017	62946867	35	13.099	0.98369	
PARAR_0140	Hawaii	IWS	2010	QDIY-072	43147787	34.8	9.11453	0.98110	
PARAR_0141	Hawaii	IWS	2010	QDIY-018	47594080	35	10.0049	0.98146	
PARAR_0142	Hawaii	IWS	2010	QDIY-054	36161635	34.9	7.704	0.97781	
PARAR_0143	Hawaii	INS	2010	QDIY-073	39450152	34.8	8.48789	0.97900	

Sample ID	Island	Morph	Year	adapter	Mapped reads	Mean quality	Mean depth	coverage	notes
PARAR_0144	Hawaii	INS	2010	QDIY-055	36693935	35	7.66227	0.97601	
PARAR_0145	Hawaii	IWS	2010	QDIY-032	32149976	34.8	6.92932	0.97554	
PARAR_0146	Hawaii	IWS	2010	QDIY-033	40601005	34.9	8.77172	0.98004	
PARAR_0147	Hawaii	IFS	2010	QDIY-056	37048745	34.7	7.98898	0.97834	
PARAR_0148	Hawaii	IFS	2012	QDIY-074	18458475	34.6	4.1294	0.94656	
PARAR_0149	Hawaii	IFS	2012	QDIY-057	97066874	34.9	20.6159	0.98621	
PARAR_0150	Hawaii	IFS	2012	QDIY-075	24754920	34.8	5.4256	0.96805	
PARAR_0151	Hawaii	IFS	2012	QDIY-076	33790448	34.9	7.26211	0.97597	
PARAR_0152	Hawaii	IFS	2012	QDIY-058	68474628	35	14.2271	0.98394	
PARAR_0153	Hawaii	IFS	2012	QDIY-034	32018373	34.7	6.90196	0.97459	
PARAR_0154	Hawaii	IFS	2012	QDIY-035	34051810	34.8	7.35478	0.97703	
PARAR_0155	Hawaii	IFS	2012	QDIY-078	48327962	34.8	10.2433	0.98146	
PARAR_0156	Hawaii	IFS	2012	QDIY-079	57173003	34.9	12.2234	0.98346	
PARAR_0157	Hawaii	IFS	2012	QDIY-059	37891790	34.8	8.21596	0.97825	
PARAR_0158	Hawaii	IFS	2012	QDIY-060	43983712	34.9	9.53126	0.98111	
PARAR_0159	Hawaii	IFS	2012	QDIY-036	47539757	34.9	10.2588	0.98196	
PARAR_0160	Guam	MEL	2021	QDIY-029	24373667	35.3	5.37691	0.96378	
PARAR_0161	Guam	MEL	2021	QDIY-047	112706804	35.2	23.9926	0.98737	
PARAR_0162	Guam	MEL	2021	QDIY-031	30329408	35.4	6.55116	0.97054	
PARAR_0163	Guam	MEL	2021	QDIY-048	71340079	35.3	15.2902	0.98528	
PARAR_0164	Guam	PWS	2021	QDIY-075	101742603	35.3	21.9572	0.98724	
PARAR_0165	Guam	MEL	2021	QDIY-039	35914346	35.3	7.57664	0.97638	
PARAR_0166	Guam	PWS	2021	QDIY-076	102955029	35.3	22.1752	0.98707	
PARAR_0167	Guam	PWS	2021	QDIY-080	94710310	35.2	20.1414	0.98713	
PARAR_0168	Guam	IWS	2021	QDIY-013	805	35.4	1.45952	0.00011	QC failed
PARAR_0169	Guam	PWS	2021	QDIY-073	50952867	35.2	10.9944	0.98344	
PARAR_0170	Guam	IWS	2021	QDIY-014	25243109	35.3	5.47735	0.96492	
PARAR_0171	Guam	PWS	2021	QDIY-078	153461134	35.3	32.8215	0.98864	
PARAR_0172	Guam	IWS	2021	QDIY-009					Sequencing failed
PARAR_0173	Guam	IWS	2021	QDIY-015	24648820	35.2	5.33795	0.95847	
PARAR_0174	Guam	IWS	2021	QDIY-016	18399846	35.3	4.11403	0.93811	
PARAR_0175	Guam	PWS	2021	QDIY-060	172576056	35.3	37.065	0.98885	
PARAR_0176	Guam	IWS	2021	QDIY-019	21992923	35.4	4.84223	0.95581	
PARAR_0177	Guam	PWS	2021	QDIY-068	101236913	35.3	21.4709	0.98714	
PARAR_0178	Guam	IWS	2021	QDIY-008	16	35.2	2.00085	1.70E-06	QC failed
PARAR_0179	Guam	MEL	2021	QDIY-037	31500049	35.4	6.70802	0.97378	
PARAR_0180	Guam	PWS	2021	QDIY-074	88767908	35.2	19.1242	0.98653	
PARAR_0181	Guam	MEL	2021	QDIY-025	49274553	35.3	10.3802	0.99284	
PARAR_0182	Guam	MEL	2021	QDIY-026	34256184	35.3	7.43706	0.97735	
PARAR_0183	Guam	MEL	2021	QDIY-046	139677250	35.3	29.6906	0.98812	

Sample ID	Island	Morph	Year	adapter	Mapped reads	Mean quality	Mean depth	coverage	notes
PARAR_0184	Guam	MEL	2021	QDIY-045	93236119	35.3	20.1049	0.98639	
PARAR_0185	Guam	MEL	2021	QDIY-038	17725185	35.2	3.999	0.94302	
PARAR_0186	Guam	MEL	2021	QDIY-024	67037349	35.2	13.8165	0.99406	
PARAR_0187	Guam	MEL	2021	QDIY-033	37388990	35.3	8.0917	0.97956	
PARAR_0188	Guam	MEL	2021	QDIY-049	97659074	35.2	21.0896	0.98694	
PARAR_0189	Guam	MEL	2021	QDIY-050	96211083	35.3	20.5714	0.98676	
PARAR_0190	Guam	MEL	2021	QDIY-043	73417771	35.2	15.6559	0.98564	
PARAR_0191	Guam	PWS	2021	QDIY-059	108067423	35.4	22.9949	0.98702	
PARAR_0192	Guam	MEL	2021	QDIY-044	117211733	35.3	25.0153	0.9879	
PARAR_0193	Guam	MEL	2021	QDIY-027	20566848	35.3	4.57233	0.94912	
PARAR_0194	Guam	MEL	2021	QDIY-034	39257120	35.3	8.34146	0.97671	
PARAR_0195	Guam	MEL	2021	QDIY-032	24618704	35.3	5.39412	0.96358	
PARAR_0196	Guam	MEL	2021	QDIY-041	69353005	35.4	14.309	0.98494	
PARAR_0197	Guam	PWS	2021	QDIY-061	130772984	35.3	28.0375	0.98762	
PARAR_0198	Guam	PWS	2021	QDIY-064	124398963	35.3	26.6834	0.98814	
PARAR_0199	Guam	MEL	2021	QDIY-035	35070805	35.3	7.5415	0.97731	
PARAR_0200	Guam	MEL	2021	QDIY-036	26544426	35.3	5.80134	0.97010	
PARAR_0201	Guam	MEL	2021	QDIY-042	44560831	35.3	9.57504	0.98217	
PARAR_0202	Guam	MEL	2021	QDIY-040	33120746	35.3	7.15292	0.97702	
PARAR_0203	Guam	IWS	2021	QDIY-020	28887311	35.2	6.28693	0.97210	
PARAR_0204	Guam	PWS	2021	QDIY-070	202691191	35.3	43.611	0.98885	
PARAR_0205	Guam	PWS	2021	QDIY-065	123849771	35.2	26.6501	0.98809	
PARAR_0206	Guam	IWS	2021	QDIY-010	29355277	35.3	6.34187	0.97330	
PARAR_0207	Guam	PWS	2021	QDIY-066	130633408	35.3	28.1278	0.98784	
PARAR_0208	Guam	IWS	2021	QDIY-011					Sequencing failed
PARAR_0209	Guam	IFS	2021	QDIY-002	18614782	35.3	4.17996	0.94342	

Table S2 Gene Lists. A) All F_{ST} outliers between morphs, Guam and Hawai'i combined. B) Only F_{ST} outliers between morphs, that are outliers in both Guam and Hawai'i C) All OUTFLANK outliers D) Only genes that are OUTFLANK outliers in both Guam and Hawai'i E) Genes identified as associated with color morph by GAPIT using the FarmCPU algorithm F) Genes identified as under selection by SaltiLASSI

