THE UNIVERSITY OF CHICAGO

PHYLOGENY, BIOGEOGRAPHY, AND EVOLUTIONARY ECOMORPHOLOGY OF THE GOATFISHES (FAMILY MULLIDAE)

A DISSERTATION SUBMITTED TO
THE FACULTY OF THE DIVISION OF THE BIOLOGICAL SCIENCES
AND THE PRITZKER SCHOOL OF MEDICINE
IN CANDIDACY FOR THE DEGREE OF
DOCTOR OF PHILOSOPHY

COMMITTEE ON EVOLUTIONARY BIOLOGY

BY
CHLOE MAE NASH

CHICAGO, ILLINOIS
JUNE 2023

For my family.

TABLE OF CONTENTS

LIST OF FIGURES

ACKNOWLEDGMENTS

My dissertation would not have been possible without the extraordinary support of many people and institutions. First and foremost, I would like to thank my advisor, Mark Westneat. From day 1, Mark has encouraged and supported me as I developed my identity and focus as a scientist. Mark always believed in me and my work, even when I didn't believe in myself. I will never forget how proud he looked after my first graduate school conference talk at the Indo-Pacific Fish Conference. Although I tend to overwhelm myself with new ideas and methods, Mark always helped me to step back, refine, improve, and eventually write up my projects. This has resulted in a body of work that we have built together that I am extremely proud of and has formed the basis of my future research interests. Mark accompanied me on three field seasons to Moorea, French Polynesia, during which I learned the skills required to lead a field team, how to set up a field experiment, how to catch goatfish using a block net, how to drive a boat, and most importantly, how much joy fresh tuna and French wine brings after a long day of diving. I treasure the many conversations we had during the time we spent in Moorea, during which Mark would remind me that I am capable scientist and that goatfishes are pretty cool. I also greatly value the time Mark took during the height of the pandemic to listen and make sure I knew I was supported. The academic freedom that Mark has provided me with both in the lab and in the field greatly increased my confidence in my ability to design and conduct independent research projects and mentor students, which makes me feel well prepared to start my postdoctoral position and my future academic career.

I would also like to thank the members of my committee for their continued support throughout my time at UChicago. My thesis chair, Michael Coates, has always supported me and provided useful and enjoyable discourse on all things fish related. Mike was always up for a chat about how my research was going and provided incredibly thoughtful context to how my research fits into the greater paleofish world. I also want to acknowledge how much I

appreciate the work that Mike did during his tenure as CEB chair, and we are all so fortunate that he was at the helm of our program. I'm happy that I got to share the joy that was Barb with Mike, and I look forward to showing him the many Japanese Barbs in Okinawa. Additionally, Graham Slater has always been a valued resource, especially with regards to phylogenetic comparative methods. I greatly appreciate my meetings with Graham where he helped me with both code and how to interpret various evolutionary statistics. I also think fondly on attending the Midwest Phylogenetic Workshop with Graham, in particular a very competitive game of ping pong. Finally, I would like to acknowledge Rick Ree. Rick always had useful and valuable feedback regarding my bioegeography related projects and helped me be more confident as a tackled the large world of bioegeographic methodology. Rick would always remind me of the importance of life history when considering these types of integrative evolutionary projects, and I enjoyed discussing the world of outrigging. I am grateful for my previous mentor, Barry Chernoff, who got me started in the world of research and showed me that this future is possible. I also want to thank Michelle Kraczkowski for her continued support, both personally and as a scientist, and for making me an Auntie.

My time at UChicago would not be complete without mentioning the staff of the Fish Collection and education department at the Field Museum of Natural History. Caleb McMahan, Susan Mochel, and Kevin Swagel made me feel welcome from the beginning, and visiting them in the collections was always a highlight of my week. They encouraged and supported my research both in terms of collections access and many, many fruitful discussions about all things fish related. They taught me how to properly respect, work in, and care for a specimen collection, in addition to how to organize data so that everyone's lives are easier. Caleb was always willing to discuss my projects with him, despite the fact I was working on reef fish, and I look forward to continuing to collaborate with him in the future. Susan made working in the collection incredibly enjoyable, and always made me laugh extensively while doing somewhat monotonous tasks, such as serenading me while we subsampled tissues in

the cryo facility. Susan was always there for me, be it a ride back to Hyde Park after a long night at Dozin' or someone to gossip with in the collection. Additionally, the opportunities to work on education events was a highlight of my time in Chicago, through events such as Dozin' with the Dinos and Members Night. These opportunities were made special by the people that I was fortunate to work with, including Caleb, Susan, Kevin, Ed Schweitzer, and Josh Mata. The Field Museum will always be a special place to me, and greatly look forward to continuing to collaborate with everyone mentioned above.

I was extremely fortunate to have been part of the most supportive CEB cohort one could ever ask for: Ryan Fuller, Peter Flynn, Brooke Weigel, Ben Otoo, and Jacob Cooper. We have all been close since our first day as graduate students, and it has been incredibly fulfilling watching us grow and continue in our careers. I want to acknowledge Peter as the best roommate and cat dad, Brooke as the best scuba diving buddy, Jacob as the best prankster and bird lover, Ben as the best dancer, and Ryan as the best, most supportive human. Bog's Dog's forever! I also want to acknowledge my fellow members of the larger Darwinian cohort, Kristina Fialko and Hannah Whitehurst. Kristina and Hannah brought so much joy into my life, helped me improve as a scientist and person, and supported me in so many ways. I appreciate you both so much and cannot thank you enough. Finally, I want to thank Heather Skeen for being incredibly supportive, giving the best advice, and being my venting buddy throughout my time here.

I was also very lucky to have been surrounded by excellent fellow graduate student and postdoc members of the Westneat lab, Charlie McCord, Andrew George, Katie Whitlow, Lily Hughes, Linnea Lungstrom, and Sam Gartner. Charlie welcomed me into the Westneat lab with open arms and made me comfortable since the beginning. She will forever be my academic sister, and I will always appreciate the time she took to collaborate with me on fishy science, show me things to do in Chicago, and our wonderful mini vacation to the east coast. Andrew has become one my closest collaborators and friends throughout our time

in graduate school. He was always the most patient and helpful, showing me the ropes for all things morphometrics and kinematics. Our many discussions broadened my horizons on how I think about and approach science, and our collaborative projects are some of the most exciting projects I worked on during my time at UChicago. I couldn't have asked for a better co-grant writer and field partner for our field work in Moorea, and I look forward to continuing these projects in the future. My friendship with Katie blossomed during the second half of my time in grad school, and I so enjoyed our late-night chats, coffee sessions, and overall sass. Our discussions around mental health really transformed how I felt about a variety of things, and I will always be thankful that I had someone who I felt comfortable talking to who also understood what I was going through. One day we will go diving together! Lily was my rock during the pandemic. Although she started her postdoc right before we were in lock down, I knew I had found a close friendship and mentor in Lily. Lily taught me everything I needed to know genomics and HPC related, and my second chapter would not exist without her. Our many walks and hang outs during the height of the pandemic kept me sane, and I am so grateful for your support and friendship. To Linnea, I consider you to be my little sister, and your growth and as a person as a scientist has made the proudest. Your optimism, excitement, and endless support has kept me afloat and means more to me than I can put into words. You have a genuine curiosity for science, and it was an honor to help you flourish and find your own path. I can't wait to see what you do in the future, and I'm so lucky to be a part of it. Additionally, I really enjoyed working with many excellent undergrads: Kelly Kang, Olivia Grobmyer, and Adera Craig.

One of the major highlights during my time at Chicago was the opportunity to TA the Biodiversity and Genomics course at MBL with Oscar Pineda-Catalan. I feel incredibly fortunate to have had the opportunity to work with Oscar as he took the time to mentor and teach me how to be a more effective instructor over our 4 years working together. Oscar gave me opportunities to contribute the course and make it our own, and I will forever be

grateful for all he taught and showed me. Oscar was also a great friend who supported me throughout the academic year, whether it was chai and biscuits at Plein Air, a good chat, or a delicious meal at his apartment with the wonderful Sonia. Oscar also recruited Ryan to co-TA with me for the 3rd and 4th years, and we truly were the sassy dream trio. I will forever treasure all the memories, students, and experiences we've had throughout the years. I also am grateful for the other MBLers who made this unique experience so meaningful: Elizabeth Lee, Betsy Baldo, Dave Remson, Michael Rossi, Nani Correa, Pancho Bezanilla, Eduardo Perozo, Eric Schwartz, Jean Enright, and Jocelyn Malamy.

Another highlight of my time at UChicago were the two field seasons to Moorea, French Polynesia, to collect kinematic and foraging behavior data from goatfishes. This work would not have been possible without the space and logistical support from Centre de Recherches Insulaires et Observatoire de l'Environnement (CRIOBE). In particular, I want to thank Yannick Chancerelle, Pascal Ung, Elina Burns, Annaïg Le Guen, and Guillaume Iwankow for all of their assistance and support during these field seasons. Additionally, I would like to thank the many Barb's that made this work possible.

Broader institutional support and funding were critical to this dissertation. To Audrey Aronowsky and Carolyn Johnson, thank you for everything. Most of my research was funded by the National Science Foundation through the Fish Life grant awarded to Mark Westneat (DEB 1541547). I am thankful for the research funds provided by the Committee on Evolutionary Biology's Henry Hinds Fund, the Lerner Gray Award through the American Museum of Natural History, the Edward C. Raney Award through the American Society of Ichthyology and Herpetology, the Steiner Award through the University of Chicago, the Graduate Student Research Award through the Society of Systematic Biology, the Buffett Family Foundation, and the France and Chicago Collaborating in the Sciences (FACCTS) research grant from the University of Chicago's France Chicago Center (French Polynesia permit #MCE/DRM 6019) for funding the research described in this thesis.

Additionally, we thank the following people and institutions for providing preserved tissues for genomic sequencing: Mark McGrouther, Amanda Hay, and Joey DiBattista from the Australian Museum (AMS), Richard Pyle and Ken Hayes from the Bishop Museum of Natural History (BPBM), Luiz Rocha from the California Academy of Sciences (CAS), Alistair Graham and Will White from CSIRO, Hiroyuki Motomura from the Kagoshima University Museum (KAUM), Michael Berumen from the King Abdullah University of Science and Technology (KAUST), Andy Bentley and Leo Smith from the University of Kansas (KU), Prosanta Chakrabarty from Louisiana State University Museum of Zoology (LSUMZ), and Amanda Gura, Wouter Holleman, and Gavin Gouws from the South African Institute of Aquatic Biodiversity (SAIAB). Additional tissues were generously donated from personal collections by Arthur Bos and Giacomo Bernardi. We especially thank Caleb McMahan, Susan Mochel, and Kevin Swagel from the Field Museum of Natural History (FMNH) for their assistance in tissue storage, preparation, and shipping. Additionally, we thank Rose Peterson (George Washington University) for assisting with DNA extractions. We also thank Franz Uiblein (Institute of Marine Research, Norway) for taxonomic input during the review process. We used the Midway2 high performance computing cluster operated by the Research Computing Center at the University of Chicago for all genomic analyses used in this research.

Finally, this thesis would not have been possible without the constant support and love from my partner Cameron, my parents, my extended family, and my wonderful friends. I am also thankful for my cat, Mr. Whiskers, for providing endless entertainment and cuddles. I unfortunately lost many family members, including all my grandparents, during my time in graduate school. This dissertation is in honor to them. I hope I've made you proud.

ABSTRACT

Fishes have evolved extraordinary morphological and sensory adaptations to best utilize their natural habitats and expand their ranges. Due to the large variability among habitats and geographic regions, an examination of the mechanisms that influence and regulate feeding mechanisms and sensory systems is essential to understanding the distribution of taxa across space and time. In particular, an examination of the feeding capabilities of diverse clades of fishes will elucidate the role of these functional traits on the range size and distribution of lineages. The overarching goal of this thesis was to explore the phylogenetics, evolution and biogeography of the goatfishes, with a focus on the unique foraging behavior associated with substrate preferences and barbel use. In Chapter 2, I explore the evolution of the goatfishes through inference of the most species rich time-calibrated phylogeny of the goatfishes to date. This phylogeny was inferred using a robust genomic dataset and is used as the basis for all evolutionary analyses used throughout this thesis and reveals novel aspects of goatfish evolution that were previously unknown. In Chapter 3, I examine the global biogeographic distribution and assemblage structure of the goatfishes using bioregionalization approaches. The boundaries between these bioregions are specific to the goatfishes and highlight global patterns of species turnover and the locations of potential barriers to dispersal. In Chapter 4, I test hypotheses about the ecomorphological relationship between head and body shape with preferential foraging on different substrate types across the goatfishes using a robust geometric morphometric dataset. I find evidence that preferential foraging on hard substrate evolved relatively recently, and is associated with changes to forehead shape, elongation ratio, and head length. This dissertation represents a major advance in the integrative examination of the evolutionary relationships, assemblage patterns, and ecomorphological associations across the goatfishes. The results of this thesis provide a comprehensive examination of the many unique characteristics associated with the goatfishes, provides the phylogenetic framework required to fully understand the biology of this family, and enables us to begin

addressing questions about how future changes in reef ecosystem health will impact the biology of this clade.

CHAPTER 1

INTRODUCTION

Coral reef ecosystems are of utmost importance to both the scientific community and the public due to their biodiversity, economic value, and cultural significance. This extraordinary biodiversity was once thought to buffer coral reef ecosystems against the loss or extinction of a few species. However, recent research has found that coral reefs, despite their high biodiversity, may be more vulnerable to species loss than previously thought due to the limited distribution of certain fish species that perform important ecological roles (Mouillot et al. 2014). The importance of understanding the relationships between species distributions and ecological functions has resulted in a novel field of research termed Functional Biogeography. Functional Biogeography, defined as "the analysis of the patterns, causes, and consequences of the geographic distribution of the diversity of form and function" has great potential to predict how species and communities will respond to climate change (Violle et al. 2014).

The Mullidae (goatfishes) are an ideal family of fishes within which to examine the complex associations among morphological adaptations, evolutionary history, and biogeographical patterns. It is a globally distributed family with 104 species found in close association with coral reefs. An innovation in goatfish morphology is a pair of highly specialized throat barbels, which are fleshy extensions capable of taste and prey excavation (Gosline 1984). These barbels are unique in both their morphology and function through their use to detect and resuspend otherwise inaccessible food sources within the ocean floor. This behavior has been shown to play an important role in maintaining coral reef diversity and community composition (Lukoschek and McCormick 2000), in addition to acting as indicators of coral reef health (Uiblein 2007; Russ et al. 2015).

This dissertation focuses on understanding the unique role of goatfishes in marine coral reef communities through exploration of their evolutionary history, biogeography, distinctive cranial and sensory mechanisms, and feeding behavior. This integrative approach enables

a direct examination of the processes that influence rates of species diversification, global patterns of dispersal, and behavioral adaptability. The central aims of this thesis are a) to resolve phylogenetics relationships among the goatfishes using phylogenomic approaches (Chapter 2), b) to examine biogeographic and phylogenetic patterns among global assemblages (Chapter 3), and c) to investigate the ecomorphological relationships between head and body morphology and substrate preference in an evolutionary context (Chapter 4).

1.1 Misfit Fishes: current status of goatfish taxonomy, phylogenetics, and biogeography

First described by Linnaeus in Systema Naturae (1758), the family now contains 104 valid species in 6 genera (Fricke et al. 2023). There have been several genus-level phylogenetic studies of goatfishes that provide the necessary framework to examine the complex genetic and morphological relationships among the many closely related species (Turan 2006; Keskin and Can 2009; Uiblein and Heemstra 2010; Uiblein 2011; Uiblein and Gouws 2015). However, despite their common occurrence on reefs worldwide, the goatfishes lack a comprehensive species-level molecular phylogeny, unlike many other charismatic reef fish families (Cowman 2014).

Their divergent morphology, up until relatively recently, made the goatfishes a phylogenetic enigma within the larger teleost tree of life. The goatfishes are characterized by having a pair of hyoid barbels; two widely separated dorsal fins, the first with seven or eight spines, and the second with nine soft rays; anal fin with a small spine and seven soft rays; deeply forked caudal fin with seven and six branched rays in upper and lower lobes, respectively; and 24 vertebrae (Kim 2002). Based on vertebral and cranial similarities, goatfish have been linked within Perciformes to the Lutjanidae and Sparidae due to shared characters of the absence of crescent expansion on the endopterygoid and the anterior expansion of adductor arcus palatini, though Gosline notes that many such characters could be the result of con-

Figure 1.1: **Six species of goatfish.** Source: FishBase.

vergent adaptation to benthic feeding (Gregory 1933; Gosline 1984, Kim 2002). Goatfish taxonomy has been primarily based on morphological characters, such as dentition, fin ray counts, and body shape (i.e. Uiblein and Heemstra 2011; Uiblein and McGrouther 2012; Uiblein and Gouws 2013; Bos 2014). These morphological traits have been used to broadly examine phylogenetic relationships among genera, however the resolution of the data resulted in several large polytomies (Kim 2002).

Recent large scale phylogenomic studies have placed the Mullidae within the Syngnathiformes, an order that also contains highly morphologically derived families such as the Syngnathidae (sea horses and pipefishes), Aulostomidae (trumpetfishes), Callionymidae (dragonets), and Dactylopteridae (flying gurnards; Betancur-R et al. 2013; Longo et al. 2017; Santaquiteria et al. 2021). Most coral reef fish families are estimated to have originated in the Eocene, with subsequent diversification having occurred through the past 50-60 million years (Cowman and Bellwood 2011; Bannikov 2014; Bellwood et al. 2017). Although

prior research on higher-level phylogenetics of fishes have estimated the origin of goatfishes to have occurred during this time period (i.e. Betancur-R et al. 2013; Near et al. 2013; Rabosky et al. 2018), recent phylogenomic studies that included a sample of goatfishes inferred a much more recent crown node age of ∼18 Ma for the Mullidae (Santaquiteria et al. 2021).

Several species of Mullidae exhibit exceptionally large range sizes; the coast of Eastern Africa to the Hawaiian Archipelago is the largest geographic distance where continued gene flow has been observed (Fernandez-Silva et al. 2013; Lessios and Robertson 2013). Interestingly, there is evidence of isolation and diversification at the extent of these Indo-Pacific ranges. For example, there are several species of endemic goatfish in the Red Sea and the Hawaiian Archipelago, which are hypothesized to be closely related to wide ranging species (Figure 3; Fernandez-Silva et al. 2013; Lessios and Robertson 2013). Range size can have a large impact on the likelihood of extinction and diversification of a lineage, especially within coral reef regions (Lester and Ruttenberg 2005; Leis 2007; Luiz et al. 2013). Goatfish disperse through a pelagic larval stage, in which the larva can remain in the plankton for over 3 weeks (McCormick and Molony 1995; Pavlov et al. 2011). Traits associated with pelagic larval dispersal have been hypothesized play a role in determining the range size of a species, although adult life history traits influence these patterns as well (Lester and Ruttenberg 2005; Mora et al. 2012; Luiz et al. 2013). Additionally, it has been shown that relatively small fluctuations in water temperature can influence the patterns of size and age at settlement of *Upeneus tragula* (McCormick and Molony 1995).

1.2 Digging and Probing: unique foraging behaviors associated with the goatfishes

Goatfishes have evolved a specialized method of hunting when compared to other fish clades, in which they use their hyoid barbels to forage for food on or within the benthic substrate (Gosline 1984). They achieve this through a large variety of behaviors that differ

Figure 1.2: **Schooling behavior of *Mulloidichthys flavolineatus* and *M. vanicolensis* in Moorea, French Polynesia.** Source: C. Nash.

among lineages, as some species use their barbels as an excavation device to turn over soft substrate (i.e. sand and mud) while other species use their barbels to probe within hard crevices in coral rubble to dislodge small animals (Gosline 1984; Platell et al. 1998; Meyer et al. 2000; Lukoschek and McCormick 2001; Krajewski et al. 2006; Galarza et al. 2009; Strübin et al. 2011; Mittelheiser et al. 2022). These behaviors have been shown to play an important role in maintaining coral reef diversity and community composition, as they often act as primary agents of mixed species foraging associations, in addition to acting as indicators of coral reef health (Lukoschek and McCormick 2000; Uiblein 2007). This is due to the disruption of the substrate during foraging that leads to a resuspension of nutrients and otherwise inaccessible detritus into the water column, which can subsequently be taken up by other fish species and nearby coral reef communities (Lukoschek and McCormick 2000; Uiblein 2007). When goatfishes forage, they maintain a horizonal body position above the substrate. This body positioning is likely advantageous for continuously foraging on large areas and allowing for a rapid escape from potential predators (Gosline 1984).

Notably, different species of goatfish are hypothesized to have a strong preference for the substrate type upon which they feed, which results in habitat partitioning (Gosline 1984; Platell et al. 1998; Uiblein 2007). Although goatfishes are well adapted for their unique foraging behaviors, they are considered dietary generalists and feed primarily on infaunal

Figure 1.3: **Variability in goatfish habitat preference.** Images of *Parupeneus multifasciatus*, *P. insularis*, *P. pleurostigma*, and *Mulloidichthys flavolineatus* foraging in Moorea, French Polynesia. Source: C. Nash.

invertebrates (Gosline 1984). When species co-occur, there is evidence of habitat and diet partitioning, most likely due to variation in feeding behavior, depth, spawning time, and ontogeny (Platell et al. 1998; Lukoschek and McCormick 2001). Although there was evidence of habitat and diet partitioning on this small spatial scale, the diets between species were statistically indistinguishable when comparing diet composition across each species range (Platell et al. 1998, ; Lukoschek and McCormick 2001). Species of goatfish also vary in their range sizes, social dynamics, circadian rhythms, and color patterning (Gosline 1984; Strübin et al. 2011; Tosetto et al. 2021).

This dissertation represents the most comprehensive examination of the phylogenetic relationships (Chapter 2), global assemblage and biogeographic patterns (Chapter 3), and ecomorphology associated with foraging on differential substrate (Chapter 4) across species within the Mullidae. The overarching goal of this dissertation is to explore the phylogenetics, evolution, and biogeography of the goatfishes, with a focus on the unique foraging behavior

associated with substrate preferences and barbel use. Prior to this thesis, little was known about their evolutionary history, spatial distribution, morphological diversity, and ecological niches in comparison to other reef associated families, despite their common occurrence in these systems. This thesis represents a major advance in the integrative examination of the evolutionary relationships, assemblage patterns, and ecomorphological associations across the Mullidae. The results of this thesis provide a comprehensive examination of the many unique characteristics associated with the goatfishes, provides the phylogenetic framework required to fully understand the biology of this family, and enables us to begin addressing questions about how future changes in reef ecosystem health will impact the biology of this clade. Expanding this approach to multiple reef fish groups with complete phylogenies is an important future goal and will enable direct comparisons of clade specific assemblage patterns and the role of functional traits on distribution patterns at different spatial scales across the marine realm.

CHAPTER 2

PHYLOGENOMICS AND BODY SHAPE MORPHOMETRICS REVEAL RECENT DIVERSIFICATION IN THE GOATFISHES

2.1 Abstract

Clades of marine fishes exhibit many patterns of diversification, ranging from relatively constant throughout time to rapid changes in the rates of speciation and extinction. The goatfishes (Syngnatharia: Mullidae) are a family of marine, reef associated fishes with a relatively recent origin, distributed globally in tropical and temperate waters. Despite their abundance and economic importance, the goatfishes remain one of the few coral reef families for which the species level relationships have not been examined using genomic techniques. Here we use phylogenomic analysis of ultra-conserved elements (UCE) and exon data to resolve a well-supported, time-calibrated phylogeny for 72 species of goatfishes, supporting a recent crown age of the goatfishes at 21.9 million years ago. We used this framework to test hypotheses about the associations among body shape morphometrics, taxonomy, and phylogeny, as well as to explore relative diversification rates across the phylogeny. Body shape was strongly associated with generic-level taxonomy of goatfishes, with morphometric analyses showing evidence for high phylogenetic signal across all morphotypes. Rates of diversification in this clade reveal a recent sharp increase in lineage accumulation, with 92% of the goatfish species sampled across all clades and major body plans having originated in just the past 5 million years. We suggest that habitat diversity in the early Pliocene oceans and the generalist ecology of goatfishes are key factors in the unusual evolutionary tempo of the family Mullidae[1].

1. This chapter is published under the following citation: Nash, C. M., Lungstrom, L. L., Hughes, L. C., & Westneat, M. W. (2022). Phylogenomics and body shape morphometrics reveal recent diversification in the goatfishes (Syngnatharia: Mullidae). Molecular Phylogenetics and Evolution, 177, 107616.

2.2 Introduction

Shifts in the tempo and mode of evolution across lineages are a major force in generating global biodiversity (Simpson 1945; Gould and Eldredge 1977). Modern approaches have yielded novel evolutionary insights in many regions of the Tree of Life through the detection of variability in patterns of species diversification across large phylogenetic scales (Pennell et al. 2012; Stadler 2013; Donoghue and Sanderson 2015; Graham et al. 2018; Rabosky et al. 2018). This variation in the timing and rates of speciation and extinction is often the result of habitat fragmentation and dispersal capabilities, which can alter the amount of gene flow between populations (Kisel et al. 2011). As the largest group of living vertebrates, fishes have been a useful system to explore the impact of variable rates of diversification on current patterns of diversity (i.e. Alfaro et al. 2009; McMahan et al. 2013; Rabosky et al. 2013; Santini et al. 2013; Melo et al. 2021).

Heightened diversification rates in temperate fishes have been linked to particular geologic or climatic events (Rabosky et al. 2018), and likewise, bursts of evolution in reef fishes are linked to temporal patterns of expansion of coral reefs (Leprieur et al. 2021). The tempo and mode of diversification in fishes varies widely across taxonomic groups. Rates of diversification may be elevated towards the crown, as in Caribbean hamlets (Hench et al. 2022); towards the root, as in the case of neotropical cichlids (López-Fernández et al. 2013); or a relatively constant rate through the phylogenetic history of a group, such as in the damselfishes (McCord et al. 2021). Our ability to infer variable rates of lineage accumulation depends on our confidence in the underlying phylogeny and estimation of divergence times. Fortunately, genome-wide datasets that can greatly improve the resolution of phylogenetic relationships are now attainable for non-model organisms.

The goatfishes (Syngnatharia: Mullidae) are a diverse, globally distributed family of fishes found in close association with coral reef ecosystems. First described by Linnaeus in Systema Naturae (1758), the family now contains 104 valid species in 6 genera (Fricke et al.

2022). There have been several genus-level phylogenetic studies of goatfishes that provide the necessary framework to examine the complex genetic and morphological relationships among the many closely related species (Turan 2006; Keskin and Can 2009; Uiblein and Heemstra 2010; Uiblein 2011; Uiblein and Gouws 2015). However, despite their common occurrence on reefs worldwide, the goatfishes lack a comprehensive species-level molecular phylogeny, unlike many other charismatic reef fish families.

Most coral reef fish families are estimated to have originated in the Eocene, with subsequent diversification having occurred through the past 50-60 million years (Cowman and Bellwood 2011; Bannikov 2014; Bellwood et al. 2017). Although prior research on higher-level phylogenetics of fishes have estimated the origin of goatfishes to have occurred during this time period (i.e. Betancur-R et al. 2013; Near et al. 2013; Rabosky et al. 2018), recent phylogenomic studies that included a sample of goatfishes inferred a much more recent crown node age of ∼18 Ma for the Mullidae (Santaquiteria et al. 2021). This raises intriguing questions about the timing of goatfish origins and the tempo of diversification that require a well-sampled time-calibrated phylogeny of the family to answer.

Goatfish taxonomy has been primarily based on morphological characters, such as dentition, fin ray counts, and body shape (i.e. Uiblein and Heemstra 2011; Uiblein and McGrouther 2012; Uiblein and Gouws 2013; Bos 2014). These morphological traits have been used to broadly examine phylogenetic relationships among genera, however the resolution of the data resulted in several large polytomies (Kim 2002). The main synapomorphy that unites goatfishes is a pair of highly specialized hyoid barbels, which are fleshy extensions capable of chemoreception and prey excavation (Gosline 1984). Goatfishes use these barbels to detect and resuspend otherwise inaccessible food sources within the benthic substrate (Gosline 1984). This unique behavior has been shown to play an important role in maintaining coral reef diversity and community composition (Lukoschek and McCormick 2000), in addition to acting as indicators of coral reef health (Uiblein 2007; Russ et al. 2015).

These traits may provide the necessary framework to examine the complex genetic and morphological relationships among many closely related species and reiterate the need for a comprehensive phylogeny to improve our understanding of goatfish evolution. The integration of phylogenomics with geometric morphometrics will allow us to examine morphological evolution associated with lineage diversification across the goatfishes.

The central aims of this study are to 1) infer the phylogenetic relationships among species, 2) test hypotheses about the congruence between body shape morphology, taxonomy, and phylogenetic placement, and 3) examine the tempo of phylogenetic and morphological diversification across species of goatfishes. To achieve the primary aim, we use a robust genomic dataset to infer the first molecular phylogeny of goatfishes using combination of ultra-conserved elements (UCEs), exons, and ribosomal sequences in order to include the most species possible from throughout the family. UCEs and exon capture have been shown to be effective in resolving evolutionary relationships at different degrees of divergence, although genome coverage inversely decreases with the amount of divergence among species (Faircloth et al. 2012; Bragg et al. 2016). Additionally, we examine the evolution of body shape using a comprehensive geometric morphometric dataset to assess the phylogenetic affinities of goatfish morphology.

2.3 Methods

2.3.1 Taxon sampling

We developed a comprehensive data matrix of 72 species of Mullidae, representing all genera in the family. We include 38 published UCE assemblies for Mullidae from two recent studies (Longo et al. 2017; Santaquiteria et al. 2021), as well as the full genome of *Mullus surmuletus* (Fietz et al. 2020). We expanded this taxon sampling by sequencing an additional 22 species of Mullidae (Table 2.S1) using both acanthomorph UCE (Alfaro et al.

2018) and sygnatharian-specific exon capture sequencing probes (Hughes et al. 2021). To further increase taxon sampling, we compiled a matrix of 7 mitochondrial loci (*12S, 16S, ATP6, COI, COII, cytB*, and *nd2*) and one nuclear locus (*rhod*) to incorporate an additional 11 species from previously published datasets available on NCBI (Table 2.S2). Sequences for an additional 14 species were additionally from Longo et al. as representative outgroup lineages within Syngnatharia. Species identification was confirmed for all samples through comparison of COI with sequences available in the BOLD database. One specimen previously identified as *Upeneus guttatus* was updated to *Upeneus willwhite* given the recent reclassification of the sampled specimen (Uiblein and Motomura 2021). Additionally, we updated the name of a specimen initially identified as *Upeneus taeniopterus* to *Upeneus suahelicus* based on taxonomic input by Franz Uiblein (Uiblein and Gouws 2015, Table 2.S1). However, given its phylogenetic distinctivness from the other *U. suahelicus* included in this study, we kept its classification as a distinct taxonomic unit.

2.3.2 Library preparation and sequencing

Genomic DNA from 22 goatfish samples was extracted from preserved tissues in a 96-well plate format on a GenePrep (Autogen) at the Smithsonian Institution Laboratory of Analytic Biology (Washington, DC), following manufacturer's instructions. Extraction quality was assessed via visual inspections of genomic DNA on a 1.5% agarose gel. Illumina library preparation and sequence capture of both UCEs (Acanthomorph 1Kv1; Alfaro et al. 2018) and single-copy exons (Hughes et al. 2021) were conducted at Arbor Biosciences (Ann Arbor, MI). Libraries were sequenced on one lane of an Illumina HiSeq 4000 at the University of Chicago Genomics Facility as 100-bp paired-end reads. The sequence reads generated here have been deposited on NCBI's SRA database under BioProject PRJNA824087.

2.3.3 UCE data assembly and bioinformatic processing

We assembled the genomic data using the phyluce 1.6.8 pipeline (Faircloth 2016). Read data was trimmed and cleaned using the default settings for illumiprocessor 2.0.9 (Faircloth 2013), which uses trimmomatic 0.39-1 (Bolger et al. 2014), to remove adapter contamination and low-quality bases. After adaptor quality and trimming of the 75 raw sequences, an average of 21,641 contigs were assembled in Trinity 2.8.5 (Grabherr et al. 2011) and were matched to the UCE Acanthomorph probe set (Faircloth et al. 2013). We recovered a total of 1,188 UCE loci with an average of 543 loci per sample. Loci were concatenated using mafft v7 (Katoh et al. 2002; Katoh and Standley 2013), and we performed edge trimming on the resulting alignments. Final filtering and concatenation of UCE loci generated matrices of 1031 (60%), 979 (70%), 763 (80%), and 198 (90%) loci. The matrix with $\geq$ 90% proportion of taxa was selected for use in downstream analyses as support did not decrease with the reduction of included loci. It recovered UCE contigs with a mean length of 532 bp and 122 informative sites.

We used the sliding-window approach and entropy site characteristic (SWSC-EN), a partitioning scheme proposed for UCEs, to determine the best fit scheme for the flanks and cores across each locus in the 90% complete matrix in PartitionFinder 2 (Tagliacollo and Lanfear 2018). We merged the right and left flanks into a single "flank" partition and kept the "core" partition independent. Using ModelFinder implemented in IQ-TREE 1.6.12, we determined the best-fit partitioning scheme (Chernomor et al. 2016; Kalyaanamoorthy et al. 2017; Minh et al. 2020). We used this partitioning scheme to estimate the concatenation-based maximum likelihood (ML) trees in IQ-TREE using 1000 Ultrafast Bootstrap replicates to asses branch support (Hoang et al. 2018). We also conducted coalescent-based species-tree analyses in ASTRAL 5.7.3 using the UCE gene trees as input (Sayyari and Mirarab 2016; Zhang et al. 2018; Rabiee et al. 2019). Gene trees were inferred using the UCE "flank" and "core" partitioning scheme. Given the minimal differences in support between the 70% and

90% complete UCE loci matrices, we used the 90% complete matrix in downstream analyses to reduce computational time (Figure 2.A.1).

2.3.4 Incorporating public sequence data

We included data from eight additional loci, downloaded from either GenBank or assembled from raw sequencing reads, in order to increase taxonomic sampling of the group by including species that did not have genomic data available (Table 2.S1). To extract mitochondrial genes and the nuclear gene *rhod*, we used the pipeline described in Hughes et al. (2021). These scripts were additionally modified to include reference sequences from GenBank for the non-coding loci, *12S* (FJ008141.1, EF095566.1, FJ008155.1, LC036890.1) and *16S* (FN688081.1, OM470924.1, EU848456.1). All sequences were aligned using mafft and visually inspected in AliView (Larsson 2014). These resulting alignments were concatenated with the 90% complete UCE alignments into a single supermatrix. We used of partitioning scheme consisting of 22 separate partitions to test differences among the rates of sequence evolution across loci. For protein-coding loci (*ATP6, COI, COII, cytB, rhod*, and *nd2*), we assigned each codon position for each gene as a potential partition. For the non-coding loci *12S* and *16S*, each locus was considered designated a single partition. To infer the appropriate placement of species where only data from GenBank was available, we used the UCE only ML topology generated previously as a constraint in order to reduce to potential attraction of lineages with missing data. We reran the ML analysis and searched for the best partitioning scheme in IQTREE using the concatenated UCE and exon sequence matrix.

2.3.5 Divergence time and rate estimation

We estimated divergence times of the Mullidae using BEAST 2.6.2 (Drummond et al. 2002; Bouckaert et al. 2014) with multiple calibration points located across the Syngnatharia

phylogeny. These included a secondary calibration and six fossil calibrations utilized in Santaquiteria et al. (2021), in addition to two fossils that provide higher resolution within our clade of interest (Appendix A). This includes the recently described crown callionymoid fossil species †*Gilmourella minuta*, which is estimated to be ∼49 Ma from Monte Bolca (Carnevale and Bannikov 2019). We also included a fossil placed within the modern genus *Mullus*, which is estimated to ∼13 Ma (Carnevale et al. 2006). Given our interest in the timing of biogeographic events in future studies, we did not include any geographical/geminate calibrations within this study. To reduce analysis runtime, we used a restricted subset of the outgroup topology by including only the most divergent pairs of lineages for each calibration node. The only secondary calibration used in this study was located at the root of Syngnatharia, based on ages obtained from other phylogenetic studies (Santaquiteria et al. 2021). Following Santaquiteria et al, we used a uniform distribution with a hard lower bound at 92 Ma and a hard upper bound at 103.5 Ma. All other calibration points used a log normal distribution, and their associated priors are defined in Appendix A.

To estimate divergence times, we used a relaxed clock of log normal distributed rates in BEAST 2.6.2 (Drummond et al. 2002; Bouckaert et al. 2014). A standard birth-death rate model was used as a tree prior. We partitioned the concatenated matrix containing UCE and GenBank data into eleven separate partitions: UCE flank, UCE core, and nine additional partitions based on the best fit partitioning scheme inferred from the initial ModelFinder search (Chernomor et al. 2016). Each partition was given their own unlinked substitution model parameters (Supplemental Material). We used the topology inferred from the ML analyses as a starting topology and we constrained the monophyly of each genus within Mullidae to reduce the erroneous placement of lineages with limited data. Tracer 1.7.1 was used to assess the convergence of two independent analyses of 300 million generations each and confirm proper mixing of all model parameters used in the final Markov Chain Monte Carlo (Rambaut et al. 2018). We used LogCombiner 2.6.3 to both combine these runs

and generate a manageable tree set using a 90% burn-in. TreeAnnotator 2.6.3 was used to calculate the maximum clade credibility tree (MCCT) from the tree set, as well as to estimate posterior probabilities at each node. This time calibrated phylogeny was used in subsequent diversification analyses.

2.3.6 Geometric morphometrics

We performed a morphometric analysis to quantify the variation in body shape across species within the Mullidae. We collected 2D lateral specimen photographs from museum collections, personal databases, primary literature, and aquarium trade, resulting in a total of 253 individuals across 6 genera within Mullidae (Appendix C). A total of 26 landmarks were placed along the specimen image using the R package StereoMorph 1.6.4 (Olsen and Westneat 2015; Figure 2.1) To adjust several specimens with preservation artifacts, such as jaws in the open position, we digitally and interactively rotated the jaw tip landmarks around a reconstructed jaw joint position to a closed orientation, using custom trigonometric code, while maintaining a fixed distance, base position and maintaining all other coordinates (Appendix B).

Using the protocol of George and Westneat (2018) as a general framework, we projected each digitized landmark into linear tangent space using the `gpagen` function in geomorph 4.0.2 (Adams and Otárola-Castillo 2013; Adams et al. 2013, 2021; Baken et al. 2021). The generalized Procrustes analysis (GPA) removes the variation in landmark position that is attributable to rotation, translation, and scaling, while preserving the relevant shape information (Gower 1975; Rohlf and Slice 1990). We then calculated species means of the Procrustes-transformed landmark coordinates to account for intraspecific variation using the `mshape` function in geomorph 4.0.1 (Adams et al. 2021; Baken et al. 2021). This protocol was iterated over all species-grouped landmark sets producing 70 shapes representing the mean body shape of each individual species. From these 70 mean shapes, another GPA was

performed and the resulting Procrustes aligned shape data was projected into linear tangent space using the `gpagen` function in geomorph 4.0.1 (Adams et al. 2021; Baken et al. 2021). Principal components analyses (PCA) were performed on the species averaged, Procrustes aligned landmark coordinates using the `gm.prcomp` function in geomorph 4.0.1 (Adams et al. 2021; Baken et al. 2021). We used back transformations to visualize the shape changes associated with the most significant axes of variation (PC1 and PC2), which plots the shape variation in its relative position in the morphospace (Olsen 2017).

We projected the time-calibrated phylogeny onto the morphospace using the `phylomorphospace` function in the R package phytools 0.7-90 (Revell 2012). Additionally, we generated a phylogenetic PCA using GLS-centering and projection using the implementation in `gm.prcomp` function in geomorph 4.0.1 (Revell 2009; Adams and Otárola-Castillo 2013). To examine the relationship between body shape and genera, we performed a MANOVA using the `procD.lm` function in geomorph 4.0.1 (Adams and Otárola-Castillo 2013). We used the broken-stick method to inform how many principal components (PC) axes to retain for subsequent analyses (Barton and David 1959; Frontier 1976).

The partial disparity for genera was calculated using the overall mean over 1000 iterations with the function `morphol.disparity` in geomorph, taking care to correct the denominator in the variance calculation from n to N-1, where n is the group size and N is the number of observations (Adams et al. 2021; Baken et al. 2021). Procrustes partial variances were reported for each genus and these values were used to calculate the proportion of the total disparity accounted for by each genus (Table 2.S3). We calculated the average disparity through time (DTT) and the mean disparity index (MDI) using the `dtt` function in geiger 2.0.7 (Harmon et al. 2008). Significance was determined using 10,000 null disparity through time simulations generated using the assumptions of Brownian Motion (BM). Additionally, we computed the weight average times, commonly referred to as the Center of Gravity, to examine trends within the distribution of disparity across subclades and time (Slater 2022).

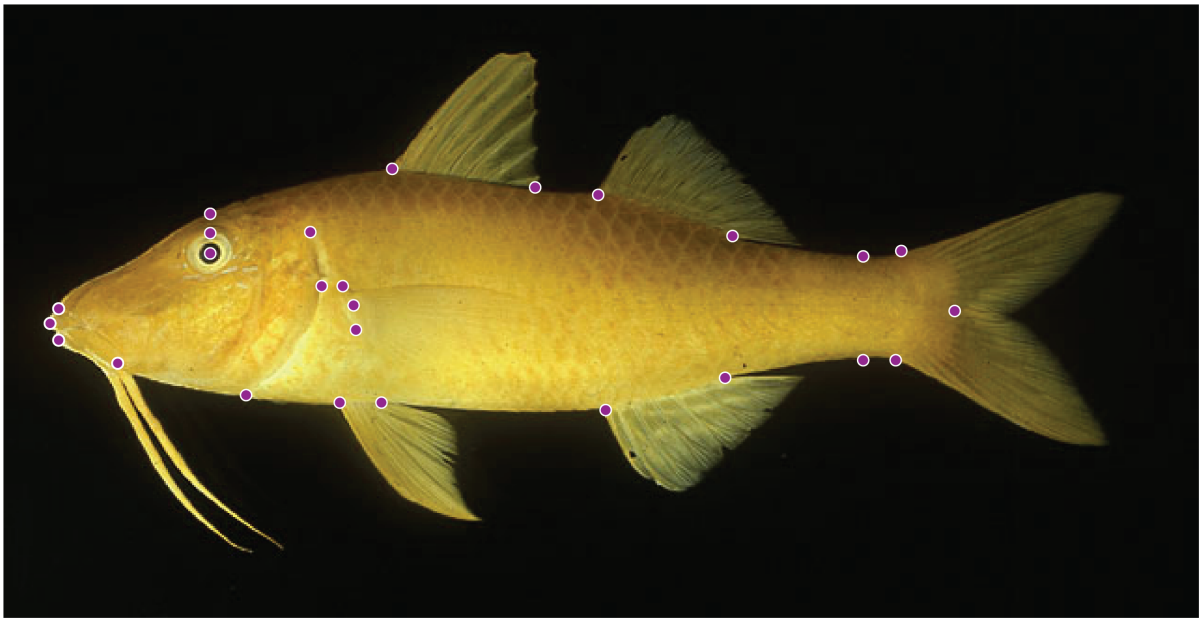

Figure 2.1: **Morphometric protocol for the goatfishes.** Full body landmark scheme for the analysis of the geometric morphometrics of goatfishes. The placement location of 26 landmarks with anatomical descriptions are listed in Appendix B.

2.3.7 Rates and pattern of lineage diversification

To test hypotheses about the relative rate of diversification across the phylogeny, we first tested deviation from a constant-rate pure-birth diversification process using the γ-statistic (Pybus and Harvey 2000). We controlled for incomplete taxon sampling using the Monte Carlo constant rates (MCCR) test with 10,000 simulations and a sampling fraction of 74.2% (72 out of 102 species), implemented using the `mccr` function in phytools 0.7-90 (Revell 2012). We compared the observed lineage through time (LTT) plot and γ-statistic to the LTT and γ-statistics of 10,000 simulated phylogenies generated under a pure-birth process using the `pbtree` function in phytools 0.7-90 to test for significance (Revell 2012).

2.4 Results

The central product of this research is a well-resolved, time calibrated phylogenetic hypothesis for the Mullidae. The phylogeny is largely congruent with previous taxonomy,

and reveals that the crown age of the family is estimated at 21.9 Ma, providing a framework for comparative analysis. Morphometric analysis of body shape strongly supports the current taxonomy and has high phylogenetic signal. Our results reveal that goatfish evolutionary history is characterized by unstable rates of lineage diversification and body shape variation, with a recent increase in lineage accumulation across all genera in just the past 5 million years.

2.4.1 Phylogenomic analyses and divergence time estimation

The phylogeny of the goatfishes (Syngnatharia: Mullidae) is well resolved and highly supported (Figure 2.2 & 2.3). We find strong support for a monophyletic group with the six described genera: *Mulloidichthys, Mullus, Pseudupeneus, Parupeneus, Upeneichthys* and *Upeneus*. The topologies derived from the ML, ASTRAL, and time calibrated BEAST analyses were well supported and species relationships, for the most part, were concordant. The Mullidae are sister to the family Callionymidae, which combined represents the suborder Callionymodei within the order Syngnatharia (Figure 2.2, Betancur et al. 2017). The divergence between the Mullidae and Callionymidae is estimated to have occurred around 58.2 Ma (48.6 – 74.4 Ma), which resulted in a relatively long branch separating these two families (Figure 2.2).

Within the Mullidae, there are two major clades, which split at the root approximately 21.9 Ma (17.0 – 27.7 Ma; Figure 2.3). The first clade is comprised of the paraphyletic genus *Upeneus*, which is subdivided into two subclades. The second major clade within the Mullidae is comprised of five genera: *Mulloidichthys, Upeneichthys, Mullus, Pseudupeneus*, and *Parupeneus*. All genera are monophyletic and relationships are well resolved within this clade. There are two main generic sister pairs in this clade, *Mullus* + *Upeneichthys* and *Pseudupeneus* + *Parupeneus*. Combined, these genera form monophyletic group with the genus *Mulloidichthys* (Figure 2.3).

Upeneus is the most species rich genus with 25 species. There are two subclades within the *Upeneus* that diverged approximately 18.3 Ma, denoted as *Upeneus 1* and *Upeneus 2* (Figure 2.3). The only taxonomic anomaly in this clade is the inclusion of *Mulloidichthys pfleugeri*, which is most closely related to the newly described species goatfish species, *Upeneus nigromarginatus* (Bos 2014). A comparison of *COI* from *M. pfleugeri* across the BOLD database indicates a correct species ID for this sample, and this node has a 100% UF bootstrap support and 1.0 Posterior probability in the ML and ASTRAL analyses, respectively (Figure 2.4). Within *Upeneus*, there are four identifiable species group, which we define as a clade of three or more species that diverged within the Pleistocene (Figure 2.A.2). There is one species group in *Upeneus 1*, which is composed of *U. quadrilineatus, U.suahelicus (taeniopterus), U. supravittatus*, and *U. suahelicus*. The remaining three species groups are found in *Upeneus 2*. Group 2.1 is composed of *U. margarethae, U. heterospinus, U. spottocaudalis*, and *U. caudofasciatus*. Group 2.2 is composed of *U. tragula, U. luzonius, U. oligospilus*, and *U. heemstra*. Group 2.3 is composed of *U. asymmetricus, U. lombok, U. seychellensis, U. itoui, U. willwhite, U. pori*, and *U. floros*. (Figure 2.A.2).

Parupeneus is the second most species rich clade with 22 species, followed by *Mulloidichthys* (6 sp.), *Mullus* (4 sp.), *Pseudupeneus* (3 sp.), and *Upeneichthys* (3 sp.). The root age of *Parupeneus* was estimated to be ∼7.1 Ma and the estimated divergence time from its sister genus, *Pseudupeneus*, was ∼12.9 Ma (Figure 2.3). There is evidence of recent divergence across *Parupeneus*, with 16 species having diverged within the Pleistocene. However, only two clades met the criteria of species group as the majority of these speciation events were among species pair. The first group is composed of *P. insularis, P. crassilbaris, P. macronemus, P. multifasciatus*, and *P. trifasciatus*. The second group is comprised of *P. spilurus, P. rubescens*, and *P. poryphyreus*.

The mean crown ages of the remaining genera were estimated to be between 6.3 Ma and 3.1 Ma (Figure 2.3). When considering *Mulloidichthys, Mullus, Pseudupeneus*, and

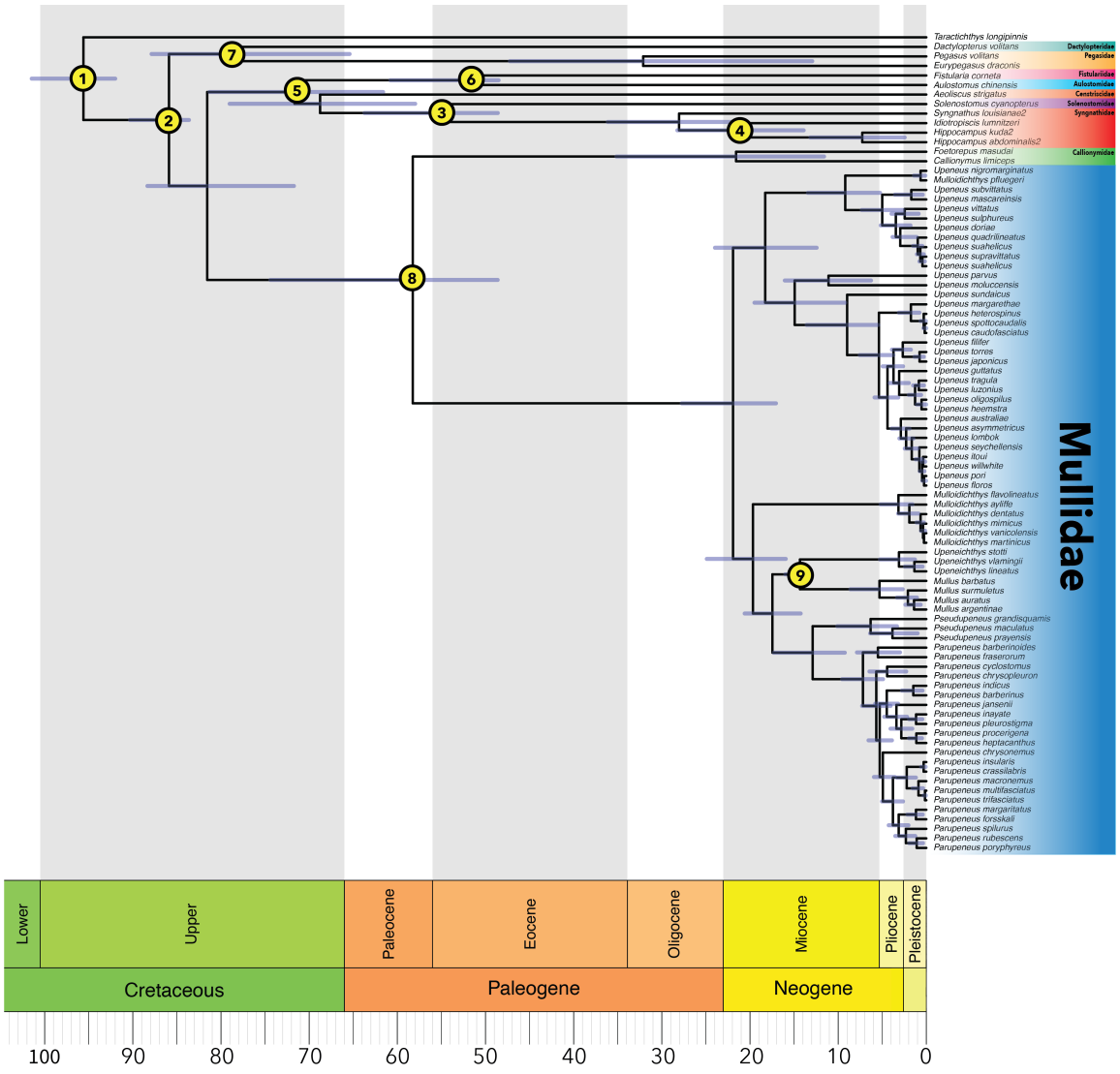

Figure 2.2: **Time-calibrated phylogenetic tree for goatfishes (Family Mullidae) with outgroups.** Divergence time estimation for the Mullidae in the context of the outgroups from the order Syngnatharia based on relaxed clock node calibration in BEAST 2. The placement of nine fossil calibrations are shown, and additional information for each calibration can be found in Appendix 1. Blue error bars at each node represent the high posterior density (HPD) of each node age estimate.

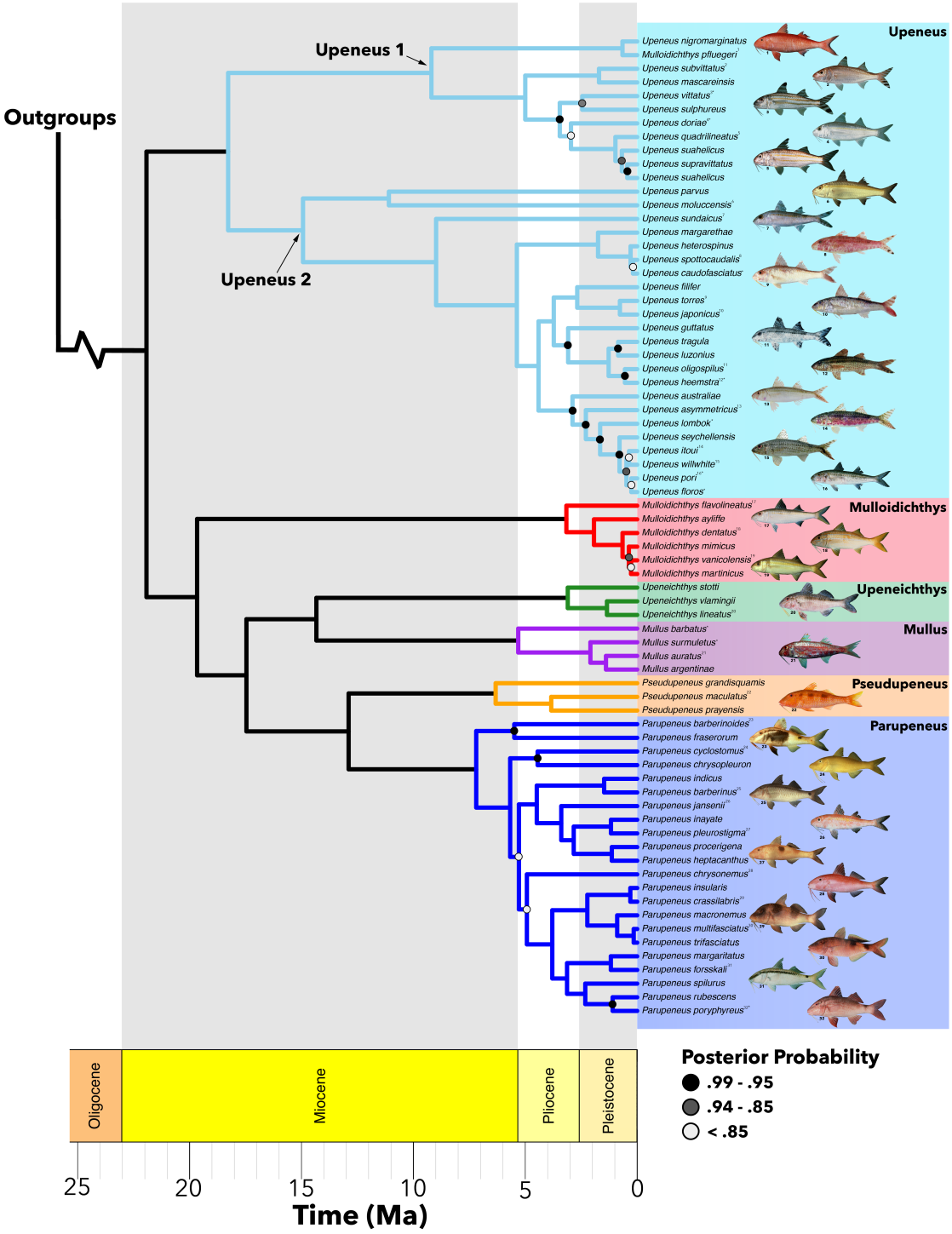

Figure 2.3: **Time-calibrated phylogenetic tree for the Mullidae.** Divergence time estimates are shown for clades within the Mullidae. Subgenera within *Upeneus* are indicated at each respective node. Nodes without a circle have posterior probability of 1.0. Tips indicated with an asterisk were analyzed with multi-locus data from GenBank only. (Image credits: John E. Randall via FishBase or referenced by number in Appendix 3).

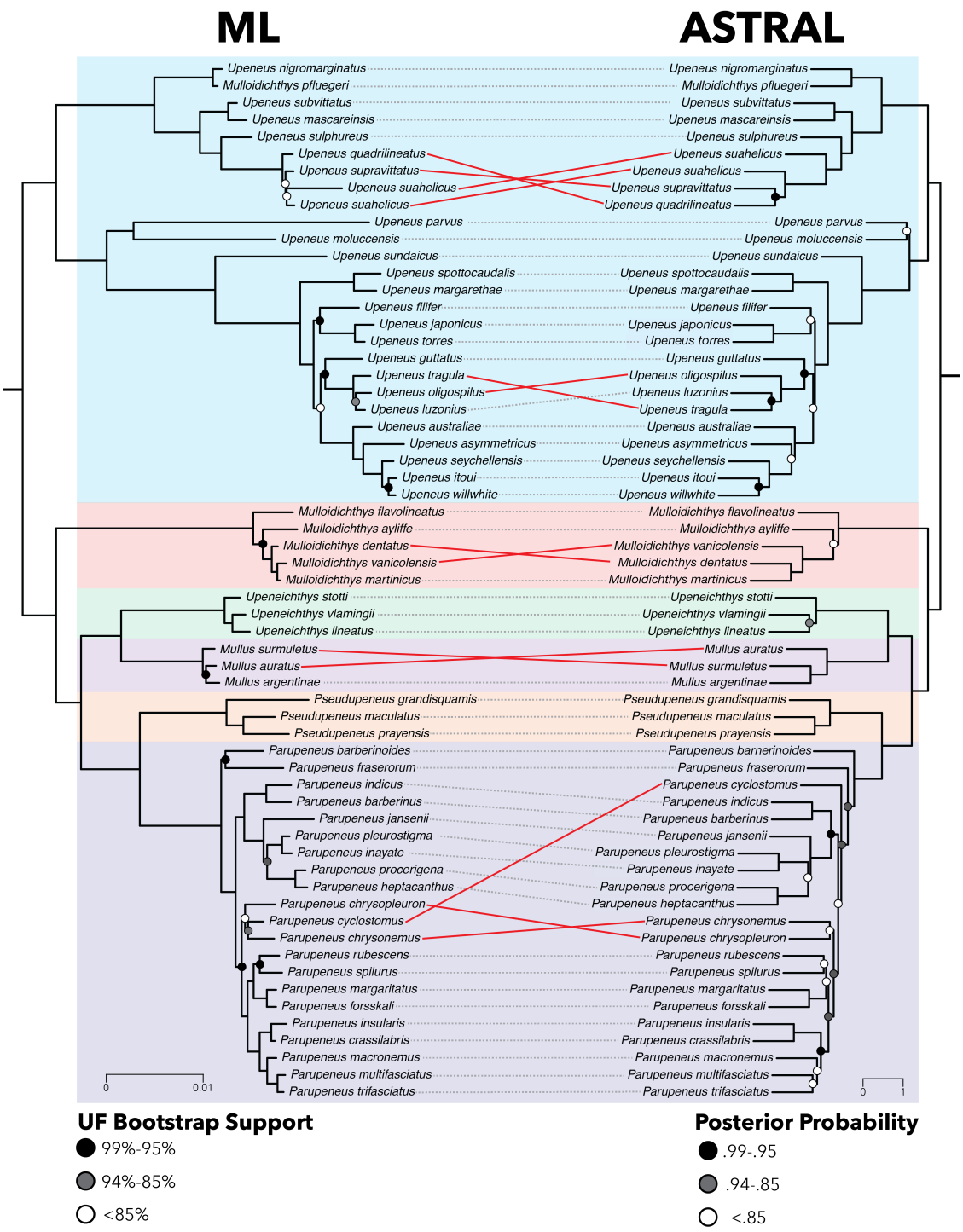

Figure 2.4: **Phylogenetic tree comparisons for the Mullidae.** Congruence between the ML tree (left) and the tree inferred using ASTRAL III (right) using the 90% complete UCE matrix. The topologies are largely congruent, with exceptions indicated by a red line between corresponding tips. Nodes without a circle have either 100% Ultrafast Bootstrap support (ML) or a local posterior probability of 1.0 (ASTRAL).

Upeneichthys, the mean difference between root age and stem age is estimated to be ∼9.4 Ma. *Mulloidichthys* has the maximum distance between the root age with a value of ∼9.2 Ma and *Pseudupeneus* has the minimum distance with ∼6.6 Ma between the root age and its divergence from *Parupeneus* (Figure 2.3). Similar to the diversification patterns observed in *Upeneus* and *Paruepenus*, the majority of divergence events among species occurred within the middle Pliocene into the Pleistocene. The majority of species within *Mulloidichthys*, with the exception of *M. flavolineatus*, diverged within the Pleistocene and can be classified as a species group. There is a similar pattern of diversification within the majority of species within *Mullus*, consisting of *Mu. surmuletus, Mu. auratus*, and *Mu. argentinae* represent a species group, with the exception of *Mu. barbatus*.

2.4.2 Lineages through time

Results of diversification analyses strongly support the hypothesis of unstable rates of diversification throughout goatfish evolution, with a notable inflection point to a sharp increase in lineage accumulation at about 5 Ma (Figure 2.5A). This pattern of an increase in lineage accumulation is observed within all genera, but is most evident in the most species rich clades, *Upeneus* and *Parupeneus* (Figure 2.3). The lineage through time (LTT) plot (Figure 2.5A) shows an exponential growth curve with a significant gamma value ($\gamma = 2.4659$), which strongly rejects the constant speciation hypothesis under a Yule pure-birth model ($p = 0.004$). This result indicates that there is a variable rate of diversification across the Mullidae phylogeny, which is evident by a reduction in lineage accumulation during the Middle to Late Miocene prior to the striking acceleration of lineage accumulation during the Miocene-Pliocene transition (Figure 2.5A). Additionally, over 50% of the mean node ages are estimated as less than 5 Ma (Figure 2.5B), with 92% of total species (66 out of 72 species) divergent from their most recent ancestor within the past 5 Ma.

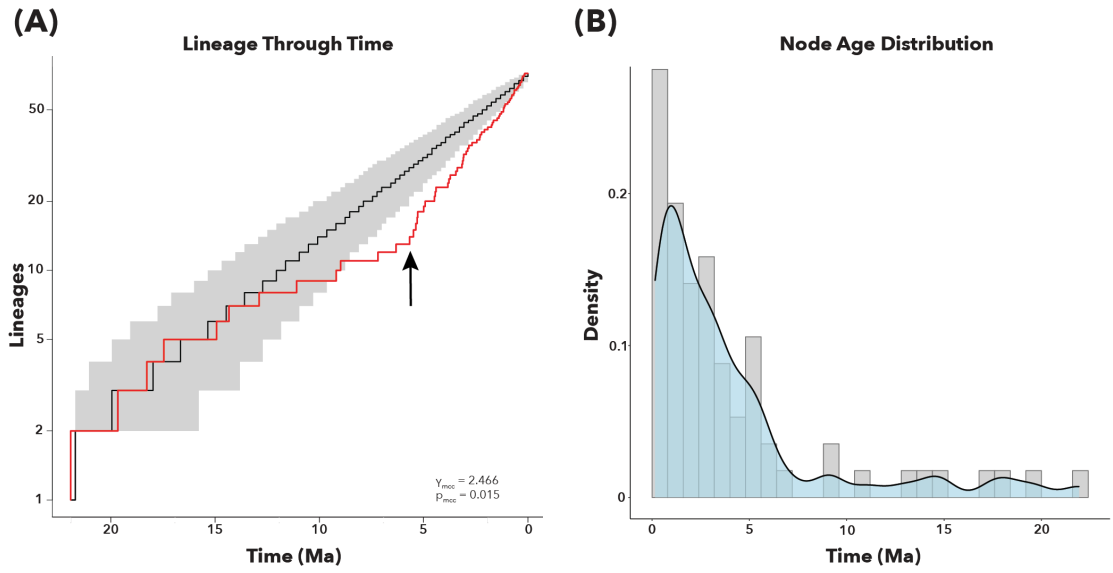

Figure 2.5: **Goatfish diversification patterns through time.** **(A)** Trends in lineage accumulation through time. The red curve indicates the data from this study, the black curved indicates the mean of the simulated phylogenies, and the grey area represents the 95% confidence interval of the simulations. Note the inflection point which indicates the increase in lineage accumulation at approximately 5 Ma. **(B)** Density of nodes with ages at specific time intervals.