FST (total)	**FST (shared)**	**OUTFLANK (shared)**	**GAPIT (total)**	**LASSI**
AANAT	Ano1	APBA1	Inpp4a	BRAP
abcd1	CALM2	APOL6	rabgef5	C6
ABI1	DSP	Ccdc138	Rnd3	C7
Acod1	TJP2	FAM219A		Chd1
ADAMTS2	ZXDC	FGA		ETV6
ADAMTS7	fgf3	FGB		GBF1
AMBRA1	slc25a1b	HMGCS1		Kcnv2
Amt	tubb2b	KLC2		Leucine-rich repeat and calponin homology domain-containing protein 1
ANKMY2		Myo10		Ltbp1
Ano1		Myo9b		NAA25
AP4S1		PKHD1L1		PARP12
APBA1		TFCP2L1		pitx3
APOB		tmem150a		REEP5
ARHGEF10L		UBXN4		RERG
ARHGEF2		UPF0687 protein C20orf27 homolog		rgs7bpb
arpc4		Ythdc2		Setd5
ARR3				Sh3bp2
ATP2A2				sh3bp4
Atp2a3				SLC16A7
AUH				Syn3
Bax				Tafa2
BCL6				timp3
BCR				tor4a
bicdl1				VDAC2
BIRC6				VLDLR
Bmper				
bop1-a				
BRAP				

FST (total)	FST (shared)	OUTFLANK (shared)	GAPIT (total)	LASSI
bzw2				
C1ql3				
CALM2				
CAMK2G				
CAPN2				
CCDC88A				
CCN1				
Celsr3				
CEP120				
Chek2				
CLN3				
cnfn				
CNOT10				
cog4				
COL6A3				
Cpsf2				
CPT1A				
crispld2				
cuta				
Cux2				
CXCR4				
Dach2				
Dapk2				
diexf				
DNPEP				
DSP				
ENAH				
EPB41				
ERBB3				
EVPL				
FADD				
Fam169a				
fdxr				
fgf3				
Frmd4a				
GIMAP8				
GJD2				
Glrb				
gnai1				
Gnal				

FST (total)	FST (shared)	OUTFLANK (shared)	GAPIT (total)	LASSI
gnb1				
Gon4l				
GRB14				
GRIA1				
GRIA2				
grin1				
H2A				
HECTD1				
Hectd3				
Herc4				
hhex				
HIPK2				
HMGCR				
hmox				
HSF1				
Ift81				
IKZF3				
IP6K2				
ism2				
ISP2				
ITPR1				
Kifc3				
KLHL25				
KMT2C				
LAMA4				
lce				
Lcp2				
LMNB1				
LRRC45				
Ltbp1				
Magi2				
MAPK11				
Mcfd2				
Mphosph8				
MRC1				
Mtmr14				
Mtmr3				
MUC2				
Mxd1				
myhb				

FST (total)	FST (shared)	OUTFLANK (shared)	GAPIT (total)	LASSI
NISCH				
nlk.2				
nol12				
Npr1				
NSF				
NUP160				
P2rx3				
PARP12				
Pck1				
pdgfrb				
PDIA3				
PFKP				
Pfn2				
PLIN3				
PLXNB1				
POLR1A				
PRKACB				
PSIP1				
PTBP1				
Ptprk				
pwwp2a				
RABGEF1				
RASAL1				
RBPMS				
REEP5				
RPL7L1				
Rps6kb1				
SCHIP1				
scn4aa				
SF1				
Slc1a2				
SLC22A6				
slc25a1b				
SLC31A1				
Slc38a3				
Slc43a1				
SLC4A8				
Slco4a1				
Smad4				
Snx19				

FST (total)	FST (shared)	OUTFLANK (shared)	GAPIT (total)	LASSI
Son				
specc1l				
STARD3NL				
STUB1				
STX12				
Syn3				
TARS3				
TESK2				
TFCP2L1				
TJP2				
TMPRSS15				
Tnfaip6				
TNXB				
tpcn2				
Ttn				
tubb2b				
Ube2h				
Ubr7				
UMOD				
USP34				
Vps13c				
Vps37c				
VPS4B				
VPS8				
zdhhc5a				
ZNF135				
ZNF385A				
ZNF622				
ZNF644				
ZNF684				
ZXDC				

Table S3 Coloration genes meeting a maximum F_{ST} threshold of 0.15 (top), or mean F_{ST} threshold of 0.10 (bottom). Note that the genome mean F_{ST} between the morphs is 0.0047.

Gene Name	function	Guam max Fst	Hawai'i Max Fst
adam17	melanophore development	0.167	0.202
egfr	melanophore development	0.156	0.177
itgb1	melanophore development	0.182	0.238
sfxn1	melanophore development	0.308	0.152

gpnmb	components of melanosomes	0.176	0.269
ap3d1	melanosome construction	0.164	0.157
fig4	melanosome construction	0.170	0.163
hps3	melanosome construction	0.177	0.178
lyst	melanosome construction	0.164	0.158
vps39	melanosome construction	0.152	0.205
vps39	iridophore development	0.152	0.205
myo5a	melanosome transport	0.157	0.177
myo7a	melanosome transport	0.185	0.160
atrn	regulation of melanogenesis	0.171	0.195
drd2	regulation of melanogenesis	0.164	0.153
mycbp2	pteridine synthesis	0.164	0.153
skiv2l2	uncategorized function	0.203	0.163

Gene Name	function	Guam Mean Fst	Hawai'i Mean Fst
sox9	melanophore development	0.139	0.114
sox9	iridophore development	0.139	0.114
gpnmb	components of melanosomes	0.104	0.124
rps20	systemic effects	0.122	0.103

APPENDIX B: SUPPLEMENTAL FIGURES AND TABLES FOR CHAPTER 3

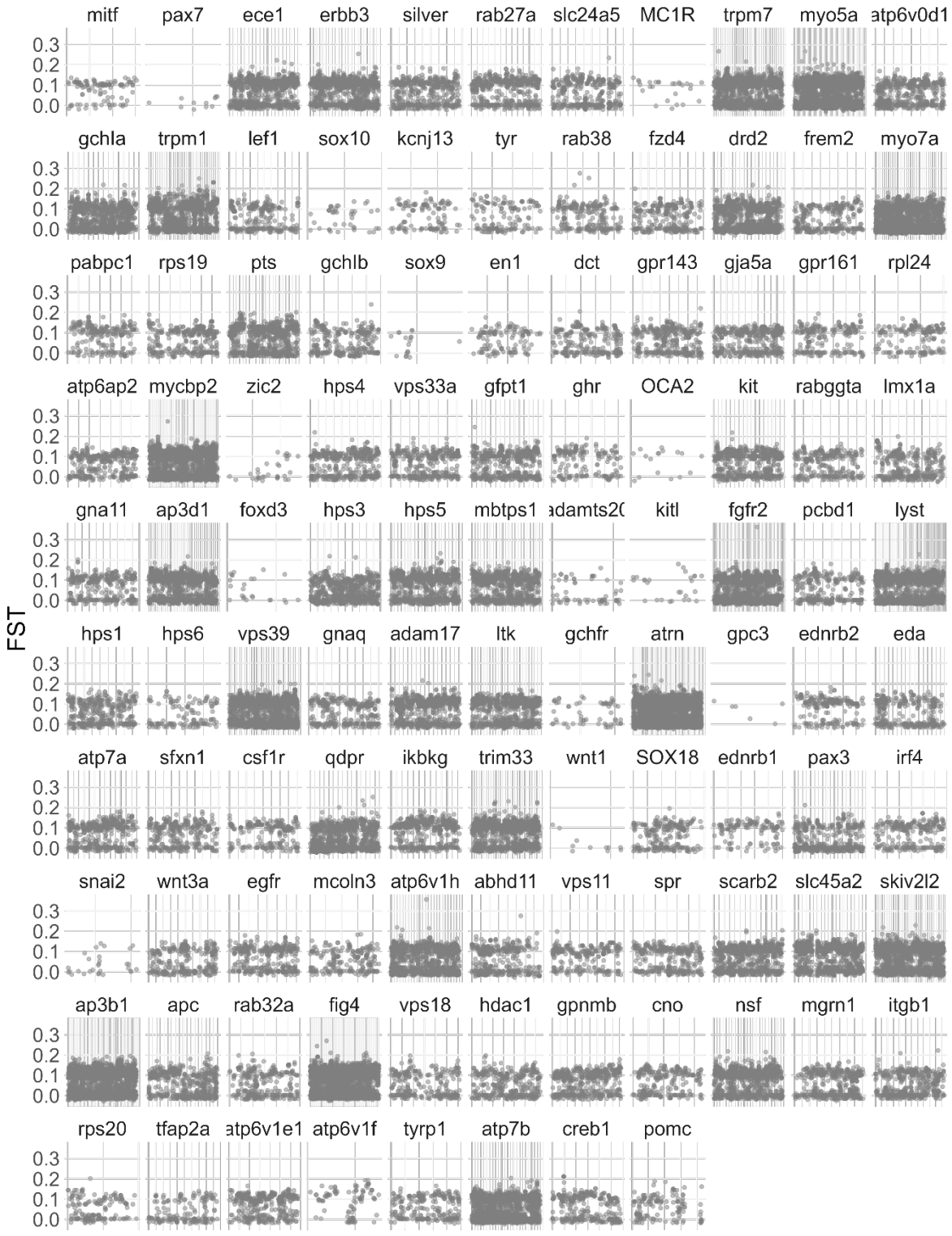


Figure S4 F_{ST} estimates calculated per-SNP for coloration genes using vcftools (Danecek et al., 2011). Outliers are identified using an FDR < 0.01 on Z-adjusted F_{ST} estimates, however, there are none.