2.4.3 Body shape morphometrics

Body shapes cluster based on genera and have high phylogenetic signal. The primary axis of variation (PC1: 45.26%) for this dataset describes the relative ratio of length to width, termed the elongation ratio (Claverie and Wainwright 2014; Friedman et al. 2019; Figure 2.6). Higher PC1 scores are associated with a lower elongation ratio, which is associated with variability on height in the across the body and to a relatively shorter distance between the tip of the mouth and edge of the caudal peduncle. This is especially evident when comparing the height and location of the base of the first dorsal fin to the height of the caudal peduncle, for example between *Parupeneus* and *Upeneus* (Figure 2.6). Lower PC1 scores are associated with a higher elongation ratio, which corresponds to species with longer, more evenly slender body shapes. The second axis of variation (Figure 2.6; PC2: 14.11%) is associated with eye size, the relative distance between the bases of the second

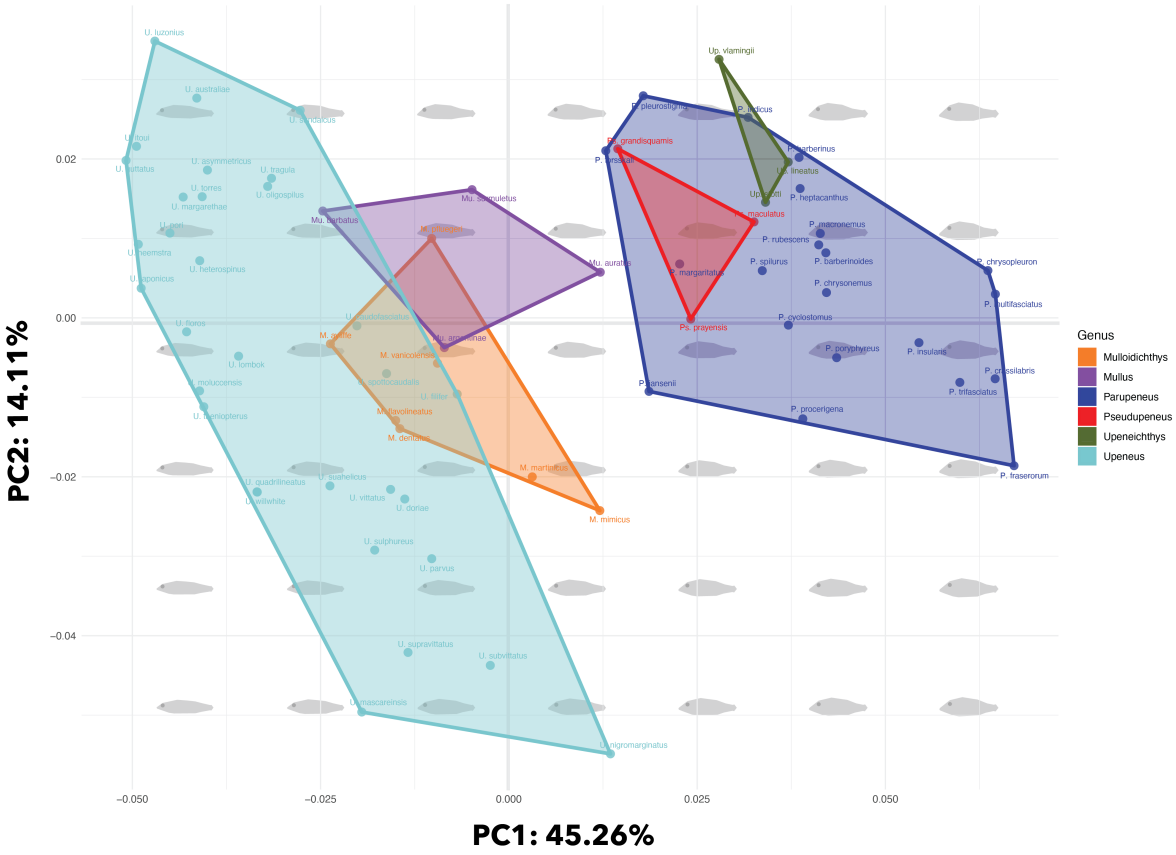

Figure 2.6: **Morphospace for the body shape of the goatfishes (Family Mullidae).** Axis 1 and 2 from Principal Component Analysis (PCA) on Procrustes aligned body shape landmarks for 70 goatfish species. Each coordinate represents the mean shape of each species. PC1 represents 45.26% of the total variation, and PC2 represents 14.11% of the total variation. Species are grouped by genus, and this is indicated by color coded hulls to show the distribution of each genus in morphospace. Gray polygons show the back transformation, which represents the hypothetical goatfish body shape at intervals along each axis.

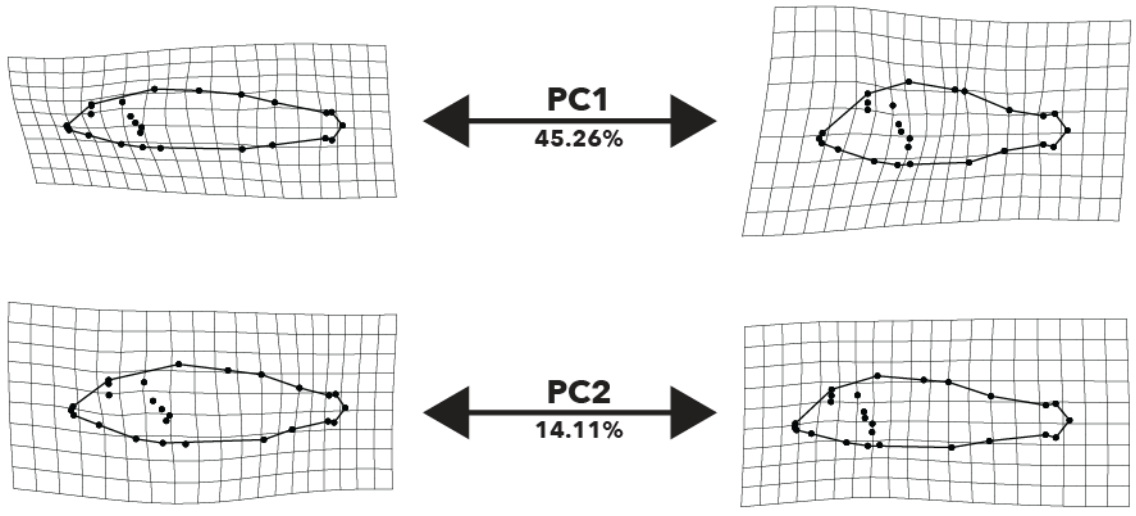

Figure 2.7: **Thin-plate spine warp grids for PC1 and PC2 of goatfish body shape.** These plots show the relative deformation associated with the extreme ends of PC1 and PC2 compared to the overall mean shape

dorsal fin and the caudal peduncle, and body depth in relation to the midline (Figure 2.7). Principal components analysis shows that 90% of the variance in body shape is summarized within PC1 through PC10.

Genera are significantly associated with body shape and segregate primarily along PC1 into two main groups. Group 1 is composed of *Upeneus, Mulloidichthys*, and *Mullus*, and Group 2 consists of *Parupeneus, Pseudupeneus*, and *Upeneichthys*. Group 1 is clustered on lower on the axis of PC1 and Group 2 is clustered on higher values of PC1. Of note, *Pseudupeneus* is fully encapsulated within the *Parupeneus* morphospace. Although genera do not cluster across PC2, there are differences among the ranges of values in which each genus occupies. Species within *Upeneus* span the entirety of PC2, which ranges highest to lowest represented by *U. luzonius* to *U. mascareinsis*, respectively (Figure 2.6). PC2 does appear to separate the two subclades within *Upeneus*, with *Upeneus 1* occupying higher values of PC2 and *Upeneus 2* occupying lower values of PC2 (Figure 2.A.4). However, these groups are not

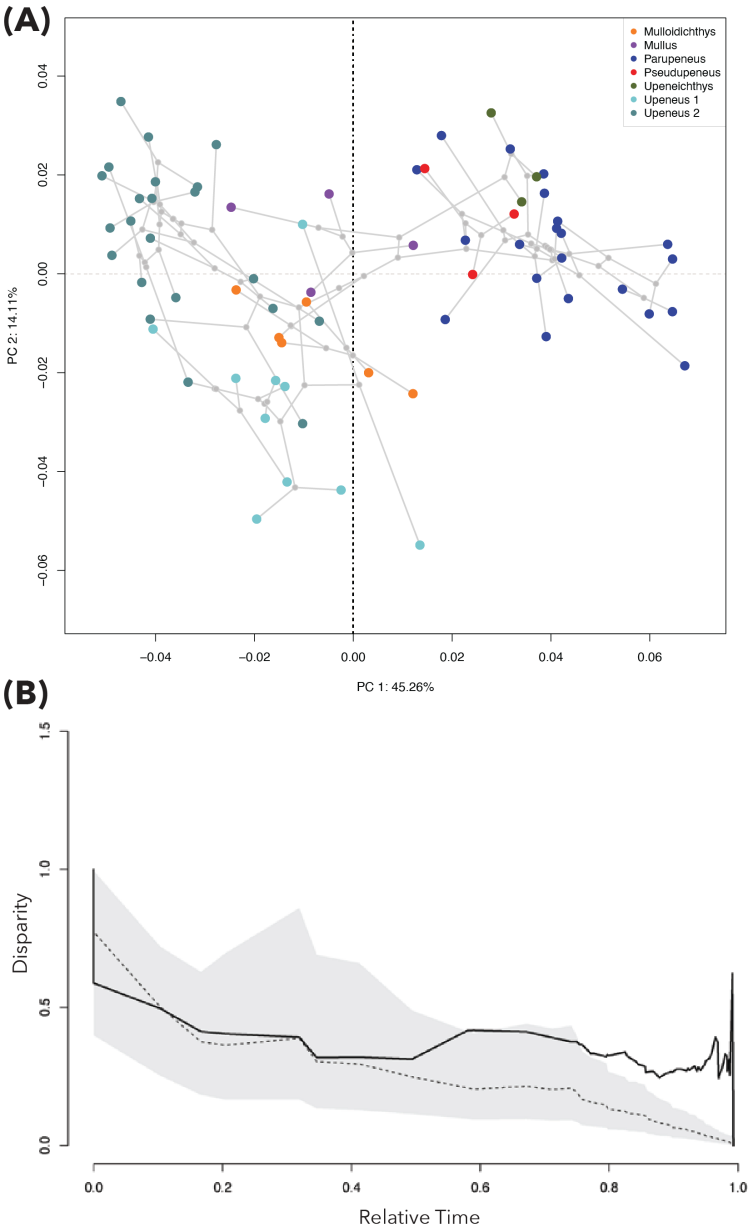

Figure 2.8: **Phylomorphospace and body shape variability through time for the goatfishes.** (A) Phylomorphospace projections of body shape. Each tip is colored according to clade membership. Grey branches indicate the internal branch structure of the superimposed phylogeny shown in Figure 3. (B) Disparity Through Time (DTT) plots for PC1 – PC3 of body shape. Disparity is the average subclade disparity divided by the total clade disparity for each internal node of the tree. The solid black curve indicates the trait disparity values for this study, and the dashed curve is trait disparity for simulated data under Brownian Motion. Grey shaded areas of each plot indicate the 95% confidence interval inferred from the simulations.

fully distinct as there is overlap within values of PC1 (Figure 2.A.4). The only other genus that has species that falls within similarly high ranges of PC2 is *Upeneichthys*, represented by *Up. vlamingii*. There are no species that have both high values of PC1 and low values of PC2, which is represented by the empty lower right quadrant in the morphospace. *Upeneus* and *Parupeneus* have the highest proportion of total variance (Table 2.S3) and a MDI of 0.098.

2.4.4 Body shape morphometrics across phylogeny

Analyses revealed a significant phylogenetic signal in the body shapes of goatfishes. The phylomorphospace estimates the root node within the lower left quadrant, indicating low values of PC1 and PC2. When considering phylogenetic relationships among the genera, the two morphotype groups are largely congruent with phylogenetic structure (Figure 2.8A & 2.A.4). The main exception to this pattern is the convergence of *Upeneichthys* with *Parupeneus* in morphospace, despite its sister genus being *Mullus* (Figure 2.8A). When accounting for the variance that is correlated with phylogenetic relationships, there is a strong overlap among the morphotype groups (Figure 2.A.5). All genera were centered around the origin, although there were differences between the distribution of the two most species rich genera, *Parupeneus* and *Upeneus*. P*arupeneus* and *Upeneus* have equivalent variance across PC1, and *Parupeneus* and *Upeneus 1* split with *Upeneus 2* along PC2. There is a significant deviation from the null expectations for disparity through time (Figure 2.8B & 2.A.6). Additionally, there is relatively late maintenance of higher within subclade variation than we would expect with BM into the Pliocene and Pleistocene (Figure 2.8B & 2.A.6).

2.5 Discussion

The addition of a robust phylogeny of goatfishes to the growing marine fish Tree of Life enriches our understanding of the tempo and mode of coral reef fish biodiversity and

evolution. The central conclusions of this study are 1) phylogenetic relationships among the species and genera of Mullidae were highly resolved, 2) body shape morphometrics revealed several diverse body profiles among the goatfishes, largely congruent with phylogeny and taxonomy position, and 3) we found strong evidence for unstable rates of phylogenetic and morphological diversification, with an unusual pattern of recent increase in lineage accumulation across all genera in the family that may be associated with habitat changes at the Miocene-Pliocene transition.

2.5.1 Highly supported phylogenetic relationships within goatfishes

Our phylogenetic analysis represents the most species rich Mullidae phylogeny to date with the inclusion of 72 species, representing just over 70% of extant goatfishes. We resolved the Mullidae into six highly supported genera, which were monophyletic with the exception of *Mulloidichthys* (Figure 2.3). The majority of internal nodes were well supported and the phylogenetic relationships were largely consistent across inferences made using ML, ASTRAL, and Bayesian approaches with only a few notable exceptions (Figure 2.3 & 2.4). We found evidence that some species groups have very short branch lengths between taxa, which made it difficult to accurately resolve the tip level relationships among some species. This is most evident with the species group in the *Upeneus 1* clade, comprised of *U. quadrilineatus, U. suahelicus (taeniopterus), U. supravittatus*, and *U. suahelicus*. The branch lengths between these species are extremely short, resulting in different topologies with low support across all three of our inference methods (Figure 2.3 & 2.4). Another interesting incongruence is *Parupeneus cyclostomus*, which has a large shift in its phylogenetic placement that is particularly evident when comparing the ML and ASTRAL topologies (Figure 2.4). This is again likely the result of short branch lengths that result in low phylogenetic support and low gene tree concordance among *P. cylcostomus, P. chrysonemus*, and *P. chrysopleuron* (Figure 2.4).

While our phylogenetic analyses generally support the results of previous studies, there are several main exceptions that can be found across the goatfish topology. In comparison with the clade-wide phylogeny inferred from solely morphological characters, our results resolve a slightly different relationship structure among genera with strong support for the sister relationship of *Mullus* and *Upeneichthys* (Kim 2002). Additionally, our inferences resolved many of the polytomies found in that study, primarily among closely related species. Interestingly, many of the phylogenetically informative morphological synapomorphies of each genus were well supported when examining the goatfishes using genomic data. The other primary source of phylogenetic incongruence with previous studies was found when examining the relationships among species within closely related species complexes, primarily within *Upeneus* and *Parupeneus* (Figure 2.A.7). These species complexes are often described based on similarities taxonomic traits, such as the number of fin rays and geographic proximity (Uiblein et al. 1998; Uiblein and Heemstra 2010, 2011; Uiblein and Gouws 2015; Uiblein and Motomura 2021). It is important to note that there is uncertainty across many of the nodes associated with species that are assigned to species groups, which reaffirms the issue of resolving the relationships among species separated by very short branch lengths (Figure 2.4 & 2.A.7). Despite this, the majority of described species complexes cluster in morphospace, which indicates some degree of similarity across body shapes (Figure 2.A.8).

2.5.2 Recent diversification of goatfishes

Recent genomic studies on the order Syngnatharia have estimated a far younger crown age of the goatfishes than previously suggested (Betancur et al. 2017; Santaquiteria et al. 2021). The first study to include an estimation of the crown age of the goatfishes was Near et al. (2013), which estimated an age of approximately 40 Ma based on ten nuclear genes for five species (Near et al. 2013). Subsequent studies resulted in a large amount of variation in the estimated crown age, with ages ranging from 20 Ma to 67 Ma depending on the type of

genomic data used and the number of species included (Betancur-R et al. 2013; Betancur et al. 2017; Alfaro et al. 2018; Hughes et al. 2018; Rabosky et al. 2018).

In this study, we estimated the crown age of goatfishes to be approximately 21.9 Ma (HPD 17.0 – 27.7 Ma) and the divergence between Mullidae and Callionymidae to have occurred approximately 58.2 Ma (HPD 48.6 – 74.4 Ma; Figure 2.2). Our results most closely resemble the findings of Santaquiteria et al. (2021), which estimated a slightly younger mean age of goatfishes at 18.0 Ma and high probability density (HPD) node age range of 14.9–21.8 Ma (Santaquiteria et al. 2021). However, our estimated divergence time between Mullidae and Callionymoidei is much younger than Santaquiteria et al. (2021), as their estimates date the split during the Late Cretaceous (Santaquiteria et al. 2021). This variation in the observed HPD ranges and outgroup divergence times could be due to diffferences in the calibration scheme. In this study, we did not use any germinate calibrations to constrain the node ages based on geologic ages, and we included two additional fossil calibrations located at the stem node and and an internal node within the Mullidae (Figure 2.2). For example, we included a fossil *Mullus* from Northern Russia dated as approximatly 13 Ma, which is relatively close in age to the lower age estimate for the root of Mullidae in Santaquiteria et al. (2021; Carnevale et al. 2006). Additionally, our inclusion of more species in the crown group could push the age range older due to the higher resolution of species level relationships.

The estimated crown age of goatfishes is notably younger than many comparable marine fish clades of this size and geographic distribution (Tedesco et al. 2017; Alfaro et al. 2018; Miller et al. 2018; Rabosky et al. 2018). For example, the majority of modern coral reef associated fish families have fossil representatives in the early Eocene deposits at Monte Bolca, with only butterflyfishes originating after the Eocene at approximately 32 Ma (Cowman and Bellwood 2011; Bannikov 2014; Bellwood et al. 2017). This young origin age is evident even among comparisons with the other suborder level clades within the Syngnatharia, as all of the other major suborder clades within the Syngnatharia diverged within the Late

Cretaceous (∼70-83 Ma; Santaquiteria et al. 2021). These divergence ages are drastically older than the estimated age of the Mulloidei, a suborder that is comprised of the Mullidae (Figure 2.2).

Additionally, our inferences resulted in a relatively long branch length between the stem and crown node of the goatfishes, which spans from the late Oligocene to the early Miocene (Figure 2.2). The Oligocene – Miocene transition was characterized by a relative decrease in global temperature and fluctuations in Antarctic ice volume, as well as significant changes in marine biogeographic regions, which resulted in the migration of the warm water faunas of the Indo - Pacific into the Mediterranean basin (Beddow et al. 2016). Given the limited availability of fossil data within the goatfishes, it is impossible to accurately estimate possible extinction events, and therefore diversification, along the stem branch (Louca and Pennell 2020). However, this transition marked the first major divergence at the root of the crown goatfishes as the circumglobally distributed genus *Upeneus* split from the remaining genera (Figure 2.2 & 2.3).

2.5.3 Morphological differentiation among genera

The geometric morphometrics of body shape among species within the Mullidae highly supports the phylogenetic placement and previous taxonomy. Our analyses reveal two main morphotypes that diverged along PC1, along which variation is primarily associated with differences in body elongation, head depth, and fin base position (Figure 2.6 & 2.7). Each of these morphotypes are significantly associated with a particular set of genera. For example, morphotype 1 is comprised of *Upeneus, Mulloidichthys*, and *Mullus* which tend to be slender and more elongated. In contrast, morphotype 2 is comprised of *Upeneichthys, Pseudupeneus*, and *Parupeneus* and is associated with truncated body length and a larger head depth. These results are consistent with findings from previous studies on body shape evolution across fishes, which demonstrated that the major axis of variation was often associated with body

elongation (Claverie and Wainwright 2014). Body elongation, often described as elongation ratio, has many functional implications and may be associated with habitat, trophic mode, and swimming performance across a wide variety of fish clades (Claverie and Wainwright 2014; Friedman et al. 2019). Another important aspect of morphological variation associated with PC1 is head shape. Head shape is often associated with trophic specialization, as the feeding mechanics that underlie head morphology often impose restrictions on the types and range of possible morphological variation (i.e. Cooper and Westneat 2009; López-Fernández et al. 2013; Mahe et al. 2014). Variation along PC2 is associated with eye diameter, ventral body depth, and the relative distance between the base of the second dorsal fin and the caudal peduncle (Figure 2.6 & 2.7). Variation across these traits is associated primarily with habitat and prey capture. Interestingly, only species within *Upeneus* occupy the portion of morphospace associated with lower values of PC2 (Figure 2.6).

All species of goatfishes are specialized benthic carnivores that use their tactile and chemosensitive barbels to disturb the substratum in search of food (Gosline 1984; McCormick 1993, 1995; Platell et al. 1998; Lukoschek and McCormick 2001). Despite their relative similarities in feeding morphology, previous research has shown that co-occurring species often differ in their substrate preferences, feeding modes, and diet compositions (Gosline 1984; Golani 1994; McCormick 1995; Platell et al. 1998; Krajewski et al. 2006). Given that the precise relationships among head and body shape, substrate preference, and feeding mode are currently unknown across all species of goatfishes, future work should test the functional and ecomorphological associations among these traits.

Our results show a highly significant association among genera, body shape, and head shape. Divergence in body and head shape is consistent with a BM model of evolution, in which lineages exploited new regions of morphospace during the early divergences among genera (Figure 2.6A). However, the distribution of PC scores across the phylogeny suggests that morphotype 2 is derived from morphotype 1, and may be associated with changes in

preferred benthic substrates and feeding behaviors (Figure 2.A.3). Interestingly, no lineages occupy the area of morphospace defined by high values of PC1 and low values of PC2, which is associated with truncated body, larger body depth, and larger eye size (Figure 2.6). Additionally, we do not find evidence for significant differences in the amount of disparity among genera. These evolutionary patterns suggest that there were early divergent strategies across morphotypes early in goatfish history, and species have subsequently radiated within their distinct areas of morphospace. Further research needs to be done to test hypotheses about potential ecological or functional constraints and whether this is an example of phylogenetic niche conservatism within genera.

2.5.4 Phylogenetic and morphological diversification associated with Miocene-Pliocene transition

The transition from the late Miocene into the early Pliocene marked noticeable changes in the rates of lineage accumulation across all genera of goatfishes. There was a pulse of rapid diversification beginning approximately 5 Ma, which is evident by the presence of short branch lengths and the high density of nodes with ages less than 5 Ma (Figure 2.3 & 2.4). This pulse follows a period of relatively slower rates of diversification during the middle to late Miocene, which is evident by long branch lengths and significantly lower rates of lineage accumulation (Figure 2.3 & 2.4). Although goatfishes have been examined in the context of large-scale marine diversification studies now (Siqueira et al. 2016; Alfaro et al. 2018; Rabosky et al. 2018), these variable diversification rates have not been observed until now. The timing of this increase in lineage accumulation during the end of the Miocene and into the Pliocene corresponds with the maintenance of greater mean subclade disparity than expected under pure BM, which could indicate that species were simultaneously undergoing changes in the rates of both phylogenetic and morphological diversification (Figure 2.4 & 2.8B). The timing of this variability in diversification rates suggests that the paleogeographic, climatic,

and sea level changes associated with the Miocene-Pliocene transition played an important role in the evolution of goatfishes. In particular, the end of the Miocene is associated with an increase in the interchange of Indo-West Pacific fauna caused by the separation of the Arabian and African plates, despite the complete separation of the Mediterranean Sea from the Red Sea and the Indian Ocean due to terrain uplift (Briggs 1995).

The timing of lineage accumulation and morphological diversification most closely resembles the trends observed in studies of adaptive radiations, which often occur in limited geographic areas and in the absence of competition (Glor 2010; Hench et al. 2022; Thacker et al. 2022). Although there is evidence for similar rates of diversification in fish clades such as the parrotfishes, hamlets, and gudgeons, it is unusual for a clade of the species richness and distribution of goatfishes to have recent bursts of speciation across all genera (Smith et al. 2008; Hench et al. 2022; Thacker et al. 2022). One example can be found in a recent study by Knudsen et al. (2019), which observed rapid diversification in the Kyphosidae (sea chubs), during the middle Miocene into the Pleistocene that was likely associated with their generalist herbivorous diet and pelagic larval dispersal (Knudsen et al. 2019). However, the species richness of Kyphosidae is much lower than the goatfishes with only 16 described species despite their circumglobal distribution (Knudsen and Clements 2016; Fricke et al. 2022).

Given our current understanding of goatfish ecology and life history, high rates of larval dispersal and specialized feeding modes likely play a large role in driving recent bursts in diversification. Water column position has been shown to correlate with rates of morphological diversification, as benthic associated species may be associated with higher diversification rates compared to their pelagic counterparts (Friedman et al. 2020; Rincon-Sandoval et al. 2020). However, this trend has been shown to vary widely across clades depending on their developmental patterns and life history traits (Claverie and Wainwright 2014; McCord et al. 2021). Population level studies have shown that species of goatfish have been able to

maintain panmictic populations across large geographic scales due to their long pelagic larval duration (Lessios and Robertson 2013; Fernandez-Silva et al. 2015). Interestingly, evidence for population isolation and differentiation was observed primarily at the periphery of the species' range (Lessios and Robertson 2013; Fernandez-Silva et al. 2015), which suggests that recent geographic changes in habitat availability can impact the ability of species to effectively disperse and subsequently diversify.

2.5.5 Conclusion

A species rich time-calibrated phylogeny of the Mullidae allows us to make important additions to our understanding of goatfish evolution and diversification. Phylogenetic relationships among species and genera are highly resolved, and corroborate with the structure among body shape morphologies. The Miocene-Pliocene boundary marked a period of both rapid changes in the rates of phylogenetic and morphological diversification, which was possibly driven by changes in the availability of suitable habitats caused by geologic changes. An increased understanding of the evolutionary relationships among key groups of organisms will continue to reveal surprising trends in the tempo and mode of diversification that resulted in the patterns of biodiversity that we observe today.

APPENDIX

2.A Supplementary Figures

Supplementary tables for this chapter can be viewed at `https://figshare.com/s/f8186af6d6faf64900e6`

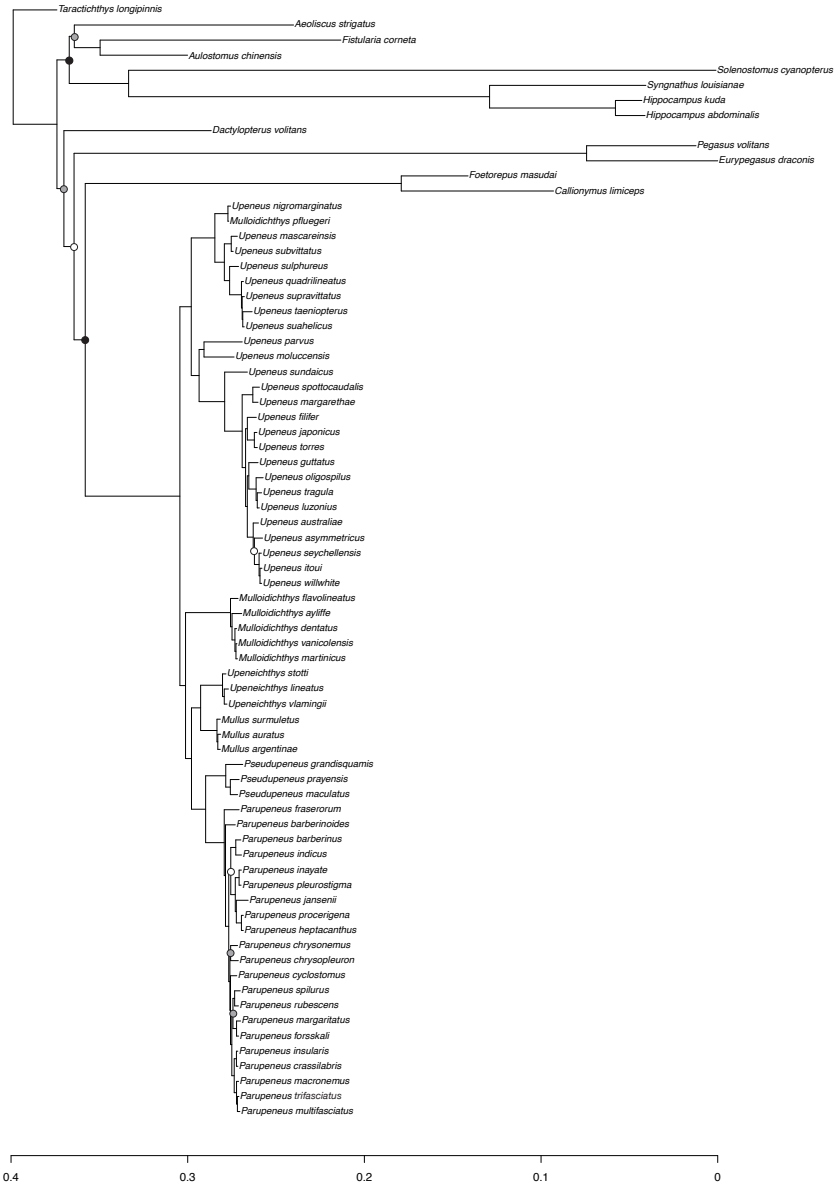

Figure 2.A.1: **Maximum likelihood phylogeny of the goatfishes inferred using a 70% complete UCE matrix in IQTREE.** UF bootstrap support values are shown using color coded circles on certain nodes. Nodes without a circle have a UF Bootstrap support of 100%, while nodes shaded black = 99% - 95%, grey = 94%-85%, and white = <85%. The relevant tree and alignment files can be found in supplementary data.

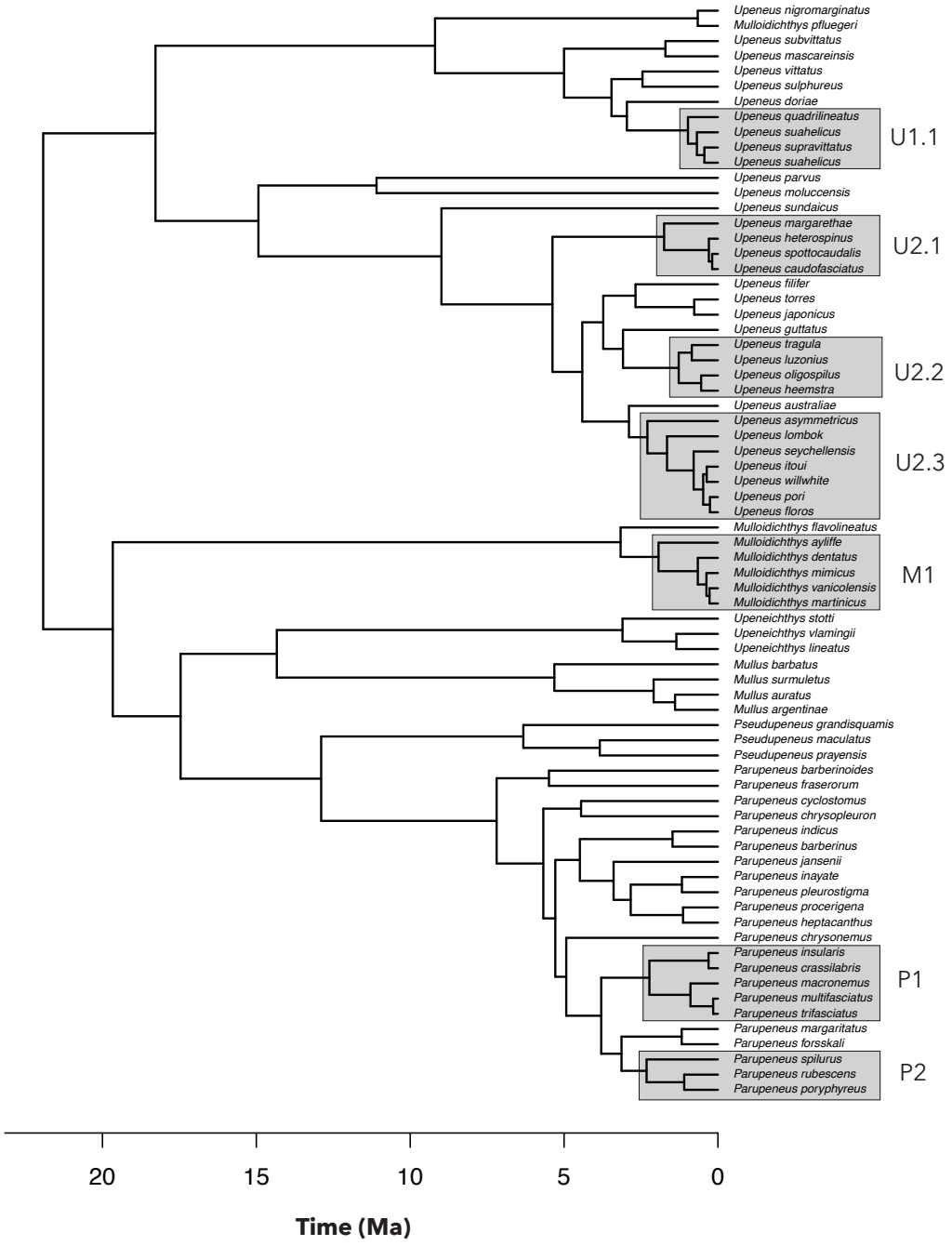

Figure 2.A.2: **Phylogenetic placement of goatfish species groups.** Location and labels of species groups within the Mullidae, shown on the time-calibrated BEAST phylogeny. We define species group as a clade of three or more species whose mean node age is within the Pleistocene (2.58 - 0.012 Ma). Posterior probability values for this phylogeny can be found in Figure 3 and the HPD values are shown in Figure 2.

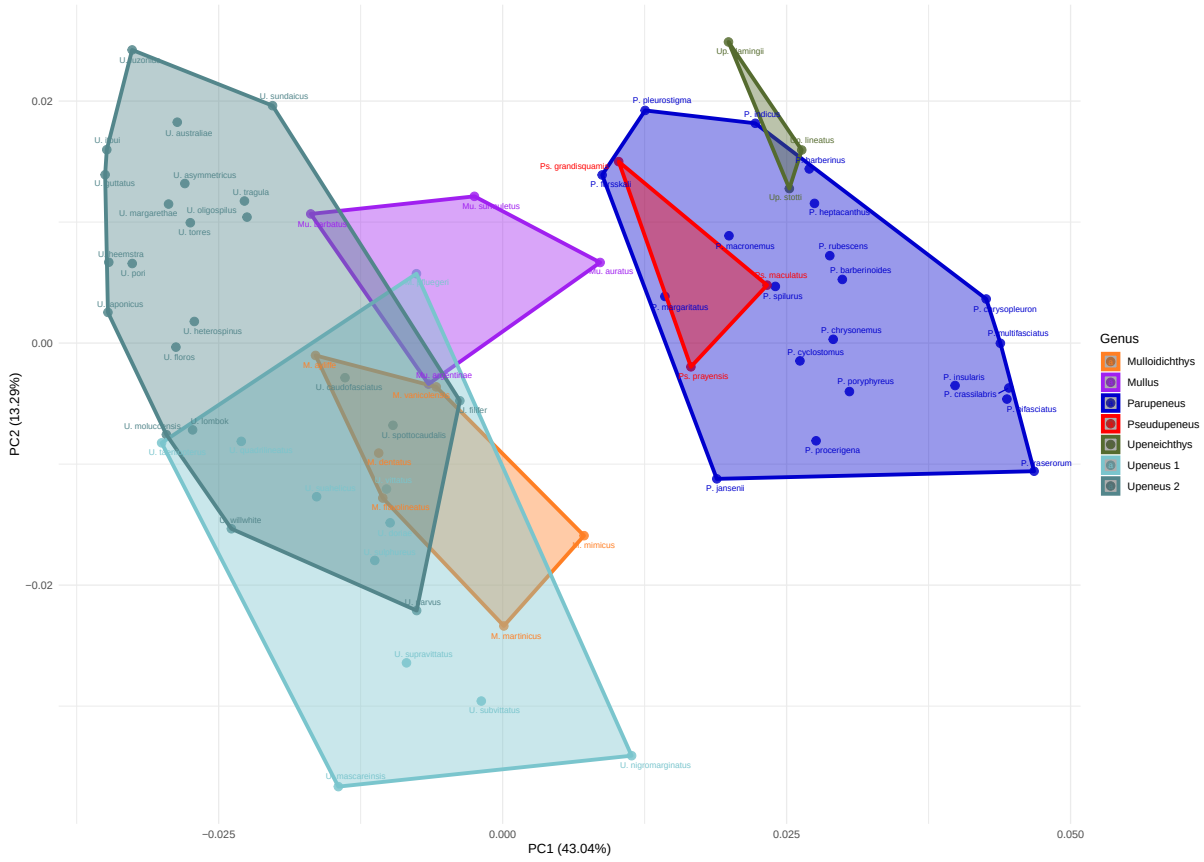

Figure 2.A.3: **Body shape morphospace for goatfishes labeled by clade membership.** Axis 1 and 2 from Principal Component Analysis (PCA) on Procrustes aligned body shape landmarks for 70 species. Each coordinate represents the mean shape of each species. PC1 represents 44.95% of the total variation, and PC2 represents 14.06% of the total variation. Species are grouped by clade, and this is indicated by color coded hulls to show the relative distribution of each clade in morphospace.

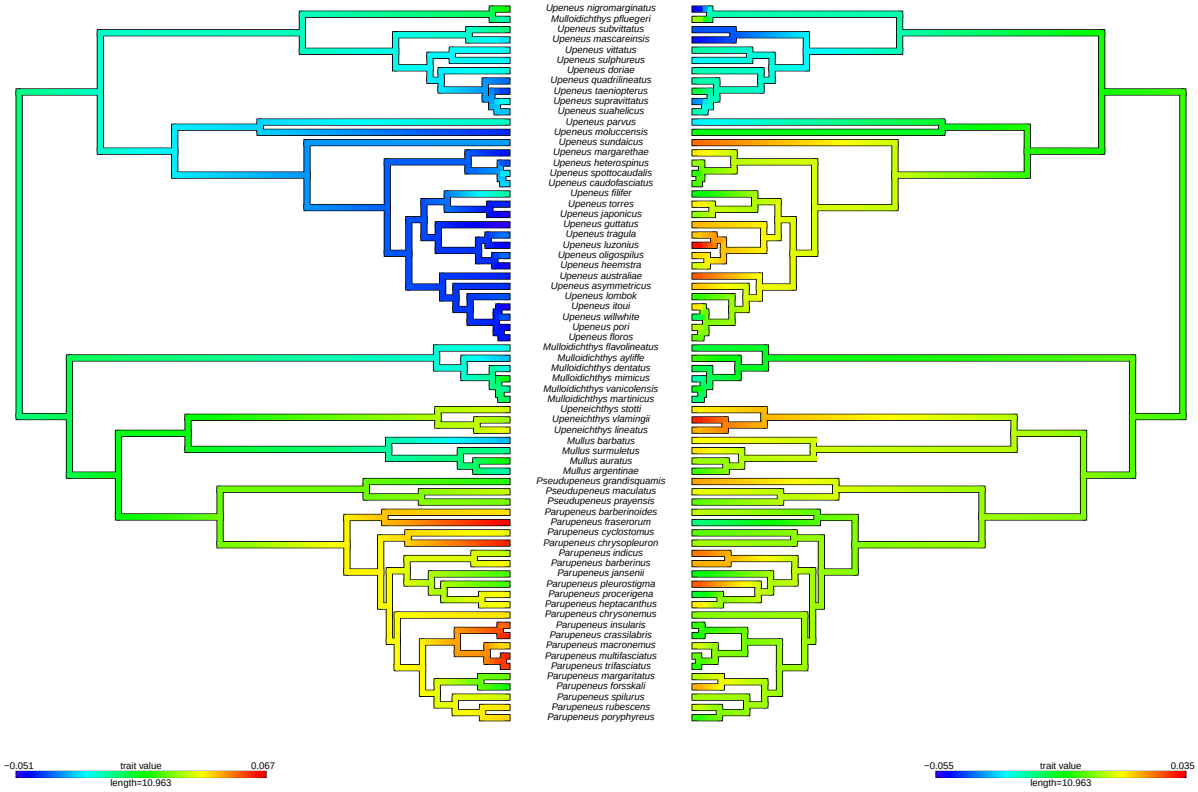

Figure 2.A.4: **ContMap showing the distribution of PC1 and PC2 scores across the phylogeny.** Higher values are indicated in warmer colors and lower values are shown in cooler colors. Figure generated using Phytools (Revell 2012).

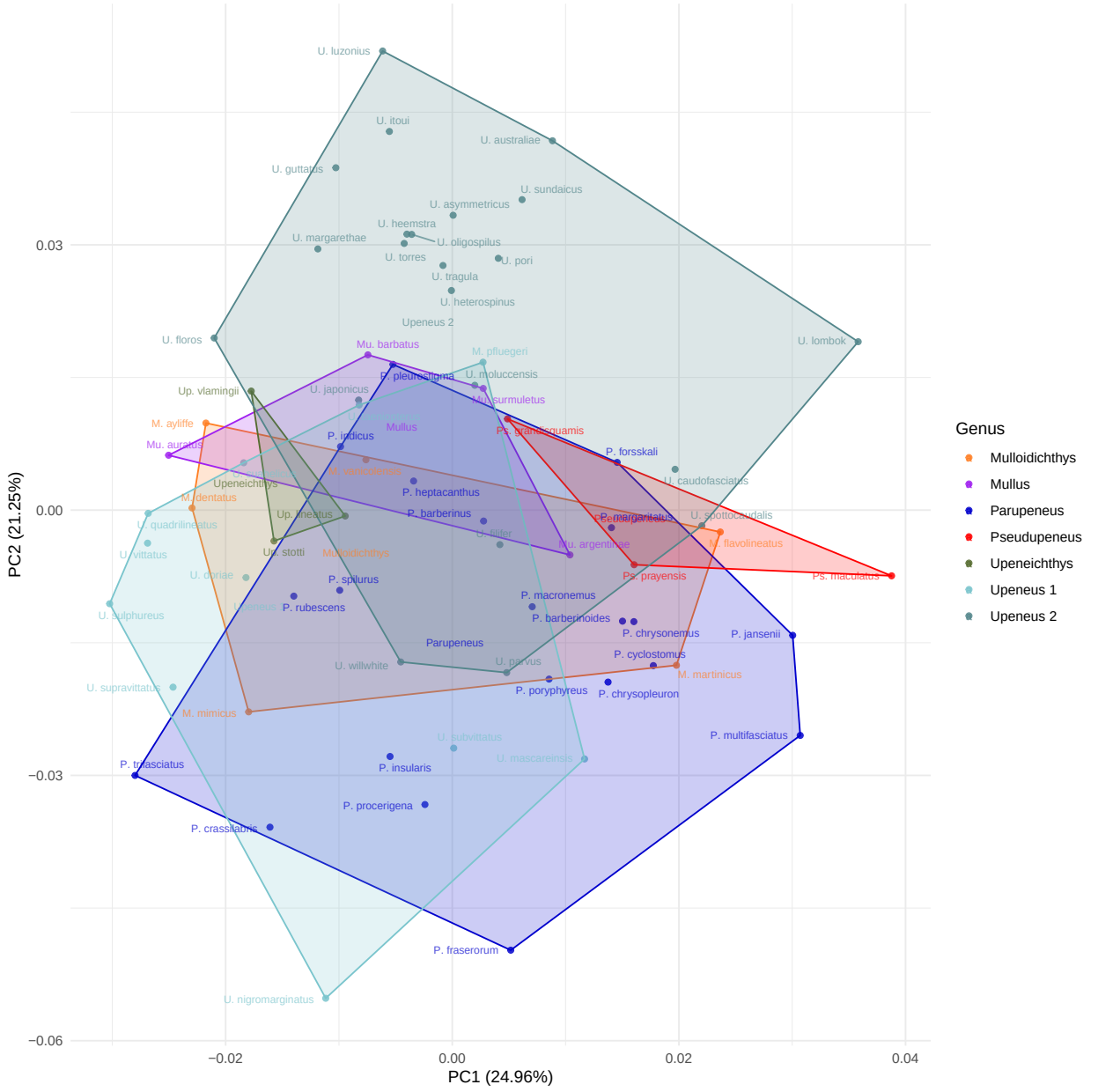

Figure 2.A.5: **Phylogenetic PCA of goatfish body shape.** Phylogenetic PCA using GLS-centering and projection implemented in the gm.prcomp function in geomorph 4.0.1 (Revell 2009; Adams and Otárola-Castillo 2013). Coordinates associated with each species and hulls are color coded by clade membership.

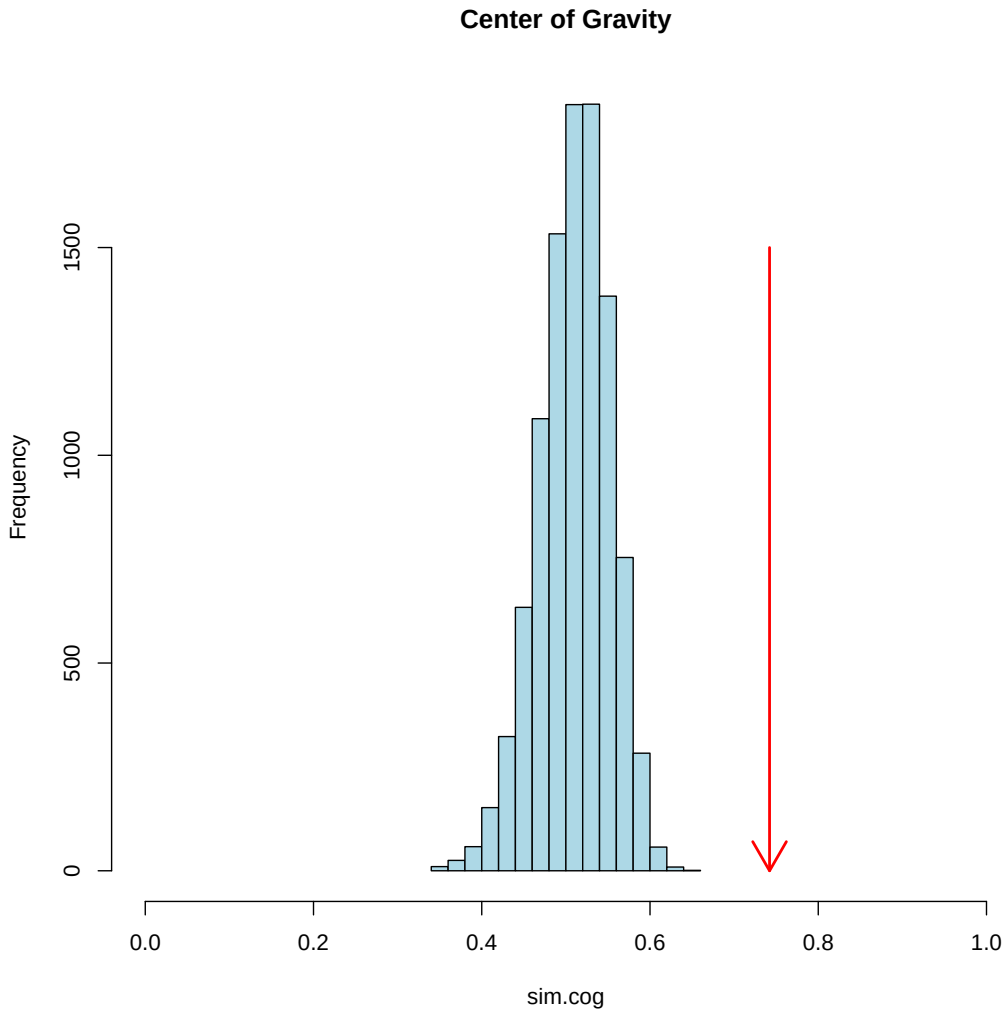

Figure 2.A.6: **Weighted average times, termed Center of Gravity (CoG).** The blue histogram are the results of 9,999 simulations generated under the assumptions of BM, and the red arrow indicates the CoG value for this study.

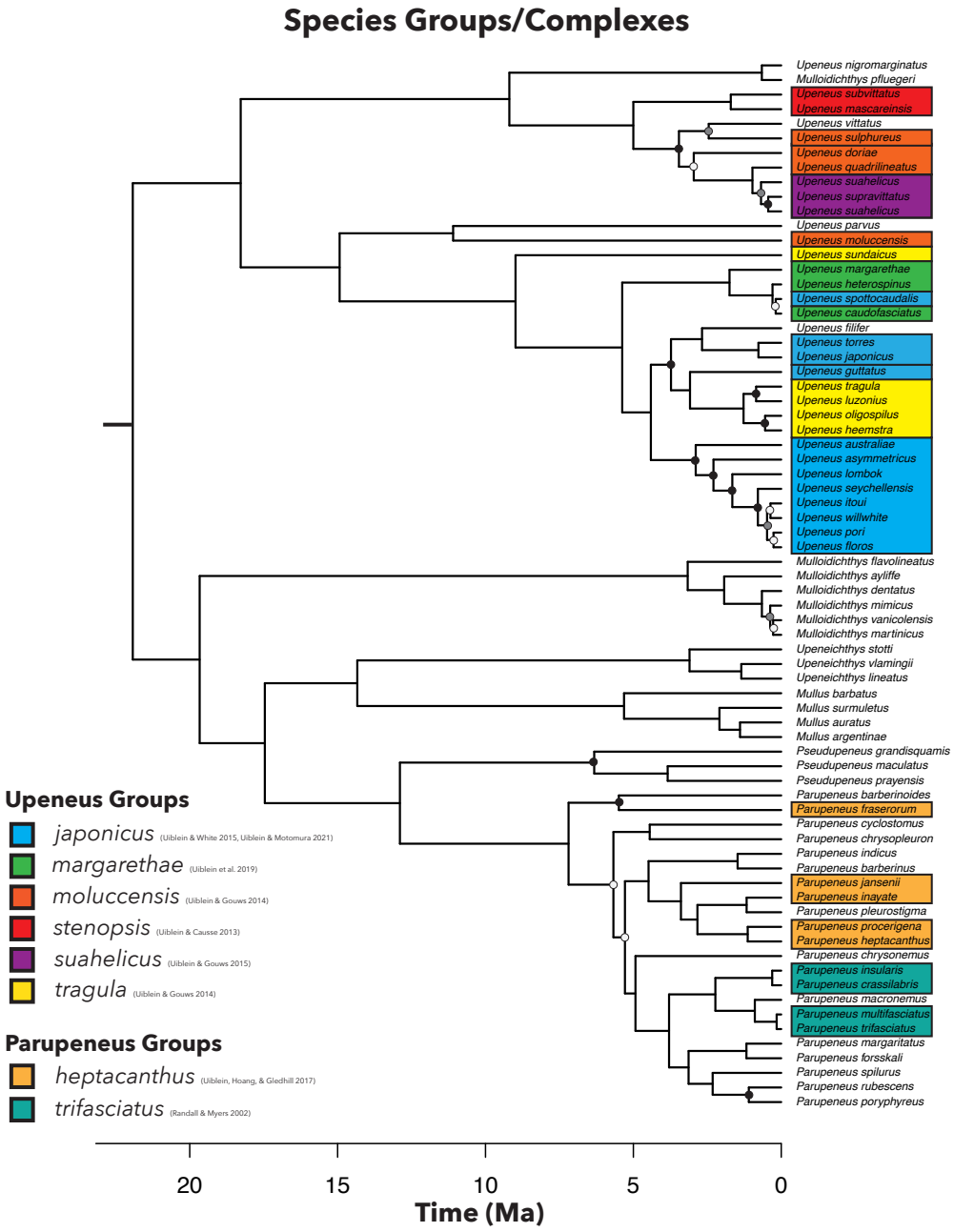

Figure 2.A.7: **Described species complexes within the Mullidae.** Distribution of previously described species groups and complexes on the time calibrated phylogeny inferred in this study. Species groups/complexes are indicated by color, and a black bar separating species within the same species group indicates that they are not monophyletic. Nodes without a circle have posterior probability of 1.0 from the BEAST analysis. Nodes with a circle are color coded by their posterior probability, with black = .99 - .95, grey = .94-.85, and white = <0.85.

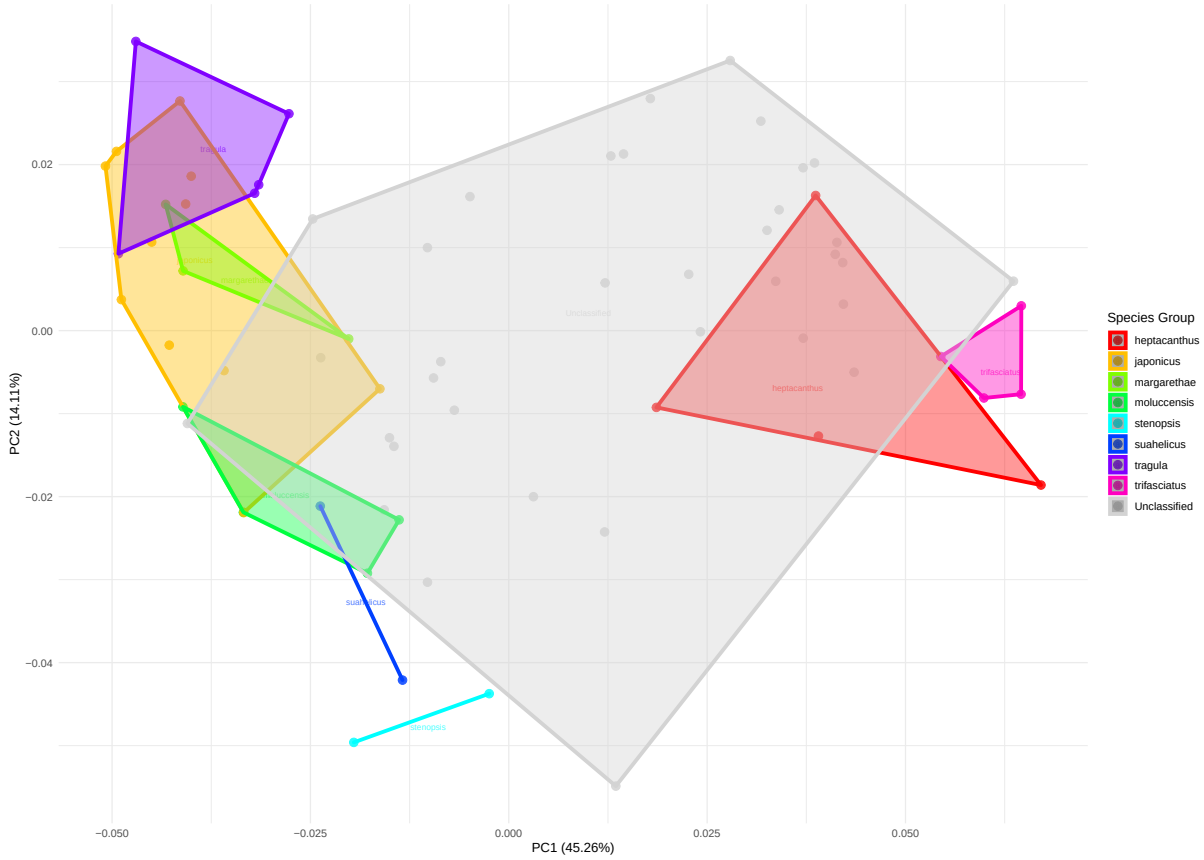

Figure 2.A.8: **Body shape morphospace for the Mullidae labeled by species complex membership.** Axis 1 and 2 from Principal Component Analysis (PCA) on Procrustes aligned body shape landmarks for 70 species. Each coordinate represents the mean shape of each species. PC1 represents 44.95% of the total variation, and PC2 represents 14.06% of the total variation. Species are grouped by membership in each species group/complex as defined by the literature (Figure 3.A.7), and this is indicated by color coded hulls to show the distribution of group in morphospace. Species that are not assigned to a complex or group are indicated in light grey.

CHAPTER 3

BIOREGIONALIZATION REVEALS VARIABILITY IN PHYLOGENETIC AND SPATIAL DIVERSITY ACROSS GOATFISH ASSEMBLAGES ON A GLOBAL SCALE

3.1 Abstract

Although the marine realm hosts an extraordinary amount of biodiversity, the identification of important barriers to dispersal is difficult due to the variability among evolutionary and ecological traits across clades. In this study, we examine the spatial distribution of the goatfishes (Mullidae) using bioregionalization methods to infer goatfish specific bioregions. We curated over 300,000 georeferenced specimen records for the family and used both hierarchical and grade-of-membership clustering approaches to identify goatfish-specific bioregions in shallow marine waters throughout the world. Using a recently published phylogeny of 72 species, we explore the evolution of species within and among bioregions to identify areas of low species turnover, assess diversification patterns, and test hypotheses about the influence of range size and dispersal mode on current distribution patterns. At large spatial scales, we recover distinct assemblages within the Indo-Pacific and the Atlantic/East Pacific oceans, which is the result of well-known vicariance events. However, at smaller scales, we observe intermixing among assemblages that likely represent the presence of semi-permeable barriers to dispersal. We find that the Central Indo-Pacific bioregion defined by high species richness and low endemism due to the overlap of species' ranges. Finally, we find limited evidence for geographic radiations, rather most bioregions trend towards phylogenetic overdispersion[1].

1. This chapter will be published under the following citation: Nash, C. M., Kang, K. J., & Westneat, M. W. Bioregionalization reveals variability in phylogenetic and spatial diversity across goatfish assemblages on a global scale (Syngnatharia: Mullidae). *In Prep.*

3.2 Introduction

The delineation of biogeographical boundaries between groups of taxa is one of the primary objectives of evolutionary biology, macroecology, and community phylogenetics. Lineages and their distributions are dynamic and heterogeneous, resulting in complex patterns of biotic assembly. The integration of species distribution with phylogenetics have been examined across numerous taxa, and yet the central question remains: why do some species co-occur while others do not? This is a deceivingly simple question, as there are many underlying processes that can influence the distribution and evolution of lineages within assemblages and communities. Only through the consideration of phylogenetics in the context of historical biogeographic and recent ecological processes, in combination with rich spatial data sets on species occurrence, can we attempt to address these questions.

Research on patterns of species co-occurrence in marine systems has lagged behind their terrestrial counterparts due to the complexity and inaccessibility to many parts of the marine environments (Pittman and Olds 2015; Cowman et al. 2017; Sambrook et al. 2019; Braga et al. 2023; Pinheiro et al. 2023). However, major insights into the patterns of biotic assembly and biogeography of fishes are now possible with the availability of highly resolved phylogenetic hypotheses and increasingly detailed spatial distribution data. The prevalence of georeferenced data enables a high-resolution approach to evaluate the factors that influence biogeographic boundaries through the fine-scale delineation of faunal groups and assemblage boundaries. Previous studies on marine biogeography typically delineate boundaries between regions based on variables such as temperature, depth, benthic complexity, algal/coral cover, and overall patterns of biodiversity (Briggs 1961, 1974; Grant and Bowen 1998; Spalding et al. 2007; Tittensor et al. 2010; Kulbicki et al. 2013; Mouillot et al. 2013; Cowman et al. 2017; Pinheiro et al. 2023). Although these metrics are useful for broad scale biodiversity analyses, they often mask taxon-specific biogeographic patterns and tend to be biased towards lineages with high dispersal capabilities due to their ability to tolerate

a variety of habitats (Kreft and Jetz 2010). The permeability of a barrier is dependent upon physiological, morphological, ecological, and behavioral traits, which can vary widely across the diversity of fishes (Robertson et al. 2004; Bradbury et al. 2008; Luiz et al. 2012). In order to determine the factors that influence the distribution of specific clades, it is essential to parse patterns of evolutionary diversification and spatial variance across temporal and spatial scales.

In this study, we examine patterns of co-occurrence, infer potential barriers to dispersal, and explore the role of geographic dispersion and community composition on the evolution of the goatfishes. The goatfishes (Syngnatharia: Mullidae) are a diverse, globally distributed family of fishes found in close association with coral reef ecosystems. Goatfishes occupy a distinct role in coral reef ecosystems due to an innovation in their feeding morphology, a pair of highly specialized hyoid barbels. These barbels are fleshy extensions located under the jaw that are capable of chemoreception and prey excavation (Gosline 1984). This unique feeding behavior of the goatfishes have been shown to play an important role in maintaining coral reef diversity and community composition, in addition to acting as an indicator of coral reef health, due to the resuspension of otherwise inaccessible food sources within various benthic substrates (Lukoschek and McCormick 2000; Uiblein 2007; Russ et al. 2015).

A recently published time-calibrated phylogeny of the goatfishes has shown that the crown ages of goatfishes is younger than most other coral reef fish families of comparable species richness and global distributions (Nash et al. 2022). Additionally, there was evidence of variability in the rates of phylogenetic and morphological diversification, with an increase in diversification rates found in association with the Miocene-Pliocene boundary (Nash et al. 2022). Interestingly, the pulse of an increase in diversification is observed across most extant lineages, which suggests that the global paleogeographic, climatic, and sea level changes associated with the Miocene-Pliocene transition (Briggs 1995) impacted the rates of goatfish evolution. Given these unique evolutionary characteristics of the goatfishes, a

direct examination of how the geographic distribution of extant lineages is associated with patterns of phylogenetic diversification will help elucidate potential influences on patterns of goatfish diversification.

The central aims of this study are to 1) delimit biogeographic regions based on patterns of species co-occurrence using two methods of bioregionalization, 2) compare levels of phylogenetic diversity across assemblages and regions, and 3) examine the role of range size and dispersal on current distribution patterns. The use of a robust georeferenced coordinate dataset allows us to examine the locations of faunal shifts within the global distribution of goatfishes. We hypothesize that the boundaries of extant assemblages of goatfishes will correspond to vicariance in association with the Miocene-Pliocene boundary. Additionally, we hypothesize that there will be phylogenetic evidence of isolation and subsequent diversification within each assemblage, which would result in a notable proportion of endemic species within each assemblage. The results of this study, integrated with an examination of the evolutionary and geographic distribution of key functional traits, forms a powerful tool that can be used to understand the processes that drive extant goatfish distributions and dispersal patterns, to explore how shifts between regions have occurred historically, and to predict where lineages will persist in the future.

3.3 Methods

3.3.1 Georeferenced coordinate acquisition and processing

We included all available geographic coordinates for species within the Mullidae, which resulted in at minimum a single occurrence for 95 out of 101 total species. Species names were checked for synonyms using the Catalog of Fishes database, and distribution data was downloaded from the following databases: Global Biodiversity and Information Systems (GBIF), VertNet, and the Ocean Biodiversity Information System (OBIS) using their respective R

interface (Guralnick and Constable 2010; Boettiger et al. 2012; Provoost and Bosch 2017; Fricke et al. 2022; Chamberlain et al. 2023; Table 3.S1) A total of 301,458 coordinates were recovered from this search, which was conducted in October 2022. Given the issue of erroneous coordinates, we used expert compiled distribution descriptions from the IUCN Red List, AquaMaps, FishBase, the California Academy of Science's Catalog of Fishes, and hundreds of published studies to curate the coordinate set for each species independently (Fricke et al. 2022). This cleaning process involved the removal of all coordinates that fell outside of described ranges which did not have a recently identified and vouchered specimen with an associated photograph or a published range extension. Our final coordinate data set consisted of a total of 109,969 unique coordinates for 95 species (Figure 3.A.1).

To account for variation in sampling effort across taxa and sites, we generated a circular buffer with a radius of 2 degrees around each coordinate point using the R packages sp 1.5-1 and rgeos 0.5-9, then merged each buffer to form a single polygon layer for each species (Pebesma and Bivand 2005; Bivand et al. 2013; Bivand and Rundel 2021). Each buffer is used to estimate the spatial area of potential home ranges of individuals within the sampled population, and to not bias areas of high sampling density over areas with low sampling density. These polygons were used to determine the presence or absence of each species within the grid cells of raster, in the WGS84 projection, with the dimensions 2 degrees by 2 degree grid. This matrix was then filtered to remove grid cells that only contained a single species, resulting in an assemblage matrix that contains 12,573 grid cells. Any endemic species that were removed from the matrix filtering were noted.

3.3.2 Bioregion delineation using hierarchical clustering

Following the methodology described in Kreft and Jetz (2010), we generated a pairwise distance matrix from the grid cell presence and absence matrix using Simpson's pair-wise dissimilarity (βsim), implemented in the R package betapart 1.5.6 (Kreft and Jetz 2010; Baselga

and Orme 2012). βsim, which is a dissimilarity metric of species turnover by calculating the ratio of shared and unique species between assemblages, is ideal for this type of study given that it is a richness independent criterion of turnover. This pairwise dissimilarity matrix was used to assign each grid cell to a particular bioregion using Ward's clustering algorithm and k-means clustering. We performed hierarchical clustering using Ward's clustering algorithm, which minimizes within cluster variance, maximizes, between cluster variance, and results in similarly sized clusters bins (Ward 1963). We used the "elbow method" heuristic was used to determine the optimal number of k-means clusters to prevent over-fitting based on explained variance (Thorndike 1953). Given that we wanted to examine global scale patterns, we restricted the maximum allowed k-means cluster to 10. Once the data was partitioned based on the optimal clustering scheme, each grid cell was assigned membership to a single cluster.

To examine the spatial distribution of the clusters, grid cells were visualized on a map projected using WGS84 and were color coded according to their bioregion (Figure 3.1). Additionally, we calculated species membership, species abundance, and relative range composition. To reduce the influence of coordinate and binning errors, we only classified a species as present within a bioregion if more than five percent of their range was distributed within that bioregion. Cluster membership is simply the presence of a species within each bioregion. We define species abundance as the percentage of the cluster that a species occupies, which was calculated by dividing the number of grid cells a species occupies within a bioregion by the total number of grid cells within that bioregion. We define relative range composition as the percentage of a species range found within each bioregion, which is computed by dividing the number of grid cells a species occupies in each bioregion by range size of the species.