Table S4 Sample information and sequencing statistics for *P. hemistictus*.

Sample ID	morph	Year Collected	adapter	mapped reads	mean quality	mean depth	coverage
PARHS_0017A_01	MEL	2021	QDIY-025	32312972	35.4	12.2585	0.9544
PARHS_0020C_01	MEL	2021	QDIY-060	28112980	35.5	10.6072	0.951582
PARHS_0021C_01	MEL	2021	QDIY-047	75040256	35.3	27.866	0.962162
PARHS_0022C_01	MEL	2021	QDIY-027	40029674	35.4	15.344	0.957121
PARHS_0023C_01	MEL	2021	QDIY-033	66991560	35.4	25.4094	0.961644
PARHS_0024C_01	MEL	2021	QDIY-031	74217978	35.6	27.4238	0.960862
PARHS_0025C_01	MEL	2021	QDIY-028	58616179	35.4	22.0395	0.960441
PARHS_0026C_01	MEL	2021	QDIY-042	51476207	35.3	19.5634	0.960015
PARHS_0027C_01	MEL	2021	QDIY-039	42803611	35.3	16.1502	0.957241
PARHS_0028C_01	MEL	2021	QDIY-056	36872260	35.4	13.7094	0.955813
PARHS_0029C_01	MEL	2021	QDIY-046	70419608	35.5	25.7685	0.961302
PARHS_0030C_02	YWS	2021	QDIY-054	48479545	35.4	17.8965	0.958738
PARHS_0031C_01	YWS	2021	QDIY-055	39883121	35.4	15.1096	0.956381
PARHS_0032C_01	YWS	2021	QDIY-065	20876888	35.3	7.98773	0.946956
PARHS_0033C_01	YWS	2021	QDIY-081	12458646	35.3	4.88917	0.925646
PARHS_0034C_01	YWS	2021	QDIY-082	15891387	35.4	6.18601	0.9375
PARHS_0035C_01	YWS	2021	QDIY-076	7498903	35.4	3.17512	0.850576
PARHS_0036C_01	YWS	2021	QDIY-073	14768314	35.4	5.7221	0.933752
PARHS_0037C_01	YWS	2021	QDIY-078	14210366	35.3	5.48036	0.932331
PARHS_0038C_01	YWS	2021	QDIY-090	14623357	35.3	5.65166	0.9342
PARHS_0039C_01	YWS	2021	QDIY-091	8090792	35.3	3.38429	0.916444
PARHS_0040C_01	YWS	2021	QDIY-079	16933506	35.4	6.43287	0.939894
PARHS_0041C_01	YWS	2021	QDIY-083	12100313	35.4	4.68586	0.905453
PARHS_0042C_01	YWS	2021	QDIY-092	8382509	35.2	3.48292	0.92272
PARHS_0043C_01	YWS	2021	QDIY-084	12521360	35.3	4.87681	0.924037
PARHS_0044C_01	YWS	2021	QDIY-074	8022747	35.3	3.29097	0.863099
PARHS_0045C_01	YWS	2021	QDIY-085	15000252	35.4	5.67048	0.934322
PARHS_0046C_01	YWS	2021	QDIY-086	15382159	35.4	5.89883	0.936228
PARHS_0047C_01	YWS	2021	QDIY-070	14633770	35.4	5.64499	0.93459
PARHS_0048C_01	YWS	2021	QDIY-066	17599463	35.4	6.68579	0.940757
PARHS_0049C_01	YWS	2021	QDIY-093	10839340	35.2	4.29751	0.954648
PARHS_0050C_01	YWS	2021	QDIY-094	7675764	35.3	3.19079	0.889524
PARHS_0051C_01	YWS	2021	QDIY-072	13263191	35.3	5.15415	0.927352
PARHS_0052C_01	YWS	2021	QDIY-075	8651832	35.4	3.54362	0.880178
PARHS_0053C_01	YWS	2021	QDIY-087	9296686	35.4	3.76207	0.89155
PARHS_0054C_01	MEL	2021	QDIY-040	70089379	35.5	26.0242	0.961643
PARHS_0055C_01	MEL	2021	QDIY-061	38124462	35.5	14.3161	0.9555
PARHS_0056C_01	MEL	2021	QDIY-052	55094658	35.5	20.5943	0.959541
PARHS_0057C_01	MEL	2021	QDIY-038	27635380	35.4	10.6421	0.952603

PARHS_0058C_01	MEL	2021	QDIY-029	36956466	35.4	14.1976	0.956093
PARHS_0059C_01	MEL	2021	QDIY-062	36809663	35.3	13.5945	0.956036
PARHS_0060C_01	MEL	2021	QDIY-057	22128396	35.4	8.5888	0.948032
PARHS_0061C_01	MEL	2021	QDIY-035	70214455	35.4	26.0971	0.961825
PARHS_0062C_01	YWS	2021	QDIY-095	14542918	35.3	5.65395	0.978024
PARHS_0063C_01	YWS	2021	QDIY-080	14391670	35.3	5.55747	0.932514
PARHS_0064C_01	MEL	2021	QDIY-063	18221090	35.5	6.95902	0.940687
PARHS_0065C_01	MEL	2021	QDIY-026	41852342	35.3	15.8048	0.957359
PARHS_0066C_01	MEL	2021	QDIY-058	31262721	35.4	11.6402	0.952942
PARHS_0067C_01	MEL	2021	QDIY-034	54657627	35.3	20.5485	0.95957
PARHS_0068C_01	MEL	2021	QDIY-043	47842724	35.4	17.8829	0.957994
PARHS_0069C_01	MEL	2021	QDIY-044	47723708	35.4	17.9087	0.958451
PARHS_0070C_01	MEL	2021	QDIY-048	55677921	35.2	20.5432	0.959932
PARHS_0071C_01	YWS	2021	QDIY-088	11478406	35.2	4.53365	0.921047
PARHS_0072C_01	MEL	2021	QDIY-041	67991013	35.5	24.2577	0.960584
PARHS_0073C_01	YWS	2021	QDIY-089	10943964	35.4	4.38313	0.915201
PARHS_0074C_01	YWS	2021	QDIY-096	13271711	35.2	5.19884	0.972265
PARHS_0075C_01	MEL	2021	QDIY-053	58089070	35.4	21.5921	0.960211
PARHS_0076C_01	MEL	2021	QDIY-024	44313093	35.3	16.5192	0.957421
PARHS_0077C_01	YWS	2021	QDIY-064	18098838	35.5	6.81049	0.940247
PARHS_0078C_01	YWS	2021	QDIY-067	11522147	35.4	4.61849	0.918403
PARHS_0079C_01	YWS	2021	QDIY-068	11297195	35.4	4.54705	0.910504
PARHS_0080C_01	MEL	2021	QDIY-049	40614049	35.3	15.3413	0.957231
PARHS_0081C_01	MEL	2021	QDIY-050	83675653	35.4	30.3229	0.963316
PARHS_0082C_01	MEL	2021	QDIY-036	49310337	35.3	18.6141	0.95954
PARHS_0083C_01	MEL	2021	QDIY-032	59517155	35.4	22.4638	0.960528
PARHS_0084C_01	MEL	2021	QDIY-037	31071455	35.5	11.8764	0.953738
PARHS_0085C_01	MEL	2021	QDIY-059	46229398	35.4	17.2126	0.957483
PARHS_0086C_01	MEL	2021	QDIY-051	49256715	35.4	18.6624	0.958491
PARHS_0087C_01	MEL	2021	QDIY-045	55549104	35.4	20.9674	0.95976