3.3.3 Clustering procedures using grade of membership models

A key limitation in using hierarchical clustering to delineate biogeographic regions is that it results in discrete boundaries between geographic units. Although these boundaries could imply a sharp change in species turnover, it is also possible that these boundaries represent areas of integration and overlap between biotas (White et al. 2019). To examine how assemblages co-occur, integrate, and contribute at the junction between bioregions, we used a grade of membership model using the R package ecostructure 0.99.1 (White et al. 2019). This method assigns species into a particular motif using a probabilistic model and then computes the contribution of each motif to a realm or local community in order to more precisely identify regions of mixture and patterns of turnover (White et al. 2019). The proportional contribution of each motif to each individual grid cell is visualized using a pie diagram on the global map. We identify biogeographical regions based on the presence of a predominant contribution from a single motif and if there is a sharp turnover associated with that region. We investigated the variability of motif schemes inferred using k values ranging from 2 to 10 to explore patterns of stability across of biogeographical regions and to examine the patterns of geographic nestedness across the distribution of motifs.

3.3.4 Range size estimation and phylogenetic distribution

Range size was calculated by summing the total number of grid cells that each species occupies. To adjust for the large amount of variation among species, we scaled range size values using z-scores. Summary statistics for the distribution of latitude and longitude were also calculated for each species using the coordinate dataset. We used the `phylosig` function in phytools 1.2-0 to compute phylogenetic signal using lambda and K, as well as to conduct hypotheses tests about their significance while accounting for sampling error (Pagel 1999; Blomberg et al. 2003; Ives et al. 2007; Revell 2012).

3.3.5 Phylogenetic diversity metrics across species, grid cells, and bioregions

Using our recent time-calibrated phylogeny as an evolutionary framework, we examined the phylogenetic structure of assemblages at a variety of spatial scales to test hypotheses patterns of species co-occurrence across species' ranges, global grid cells, and our delineated bioregions (Nash et al. 2022). We used Phylogenetic Species Variability (PSV) index to measure and summarize the degree to which species are phylogenetically related in an assemblage (Helmus et al. 2007). PSV is directly related to the mean phylogenetic distance, with the exception that it uses a scaled phylogenetic covariance matrix (Helmus et al. 2007). The expected value of PSV is statistically independent from species richness, and "species effects" can be parsed through an examination of the signed proportion of the total derivation of PSV when each species is removed from the dataset. The PSV index ranges from zero to one, with a value of one indicating that species are unrelated and approaches zero as species become more related.

To examine the phylogenetic structure of the goatfishes globally, we computed PSV values for individual grid cells and bioregions by using the subtree that represents all the species found within each unit (Daru et al. 2017). Additionally, we used a Phylogenetic Field approach, also known as PhyloFields, to examine the overall phylogenetic structure contained within the range of each species. This is computed by calculating the PSV value of the subtree representing the species co-occurring within the range of a focal species, resulting in a single value that summarizes the phylogenetic structure across the range of the focal species (Villalobos et al. 2013). Using adapted code from Ceccarelli et al (2020), we computed a PhyloField for all 95 species of goatfish in our dataset. We used a randomized approach to test whether the observed PhyloField for each species is significantly "clustered" or "over dispersed", which we define as the observed value falling within the lower 0.25% or upper 0.75% compared to the null distribution, respectively (Ceccarelli et al. 2020) . Additionally, we evaluated the phylogenetic alpha diversity within each bioregion using the

Net Relatedness Index (NRI) for both the "taxon shuffling" and "phylogenetic pool" null models to test for significance using the `ses.mpd` function in the R package picante 1.8.2 (Webb et al. 2008; Kembel et al. 2010). To compare levels of beta diversity across clusters, we used the PhyloSor Index, which is based on the total shared branch length between shared and unshared taxa among bioregions (Bryant et al. 2008).

3.4 Results

The central results of this study are that goatfish specific bioregions inferred using bioregionalization approaches differ in their geographic continuity, species richness, and phylogenetic structuring. Although there is overlap in species composition among bioregions, we found variable rates of endemism across regions, with the major distinction found between regions in the Atlantic/East Pacific compared to those in the Indo-Pacific oceanic basins.

3.4.1 Global assemblage structure using hierarchical clustering

The optimal number of assemblage clusters determined using βsim are k = 10, 13, and 17 (assuming that a maximum number of allowed values of k = 10, 15, and 20 clusters, respectively; Figure 3.A.2 & Figure 3.A.3). Focusing on the cluster scheme with k = 10 to reduce complexity, our results show separation among groupings of goatfish bioregions across major ocean realms, such as the Atlantic Ocean, Eastern Tropical Pacific, Indo-Pacific, and temperate Australasia, as well as across tropical and temperate regions (Figure 3.1 & Figure 3.A.4). Notably, there is variability in the degree of geographic continuity and clustering within and among bioregions. The Atlantic Ocean is comprised of four assemblage clusters distributed across A) the Western Tropical Atlantic, B) the Southwest Temperate Atlantic (primarily southern Brazil and Argentina), C) the Temperate Northeastern Atlantic (including the Azores) and Mediterranean, and D) the Eastern Tropical Atlantic (Figure 3.1). There is an additional bioregion in the Eastern Tropical Pacific (E), which borders southern

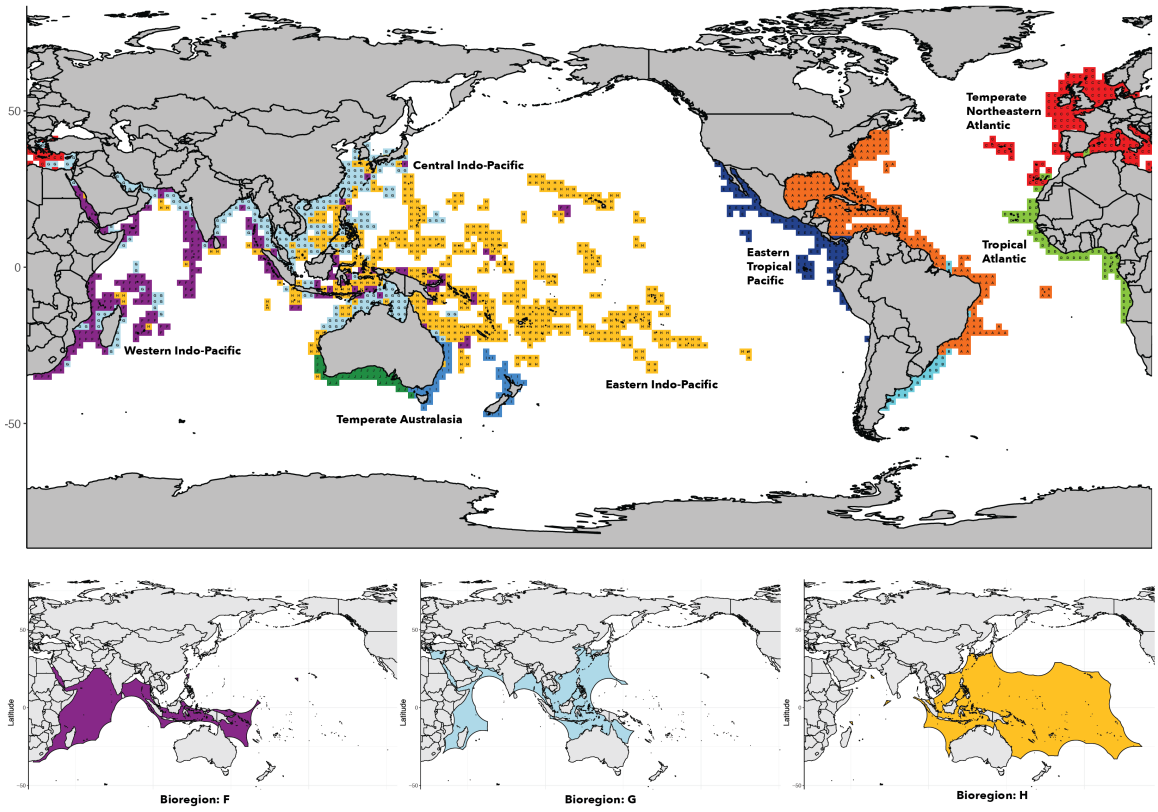

Figure 3.1: **Map of bioregions resulting from Ward's hierarchical clustering of grid cell assemblages with k = 10 based on βsim dissimilarity matrices for species of goatfish.** Colors represent the assigned cluster group to each 2° by 2° grid cell, with letters added for additional clarity. Area names are based on the relative corresponding regions defined by Spalding et al. (2007). Smaller inset maps show the border of three representative bioregions found in the Indo-Pacific. These maps were created using convex hulls with alpha values of 7, and their colors are identical to their corresponding bioregion in the global map to facilitate comparison.

California south to Peru (Figure 3.1). Overall, the boundaries between the bioregions in the Atlantic and East Pacific indicate that there is a distinct geographic area of separation with minimal overlap between bioregions (Figure 3.1). This geographic continuity and distinct boundaries between bioregions can also be observed in Temperate Australasia, in which the bioregions cluster at I) the Southeast Australian Shelf and New Zealand, and J) Southwest Australian Shelf (Figure 3.1).

The three bioregions in the Indo-Pacific span a larger geographic area and distribute

with considerable overlap among each region (Figure 3.1; Table 3.S2). Despite the patchwork pattern, each region can be broadly classified as belonging into three main realms using Spalding's regionalization scheme: F) Western Indo-Pacific, including the Red Sea and the Seychelles, to Southcentral Indo-Pacific, G) Coastal Western Indo-Pacific, including the Red Sea and the Gulf of Oman, to Central Indo-Pacific, including the South China Sea and Northern Australia, and H) Central Indo-Pacific to Eastern Indo-Pacific, including the tropical northwestern and southwestern Pacific, the Hawaiian Islands, the Marshall Islands, Central and Southeast Polynesia, and the Marquesas (Figure 3.1). The primary location of overlap is within the Central Indo-Pacific, often termed the "Coral Triangle", which spans from Western Indonesia east to Papua New Guinea, north to the Philippines, and south to Timor Leste (Figure 3.1).

Given the nature of hierarchical clustering, species are often shared among clusters, and the way that this occurs can be informative to the processes that create each assemblage group. There is a considerable range of species richness across each bioregion. Assemblage groups within the Atlantic Ocean and Temperate Australasian realms have relatively lower species richness, ranging from a single species, as shown in groups B and D, to eight species found in I. On the contrary, the species richness in clusters associated with the Indo-Pacific realm are several magnitudes larger, ranging from 52 to 66 total species. Notably, the percentage of endemic species, in this case defined as species that have more than 95% of their range in a particular cluster, varies across clusters and within realms (Table 3.S3). For example, 100% of species in the Bioregions A and D, the Western Tropical Atlantic and Eastern Pacific, respectively, are endemic to these bioregions (Table 3.S3). This is in contrast with clusters B and C, in which there are no endemic species (Table 3.S3). Despite the larger observed species richness in bioregions within the Indo-Pacific, only a small proportion of species are endemic in these bioregions (Table 3.S3 & 3.S4). There is a minimum of one species from each genus that can be considered an endemic when considering

all the bioregions (Figure 3.A.5; Table 3.S3).

The species motifs inferred using a grade of membership model with $K < 5$ separates biogeographical regions associated primarily with the geological barriers between major ocean basins, which is comparable to the results of hierarchical clustering (Figure 3.2). Additional species motifs based on increasing values of K represent further delineations within these major ocean basins and represent various levels of cohesiveness and integration across geographic locations. Results for $K = 10$ most closely resemble the delineated bioregions inferred using hierarchical clustering (Figure 3.1). Results show that there is considerable overlap and interchange of species motifs across the Central Indo-Pacific, represented by the lack of a single dominant motif in most grid cells in this region (Figure 3.2). Additionally, there is a higher density of clusters recovered within the Central Indo-Pacific and a distinct species motif grouping a marine region often termed the Coral Triangle (Figure 3.2).

3.4.2 Phylogenetic diversity within and among assemblages

Using a presence/absence matrix that excludes for species that have less than 5% of their range in a particular bioregion, our analyses show that the bioregions defined by goatfish distribution within the Atlantic/East Pacific have the highest amount of Phylogenetic Species Variability, which indicates a higher mean phylogenetic distance among species within these assemblages (Figure 3.3, Table 3.S4 & 3.S5). Bioregions within the Indo-Pacific, specifically Bioregions F, G, H, and I, have lower PSV values with little difference among each other (Table 3.S5). Bioregion J has the lowest PSV value across all assemblages, indicating that species within this bioregion are more closely related to each other (Table 3.S5).

The PSV values across bioregions generally corroborate the results of the Mean Pairwise Distance metrics. Our results show that lineages within bioregions geographically located within the Atlantic have positive observed z-values when examining the standardized effect size of the mean phylogenetic distance within bioregions, which indicates that there is some

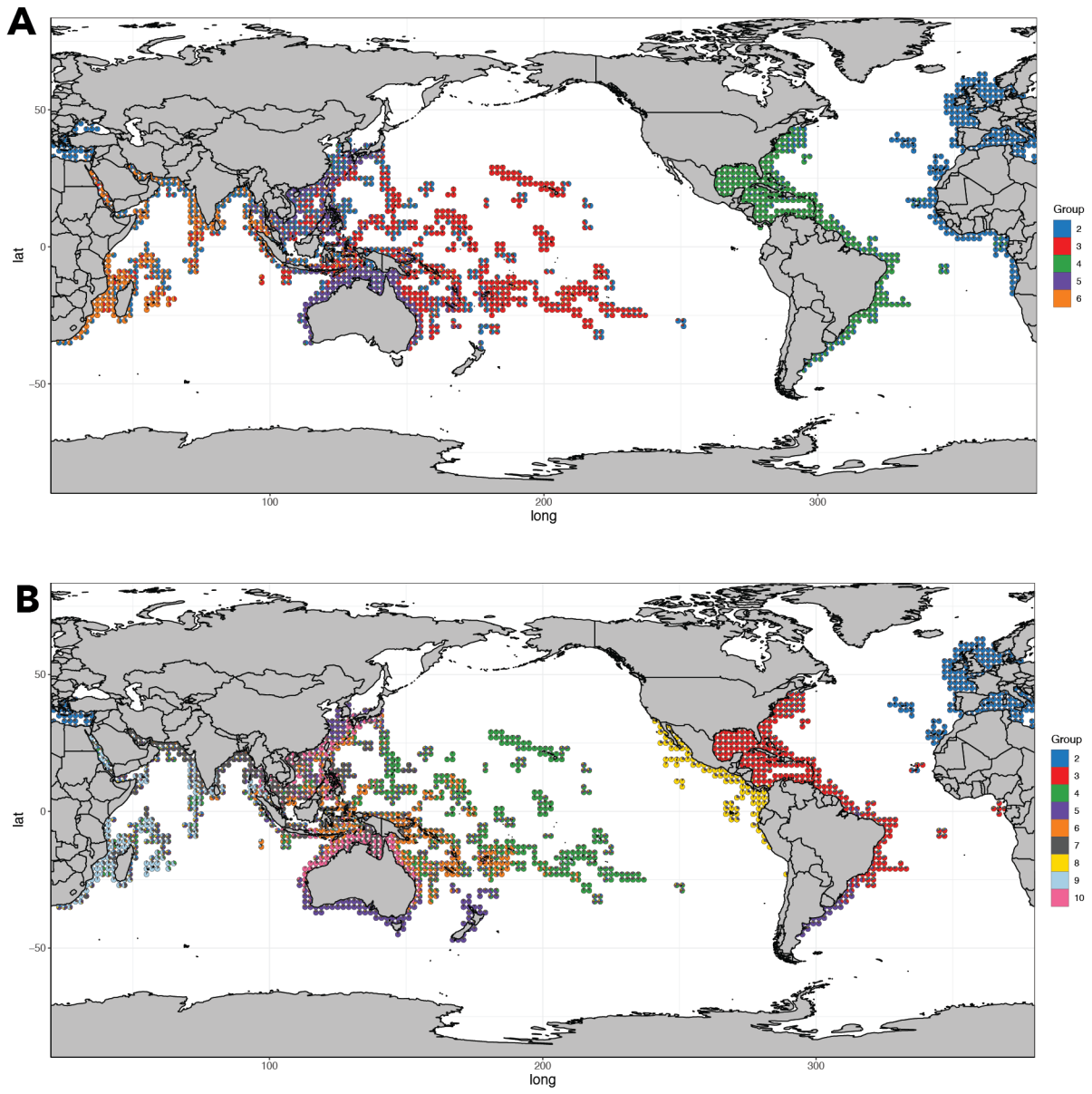

Figure 3.2: **Species motifs for goatfishes with A) k = 6 and B) k = 10.** Pie charts within each 2° by 2° grid cell represent the relative proportion of each species motif within that particular area, with colors representing the contributing motifs. Empty grid cells either had no species of goatfish present or relatively few species, which resulted in the dominance of a species motif that was driven by species with low richness and not patterns of co-occurrence.

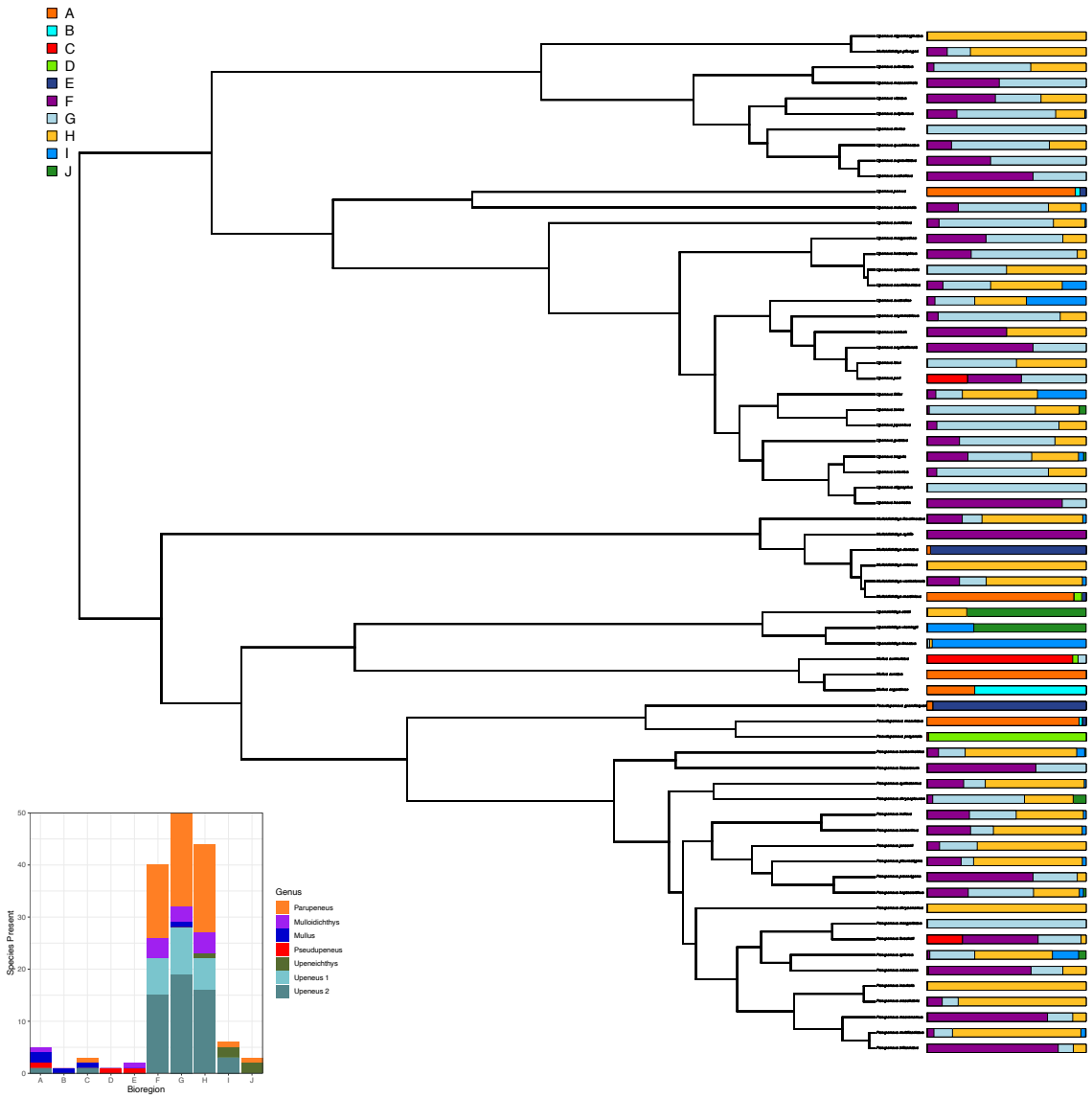

Figure 3.3: **Percentage of each species' range distributed within each bioregion.** Based on the geographic delineation of bioregions shown in Figure 3.1, the colored bars plotted on the tips of the time-calibrated phylogeny of the goatfishes represent the relative proportion of each species' range that is present within a particular bioregion. Colors shown are identical to the colors used in Figure 3.1 for useful comparison.

evidence that the bioregions tend to be more phylogenetically over dispersed, although we cannot reject random distributions (Table 3.S5). Additionally, our results show significant negative observed z-values in Indo-Pacific bioregions F, G, and H, which suggests there is less phylogenetic distance than expected by random sampling (Table 3.S5).

The results of computing PSV values for each grid cell reiterate the geographic trends found across bioregions, with grid cells in the Atlantic region consisting of on average higher PSV values compared to grid cells across the Indo-Pacific (Figure 3.4). Species richness across grid cells is not correlated with PSV values. When sub-setting the data to species within each genus, our results show that the PSV for all genera are randomly distributed except for *Mulloidichthys*, which is statistically over-dispersed (Figure 3.A.10). Our analysis of phylogenetic fields (PhyloFields) for each species showed that 26 species were phylogenetically clustered, which indicates that these species tend to occur with more closely related species across their range than expected by random sampling (Table 3.S6). The PhyloFields for the remaining species could not be rejected from statistically random (Table 3.S6).

3.4.3 Distribution patterns and range size evolution

Analyses revealed that the distribution of range sizes is skewed towards relatively smaller range sizes across species within the Mullidae, with a mean of 113 cells (range = 3 – 668, s.d. = 158.7, Figure 3.A.6). There are six species across two genera that have the minimum observed range size of three grid cells: *Parupeneus louise*, *P. orientalis*, *Upeneus dimipavlov*, *U. nirgromarginatus, U.randalli*, and *U. seychellensis*. On the contrary, there are five species across two genera that have range sizes larger than 500 cells: *Mulloidichthys flavolineatus, M. vanicolensis, P. barberinus, P. cyclostomus*, and *P. multifaciatus*, with *M. flavolineatus* maintaining the largest range size (Table 3.S7). Notably, the maximum range observed for a species with the genus *Upeneus* is 439 grid cells and the genus *Mullus* is 237 (Table 3.S7). We further examined the distribution of range size through focusing on the latitudinal and

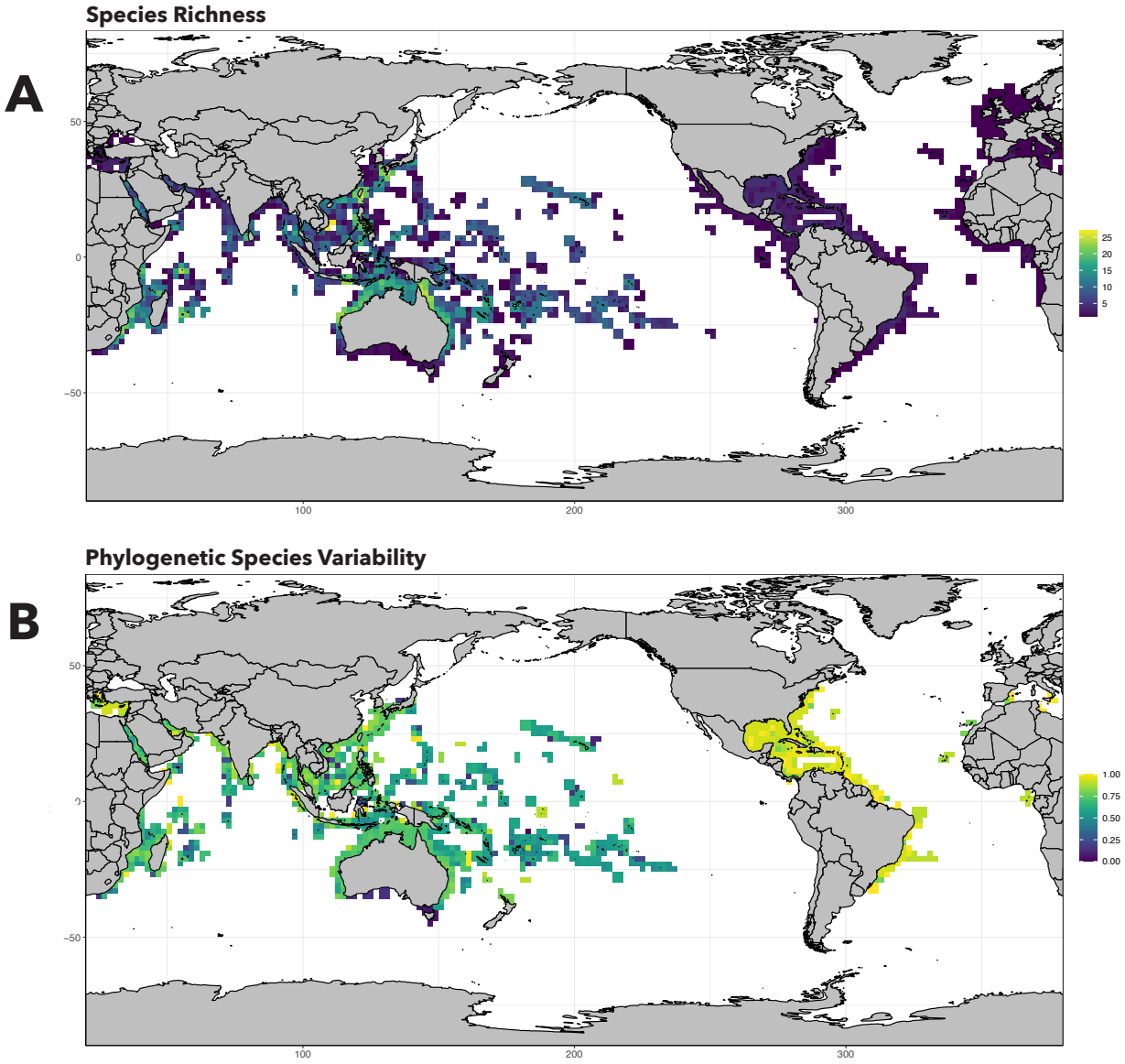

Figure 3.4: **Global map of species richness and phylogenetic species variability of the goatfishes.** This figure shows A) the total number species and B) the phylogenetic species variability (PSV) for all species present in each 2° by 2° grid cell. Higher values are shown in warmer colors and lower values are shown in cooler colors. Higher values of PSV correspond to grid cells with assemblages that are more phylogenetically distinct than would be expected by random sampling from the phylogeny.

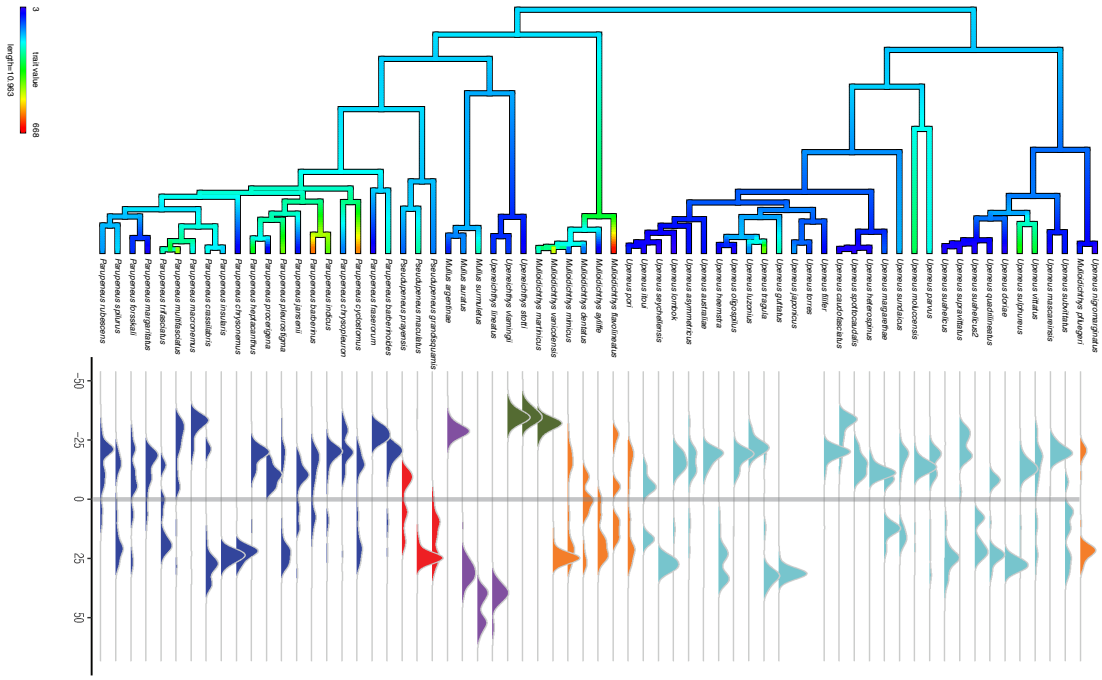

Figure 3.5: **Evolution of range size and latitudinal distribution of georeferenced coordinates across species of goatfish.** Branch colors correspond to the estimated range size, which is computed as the total number of 2° by 2° grid cells, along the branch lengths, with cooler colors representing relatively smaller ranges and warner colors representing larger range sizes. Density plots at the tips show the relative distribution of latitude of georeferenced coordinates for each species and are color coded by genus, with latitude 0 (equator) indicated by the gray line.

longitudinal gradients independently (Figure 3.5). We find evidence that the density of coordinates across species is anti-equatorial, defined here as the observation that the peak of coordinate point density is clustered either at above the equator, below the equator, or multiple peaks not centered on the equator (Figure 3.5). Individual species span on average 32.9 of latitudinal degrees and 91.8 of longitudinal degrees (Table 3.S7).

To test hypotheses about the evolution of range size in the context of the phylogeny, we find evidence that range size has a relatively low phylogenetic signal ($\lambda = 0.115$, $p = 0.10$; K = 0.072, $p = 0.51$) and does not fit the assumptions of Brownian motion. Rather, we find higher likelihood support for models of trait evolution that follow the assumptions of the Trend

Model and the OU model (Table 3.S8). However, given the stochastic nature of range size as a trait, we do not put further emphasis on explaining the underlying evolutionary process that led to the current distribution of range size across extant species. The distribution of range sizes for species within each bioregion are not statistically significant across regions, indicating that we cannot reject that species are spread randomly across groups. Despite this, we find evidence that the percent of a species range that falls within each cluster differs among species and genera (Figure 3.3; Table 3.S4).

3.5 Discussion

We used metrics of species turnover and grade of membership models to delineate distinct biogeographic regions for the global distribution of the goatfishes. A central conclusion of this study is that bioregionalization approaches identified ten distinct goatfish global bioregions, which broadly match other proposed regionalization schemes of global marine fauna (Spalding et al. 2007; Guillemot et al. 2011; Kulbicki et al. 2013; Mouillot et al. 2013; Parravicini et al. 2013; Leprieur et al. 2016; Cowman et al. 2017). However, we conclude that global goatfish bioregions contain varying levels of complexity, representing the influence of various biogeographical and ecological processes that drive patterns of co-occurrence. We find that the boundaries between goatfish assemblages differ in their amount of ambiguity, geographic continuity, and species turn-over, thus highlighting the importance of evaluating taxon-specific biogeographic regionalization patterns.

3.5.1 Bioregions at large spatial scales corroborate known biogeographical boundaries

The largest divergence in goatfish species turnover between bioregions is found between species that occur in the Indo-West Pacific compared to those that occur in the Atlantic/East

Pacific. The occurrence of barriers to dispersal between these regions for marine fauna has been well documented, with direct associations to historical vicariance events such the closure of the Tethys sea ∼12 Ma and the East Pacific Barrier (Bellwood and Wainwright 2002; Hughes et al. 2002; Barber and Bellwood 2005; Renema et al. 2008; Cowman et al. 2009). Within the Atlantic, we find evidence for four distinct bioregions, which can broadly be defined as the Western Atlantic, Southern Brazil, Northeastern Atlantic/Mediterranean, and Eastern Atlantic (Figure 3.1). These bioregions corroborate with the four primary provinces described within the Atlantic, which are separated by geographic barriers such as the Mid-Atlantic Barrier, the Terminal Tethyan Event, Amazon-Orinoco Plume, closure of the Isthmus of Panama, and the Benguela barrier (Floeter et al. 2008; Luiz et al. 2012). Our bioregionalization approach also supports the presence of bioregions that delimit the Western Indo-Pacific, Temperate Australasia, and the Eastern Tropical Pacific (Figure 3.1). The presence of barriers to dispersal that result in the delineation of these regions have been well documented, which limit the dispersal of planktonic larva due to variability in current patterns and water temperature (Briggs 1961, 1995; Bellwood and Wainwright 2002; Spalding et al. 2007; Gaither et al. 2016).

3.5.2 Central Indo-pacific assemblages are a mixing point of species with overlapping ranges

The advantage in using regionalization-based methods is the ability infer potential areas of dispersal barriers unique to the taxonomic group of interest (Kreft and Jetz 2010; Daru et al. 2017). We delineate goatfish bioregions in the Atlantic/East Pacific as geographically contiguous areas with clear barriers between each region (Figure 3.1). Although Mullidae species richness is comparatively low in this hemisphere, species' ranges have considerable overlap, resulting in distinct assemblage boundaries that are defined by shared species range boundaries and relatively higher amounts of endemism within each bioregion (Fig.1; Table

3.S3). For example, Bioregion A, which is distributed across the Western Atlantic and the Caribbean, has the highest percentage of endemism across all global bioregions (Table 3.S3). These patterns of geographic affinities are not observed in the goatfish bioregions found in the Indo-Pacific. Although it is possible to categorize most of the bioregions into a general geographic area, the boundaries between each bioregion are relatively ambiguous due to the high degree of mixing among grid-cell identities (Figure 3.1). Species richness in these areas are considerably higher, with the center of goatfish species diversity found within the center of the Indo-Pacific (Figure 3.4). This high degree of species richness is the result of overlapping ranges of species the primarily inhabit the Indian Ocean or Western Pacific bioregions (Table 3.S3). Notably, less that 10% of species within each Indo-Pacific bioregions are endemic, despite a 10-fold increase in the total number of species when compared to the Atlantic and East Pacific bioregions (Table 3.S3).

There are large-scale regional and historical factors that influence patterns of species richness across global marine assemblages and communities. Although these factors are often scale-dependent, it represents the complex influence of historical, biogeographical, and environmental processes on patterns of lineage diversification. The relatively depauperate species richness in the Atlantic and Eastern Pacific have been documented across various reef fish clades and reflects increasing isolation and lineage extinction within these regions (Floeter et al. 2008; Luiz et al. 2012; Cowman and Bellwood 2013; Parravicini et al. 2013; Cowman et al. 2017; Sambrook et al. 2019). Although this pattern could be exacerbated by the smaller geographic scale of this region, it is likely that several wide-scale disturbances and episodic extinctions have influenced current biodiversity patterns. For example, fossil evidence suggests a 20-30% genus-level extinction in the Western Atlantic during the Pliocene due to wide-scale disturbances and episodic extinctions (Vermeij 2001, 2005; Floeter et al. 2008). Additionally, a reduction in the availability of shallow water habitats as a result of Pleistocene glaciations are predicted to have impacted rates of diversification within the

tropical Atlantic (Floeter et al. 2008). However, evidence suggests that Indo-Pacific is a geographic area defined by connectivity and the presence of semi-permeable barriers to dispersal (Cowman and Bellwood 2013; Parravicini et al. 2013; Cowman et al. 2017) . Patterns of dispersal across several reef associated families indicate that the Central Indo-Pacific was historically associated with higher rates of lineage survival, with the adjacent Indian Ocean and Western Pacific acting as macroevolutionary sinks since the Miocene (Cowman and Bellwood 2013). Additionally, coral reef benthic complexity is highest in the center of the Indo-Pacific when compared to other reef systems (Li and Asner 2023).

The complex patterns of species connectivity in the Indo-Pacific are evident when examining the distribution data through the lens of species motifs. Although the Indo-Pacific is defined by three major bioregions, the intermixing of species motifs demonstrates that these boundaries in turnover are permeable (Figure 3.2). Although several species of goatfish can maintain large panmictic distributions across ocean basins, this pattern of delineation does not appear to be driven by these wide-ranging species (Lessios and Robertson 2013; Szabó et al. 2014). Rather, these regions within the Central Indo-Pacific, often termed the Coral Triangle, represents the nexus of species' ranges that span some combination of west into the Indian Ocean, east into the Pacific Ocean, north into the South China Sea, or south towards the Coral Sea (Hoeksema 2007). This is also evident by the increasing degree of geographic continuity across the boundaries between bioregions as you move away from this central point (Figure 3.1 & Figure 3.2). It is likely that the relatively recent changes in sea level and shallow coastline availability in the Coral Triangle have played a role in shaping the patterns of co-occurrence of goatfishes in this region (Briggs 1995). The end of the Miocene is associated with an increase in the interchange of Indo-West Pacific fauna caused by the separation of the Arabian and African plates, despite the complete separation of the Mediterranean Sea from the Red Sea and the Indian Ocean due to terrain uplift (Briggs 1995).

3.5.3 Patterns of phylogenetic diversity differ among assemblages of goatfish

Species within bioregions located in the Atlantic are phylogenetically over-dispersed. This result is well supported across all phylogenetic diversity metrics used in this study (Table 3.S5). Although the Atlantic hosts the lowest species richness, the species that are present represent a minimum of two genera in each bioregion. Given that these species are primarily associated with longer branch lengths than average, it is likely that they did not migrate into the Atlantic recently. Additionally, given the low species richness particularly within a genus, we find limited evidence to support speciation occurring in these regions. Rather, it is more likely that there has been substantial extinction across the Atlantic bioregions, which has been noted by the variable diversification rates of goatfishes throughout their evolutionary history (Nash et al. 2022). Reduced rates of diversification have been observed across several fish clades given the relatively unstable geologic history within the Atlantic Ocean and East Pacific (Briggs 1961; Floeter and Gasparini 2000; Lessios and Robertson 2006; Bender et al. 2013; McCord et al. 2021; Waechter et al. 2021). The only exception can be found within the genus *Mullus*, in which all five of the described species occur within the Atlantic Ocean or the Mediterranean Sea. Despite this, only two species within this genus marginally co-occur within this geographic area, as *Mullus auratus* is distributed along the eastern coast of North America south to French Guiana and *M. argentinae* is primarily distributed from central Brazil south to southern Argentina, although there is a record of this species off the coast of French Guiana.

Our results show that bioregions within the Indo-Pacific are phylogenetically clustered with relatively similar amounts of phylogenetic diversity represented across the area (Table 3.S5). This finding is a bit counter intuitive given that rates of endemism within the bioregions in the Indo-Pacific are comparatively low despite higher levels of species richness (Table 3.S3). However, it is important to note that the Indo-Pacific is primarily occupied by taxa within the most species rich genera: *Parupeneus* and *Upeneus*, whose presence

would lead to a higher degree of phylogenetic relatedness (Table 3.S4). Bioregion J has the lowest PSV values, which makes sense given that this bioregion is geographically restricted to the southwest coastline of Australia and is primarily occupied by species within the genus *Upeneichthys* (Figure 3.1; Table 3.S3). This genus is unique among other genera, in which species occur only in association with the temperate waters of Southern Australia and New Zealand (Gosline 1984; Platell et al. 1998).

Although we do find evidence that many species are shared amongst at least two of the bioregions in the Indo-Pacific, there are still considerable amounts of species turnover and variability in range size across the entire area (Table 3.S3 & 3.S7). One possible explanation is that our bioregion delimitation is sensitive to wide-spread species, which results in high numbers of shared species across regions (Kreft and Jetz 2010; Tittensor et al. 2010). Despite this, there is notable variation among the relative proportion of a species' range in a bioregion even if they are present across multiple bioregions (Figure 3.3). For example, we find that an average of 74% of the ranges of *Mulloidichthys* are located within Bioregion H compared to only 39% for species within *Upeneus* (Figure 3.3). Although this level of variability is not directly captured by the delimitation of boundaries between bioregions at this relatively broad geographic scale, it is captured at higher cluster numbers with both hierarchical clustering and grade of membership methods (Figure 3.2 & 3.S4).

3.5.4 Implications of range size evolution and historic dispersal patterns

The phylogenetic structuring across the Indo-Pacific is of particular interest given that most species of goatfishes diverged within the past ∼5 Ma (Nash et al. 2022). Despite evidence of phylogenetic clustering on a phylogeny-wide scale, our findings show these recent divergences are likely not the result of radiations within a particular bioregion, which is unusual for a clade with this pattern of diversification (Nash et al. 2022). Population level studies have shown that species of goatfish have been able to maintain panmictic popula-

tions across large geographic scales due to their mass spawning behavior and long pelagic larval duration (Lessios and Robertson 2013; Fernandez-Silva et al. 2015). Interestingly, evidence for population isolation and differentiation was observed primarily at the periphery of the species' range (Fernandez-Silva et al. 2013, 2015; Lessios and Robertson 2013), which suggests that speciation is more likely to occur at the extent of species ranges due to high rates of dispersal.

The range size of a species can influence the likelihood of extinction and diversification of a lineage, especially within coral reef regions (Lester and Ruttenberg 2005; Leis 2007; Luiz et al. 2013). We find evidence that the distribution of range size across the goatfishes has low phylogenetic signal and does not follow the assumptions of Brownian Motion (Figure 3.5, Table 3.S8). Goatfish disperse through a pelagic larval stage, in which the larva can remain in the plankton for relatively long amounts of time and large quantities (McCormick and Molony 1995; Pavlov et al. 2011; Hernández et al. 2023). Traits associated with pelagic larval dispersal have been hypothesized to influence the range size of a species, although adult life history traits affect these patterns as well (Lester and Ruttenberg 2005; Mora et al. 2012; Luiz et al. 2013). Several species of Mullidae exhibit exceptionally large range sizes (Figure 3.5, Table 3.S7). For example, *Mulloidichthys flavolineatus* has been shown to maintain gene flow from the coast of Eastern Africa to the Hawaiian Archipelago, which is an incredibly large geographic distance for a non-migratory species of this size (Fernandez-Silva et al. 2013; Lessios and Robertson 2013). However, the maintenance of these large geographic ranges is likely associated with dispersal during the larval phase in conjunction with other ecological factors such as habitat availability and competition, given that the post-settlement individuals likely inhabit a relatively small home range (Meyer et al. 2000; Mora et al. 2012).

There have been several hypothesized pathways of historic dispersal among the goatfishes, from vicariance events that isolate circumglobally distributed lineages to more recent

migration across the East Pacific Barrier, Tethyan Sea, and Isthmus of Panama (Lessios and Robertson 2013; Santaquiteria et al. 2021). However, there is still uncertainty regarding dispersal patterns given the variable diversification rates, asynchronous geographical distributions of most genera, and a lack of comprehensive fossil deposits (Nash et al. 2022). It is well understood that variation in community structure is driven by the complex interactions among ecological and evolutionary processes, such as diversification rates, dispersal capabilities, and species interactions (Bradbury et al. 2008; Luiz et al. 2012; Cowman and Bellwood 2013; Kulbicki et al. 2013; Cowman et al. 2017; Braga et al. 2023; Pinheiro et al. 2023). However, these processes are scale dependent and therefore may alter the interpretation of how current distribution patterns arose. In the goatfishes, we find evidence that the primary boundary associated with geologic events is the separation of the Indo-Pacific and Atlantic/East Pacific regions (Figure 3.1 & Figure 3.2). However, at smaller spatial scales within the Indo-Pacific, the overlap in species composition between bioregions suggests that any geographic barriers are semi-permeable and ecological process, such as species interactions and habitat composition, are at play (Figure 3.1 & Figure 3.2). These ecological processes that may influence goatfish distributional patterns, such as coral reef health, competitive exclusion, and niche partitioning, have not been extensively studied across the family (Pulliam 2000; Wiens and Graham 2005; Ackerly et al. 2006; Pagel and Schurr 2012; Winemiller et al. 2015). However, evidence suggests that species use habitat partitioning to reduce competition in areas of co-occurrence (Platell et al. 1998; Krajewski et al. 2006; Galarza et al. 2009; Mittelheiser et al. 2022).

In conclusion, the bioregionalization of goatfish co-occurrence patterns reveals the impact of vicariance and dispersal on the current distribution and evolution of species within the Mullidae. The boundaries between these assemblages can be used to better understand potential barriers to dispersal for species of goatfish, which can then be used to extrapolate the impact of current ecological and geological changes to species distributions across

their ranges. The comparison of dispersal barriers among different taxa is crucial for the proper monitoring and conserving of lineages and will help us to better predict future species distributions and evolution.

APPENDIX

3.A Supplementary Figures

Supplementary tables for this chapter can be viewed at `https://figshare.com/s/b93c70b7b67fccb3c82f`

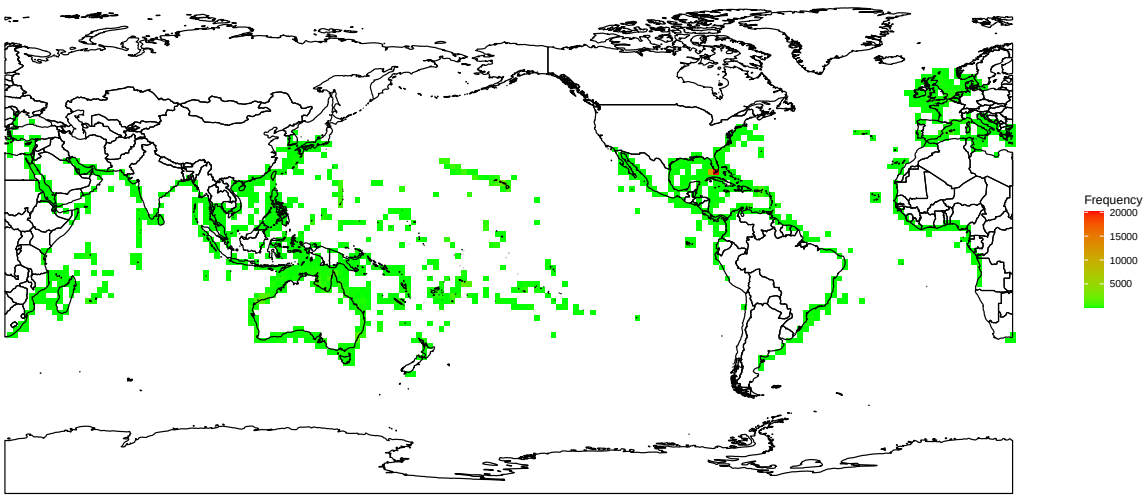

Figure 3.A.1: **Geographic distribution and frequency of all georeferenced coordinates used in this study.**

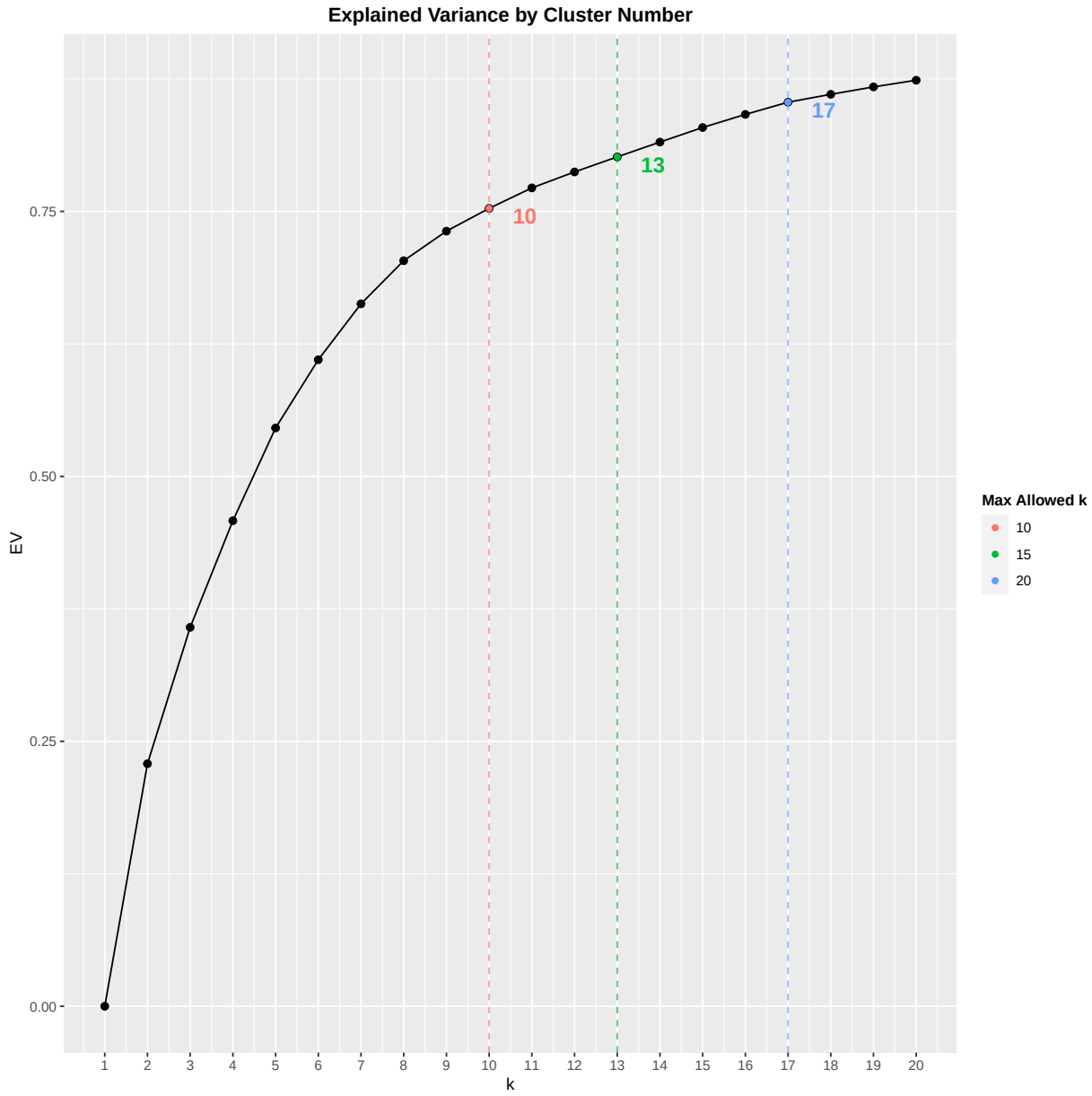

Figure 3.A.2: **The amount of explained variance by increasing the number of assigned clusters.** Using maximum k values of 10, 15, and 20, colored values and dashed vertical lines show the optimal number of clusters using the elbow method based on each maximum allowed value, respectively.

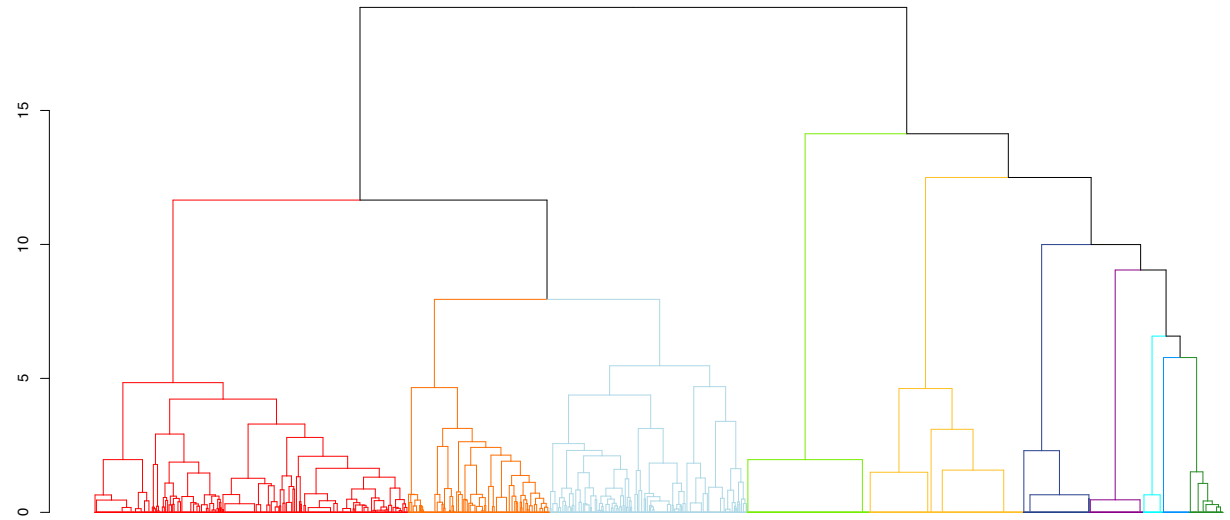

Figure 3.A.3: **Dendrogram of bioregions resulting from Ward's hierarchical clustering of grid cell assemblages with k = 10 based on βsim dissimilarity matrices for species of goatfish.** Colors represent the assigned cluster group to each 2° by 2° grid cell, which are represented by each tip on the dendrogram. Colors are identical to those used in Figure 3.1 for comparison.

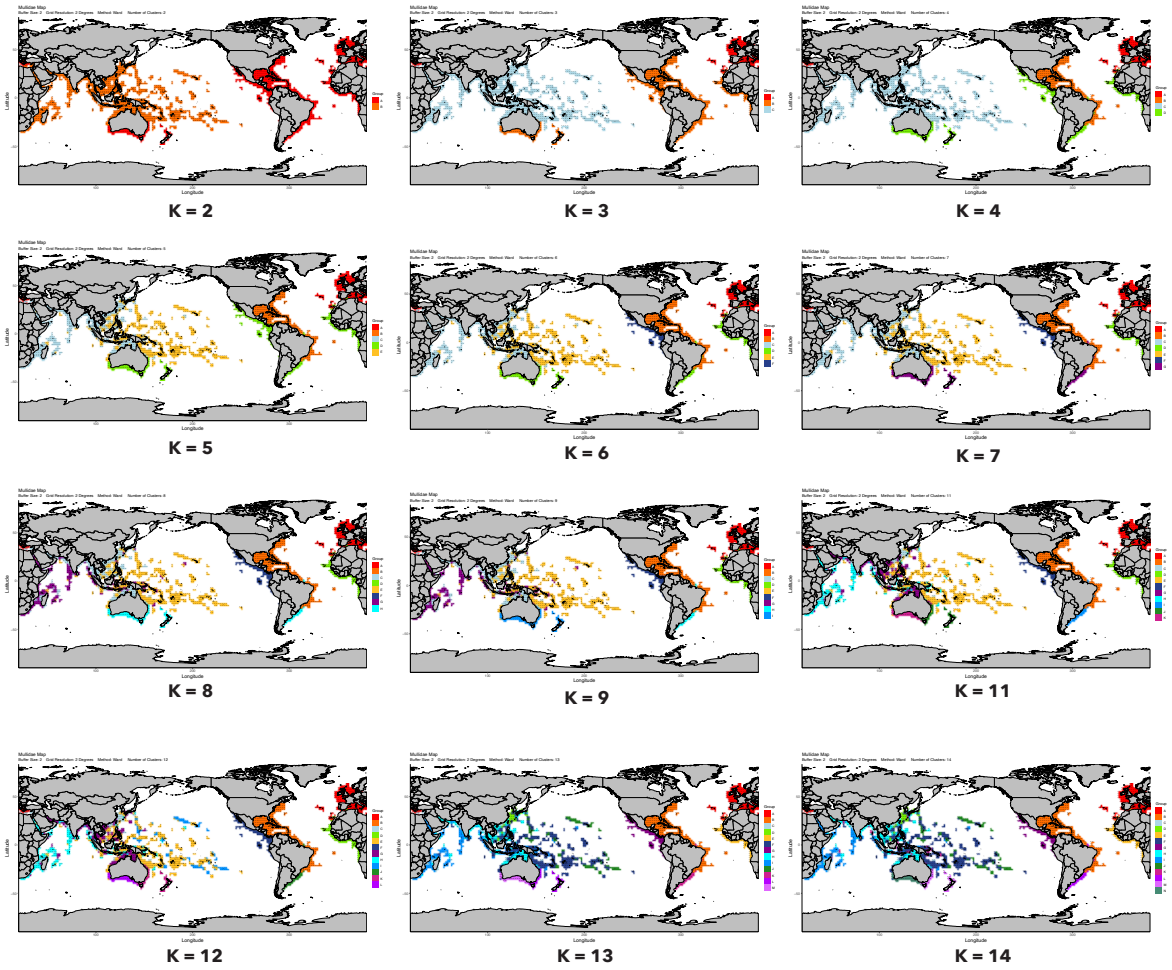

Figure 3.A.4: **Map of bioregions resulting from Ward's hierarchical clustering of grid cell assemblages with k = 2 to k = 14 based on βsim dissimilarity matrices for species of goatfish.** Colors represent the assigned cluster group to each 2° by 2° grid cell, with letters added for additional clarity.

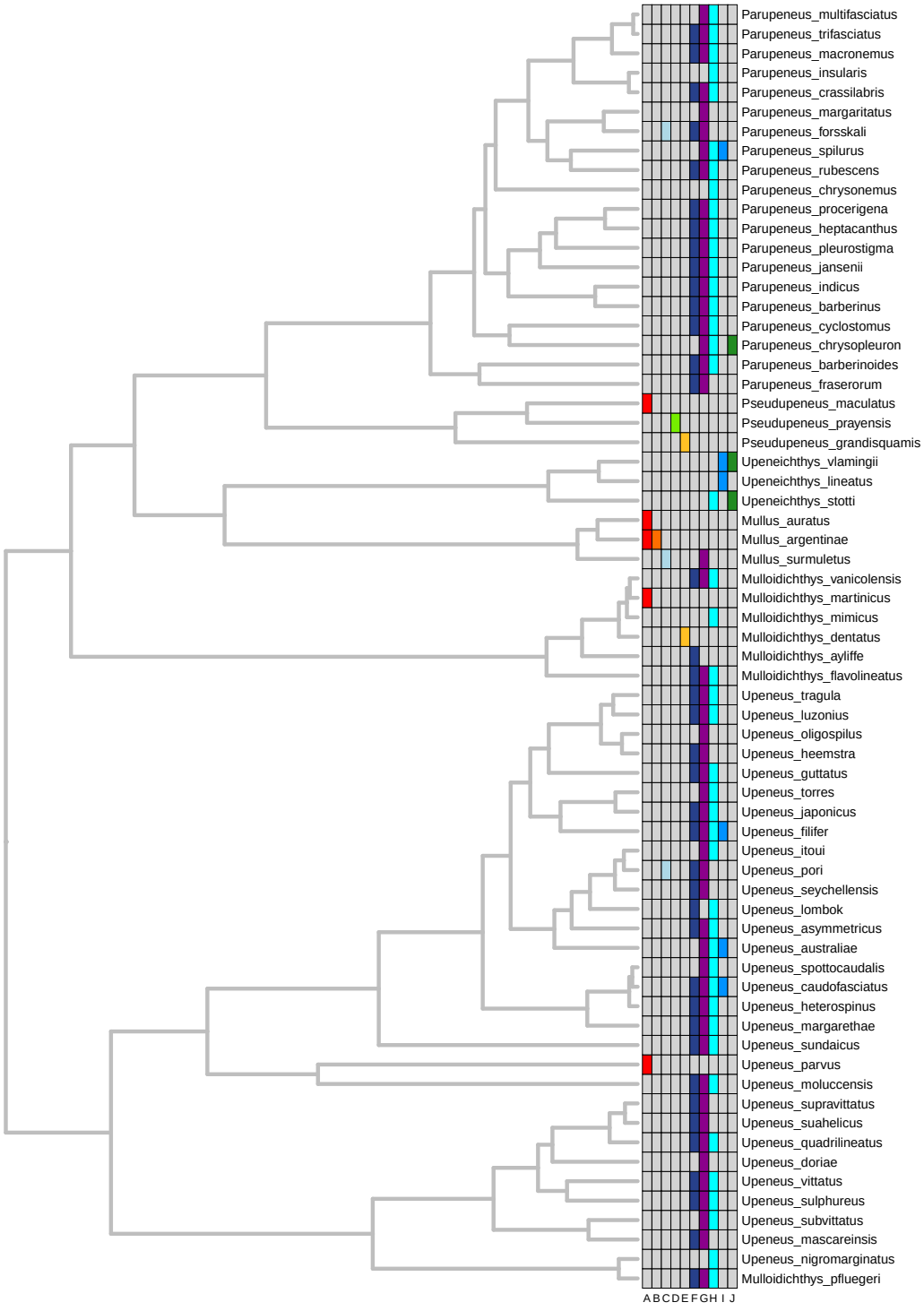

Figure 3.A.5: **Presence or absence of species of goatfish within each bioregion.** Using the geographic delineation of bioregions shown in Figure 3.1, colored in boxes, which are identical to colors used in Figure 3.1 indicate present, whereas a gray shaded box indicates the absence of a species.

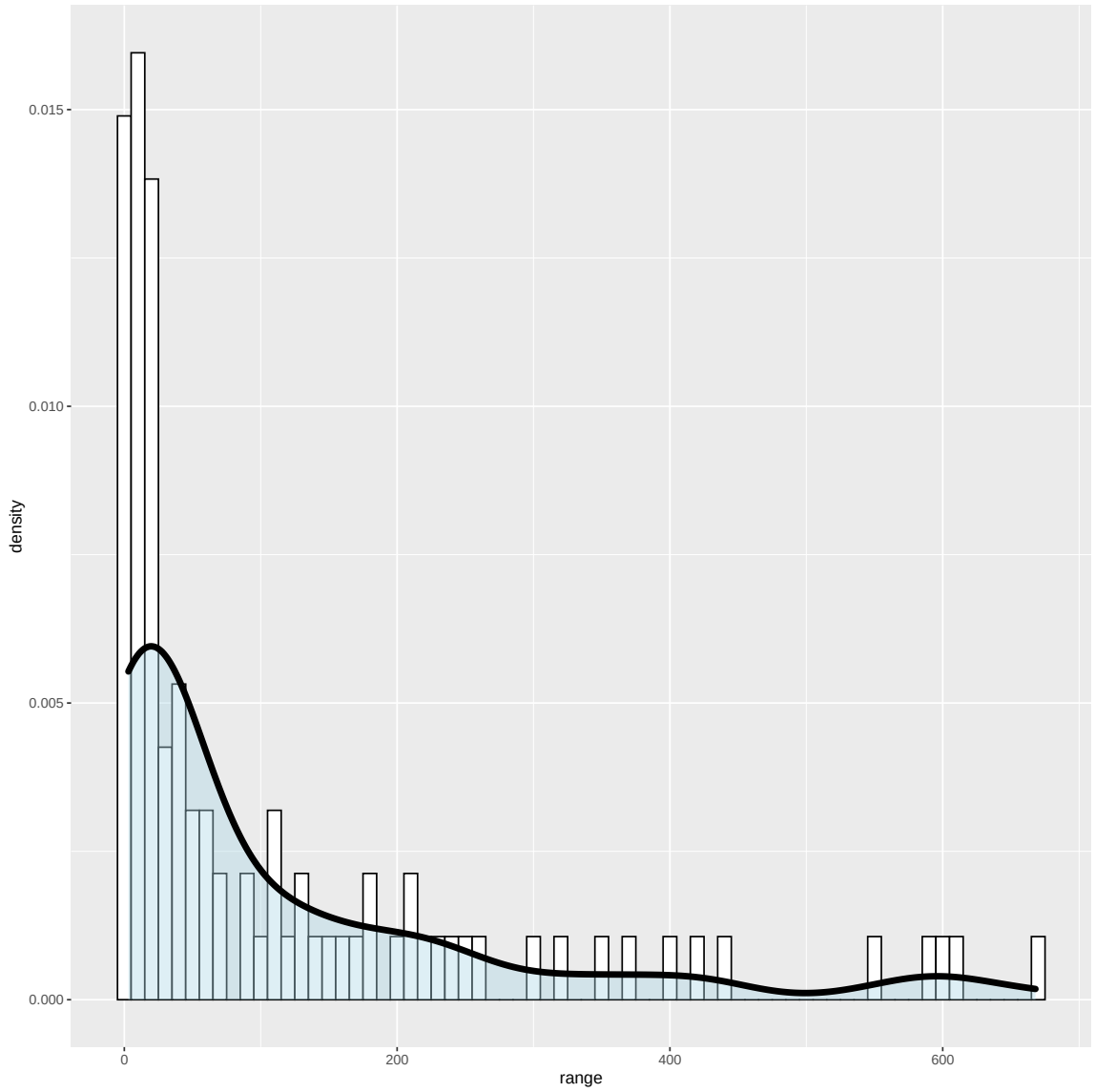

Figure 3.A.6: **Distribution of range size across species of goatfishes.** White bars represent the number of species with a particular range size, and the black line and blue shaded area indicates the density plot of the same dataset.