LIST OF REFERENCES

Allen, J. A., & Clarke, B. (1968). Evidence for apostatic selection by wild passerines. *Nature*, *220*(5166), 501–502. https://doi.org/10.1038/220501a0

Allender, C. J., Seehausen, O., Knight, M. E., Turner, G. F., & Maclean, N. (2003). Divergent selection during speciation of lake malawi cichlid fishes inferred from parallel radiations in nuptial coloration. *Proceedings of the National Academy of Sciences*, *100*(24), 14074. https://doi.org/10.1073/pnas.2332665100

Altschul, S. F., Gish, W., Miller, W., Myers, E. W., & Lipman, D. J. (1990). Basic local alignment search tool. *J Mol Biol*, *215*(3), 403–410. https://doi.org/https://doi.org/10.1016/S0022-2836(05)80360-2

Bakker, T. C. M. (1993). Positive genetic correlation between female preference and preferred male ornament in sticklebacks. *Nature*, *363*(6426), 255–257. https://doi.org/10.1038/363255a0

Bao, W., Kojima, K. K., & Kohany, O. (2015). Repbase update, a database of repetitive elements in eukaryotic genomes. *Mobile DNA*, *6*(1). https://doi.org/10.1186/s13100-015-0041-9

Barlow, G. W. (1973). Competition between color morphs of the polychromatic Midas cichlid *Cichlasoma citrinellum*. *Science*, *179*(4075), 806–807. https://doi.org/doi:10.1126/science.179.4075.806

Bezault, E., Mwaiko, S., & Seehausen, O. (2011). Population genomic tests of models of adaptive radiation in Lake Victoria region cichlid fish. *Evolution*, *65*(12), 3381–3397. https://doi.org/10.1111/j.1558-5646.2011.01417.x

Bhattacharya, D., Stephens, T. G., Tinoco, A. I., Richmond, R. H., & Cleves, P. A. (2022). Life on the edge: Hawaiian model for coral evolution. *Limnology and Oceanography*, *67*(9), 1976–1985. https://doi.org/10.1002/lno.12181

Bickhart, D. M., Rosen, B. D., Koren, S., Sayre, B. L., Hastie, A. R., Chan, S., Lee, J., Lam, E. T., Liachko, I., Sullivan, S. T., Burton, J. N., Huson, H. J., Nystrom, J. C., Kelley, C. M., Hutchison, J. L., Zhou, Y., Sun, J., Crisà, A., Ponce De León, F. A.,…Smith, T. P. L. (2017). Single-molecule sequencing and chromatin conformation capture enable de novo reference assembly of the domestic goat genome. *Nature Genetics*, *49*(4), 643–650. https://doi.org/10.1038/ng.3802

Bond, A. B., & Kamil, A. C. (2006). Spatial heterogeneity, predator cognition, and the evolution of color polymorphism in virtual prey. *Proceedings of the National Academy of Sciences*, *103*(9), 3214–3219. https://doi.org/10.1073/pnas.0509963103

Boutet, E., Lieberherr, D., Tognolli, M., Schneider, M., & Bairoch, A. (2007). UniProtKB/Swiss-Prot. In D. Edwards (Ed.), *Plant bioinformatics: Methods and protocols* (pp. 89–112). Humana Press. https://doi.org/10.1007/978-1-59745-535-0_4

Braasch, I., Brunet, F., Volff, J.-N., & Schartl, M. (2009). Pigmentation pathway evolution after whole-genome duplication in fish. *Genome Biology and Evolution*, *1*(0), 479–493. https://doi.org/10.1093/gbe/evp050

Bradbury, P. J., Zhang, Z., Kroon, D. E., Casstevens, T. M., Ramdoss, Y., & Buckler, E. S. (2007). Tassel: Software for association mapping of complex traits in diverse samples. *Bioinformatics*, *23*(19), 2633–2635. https://doi.org/10.1093/bioinformatics/btm308

Burton, J. N., Adey, A., Patwardhan, R. P., Qiu, R., Kitzman, J. O., & Shendure, J. (2013). Chromosome-scale scaffolding of de novo genome assemblies based on chromatin interactions. *Nature Biotechnology*, *31*(12), 1119–1125. https://doi.org/10.1038/nbt.2727

Calenge, C. (2006). The package “adehabitat” for the R software: A tool for the analysis of space and habitat use by animals. *Ecological Modelling*, *197*(3-4), 516–519.

Camacho, C., Coulouris, G., Avagyan, V., Ma, N., Papadopoulos, J., Bealer, K., & Madden, T. L. (2009). BLAST+: Architecture and applications. *BMC Bioinformatics*, *10*(1), 421. https://doi.org/10.1186/1471-2105-10-421

Carleton, K. L., Parry, J. W., Bowmaker, J. K., Hunt, D. M., & Seehausen, O. (2005). Colour vision and speciation in Lake Victoria cichlids of the genus *Pundamilia*. *Molecular Ecology*, *14*(14), 4341–4353.

Casas, L., Saenz-Agudelo, P., Villegas-Ríos, D., Irigoien, X., & Saborido-Rey, F. (2021). Genomic landscape of geographically structured colour polymorphism in a temperate marine fish. *Molecular Ecology*, *30*(5), 1281–1296. https://doi.org/10.1111/mec.15805

Chang, M., Shi, K., Li, W., Ma, W., Chen, S., & Sha, Z. (2025). Chromosome-level genome assembly and annotation of the spottedtail morwong (*Cheilodactylus zonatus*). *Scientific Data*, *12*(1). https://doi.org/10.1038/s41597-025-06029-x

Chen, S., Zhou, Y., Chen, Y., & Gu, J. (2018). Fastp: An ultra-fast all-in-one FASTQ preprocessor. *Bioinformatics*, *34*(17), i884–i890. https://doi.org/10.1093/bioinformatics/bty560

Clague, D. A., & Dalrymple, G. B. (1987). The Hawaiian-emperor volcanic chain. *US Geologcal Survey Professional Paper*, *1350*, 5–54.

Coker, D. J., Chaidez, V., & Berumen, M. L. (2017). Habitat use and spatial variability of hawkfishes with a focus on colour polymorphism in *Paracirrhites forsteri*. *PLoS ONE*, *12*(1), e0169079. https://doi.org/10.1371/journal.pone.0169079

Colin, P. L. (2012). Studying and monitoring aggregating species. In Y. Sadovy de Mitcheson & P. L. Colin (Eds.), *Reef fish spawning aggregations: Biology, research and management* (pp. 285–329). Springer Netherlands. https://doi.org/10.1007/978-94-007-1980-4_9

Cook, L. M., Grant, B. S., Saccheri, I. J., & Mallet, J. (2012). Selective bird predation on the peppered moth: The last experiment of Michael Majerus. *Biology Letters*, *8*(4), 609–612.

Coombe, L., Kazemi, P., Wong, J., Birol, I., & Warren, R. L. (2024). Multi-genome synteny detection using minimizer graph mappings. *bioRxiv*, 2024.2002.2007.579356. https://doi.org/10.1101/2024.02.07.579356

Coombe, L., Warren, R. L., & Birol, I. (2025). ntSynt-viz: Visualizing synteny patterns across multiple genomes. *bioRxiv*, 2025.2001.2015.633221. https://doi.org/10.1101/2025.01.15.633221

Corl, A., Davis, A. R., Kuchta, S. R., & Sinervo, B. (2010a). Selective loss of polymorphic mating types is associated with rapid phenotypic evolution during morphic speciation. *Proceedings of the National Academy of Sciences of the United States of America*, *107*(9), 4254–4259. https://doi.org/10.1073/pnas.0909480107

Corl, A., Davis, A. R., Kuchta, S. R., & Sinervo, B. (2010b). Selective loss of polymorphic mating types is associated with rapid phenotypic evolution during morphic speciation. *Proceedings of the National Academy of Sciences*, *107*(9), 4254–4259. https://doi.org/10.1073/pnas.0909480107

Cuvier, G., & Valenciennes, A. (1829). *Histoire naturelle des poissons. Tome troisième. Suite du livre troisième. Des percoïdes à dorsale unique à sept rayons branchiaux et à dents en velours ou en cardes*. F. G. Levrault, Paris.

Danecek, P., Auton, A., Abecasis, G., Albers, C. A., Banks, E., DePristo, M. A., Handsaker, R. E., Lunter, G., Marth, G. T., Sherry, S. T., McVean, G., Durbin, R., & Group, G. P. A. (2011). The variant call format and VCFtools. *Bioinformatics*, *27*(15), 2156–2158. https://doi.org/10.1093/bioinformatics/btr330

Danecek, P., Bonfield, J. K., Liddle, J., Marshall, J., Ohan, V., Pollard, M. O., Whitwham, A., Keane, T., Mccarthy, S. A., Davies, R. M., & Li, H. (2021). Twelve years of SAMtools and BCFtools. *GigaScience*, *10*(2). https://doi.org/10.1093/gigascience/giab008

DeGiorgio, M., Lohmueller, K. E., & Nielsen, R. (2014). A model-based approach for identifying signatures of ancient balancing selection in genetic data. *PLoS Genetics*, *10*(8), e1004561. https://doi.org/10.1371/journal.pgen.1004561

DeGiorgio, M., & Szpiech, Z. A. (2022). A spatially aware likelihood test to detect sweeps from haplotype distributions. *PLoS Genetics*, *18*(4), e1010134. https://doi.org/10.1371/journal.pgen.1010134
DeMartini, E. E. (1996). Sheltering and foraging substrate uses of the arc-eye hawkfish Paracirrhites arcatus (Pisces: Cirrhitidae). *Bulletin of Marine Science*, *58*(3), 826–837.
DeMartini, E. E., & Donaldson, T. J. (1996). Color morph-habitat relations in the arc-eye hawkfish *Paracirrhites arcatus* (Pisces: Cirrhitidae). *Copeia*, 362–371. https://doi.org/10.2307/1446852
Donaldson, T. J. (1986). Distribution and species richness patterns of Indo-West Pacific Cirrhitidae: Support for Woodland's hypothesis. Proceedings of the Second International Conference on Indo-Pacific Fishes,
Donaldson, T. J. (1990). Reproductive behavior and social organization of some Pacific hawkfishes (Cirrhitidae). *Japanese Journal of Ichthyology*, *36*(4), 439–458. https://doi.org/10.1007/bf02905464
Durand, N. C., Robinson, J. T., Shamim, M. S., Machol, I., Mesirov, J. P., Lander, E. S., & Aiden, E. L. (2016). Juicebox provides a visualization system for Hi-C contact maps with unlimited zoom. *Cell Systems*, *3*(1), 99–101. https://doi.org/10.1016/j.cels.2015.07.012
Durinck, S., Moreau, Y., Kasprzyk, A., Davis, S., De Moor, B., Brazma, A., & Huber, W. (2005). BioMart and Bioconductor: A powerful link between biological databases and microarray data analysis. *Bioinformatics*, *21*(16), 3439–3440. https://doi.org/10.1093/bioinformatics/bti525
Durinck, S., Spellman, P. T., Birney, E., & Huber, W. (2009). Mapping identifiers for the integration of genomic datasets with the R/Bioconductor package biomart. *Nature Protocols*, *4*(8), 1184–1191. https://doi.org/10.1038/nprot.2009.97
Dyer, S. C., Austine-Orimoloye, O., Azov, A. G., Barba, M., Barnes, I., Barrera-Enriquez, V. P., Becker, A., Bennett, R., Beracochea, M., Berry, A., Bhai, J., Bhurji, S. K., Boddu, S., Branco Lins, P. R., Brooks, L., Ramaraju, S. B., Campbell, L. I., Martinez, M. C., Charkhchi, M.,…Yates, A. D. (2024). Ensembl 2025. *Nucleic Acids Research*, *53*(D1), D948–D957. https://doi.org/10.1093/nar/gkae1071
Elmer, K. R., Lehtonen, T. K., & Meyer, A. (2009). Color assortative mating contributes to sympatric divergence of neotropical cichlid fish. *Evolution*, *63*(10), 2750–2757.
Endler, J. A. (1988). Frequency-dependent predation, crypsis and aposematic coloration. *Philosophical Transactions of the Royal Society of London. B, Biological Sciences*, *319*(1196), 505–523. https://doi.org/10.1098/rstb.1988.0062
Fang, B., Merilä, J., Ribeiro, F., Alexandre, C. M., & Momigliano, P. (2018). Worldwide phylogeny of three-spined sticklebacks. *Molecular Phylogenetics and Evolution*, *127*, 613–625. https://doi.org/https://doi.org/10.1016/j.ympev.2018.06.008
Faust, G. G., & Hall, I. M. (2014). SAMBLASTER: Fast duplicate marking and structural variant read extraction. *Bioinformatics*, *30*(17), 2503–2505. https://doi.org/10.1093/bioinformatics/btu314
Finke, D. L., & Snyder, W. E. (2008). Niche partitioning increases resource exploitation by diverse communities. *Science*, *321*(5895), 1488–1490. https://doi.org/doi:10.1126/science.1160854
FishBase. (2026). *Paracirrhites* Bleeker, 1874. In R. Froese & D. Pauly (Eds.). Retrieved 2026-03-23 through World Register of Marine Species at: https://www.marinespecies.org/aphia.php?p=taxdetails&id=204162.
Foll, M., & Gaggiotti, O. (2008). A genome-scan method to identify selected loci appropriate for both dominant and codominant markers: A Bayesian perspective. *Genetics*, *180*(2), 977–993. https://doi.org/10.1534/genetics.108.092221
Ford, E. B. (1945). Polymorphism. *Biological Reviews*, *20*(2), 73–88. https://doi.org/10.1111/j.1469-185X.1945.tb00315.x
Gaither, M. R., Coker, D. J., Greaves, S., Sarigol, F., Payet, S. D., Chaidez, V., Sinclair-Taylor, T. H., DiBattista, J. D., & Berumen, M. L. (2020). Does color matter? Molecular and ecological divergence in four

sympatric color morphs of a coral reef fish. *Ecology and Evolution*, *10*(18), 9663–9681. https://doi.org/10.1002/ece3.6566

Gaither, M. R., & Randall, J. E. (2012). On the validity of the cirrhitid fish genus *Itycirrhitus*. *Aqua*, *18*(4-15), 219–226.

Galeotti, P., & Rubolini, D. (2004). The niche variation hypothesis and the evolution of colour polymorphism in birds: A comparative study of owls, nightjars and raptors. *Biological Journal of the Linnean Society*, *82*(2), 237–248. https://doi.org/https://doi.org/10.1111/j.1095-8312.2004.00355.x

Galeotti, P., Sacchi, R., Pellitteri-Rosa, D., Bellati, A., Cocca, W., Gentilli, A., Scali, S., & Fasola, M. (2013). Colour polymorphism and alternative breeding strategies: Effects of parent's colour morph on fitness traits in the common wall lizard. *Evolutionary Biology*, *40*(3), 385–394. https://doi.org/10.1007/s11692-012-9222-3

Geange, S. W., Pledger, S., Burns, K. C., & Shima, J. S. (2011). A unified analysis of niche overlap incorporating data of different types. *Methods in Ecology and Evolution*, *2*(2), 175–184. https://doi.org/10.1111/j.2041-210x.2010.00070.x

Grabherr, M. G., Haas, B. J., Yassour, M., Levin, J. Z., Thompson, D. A., Amit, I., Adiconis, X., Fan, L., Raychowdhury, R., Zeng, Q., Chen, Z., Mauceli, E., Hacohen, N., Gnirke, A., Rhind, N., Di Palma, F., Birren, B. W., Nusbaum, C., Lindblad-Toh, K.,…Regev, A. (2011). Full-length transcriptome assembly from RNA-seq data without a reference genome. *Nature Biotechnology*, *29*(7), 644–652. https://doi.org/10.1038/nbt.1883

Gray, S. M., & McKinnon, J. S. (2007). Linking color polymorphism maintenance and speciation. *Trends in Ecology & Evolution*, *22*(2), 71–79. https://doi.org/https://doi.org/10.1016/j.tree.2006.10.005

Gross, J. B., Borowsky, R., & Tabin, C. J. (2009). A novel role for Mc1r in the parallel evolution of depigmentation in independent populations of the cavefish *Astyanax mexicanus*. *PLoS Genetics*, *5*(1), e1000326. https://doi.org/10.1371/journal.pgen.1000326

Günther, A. (1874). *Andrew garrett's fische der südsee.* (Vol. Band I Heft III). F. Friederichsen & Co.