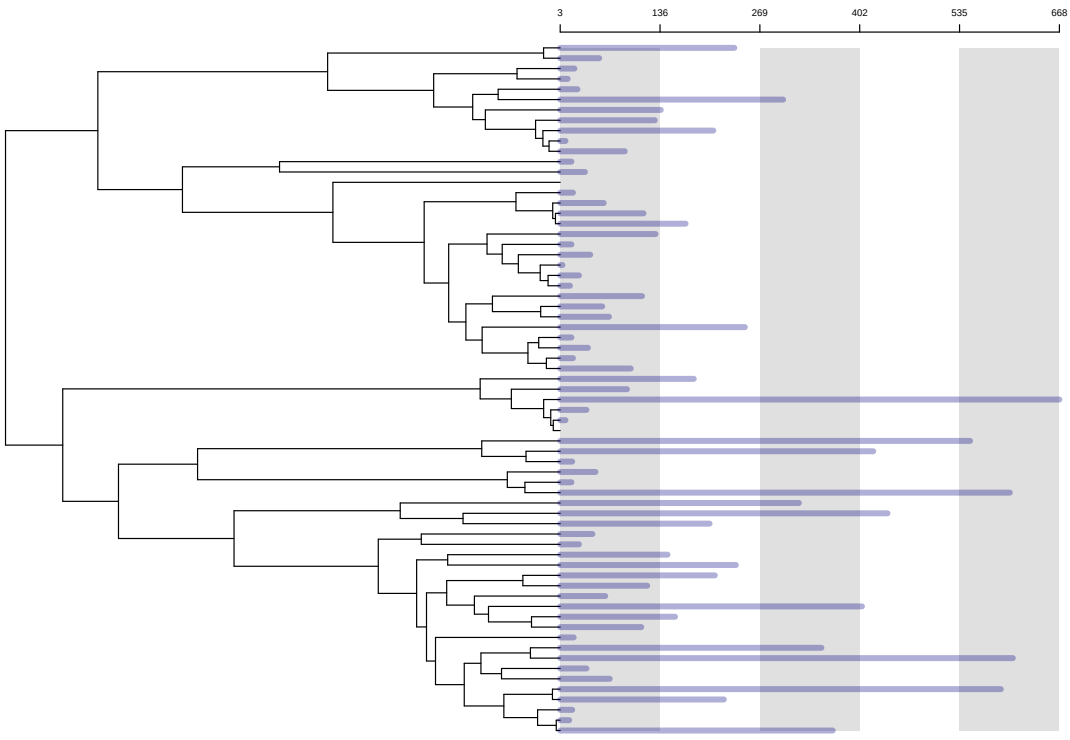

Figure 3.A.7: **Distribution of range size across the time calibrated phylogeny of goatfishes.** Range size is indicated by the the relative size of the blue bar.

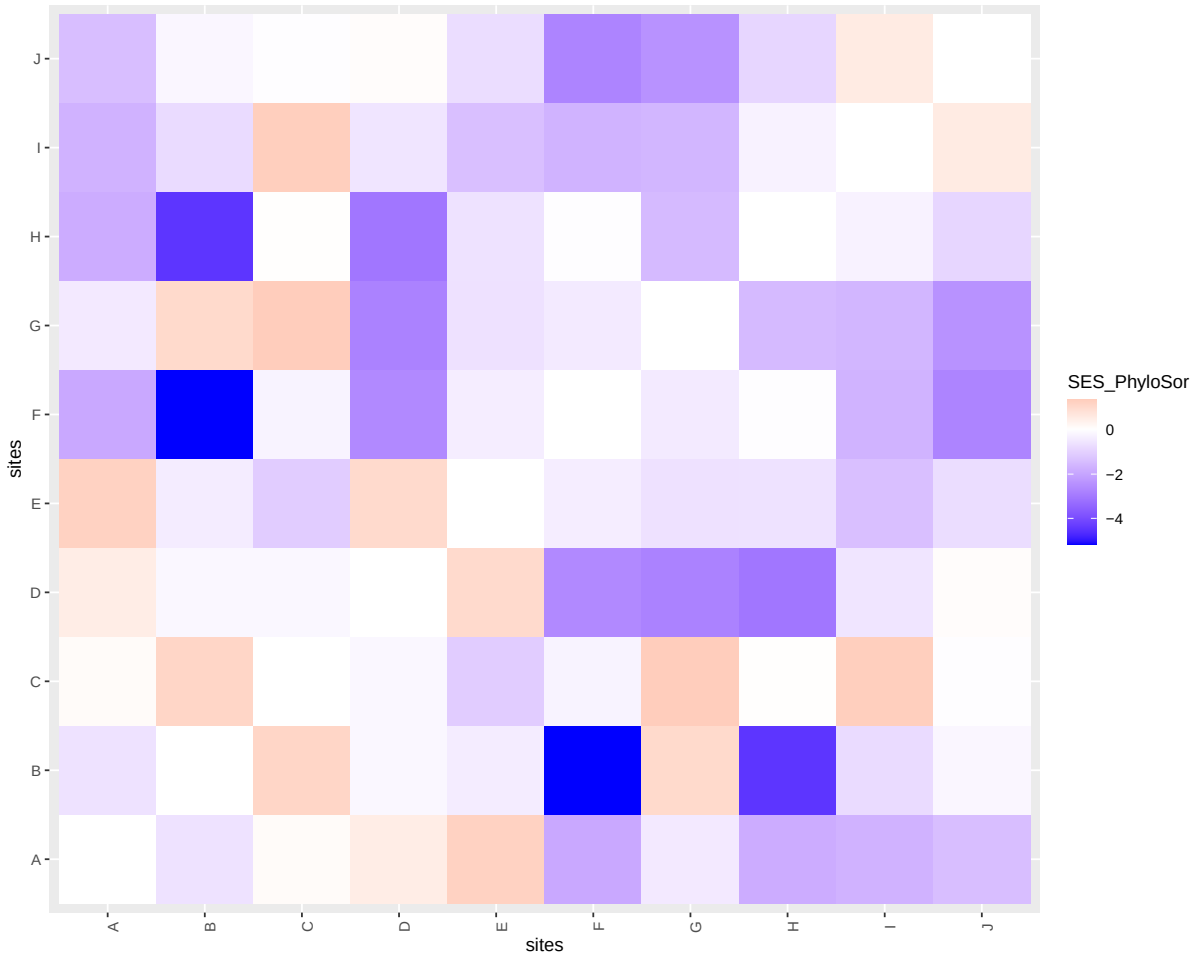

Figure 3.A.8: **Heatmap showing the pairwise comparisons of phylogenetic similarity between bioregions using Sorenson's Index.** Cooler colors indicate more phylogenetically similar and warmed colors indicated more phylogenetically distinct.

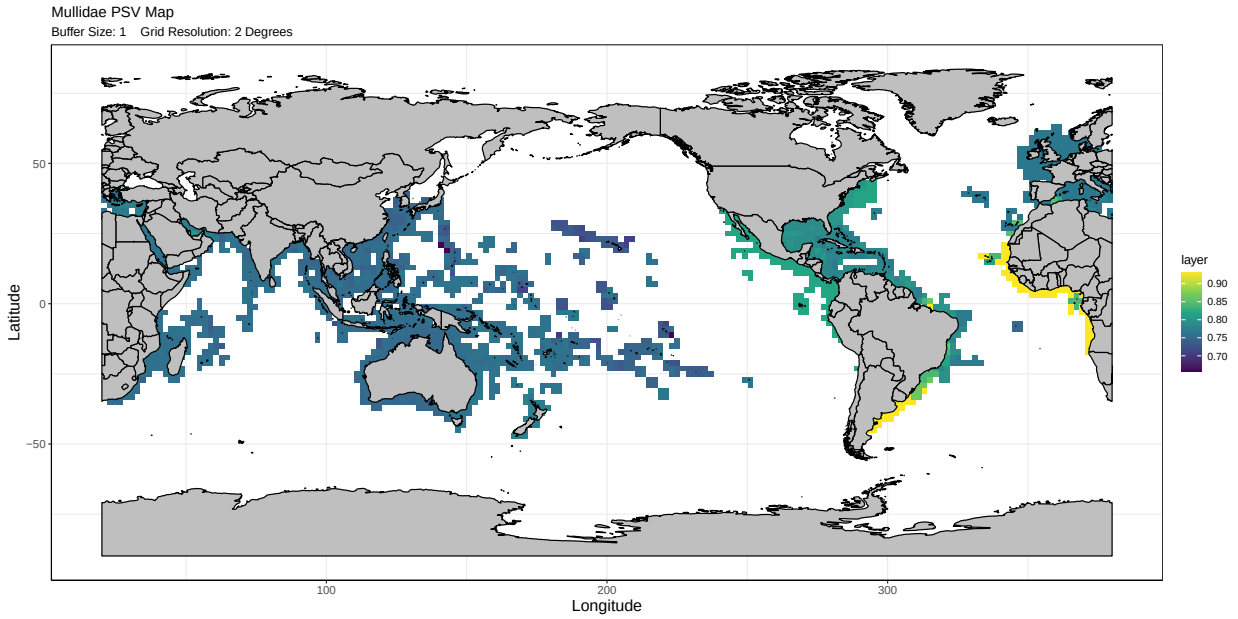

Figure 3.A.9: **Geographical representation of the mean phylogenetic field values (PSVsp) for each focal species present within a 2° by 2° grid cell.** Warmer colors indicated higher mean values of PSVsp, which indicates higher levels of phylogenetic distinctiveness across the focal species' ranges. Cooler colors represent the converse.

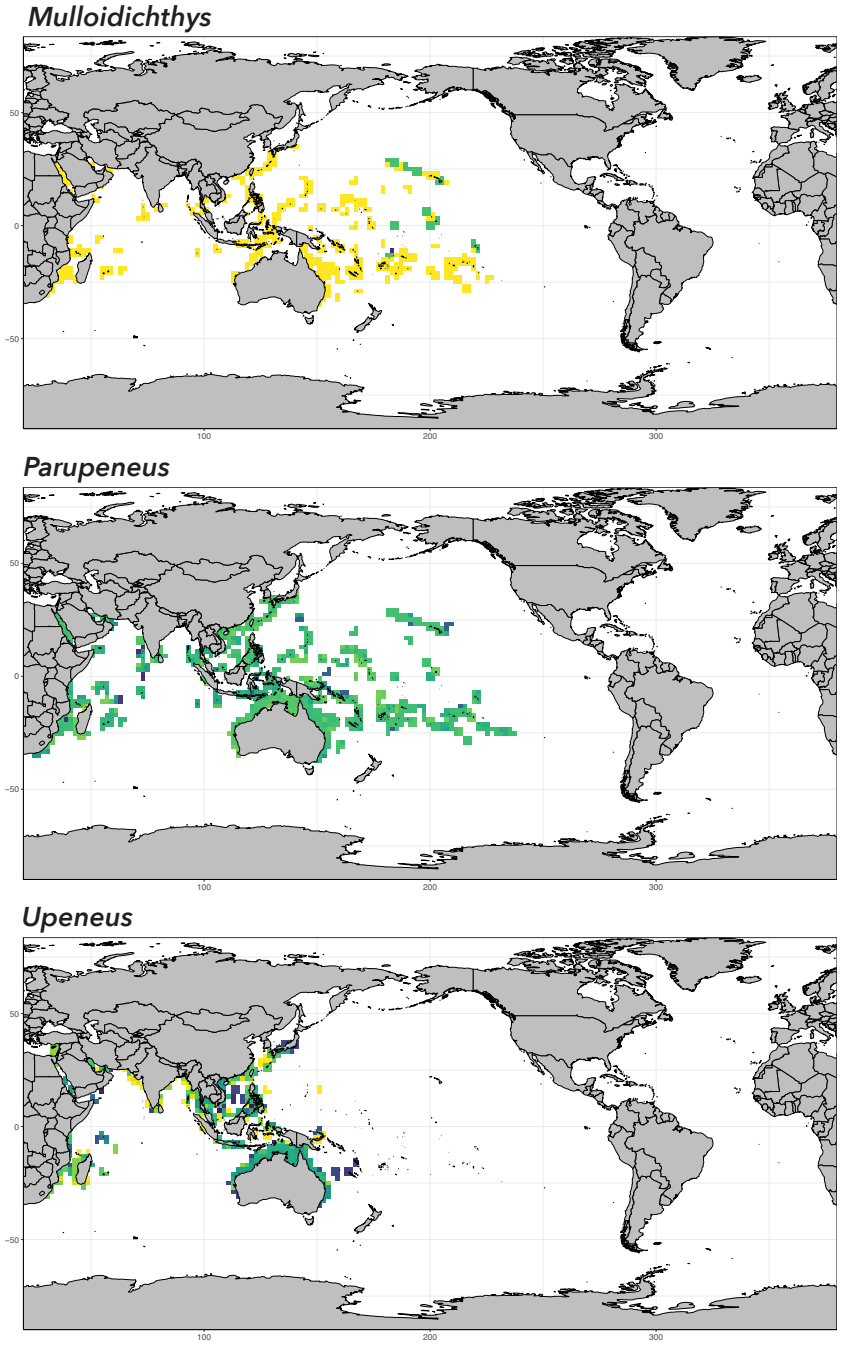

Figure 3.A.10: **Geographical representation of the phylogenetic species variability (PSV) values by genus.** Maps were generated by subsetting the dataset to species within each genus and computing the relative PSV values for species present within a 2° by 2° grid cell. Warmer colors indicated higher mean values of PSV, which indicates higher levels of phylogenetic distinctiveness across the focal species' ranges. Cooler colors represent the converse. Species within the genus *Mulloidichthys* have statistically significant values of phylogenetic overdispersion when compared to random sampling from the phylogeny.

CHAPTER 4

GEOMETRIC MORPHOMETRICS AND THE EVOLUTIONARY PATTERNS OF SUBSTRATE FORAGING IN THE GOATFISHES

4.1 Abstract

Tropical marine fishes are a spectacular case study in how the evolution of specialized ecological niches can greatly impact rates of phylogenetic and morphological change. The goatfishes (Family Mullidae) are diverse in body form and foraging behavior, yet remain relatively unexplored in their evolutionary ecomorphology. The synapomorphy that unites the goatfishes is the presence of hyoid barbels, a novel exploratory and excavation tool with important sensory abilities, enabling goatfish species to employ the barbels and other behaviors to access food within benthic substrates. Goatfishes preferentially forage on different substrate types, which results in habitat partitioning across reef systems. In this study, we integrated a recent phylogeny of the goatfishes with detailed morphometric and ecotype trait analyses to test hypotheses about the evolution of ecomorphological associations with substrate preference. We performed a comprehensive geometric morphometric analysis on various morphological subsets, such as forehead shape, body shape, and fin shape, across 72 species of goatfish. We used a novel computational approach to standardize the positions of fins and barbels, allowing detailed morphometric analysis of the entire goatfish body plan. We find evidence that body shape, head shape, and fin profiles are significantly associated with substrate preference across the family. We show that preferential foraging on hard substrates evolved relatively recently within the genus *Parupeneus*, and is associated with an elongated snout with a prominent rostrum, shorter body lengths, greater body depths, longer barbels, and larger head sizes. Within this genus, full body shape and caudal fin shape are highly significantly associated with hard substrate preference. Additionally, we find evi-

dence that the rates of morphological diversification are significantly faster in hard substrate species. In conclusion, we find robust evidence for an ecomorphological relationship between substate preference and morphological diversity[1].

4.2 Introduction

One of the most important ecological factors that drive the exceptional diversity of form and function within communities is trophic dynamics, in which lineages often evolve adaptative characteristics that exploit prey capture and predation avoidance (Hobson and Chess 2001). Due to the large variability within and across habitats, a comprehensive examination of the evolution of morphological diversity through the lens of ecological adaptations is essential to understanding the distribution of taxa across space and time (Lombarte and Aguirre 1997). The interactions between morphological traits and ecological specialization have been extensively examined across a wide range of fishes due to their extraordinary diversity (Westneat 1995; Antonucci et al. 2009; Davis et al. 2016; Davis and Betancur-R 2017; Cardozo-Ferreira et al. 2018; Rincon-Sandoval et al. 2020; Wilson et al. 2020; McCord et al. 2021; Evans et al. 2022; Mittelheiser et al. 2022). Coral reef fishes have been a particularly useful system in which to examine ecological associations of morphological diversity, with ecomorphological trends being revealed with an integrative approach involving phylogenetics, morphometrics, and ecological trait analysis (Bellwood and Wainwright 2002; Rocha et al. 2007; Brandl et al. 2015; Bellwood et al. 2016; Bridge et al. 2016; Grenié et al. 2018).

This biodiversity across reef systems was historically thought to act as a buffer against the loss or extinction of species. However, variability in the spatial distribution of lineages that occupy important ecological niches may result in greater rates of extinction than previ-

1. This chapter will be published under the following citation: Nash, C. M., Lungstrom, L. L., & Westneat, M. W. Geometric Morphometrics and the Evolutionary Patterns of Substrate Foraging in the Goatfishes (Syngnatharia: Mullidae). *In Prep.*

ously thought (Villéger et al. 2010, 2017; Mouillot et al. 2014; Grenié et al. 2018; Leprieur et al. 2021). Divergent morphological forms in closely related, sympatric species often evolve to facilitate habitat partitioning along a gradient, such as depth or substrate type, or to reduce interspecific competition through differing feeding strategies (Via et al. 1995; Bouton et al. 2002).

In ray-finned fishes, a major axis of morphological variability can be found within the craniofacial region, as these structures are essential for niche specialization as a response to shifts in resource availability (Gans and Northcutt 1983; Wainwright et al. 2004; Westneat et al. 2005; Roberts et al. 2011; McCord and Westneat 2016; Martinez et al. 2018). This association between niche partitioning and cranial evolution forms the basis of the concept of ecomorphology, in which variation in cranial morphologies will result in variation in feeding ability that will subsequently influence performance characteristics for predatory behaviors and diet (Wainwright and Richard 1995). Although the goatfishes represent a key part of reef ecosystems, few studies have examined their ecomorphological diversity and its implications on the evolution of this clade (Mittelheiser et al. 2022).

The goatfishes (Syngnatharia: Mullidae) are a diverse, globally distributed family with 104 described species found in close association with coral reef ecosystems (Santaquiteria et al. 2021; Fricke et al. 2022; Nash et al. 2022). Goatfishes occupy a distinct role in coral reef ecosystems due to an innovation in their feeding morphology, a pair of highly specialized hyoid barbels. These barbels are fleshy extensions located under the jaw that are capable of chemoreception and prey excavation through the resuspension of otherwise inaccessible food sources within various benthic substrates (Gosline 1984). These barbels are unique in both their morphology and function through a variety of behaviors (Gosline 1984; McCormick 1993; Lombarte and Aguirre 1997; Kim et al. 2001; Lukoschek and McCormick 2001; Kirino et al. 2006; Krajewski et al. 2006).

Different species of goatfish have been shown to preferentially forage on a particular

substrate type (Gosline 1984; Mittelheiser et al. 2022). For example, some species use their barbels as an excavation device to turn over soft substrate, such as sand, mud, and gravel, while other species prefer to forage on harder substrates, such as coral rubble, using their barbels as probes within crevices to dislodge prey (Sierra et al. 1994; Lukoschek and McCormick 2001; Nakamura and Sano 2004; Krajewski and Bonaldo 2006; El Bakali et al. 2010; Mittelheiser et al. 2022). Interestingly, despite evidence of habitat and niche partitioning, the diets of sympatric species overlap to a large extent and consist primarily of infaunal invertebrates (Uiblein 1991; Platell et al. 1998; Uiblein et al. 1998; Lukoschek and McCormick 2001; Kolasinski et al. 2009; Wismer et al. 2009). These behaviors have been shown to play an important role in maintaining coral reef diversity and community composition, in addition to acting as potential indicators of coral reef health (Lukoschek and McCormick 2000; Uiblein 2007; Russ et al. 2015).

In this study, we test hypotheses about the relationship of preferential foraging on different substrate types with morphometric variation of the head, body and fins across the phylogeny of goatfishes. Specifically, this paper has two central goals: test the association of head, body and fin shape with substrate preference and examine evolutionary trends within each foraging mode across the goatfishes. We use a robust geometric morphometric approach to quantify body, cranial, barbel, and fin shapes. To examine evolutionary patterns of shape change in the fins and barbels, we developed a new approach to standardizing morphometric positions of soft tissues in morphometric analysis. We hypothesized that body shape, in particular the barbel and cranial shape, would differ between ecotypes that primarily forage on sand versus those that fed primarily on hard substrates (coral reef and rubble) given that that the barbel and cranium interact directly with the substrate during foraging and feeding. We use an integrative approach to examine the evolution of substrate preference, patterns of morphological diversification, and implications for evolutionary ecomorphology. This exploration of the complex interrelationships among morphological diversity, foraging behavior,

and evolutionary trends across species of goatfish can be used to inform the potential for adaptability to various habitats due to changing environments.

4.3 Methods

4.3.1 Taxon Sampling and image analysis

In order to explore the geometric morphometrics and ecomorphological relationships of goatfishes in a phylogenetic context, we examined 72 species of goatfishes that were included in the most recent time-calibrated phylogeny from Nash et al. (2022). This represents taxonomic coverage of 69% of the 104 described species described within Mullidae. We collected 2D lateral specimen photographs from museum collections, personal databases, primary literature, and aquarium trade, resulting in a total of 494 adult individuals across 6 genera within Mullidae. A total of 83 landmarks (47 fixed landmarks and 45 semi-landmarks) were placed along the specimen image using the R package StereoMorph 1.6.5 (Olsen and Westneat 2015).

4.3.2 Computational biomechanical rotations of barbel, jaws, body and fins

Specimen photos were often variable in the fin angles, the tail spread, the barbel curvature and angle, and the degree to which the lower jaw was open. This variability is an artifact of preservation due to the fin or barbel positions when the photograph was taken. We developed a desktop app (GoatPro, available for download on github; `https://github.com/mwestneat/GoatPro`; Figure 4.A.1) to adjust each of these elements to a standard position. All standard positions are readily attained with a fresh fish and are within the behavioral and biomechanical range of motion for all species. The app rotates each fish to the horizontal axis and enables a set of batch processing (all specimens) and custom functions (individual specimen) functions to biomechanically rotate peripheral

anatomical elements. Batch functions straighten the barbels and set them to a posteroventral 45-degree angle, rotate all dorsal, anal, and pelvic fin leading edges to a posteroventral 45-degree angle, and all trailing edges to a 10-degree angle to the horizontal. Individuals with caudal peduncle or caudal fin slightly displaced dorsally or ventrally were straightened while maintaining all other coordinate positions. The caudal fin was spread by rotating dorsal and ventral lobes independently, to attain fin edge angles of 20-degrees to the horizontal. Finally, for several specimens with the mouth open, the lower jaw was rotated dorsally to close the mouth, through biomechanically computed jaw rotations, with the barbel and barbel base tracking the jaw as it closed. The app then exported coordinate sets for all specimens with their adjusted positions to a file used for morphometric analysis.

4.3.3 Morphometric analyses

Landmarks were subdivided into 5 separate datasets for independent morphometric analyses across different aspects of goatfish morphology. The landmark subsets are defined as follows: Forehead Curve, which contains 15 semi-landmarks that extend from rostral tip of the nasal bone to the anterior base of the first dorsal fin; Dorsal Curve, which consists of 5 semi-landmarks that extend from the anterior base of the first dorsal fin to the insertion of the first caudal fin ray; Caudal Fin Curve, which consists of 17 landmarks that extend from the insertion of the first dorsal caudal fin ray to the start of the ventral caudal peduncle; Ventral Curve, which consists of five semi-landmarks that extends from the tip of the ventral caudal peduncle to the anterior base of the pectoral fin; and Full Body, which includes all curves mentioned previously, in addition to 5 semi-landmarks along the barbel and all 47 fixed landmarks (Figure 4.1).

We performed a landmark based geometric morphometric analysis to quantify the variation in head and body shape across species within the Mullidae. Using the protocol of George and Westneat (2019) and Nash et al. (2022) as a general framework, we projected

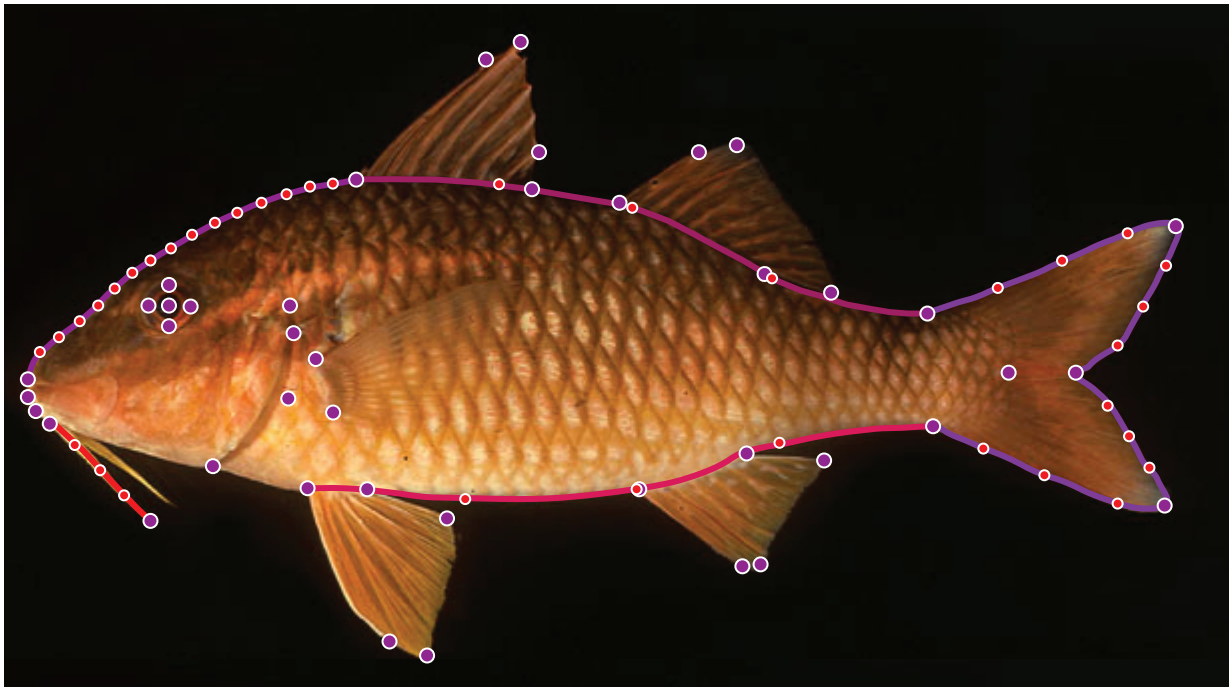

Figure 4.1: **Geometric morphometric landmark scheme with curves.** Purple circles represent fixed landmarks. Curve subsets indicated and smaller, red circles represent the approximate location of the sliding semi-landmarks.

each digitized landmark into linear tangent space using the `gpagen` function in geomorph 4.0.5 (Adams and Otárola-Castillo 2013; Adams et al. 2013, 2021; Baken et al. 2021). The generalized Procrustes analysis (GPA) removes the variation in landmark position that is attributable to rotation, translation, and scaling, while preserving the relevant shape information (Gower 1975; Rohlf and Slice 1990). We then calculated species means of the Procrustes-transformed landmark coordinates to account for intraspecific variation using the `mshape` function in geomorph (Adams et al. 2021; Baken et al. 2021). This protocol was iterated over all species-grouped landmark sets producing 70 shapes representing the mean body shape of each individual species. From these 70 mean shapes, another GPA was performed, and the resulting Procrustes aligned shape data was projected into linear tangent space using the `gpagen` function in geomorph (Adams et al. 2021; Baken et al. 2021).

Principal components analyses (PCA) were performed on the species averaged, Procrustes aligned landmark coordinates using the `gm.prcomp` function in geomorph (Adams

et al. 2021; Baken et al. 2021). We use the `shape.space` function from the R package borealis 2022.10.27 to visualize our PC axes, in addition to generating convex hulls to show the relative placement of species within each substrate preference group within the PC plot (Angelini 2023). To visualize the shape changes associated with each PC axes, we generated thin-plate splines of mean shapes using the function `plotRefToTarget` in geomorph as a deformation of the average shape of all specimens (Baken et al., 2021; Adams et al., 2021). Vector deformation plots were also produced using the function `plotRefToTarget` in geomorph comparing the mean shapes of each feeding mode group in reference to another group (Baken et al., 2021; Adams et al., 2021).

We used backtransformations to visualize the shape changes associated with the most significant axes of variation (PC1 and PC2), which plots the shape variation in its relative position in the morphospace (Olsen 2017). We projected the time-calibrated phylogeny onto the morphospace using the `phylomorphospace` function in the R package phytools 1.2-0 (Revell 2012). Additionally, we generated a phylogenetic PCA using GLS-centering and projection using the implementation in `gm.prcomp` function in geomorph (Revell 2009; Adams and Otárola-Castillo 2013). We tested hypotheses on the influence of allometry on overall shape using centroid size through the `procD.lm` function in geomorph (Adams et al. 2021). To further characterize aspects of body and head morphology, we calculated the Euclidean distance between sets of fixed landmarks to compute length measurements for each image using "GoatPro" (Figure 4.1; Westneat 2023). These measurements include Barbel Length, Head Length, Body Depth, and Total Length (Table 4.S2 & 4.S3). Barbel Length (BL) is defined as the distance from the tip of the barbel to the insertion point on the mandible. Head Length (HL) corresponds the distance between the tip of the rostrum to the farthest posterior point of the operculum (Figure 4.1). Finally, Total Length (TL) was computed as the rostral extremity to the tip of the caudal fin. We accounted for allometric variation by computing three ratios: the barbel ratio, which is defined as BL/TL, the Elongation Ratio,

which is defined as BD/TL, and Head Length, defined as HH/TL (Table 4.S2 & 4.S3). Additionally, we computed the aspect ratio (AR) for the caudal fin using the following equation: $AR = \frac{b^2}{A}$, where b is the maximum span of the caudal fin and A is the surface area of the fin (Figure 4.1; George and Westneat 2019). These ratios were used in subsequent analyses.

4.3.4 Classification of foraging substrate preference

Previous research has indicated that species of goatfish tend to preferentially forage on either soft substrates, such as sand or mud, or hard substrates, such as coral rubble (Gosline 1984; Platell et al. 1998; Meyer et al. 2000; Lukoschek and McCormick 2001; Krajewski et al. 2006; Galarza et al. 2009; Strübin et al. 2011; Mittelheiser et al. 2022). We conducted a comprehensive survey of the published literature, supplemented by personal observations, to designate species of goatfish as preferentially foraging on either "soft" or "hard" substrates. We define these delineations based on relative size and complexity of the preferred substrate composition that each species forages on, with preference is defined as more than 50% of observations (Table 4.S1). Species without any available information about their foraging behavior were given an "unknown" designation and were removed from the analyses when appropriate.

4.3.5 Ecomorphological analyses

To examine the relationship between body and head shape with substrate preference, we performed a MANOVA using the `procD.lm` function in geomorph (Adams and Otárola-Castillo 2013). The partial disparity for genera was calculated using the overall mean over 1000 iterations with the function `morphol.disparity` in geomorph, taking care to correct the denominator in the variance calculation from n to N-1, where n is the group size and N is the number of observations (Adams et al. 2021; Baken et al. 2021). For the measurement ratios,

we tested the normality and variance homogeneity using the Shapiro-Wilk and Bartlett's test, respectively (Bartlett 1937; Royston 1982). If either parametric requirement were not met, we performed a Kruskal-Wallis test to examine the differences across substrate preferences and morphological ratios (Hollander et al. 2013). Additionally, a two-sample Wilcoxon test with a Bonferroni correction was used to determine assess significance between the morphological ratios associated with each substrate preference (Bauer 1972).

In addition to these analyses, the following analyses were performed on a subset of the data that only included the genus *Parupeneus* to further examine variation in this genus. Using the time-calibrated phylogeny from Nash et al (2022), we estimated phylogenetic signal for the Procrustes shape variables using the function `physignal` from geomorph with 10,000 iterations, which uses the multidimensional equivalent of Blomberg's K for each morphometric subset (Adams and Otárola-Castillo 2013, Adams 2014). We also performed a phylogenetic 2-blocks Partial Least Squares analyses using the `phylo.integration` function from geomorph to investigate the ecomorphological relationship between substrate preference and morphological characteristics (Adams and Otárola-Castillo 2013; Adams et al. 2021). Additionally, we examined these ecomorphological trends in a univariate context using phylogenetic generalized least squares (PGLS) method implemented with the function `pgls` from the R package caper 1.0.1, in which phylogenetic signal is estimated simultaneously with the regression model (Freckleton et al. 2002). To test whether there are significant differences in the rate of shape evolution between substrate preferences in a phylogenetic context, we used the `compare.evol.rates` function in geomorph using the permutation approach with 10,000 iterations to test for significance (Adams 2014; Denton and Adams 2015; Adams and Collyer 2018, 2019).