Harris, R. B., Irwin, K., Jones, M. R., Laurent, S., Barrett, R. D. H., Nachman, M. W., Good, J. M., Linnen, C. R., Jensen, J. D., & Pfeifer, S. P. (2020). The population genetics of crypsis in vertebrates: Recent insights from mice, hares, and lizards. *Heredity*, *124*(1), 1–14. https://doi.org/10.1038/s41437-019-0257-4

Helmkampf, M., Coulmance, F., Heckwolf, M. J., Acero P., A., Balard, A., Bista, I., Dominguez, O. D., Frandsen, P. B., Torres-Oliva, M., Santaquiteria, A., Tavera, J., Victor, B. C., Robertson, D. R., Betancur-R., R., McMillan, W. O., & Puebla, O. (2025). Radiation with reproductive isolation in the near-absence of phylogenetic signal. *Science Advances*, *11*(30), eadt0973. https://doi.org/doi:10.1126/sciadv.adt0973

Hench, K., Helmkampf, M., McMillan, W. O., & Puebla, O. (2022). Rapid radiation in a highly diverse marine environment. *Proceedings of the National Academy of Sciences*, *119*(4), e2020457119. https://doi.org/doi:10.1073/pnas.2020457119

Hoekstra, H. E., Drumm, K. E., & Nachman, M. W. (2004). Ecological genetics of adaptive color polymorphism in pocket mice: Geographic variation in selected and neutral genes. *Evolution*, *58*(6), 1329–1341. https://doi.org/10.1111/j.0014-3820.2004.tb01711.x

Holt, C., & Yandell, M. (2011). MAKER2: An annotation pipeline and genome-database management tool for second-generation genome projects. *BMC Bioinformatics*, *12*(1), 491. https://doi.org/10.1186/1471-2105-12-491

Hou, L., Panthier, J.-J., & Arnheiter, H. (2000). Signaling and transcriptional regulation in the neural crest-derived melanocyte lineage: Interactions between KIT and MITF. *Development*, *127*(24), 5379–5389.

Huang, M., Liu, X., Zhou, Y., Summers, R. M., & Zhang, Z. (2019). BLINK: A package for the next level of genome-wide association studies with both individuals and markers in the millions. *GigaScience*, *8*(2). https://doi.org/10.1093/gigascience/giy154

Hubbard, J. K., Uy, J. A., Hauber, M. E., Hoekstra, H. E., & Safran, R. J. (2010). Vertebrate pigmentation: From underlying genes to adaptive function. *Trends in Genetics*, *26*(5), 231–239. https://doi.org/10.1016/j.tig.2010.02.002

Hugall, A. F., & Stuart-Fox, D. (2012). Accelerated speciation in colour-polymorphic birds. *Nature*, *485*(7400), 631–634. https://doi.org/10.1038/nature11050

Iserbyt, A., Bots, J., Van Gossum, H., & Jordaens, K. (2010). Did historical events shape current geographic variation in morph frequencies of a polymorphic damselfly? *Journal of Zoology*, *282*(4), 256–265. https://doi.org/https://doi.org/10.1111/j.1469-7998.2010.00735.x

Jamie, G. A., & Meier, J. I. (2020). The persistence of polymorphisms across species radiations. *Trends in Ecology & Evolution*, *35*(9), 795–808. https://doi.org/10.1016/j.tree.2020.04.007

Jenck, C. S., Lehto, W. R., Hunnicutt, K. E., Murphy, S. M., Quinn, T. W., Larson, E. L., & Tinghitella, R. M. (2022). Genetic divergence among threespine stickleback that differ in nuptial coloration. *Journal of Evolutionary Biology*, *35*(7), 934–947. https://doi.org/10.1111/jeb.14035

Jones, Felicity C., Chan, Yingguang F., Schmutz, J., Grimwood, J., Brady, Shannon D., Southwick, Audrey M., Absher, Devin M., Myers, Richard M., Reimchen, Thomas E., Deagle, Bruce E., Schluter, D., & Kingsley, David M. (2012). A genome-wide SNP genotyping array reveals patterns of global and repeated species-pair divergence in sticklebacks. *Current Biology*, *22*(1), 83–90. https://doi.org/10.1016/j.cub.2011.11.045

Kane, C. N., Brooks, A. J., Holbrook, S. J., & Schmitt, R. J. (2009). The role of microhabitat preference and social organization in determining the spatial distribution of a coral reef fish. *Environmental Biology of Fishes*, *84*(1), 1–10. https://doi.org/10.1007/s10641-008-9377-z

Kautt, A. F., Kratochwil, C. F., Nater, A., Machado-Schiaffino, G., Olave, M., Henning, F., Torres-Dowdall, J., Harer, A., Hulsey, C. D., Franchini, P., Pippel, M., Myers, E. W., & Meyer, A. (2020). Contrasting signatures of genomic divergence during sympatric speciation. *Nature*. https://doi.org/10.1038/s41586-020-2845-0

Keller, I., Wagner, C. E., Greuter, L., Mwaiko, S., Selz, O. M., Sivasundar, A., Wittwer, S., & Seehausen, O. (2013). Population genomic signatures of divergent adaptation, gene flow and hybrid speciation in the rapid radiation of Lake Victoria cichlid fishes. *Molecular Ecology*, *22*(11), 2848–2863. https://doi.org/10.1111/mec.12083

Kratochwil, C. F., Kautt, A. F., Nater, A., Härer, A., Liang, Y., Henning, F., & Meyer, A. (2022). An intronic transposon insertion associates with a trans-species color polymorphism in Midas cichlid fishes. *Nature Communications*, *13*(1). https://doi.org/10.1038/s41467-021-27685-8

Kronenberg, Z. N., Rhie, A., Koren, S., Concepcion, G. T., Peluso, P., Munson, K. M., Porubsky, D., Kuhn, K., Mueller, K. A., Low, W. Y., Hiendleder, S., Fedrigo, O., Liachko, I., Hall, R. J., Phillippy, A. M., Eichler, E. E., Williams, J. L., Smith, T. P. L., Jarvis, E. D.,…Kingan, S. B. (2021). Extended haplotype-phasing of long-read de novo genome assemblies using Hi-C. *Nature Communications*, *12*(1935). https://doi.org/10.1038/s41467-020-20536-y (10.1101/327064)

Lai, F.-N., Zhai, H.-L., Cheng, M., Ma, J.-Y., Cheng, S.-F., Ge, W., Zhang, G.-L., Wang, J.-J., Zhang, R.-Q., Wang, X., Min, L.-J., Song, J.-Z., & Shen, W. (2016). Whole-genome scanning for the litter size trait associated genes and SNPs under selection in dairy goat (*Capra hircus*). *Scientific Reports*, *6*(1), 38096. https://doi.org/10.1038/srep38096

Lande, R. (1980). Sexual dimorphism, sexual selection, and adaptation in polygenic characters. *Evolution*, *34*(2), 292–305. https://doi.org/10.2307/2407393

Leray, M., Boehm, J., Mills, S. C., & Meyer, C. (2012). Moorea BIOCODE barcode library as a tool for understanding predator–prey interactions: Insights into the diet of common predatory coral reef fishes. *Coral Reefs*, *31*(2), 383–388. https://doi.org/10.1007/s00338-011-0845-0

Levins, R. (1968). *Evolution in changing environments some theoretical explorations.* Princeton University Press. https://doi.org/10.2307/j.ctvx5wbbh

Li, H. (2013). Aligning sequence reads, clone sequences and assembly contigs with BWA-MEM [Preprint]. *arXiv*, 1303.3997. https://doi.org/arXiv:1303.3997v2

Li, H. (2018). Minimap2: Pairwise alignment for nucleotide sequences. *Bioinformatics*, *34*(18), 3094–3100. https://doi.org/10.1093/bioinformatics/bty191

Li, H., Handsaker, B., Wysoker, A., Fennell, T., Ruan, J., Homer, N., Marth, G., Abecasis, G., & Durbin, R. (2009). The sequence alignment/map format and SAMtools. *Bioinformatics*, *25*(16), 2078–2079. https://doi.org/https://doi.org/10.1093/bioinformatics/btp352

Lobel, P. S. (2011). A review of the caribbean hamlets (Serranidae, *Hypoplectrus*) with description of two new species. *Zootaxa*, *3096*(1), 1. https://doi.org/10.11646/zootaxa.3096.1.1

Lorenz, K. (1962). Function of colour in coral reef fishes. *Proceedings of the Royal Institution of Great Britain*, *39*, 282.

Manceau, M., Domingues, V. S., Linnen, C. R., Rosenblum, E. B., & Hoekstra, H. E. (2010). Convergence in pigmentation at multiple levels: Mutations, genes and function. *Philosophical Transactions of the Royal Society B*, *365*(1552), 2439–2450. https://doi.org/doi:10.1098/rstb.2010.0104

Manly, B., McDonald, L., Thomas, D. L., McDonald, T. L., & Erickson, W. P. (2002). *Resource selection by animals: Statistical design and analysis for field studies*. Springer Science & Business Media.

Marshall, N. B. (1950). Fishes from the Cocos-Keeling islands. *Bulletin of the Raffles Museum*, *22*, 166–205.

Matsubayashi, K. W., & Yamaguchi, R. (2022). The speciation view: Disentangling multiple causes of adaptive and non-adaptive radiation in terms of speciation. *Population Ecology*, *64*(2), 95–107. https://doi.org/https://doi.org/10.1002/1438-390X.12103

McKaye, K. R., & Barlow, G. W. (1976). Competition between color morphs of the Midas cichlid, *Cichlasoma citrinellum*, in Lake Jiloá, Nicaragua. In T. B. Thorson (Ed.), *Investigations of the ichthyofauna of Nicaraguan lakes* (pp. 34).

McLean, C. A., & Stuart-Fox, D. (2014). Geographic variation in animal colour polymorphisms and its role in speciation. *Biological Reviews*, *89*(4), 860–873. https://doi.org/10.1111/brv.12083

McMillan, W. O., Weigt, L. A., & Palumbi, S. R. (1999). Color pattern evolution, assortative mating, and genetic differentiation in brightly colored butterflyfishes (Chaetodontidae). *Evolution*, *53*(1), 247–260. https://doi.org/10.1111/j.1558-5646.1999.tb05350.x

McNamara, M. E., Rossi, V., Slater, T. S., Rogers, C. S., Ducrest, A. L., Dubey, S., & Roulin, A. (2021). Decoding the evolution of melanin in vertebrates. *Trends in Ecology & Evolution*, *36*(5), 430–443. https://doi.org/https://doi.org/10.1016/j.tree.2020.12.012

Mi, H., Ebert, D., Muruganujan, A., Mills, C., Albou, L.-P., Mushayamaha, T., & Thomas, P. D. (2021). PANTHER version 16: A revised family classification, tree-based classification tool, enhancer regions and extensive API. *Nucleic Acids Research*, *49*(D1), D394–D403. https://doi.org/10.1093/nar/gkaa1106

Mi, H., Muruganujan, A., Huang, X., Ebert, D., Mills, C., Guo, X., & Thomas, P. D. (2019). Protocol update for large-scale genome and gene function analysis with the PANTHER classification system (v.14.0). *Nature Protocols*, *14*(3), 703–721. https://doi.org/10.1038/s41596-019-0128-8

Mundy, N. I., & Kelly, J. (2006). Investigation of the role of the agouti signaling protein gene (*asip*) in coat color evolution in primates. *Mammalian Genome*, *17*(12), 1205–1213. https://doi.org/10.1007/s00335-006-0056-0

Myers, R., & Donaldson, T. (2003). The fishes of the Mariana Islands. *Micronesica*, *3536*, 594–648.

Oksanen, J., Simpson, G. L., Blanchet, F. G., Kindt, R., Legendre, P., Minchin, P. R., O'Hara, R. B., Solymos, P., Stevens, M. H. H., Szoecs, E., Wagner, H., Barbour, M., Bedward, M., Bolker, B., Borcard, D., Carvalho, G., Chirico, M., De Caceres, M., Durand, S.,…Weedon, J. (2023). *Vegan: Community ecology package*. In (Version 2.6-5) https://github.com/vegandevs/vegan

Olendorf, R., Rodd, F. H., Punzalan, D., Houde, A. E., Hurt, C., Reznick, D. N., & Hughes, K. A. (2006). Frequency-dependent survival in natural guppy populations. *Nature*, *441*(7093), 633–636. https://doi.org/10.1038/nature04646

Orteu, A., & Jiggins, C. D. (2020). The genomics of coloration provides insights into adaptive evolution. *Nature Reviews Genetics*, *21*(8), 461–475. https://doi.org/10.1038/s41576-020-0234-z

Oxford, G. S., & Gillespie, R. G. (2001). Portraits of evolution: Studies of coloration in Hawaiian spiders: The discrete color polymorphisms in spiders allow the study of evolution "in action". *BioScience*, *51*(7), 521–528. https://doi.org/10.1641/0006-3568(2001)051[0521:Poesoc]2.0.Co;2

Picq, S., McMillan, W. O., & Puebla, O. (2016). Population genomics of local adaptation versus speciation in coral reef fishes (*Hypoplectrus spp*, Serranidae). *Ecology and Evolution*, *6*(7), 2109–2124. https://doi.org/10.1002/ece3.2028

Price, T., & Birch, G. L. (1996). Repeated evolution of sexual color dimorphism in passerine birds. *The Auk*, *113*(4), 842–848. https://doi.org/10.2307/4088862

Protas, M. E., & Patel, N. H. (2008). Evolution of coloration patterns. *Annual Review of Cell and Developmental Biology*, *24*(1), 425–446. https://doi.org/10.1146/annurev.cellbio.24.110707.175302

Puebla, O., Bermingham, E., & Guichard, F. (2008). Population genetic analyses of *Hypoplectrus* coral reef fishes provide evidence that local processes are operating during the early stages of marine adaptive radiations. *Molecular Ecology*, *17*(6), 1405–1415. https://doi.org/https://doi.org/10.1111/j.1365-294X.2007.03654.x

Puebla, O., Bermingham, E., Guichard, F., & Whiteman, E. (2007). Colour pattern as a single trait driving speciation in *Hypoplectrus* coral reef fishes? *Proceedings of the Royal Society B: Biological Sciences*, *274*(1615), 1265–1271. https://doi.org/10.1098/rspb.2006.0435

Puebla, O., Bermingham, E., & McMillan, W. O. (2014). Genomic atolls of differentiation in coral reef fishes (*Hypoplectrus spp.,* Serranidae). *Molecular Ecology*, *23*(21), 5291–5303. https://doi.org/https://doi.org/10.1111/mec.12926

Puebla, O., Coulmance, F., Estap, C. J., Estap, A. M., & Robertson, D. R. (2022). A review of 263 years of taxonomic research on *Hypoplectrus* (Perciformes: Serranidae), with a redescription of *Hypoplectrus affinis* (poey, 1861). *Zootaxa*, *5093*(2), 101–141. Retrieved 2022/01//, from https://doi.org/10.11646/zootaxa.5093.2.1

R Core Team. (2019). *R: A language and environment for statistical computing*. In *R: A Language and Environment for Statistical Computing* (Version 3.6.0) R Foundation for Statistical Computing. https://www.R-project.org/

Randall, J. E. (1963). Review of the hawkfishes (family Cirrhitidae). *Proceedings of the United States National Museum*.

Randall, J. E. (1997). *Randall's tank photos*.

Randall, J. E. (2005). Reef and shore fishes of the South Pacific. *University of Hawai'i Press, Honolulu*.

Rao, S. P., Suhas, Huntley, H., Miriam, Durand, C., Neva, Stamenova, K., Elena, Bochkov, D., Ivan, Robinson, T., James, Sanborn, L., Adrian, Machol, I., Omer, D., Arina, Lander, S., Eric, & Aiden, L., Erez. (2014). A 3D map of the human genome at kilobase resolution reveals principles of chromatin looping. *Cell*, *159*(7), 1665–1680. https://doi.org/10.1016/j.cell.2014.11.021

Recknagel, H., Kusche, H., Elmer, K. R., & Meyer, A. (2013). Two new endemic species in the Midas cichlid species complex from Nicaraguan crater lakes: *Amphilophus tolteca* and *Amphilophus viridis* (Perciformes, Cichlidae). *Aqua*, *19*(4), 207–224.

Riegl, B. M., Purkis, S. J., Houk, P., Cabrera, G., & Dodge, R. E. (2008). Geologic setting and geomorphology of coral reefs in the Mariana Islands (Guam and Commonwealth of the Northern Mariana Islands). In *Coral reefs of the USA* (pp. 691–718). Springer.

Robinson, B. W., & Wilson, D. S. (1994). Character release and displacement in fishes: A neglected literature. *The American Naturalist*, *144*(4), 596–627. https://doi.org/10.1086/285696

Rosenblum, E. B., Hoekstra, H. E., & Nachman, M. W. (2004). Adaptive reptile color variation and the evolution of the *mc1r* gene. *Evolution*, *58*(8), 1794–1808. https://doi.org/10.1111/j.0014-3820.2004.tb00462.x

Sánchez-Guillén, R. A., Ceccarelli, S., Villalobos, F., Neupane, S., Rivas-Torres, A., Sanmartín-Villar, I., Wellenreuther, M., Bybee, S. M., Velásquez-Vélez, M. I., Realpe, E., Chávez-Ríos, J. R., Dumont, H. J., & Cordero-Rivera, A. (2020). The evolutionary history of colour polymorphism in *Ischnura* damselflies (Odonata: Coenagrionidae). *Odonatologica*, *49*(3-4), 333–370, 338.

Seehausen, O., Mayhew, P. J., & Alphen, J. J. M. V. (1999). Evolution of colour patterns in east African cichlid fish. *Journal of Evolutionary Biology*, *12*(3), 514–534. https://doi.org/10.1046/j.1420-9101.1999.00055.x

Seehausen, O., Terai, Y., Magalhaes, I. S., Carleton, K. L., Mrosso, H. D. J., Miyagi, R., van der Sluijs, I., Schneider, M. V., Maan, M. E., Tachida, H., Imai, H., & Okada, N. (2008). Speciation through sensory drive in cichlid fish. *Nature*, *455*(7213), 620–626. https://doi.org/10.1038/nature07285

Sergey, G. (2003). Perspective: Models of speciation: What have we learned in 40 years? *Evolution*, *57*(10), 2197–2215. https://doi.org/10.1554/02-727

Servedio, M. R., Van Doorn, G. S., Kopp, M., Frame, A. M., & Nosil, P. (2011). Magic traits in speciation: 'Magic' but not rare? *Trends in Ecology & Evolution*, *26*(8), 389–397. https://doi.org/10.1016/j.tree.2011.04.005

Simão, F. A., Waterhouse, R. M., Ioannidis, P., Kriventseva, E. V., & Zdobnov, E. M. (2015). BUSCO: Assessing genome assembly and annotation completeness with single-copy orthologs. *Bioinformatics*, *31*(19), 3210–3212. https://doi.org/10.1093/bioinformatics/btv351

Smit, A., & Hubley, R. (2017). *Repeatmodeler open-1.0.2008-2015*. In (Version 1.0.11) http://www.repeatmasker.org/

Smit, A., Hubley, R., & Green, P. (2019). *Repeatmasker open-4.0.2013-2015*. In (Version 4.0.9) http://www.repeatmasker.org.

Sovic, I., Kronenberg, Z., Dunn, C., Barnett, D., Kingan, S., & Drake, J. (2020). *Improved phased assembler*. In (Version 1.0.4) Pacific Biosciences. https://github.com/PacificBiosciences/pbipa

Stanke, M., Keller, O., Gunduz, I., Hayes, A., Waack, S., & Morgenstern, B. (2006). Augustus: *Ab initio* prediction of alternative transcripts. *Nucleic Acids Research*, *34*(suppl_2), W435–W439. https://doi.org/10.1093/nar/gkl200

Streelman, J. T., Alfaro, M., Westneat, M. W., Bellwood, D. R., & Karl, S. A. (2002). Evolutionary history of the parrotfishes: Biogeography, ecomorphology, and comparative diversity. *Evolution*, *56*(5), 961–971. https://doi.org/10.1111/j.0014-3820.2002.tb01408.x

Sturmbauer, C., Baric, S., Salzburger, W., Rüber, L., & Verheyen, E. (2001). Lake level fluctuations synchronize genetic divergences of cichlid fishes in African lakes. *Molecular Biology and Evolution*, *18*(2), 144–154. https://doi.org/10.1093/oxfordjournals.molbev.a003788

Switonski, M., Mankowska, M., & Salamon, S. (2013). Family of melanocortin receptor (MCR) genes in mammals—mutations, polymorphisms and phenotypic effects. *Journal of Applied Genetics*, *54*(4), 461–472. https://doi.org/10.1007/s13353-013-0163-z

The UniProt Consortium. (2018). Uniprot: A worldwide hub of protein knowledge. *Nucleic Acids Research*, *47*(D1), D506–D515. https://doi.org/10.1093/nar/gky1049

Toonen, R., Bowen, B., Iacchei, M., & Briggs, J. (2016). Biogeography, marine. *Encyclopedia of Evolutionary Biology*. https://doi.org/10.1016/B978-0-12-800049-6.00120-7

Torres-Dowdall, J., Pierotti, M. E. R., Härer, A., Karagic, N., Woltering, J. M., Henning, F., Elmer, K. R., & Meyer, A. (2017). Rapid and parallel adaptive evolution of the visual system of neotropical Midas cichlid fishes. *Molecular Biology and Evolution*, *34*(10), 2469–2485. https://doi.org/10.1093/molbev/msx143

Trygonis, V., & Sini, M. (2012). Photoquad: A dedicated seabed image processing software, and a comparative error analysis of four photoquadrat methods. *Journal of Experimental Marine Biology and Ecology*, *424-425*, 99–108. https://doi.org/10.1016/j.jembe.2012.04.018

Twomey, E., Melo-Sampaio, P., Schulte, L. M., Bossuyt, F., Brown, J. L., & Castroviejo-Fisher, S. (2023). Multiple routes to color convergence in a radiation of neotropical poison frogs. *Systematic Biology*, *72*(6), 1247–1261. https://doi.org/10.1093/sysbio/syad051

Uller, T., Feiner, N., Sacchi, R., Zuffi, M., Scali, S., Pafilis, P., Plavos, K., Abalos, J., Andrade, P., Aguilar, P., Salvi, D., & While, G. M. (2026). Adaptive spread of a sexually selected syndrome eliminates an ancient color polymorphism in wall lizards. *Science*, *391*(6780), 64–68. https://doi.org/doi:10.1126/science.adx3708

Vasimuddin, M., Misra, S., Li, H., & Aluru, S. (2019, 20–24 May 2019). Efficient architecture-aware acceleration of BWA-MEM for multicore systems. 2019 IEEE International Parallel and Distributed Processing Symposium (IPDPS),

Veron, J. E. N., Stafford-Smith, M. G., Turak, E., & DeVantier, L. M. (2016). *Corals of the world*. https://www.coralsoftheworld.org/coral_geographic/interactive_map/results/?version=0.01

Wagner, C. E., Harmon, L. J., & Seehausen, O. (2012). Ecological opportunity and sexual selection together predict adaptive radiation. *Nature*, *487*(7407), 366–369. https://doi.org/10.1038/nature11144

Wang, J., & Zhang, Z. (2021). GAPIT version 3: Boosting power and accuracy for genomic association and prediction. *Genomics Proteomics Bioinformatics*, *19*(4), 629–640. https://doi.org/10.1016/j.gpb.2021.08.005

Wellenreuther, M., Svensson, E. I., & Hansson, B. (2014). Sexual selection and genetic colour polymorphisms in animals. *Molecular Ecology*, *23*(22), 5398–5414. https://doi.org/10.1111/mec.12935

West-Eberhard, M. J. (1986). Alternative adaptations, speciation, and phylogeny (a review). *Proceedings of the National Academy of Sciences*, *83*(5), 1388–1392. https://doi.org/10.1073/pnas.83.5.1388

White, T. E., & Kemp, D. J. (2016). Colour polymorphism. *Current Biology*, *26*(13), R517–R518. https://doi.org/10.1016/j.cub.2016.03.017

Whitlock, M. C., & Lotterhos, K. E. (2015). Reliable detection of loci responsible for local adaptation: Inference of a null model through trimming the distribution of FST. *The American Naturalist*, *186*(S1), S24–S36. https://doi.org/10.1086/682949

Whitney, J. L., Bowen, B. W., & Karl, S. A. (2018). Flickers of speciation: Sympatric colour morphs of the arc-eye hawkfish, *Paracirrhites arcatus*, reveal key elements of divergence with gene flow. *Molecular Ecology*, *27*(6), 1479–1493. https://doi.org/10.1111/mec.14527

Whitney, J. L., Donahue, M. J., & Karl, S. A. (2018). Niche divergence along a fine-scale ecological gradient in sympatric color morphs of a coral reef fish. *Ecosphere*, *9*(1).

Williams, P. H., Byvaltsev, A. M., Cederberg, B., Berezin, M. V., Ødegaard, F., Rasmussen, C., Richardson, L. L., Huang, J., Sheffield, C. S., & Williams, S. T. (2015). Genes suggest ancestral colour polymorphisms are shared across morphologically cryptic species in arctic bumblebees. *PLoS ONE*, *10*(12), e0144544. https://doi.org/10.1371/journal.pone.0144544

Wright, D. S., Eijk, R., Schuart, L., Seehausen, O., Groothuis, T. G. G., & Maan, M. E. (2020). Testing sensory drive speciation in cichlid fish: Linking light conditions to opsin expression, opsin genotype and female mate preference. *Journal of Evolutionary Biology*, *33*(4), 422–434.

Wright, D. S., Pierotti, M. E., Rundle, H. D., & McKinnon, J. S. (2015). Conspicuous female ornamentation and tests of male mate preference in threespine sticklebacks (*Gasterosteus aculeatus*). *PLoS ONE*, *10*(3), e0120723. https://doi.org/10.1371/journal.pone.0120723

Zhang, J., & Ma, K. (2013). spaa: An R package for computing species association and niche overlap. Research Progress of Biodiversity Conservation in China,