4.3.6 Ancestral state estimation of substrate preference and morphological metrics

We used the `ace` function in the R package ape 5.6-2 to estimate ancestral character states for substrate preference across the phylogeny using the transition rate models all rates different (ARD) and equal rates (ER; Paradis and Schliep 2019). Tip states were mapped onto the time-calibrated phylogeny, and the estimated likelihoods of each ancestral state were mapped onto internal nodes within the phylogeny, allowing visualization of independent trait origins and convergence across clades. To examine the distribution of PC axes on the phylogeny, we simulated stochastic character evolution of substrate preference using the `make.simmap` and projected the phylogenetic tree in relation to PC scores using the `phenogram` functions from phytools (Evans et al. 2009; Revell 2012). For continuous morphological measurements and ratios, we used the `contMap` function from phytools to estimate states at internal nodes and to map continuous characters along branch lengths (Revell 2012).

4.4 Results

The central results of this study are that forehead shape, dorsal/ventral shape, caudal fin shape, and full body shape are associated with both generic membership and substrate preference. Preferential feeding on hard substrates evolved only in several species within *Parupeneus* and is associated with unique morphological characteristics across morphological subsets and metrics. Although hard substrate foraging derived relatively recently within goatfish evolutionary history, morphological diversification associated with this foraging mode are associated with a rapid increase in the rate of change.

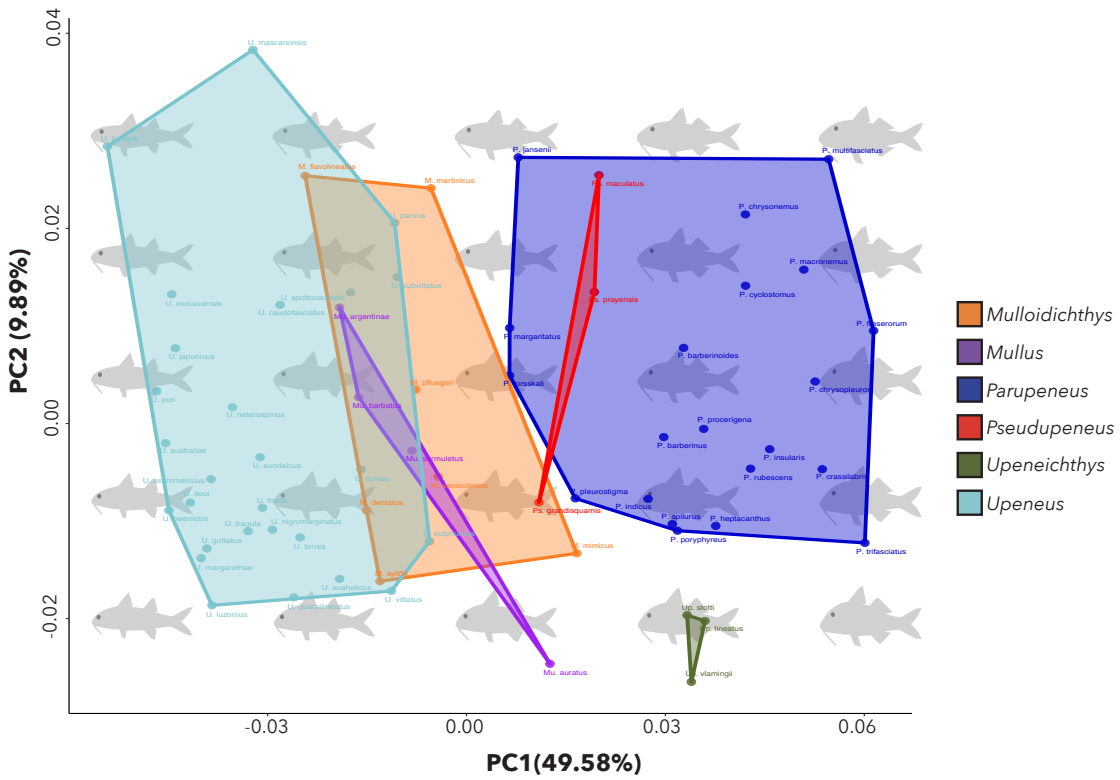

Figure 4.2: **Mullidae full shape morphospace.** PCA plot showing PC1 (49.58%) vs PC2 (9.89%) for semi-landmarks placed along the head curve for 61 species of goatfish. Genera is indicated by convex hulls. Backtransformation icons represent the estimated shape of a species in that particular portion of morphospace.

4.4.1 Morphometric Diversity

We parsed the variation in body and head shape across all species of goatfish through a comprehensive examination of four aspects of goatfish morphology: Full Body, Forehead, Dorsal/Ventral curves, and Caudal Fin Figure 4.A.10. All morphological subsets were significantly different among genera, and all subsets, except for the caudal fin, had a significant phylogenetic signal. We found no evidence for an allometric effect as centroid size was found to be non-significant across subsets.

The first two PC axes of variation associated with full body shape, which includes all fins and barbels, across all species of Mullidae account for a combined 56.47% percent of

the total variation (49.58% and 9.89%, respectively; Figure 4.2). Values of PC1 correspond to changes in head depth, barbel length, fin position, body elongation, and eye size, with higher PC1 values being associated with stockier body shapes with larger fins and barbels and smaller eye sizes (Figure 4.2). The second axis of variation reflects the forehead curve and head height (Figure 4.2). Genera cluster into two groups primarily along PC1, with species within *Mulloidichthys*, *Mullus*, and *Upeneus* found within the lower quadrants of PC1 and *Parupeneus*, *Upeneichthys*, and *Pseudupeneus* primarily distributed within the higher quadrants (Fig 2). Of note, *Upeneichthys* occupies its own portion of morphospace and does not intersect with any other genus (Figure 4.2).

The first two PC axes of variation associated with forehead shape across all species of Mullidae account for a combined 89% percent of the total variation (50.52% and 38%, respectively; Figure 4.A.3). Values of PC1 correspond to the relative convexity of the curve, with higher PC1 values being associated with reduced convexity and flatter curve (Figure 4.A.3). The second axis of variation reflects the presence and degree of an inflection point located near the rostral side of the curve, which creates well defined rostral region (Figure 4.A.3). Genera cluster into two groups primarily along PC2, with species within *Mulloidichthys*, *Mullus*, and *Upeneus* found within the lower quadrants of PC2 and *Parupeneus*, *Upeneichthys*, and *Pseudupeneus* primarily distributed within the upper quadrants (Figure 4.A.3). Only species within *Parupeneus* span into the upper values of PC2, which corresponds to degree of an inflection point in the head curve (Figure 4.A.3). Most of the proportion of total disparity is found within *Parupeneus* (48%) and *Upeneus* (22.3%), followed by *Mulloidichthys* (20.0%), *Mullus* (4.2%), *Upeneichthys* (3.8%), and *Pseudupeneus* (1.5%).

The first two PC axes of variation associated with the dorsal and ventral body curves across all species of Mullidae account for a combined 72.25% percent of the total variation (57.73% and 14.62%, respectively; Figure 4.A.4). Values of PC1 correspond to the body depth, with higher PC1 values being associated with larger body depth (Figure 4.A.4). The

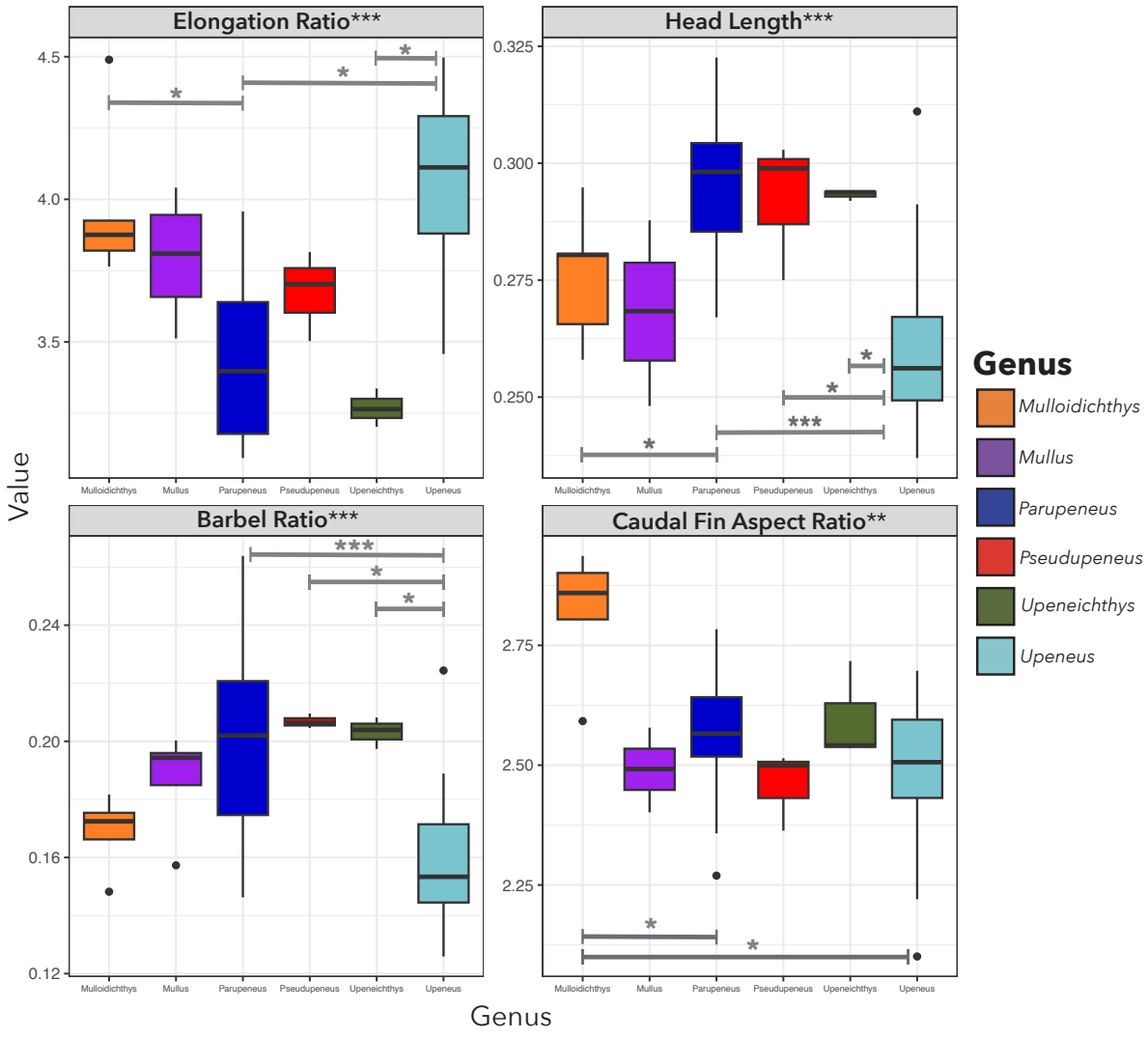

Figure 4.3: **Boxplots of morphometric ratios by genus.** Boxplots summarizing the distribution of **(A)** Elongation Ratio, **(B)** Head Length, **(C)** Barbel Ratio, and **(D)** Caudal Fin Aspect Ratio (AR) between species within each each genus. Significant pairwise comparisons are indicated by a grey bar and standard significance notation, in which * corresponds to p-values < 0.05, ** corresponds to p-values < 0.01, and *** corresponds to p-values < 0.001.

convex hulls of genera overlap quite a bit, however there is limited clustering primarily along PC1. Only species within *Upeneus* occupy the lower values of PC2, which corresponds to a rounder ventral curve compared to the dorsal curve (Figure 4.A.4). The first two PC axes of variation associated with the caudal fin curves across all species of Mullidae account for a combined 55.68% percent of the total variation (34.22% and 21.46%, respectively; Figure 4.A.6). Values of PC1 correspond to relative fin size, with higher PC1 values being associated with caudal fins that have both shorter lobe lengths and relatively rounder lobes (Figure 4.A.6). The convex hulls of genera also overlap quite a bit, however there is minimal clustering primarily along PC1. Only species within *Upeneus* occupy the highest values of PC2, which corresponds slightly rounder forked fins (Figure 4.A.6).

All morphological measurements and metrics differed significantly among genera. Although species within *Upeneus* have the highest mean Elongation Ratio (mean = 4.06, sd = 0.31), these species also have the smallest mean Head Length (mean = 0.26, sd = 0.02) and the lowest mean Barbel Ratio (mean = 0.16, sd = 0.02; Figure 4.3). On the contrary, species within *Parupeneus* have the lowest mean Elongation Ratio (mean = 3.45, sd = 0.28) and the largest Head Length (mean = 0.30, sd = 0.01; Figure 4.3). Additionally, species within Mulloidithchys have the highest Caudal Fin Aspect Ratio (mean = 2.75, sd = 0.12). Given the non-normality of the data, the non-parametric pairwise comparisons showed that *Parupeneus* was significantly different from *Upeneus* for all metrics except for Caudal Fin AR (Figure 4.3). Additionally, *Parupeneus* was significantly different from *Mullus* apart from Barbel Ratio (Figure 4.3).

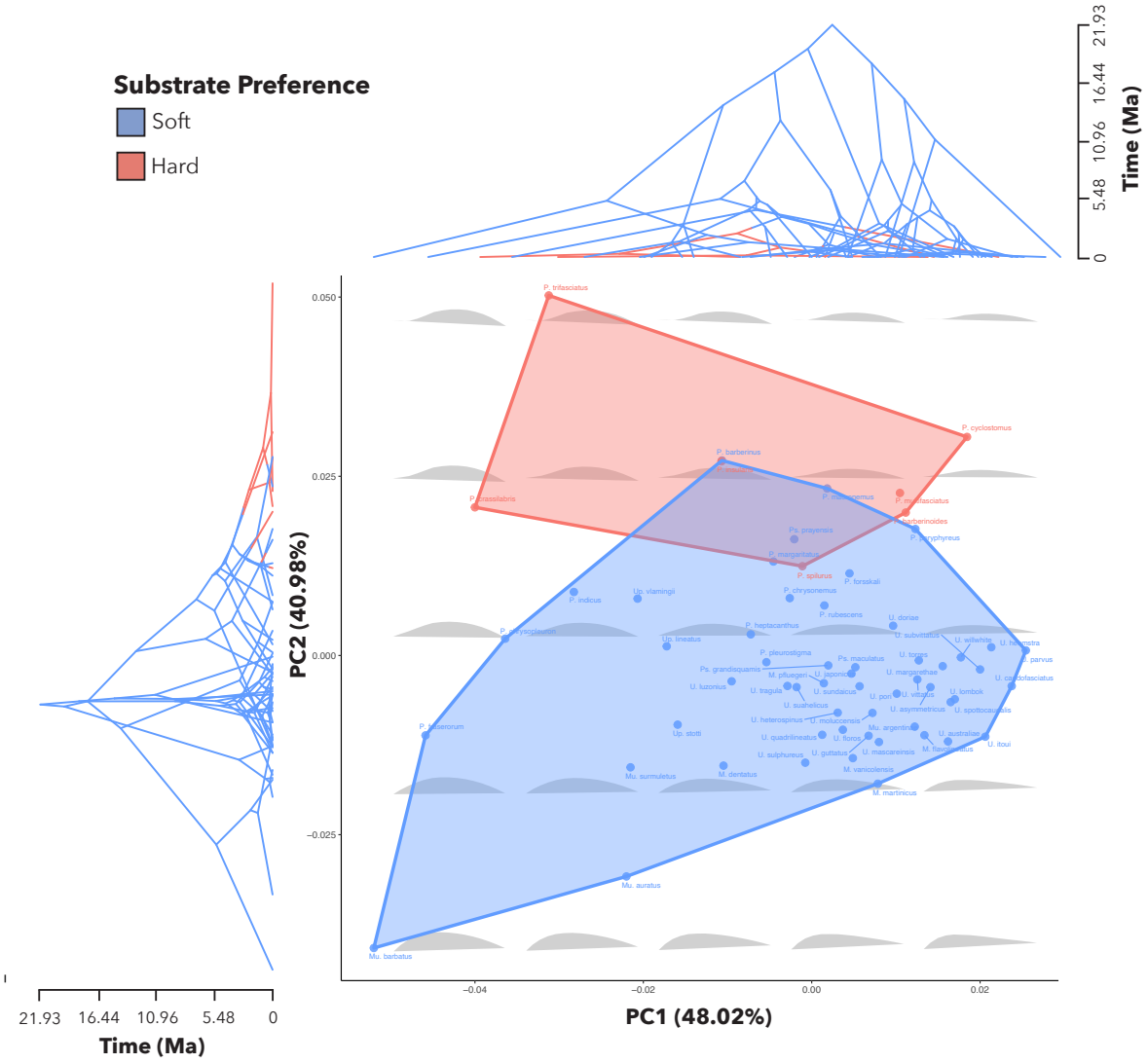

Figure 4.4: **Mullidae forehead shape morphospace by substrate preference.** PCA plot showing PC1 (48.02%) vs PC2 (40.98%) for semi-landmarks placed along the forehead curve for 61 species of goatfish. Substrate preference is indicated by convex hulls. Backtransformation icons represent the estimated shape of a species in that corresponding location in morphospace. Phenograms showing projection of the phylogenetic tree in a space defined by each PC axes. Branches are color coded according to substrate preference type.

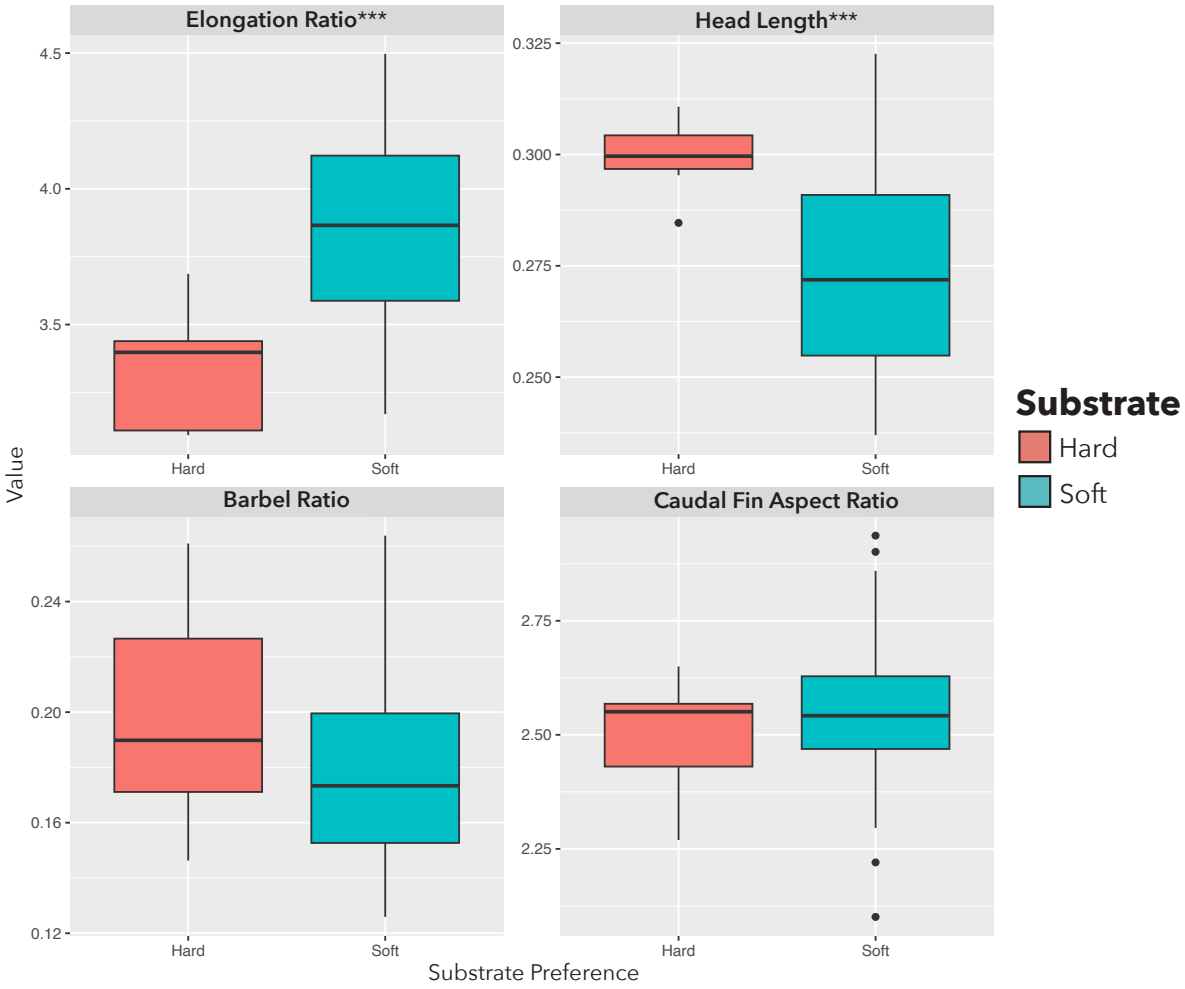

Figure 4.5: **Boxplots of substrate preferences across morphometric ratios.** Boxplots summarizing the distribution of **(A)** Elongation Ratio, **(B)** Head Length, **(C)** Barbel Ratio, and **(D)** Caudal Fin Aspect Ratio (AR) between species within each substrate preference type. Significant pairwise comparisons are indicated by ***, which corresponds to p-values $<$ 0.001. Pairwise comparisons between substrate preference within Barbel Ratio and Caudal Fin AR were non-significant.

4.4.2 Ecomorphological relationships among morphology and substrate preference across the Mullidae

Substrate preference is significantly associated with all morphometric subsets, with a significant value of $p = 0.001$ for all subsets respectively. Several species that preferentially forage on hard substrates occupy a distinct portion of morphospace across subsets, however there is always some overlap (Figure 4.4, Figure 4.A.2, Figure 4.A.5, & Figure 4.A.7). For example, soft substrate species cluster within the mid- to lower values of PC2 of the forehead shape subset, and is composed of all species within *Upeneus*, *Mulloidichthys*, *Mullus*, *Pseudupeneus*, and *Upeneichthys*, in addition to twelve species of *Parupeneus* (Figure 4.4). Hard substrate species primarily cluster within upper values of PC2 and consist of eight species within *Parupeneus* (Figure 4.4). Although soft substrate species have a higher proportion of total disparity (0.28 and 0.72, respectively), the observed differences in the disparity between groups is not statistically significant (p= 0.93). Although we find evidence that substrate preference is significantly associated with all morphometric subsets, none of the subsets show a significant association with substrate preference when evaluated using a phylogenetic framework.

Pairwise comparisons between substrate preference types were significant when considering Elongation Ratio and Head Length ($p = 0.001$, Figure 4.5). Species associated with soft substrate feeding have a relatively higher mean elongation ratio (mean = 3.85, sd = 0.39) and smaller mean head length (mean = 0.27, sd = 0.02) when compared to hard substrate species (Figure 4.5). However, differences in barbel ratio and caudal fin AR associated with substrate preference were found to be non-significant ($p = 0.2$).

4.4.3 Ancestral state estimation of substrate preference

Ancestral state estimation of substrate preference showed that the presence of hard substrate preference is restricted to species within *Parupeneus* (Figure 4.6). Within *Paru-*

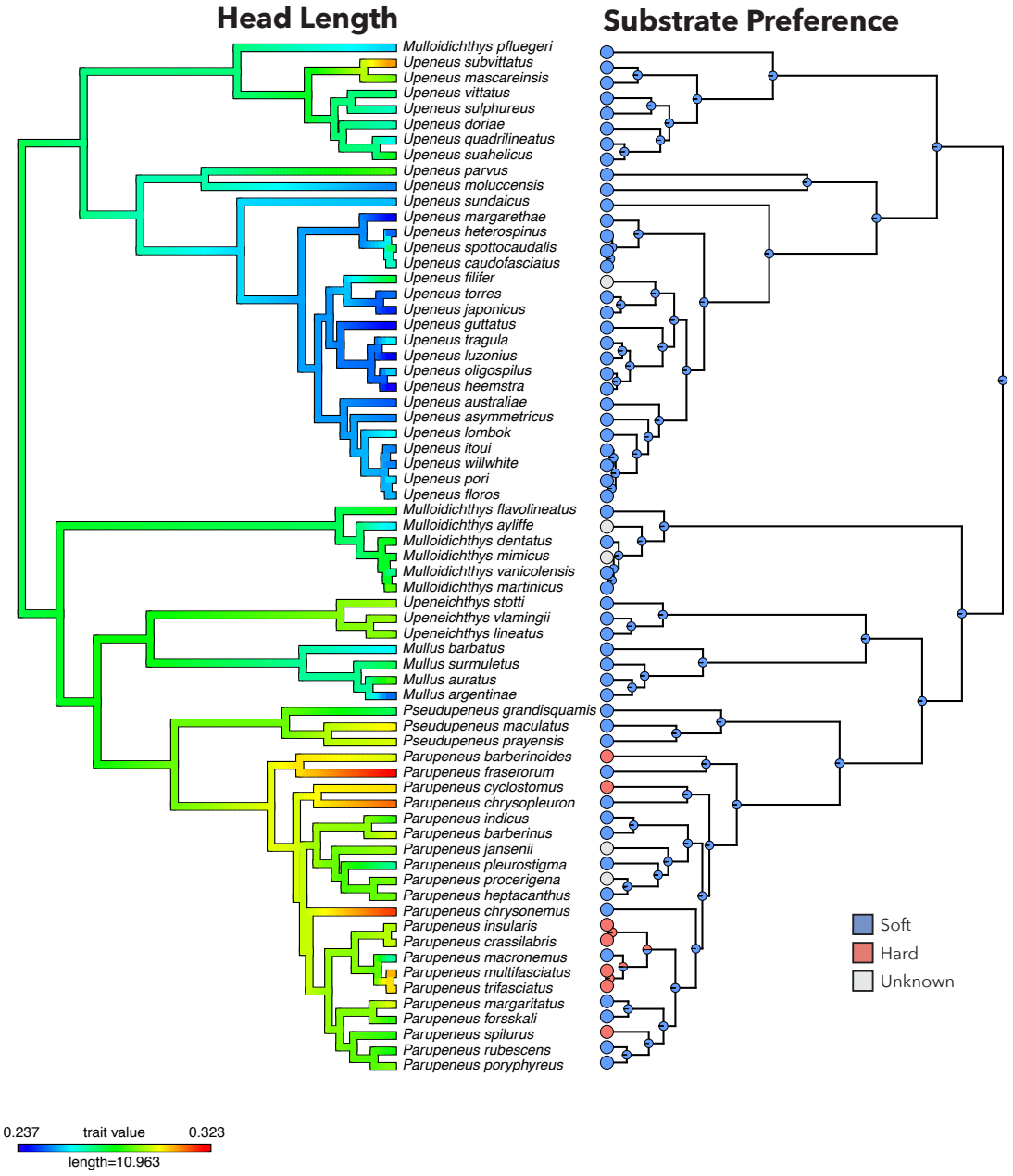

Figure 4.6: **Mirror phylogeny of head size and substrate preference.** A comparison of the mapped ancestral state estimation for A) head size and B) substrate preference across the phylogeny. Head size estimations were inferred using ContMap, in which branch lengths are color coded in associated with the mapped character state. Substrate preference is indicated by colored pie graphs, representing the likelihood of that each state across the nodes of the phylogeny.

peneus, there are five independent transitions to hard substrate preference, with a total of seven species (Figure 4.6). The best supported Markov model of state transition was "ARD", in which there is a higher rate of transition associated with shifts from soft substrate into hard substrate. Given the distribution of substrate preference across extant species, the estimated ancestral state of the root node is associated with soft substrate foraging (Figure 4.6).

4.4.4 Morphological variation and substrate preference within the genus Parupeneus

We examined morphological shape variation for four subsets of data for species within *Parupeneus*: Full Body, Forehead, Dorsal/Ventral Body, and Caudal Fin. PC1 for each subset explains 29.17%, 57.05%, 54.31%, and 42.81% of the total variation, respectively (Figure 4.7). PC2 for each subset explains 16.42%, 34.57%, 15.69%, and 19.7% of the total variation, respectively (Figure 4.7). When considering the Full Body dataset, changes along PC1 are primarily associated with body depth and barbel length, with higher values of PC1 being associated with a larger body depth and longer barbels (Figure 4.7). The variation captured along PC2 is associated with head convexity, with higher values of PC2 corresponding to a higher degree of head convexity (Figure 4.7). Comparable with the family-wide analysis, the variation along PC1 and PC2 of the forehead subset is associated with degree of head convexity and presence and degree of an inflection point located near the rostral side of the curve, respectively (Figure 4.7).

Variation along PC1 for the Dorsal/Ventral subset is primarily associated with changes in body depth and variation along PC1 for the Caudal Fin subset is related to the relative rounding of the lobes of the caudal fin (Figure 4.7). Differences in variation of the full body and caudal fin shape subsets are significantly associated with substrate preference ($p = 0.04$), although forehead shape is marginally considered to be non-significant with a p-value of 0.06. However, there are noticeable trends associated with other aspects of morphology

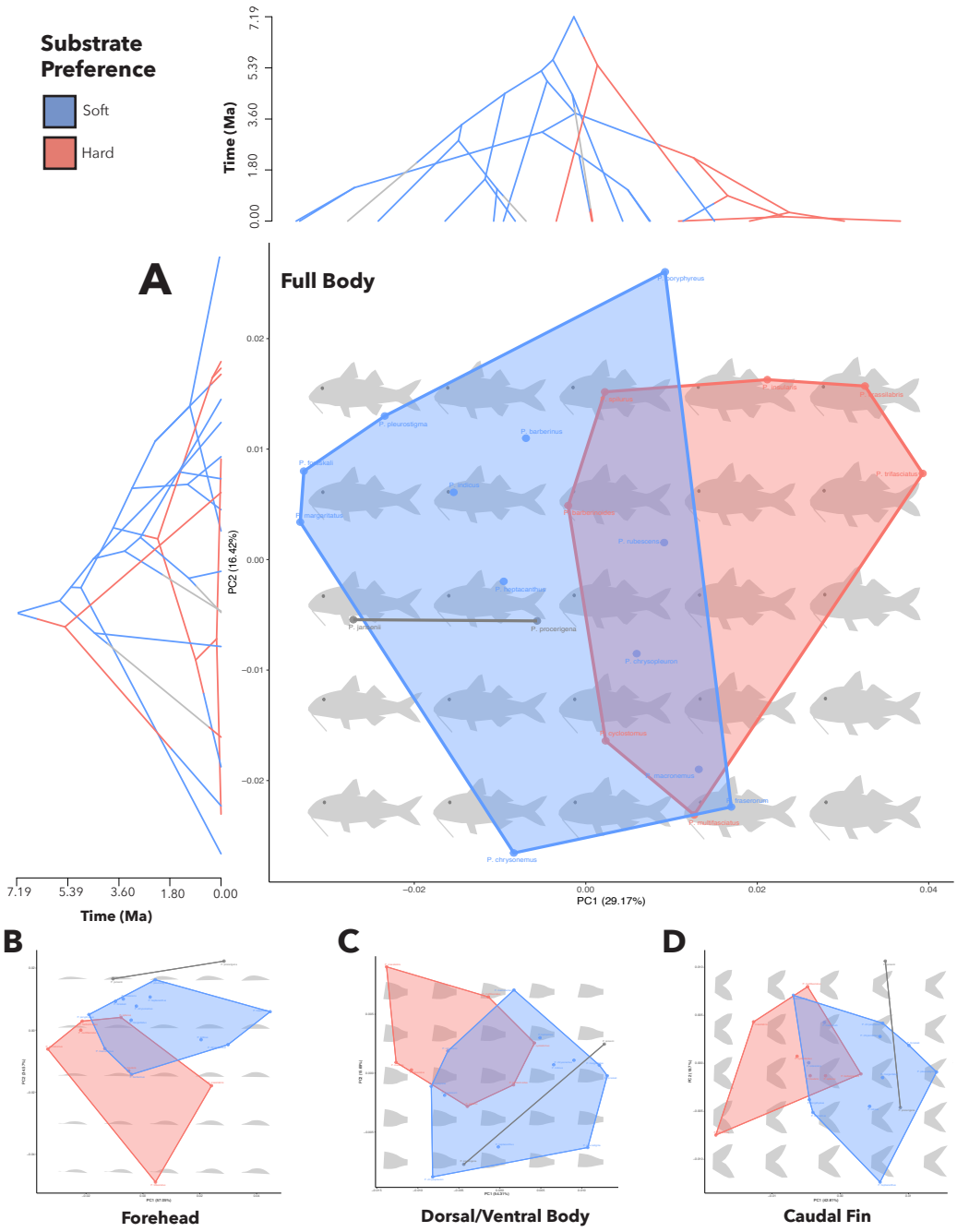

Figure 4.7: **Morphospaces of morphological subsets within *Parupeneus*.** Morphospace plots showing axes PC1 and PC2 of species within the genus *Parupeneus* using different subsets of morphometric data: **(A)** Full Body, **(B)** Forehead, **(C)** Dorsal/Ventral Body, and **(D)** Caudal Fin. Substrate preference is indicated by color coded convex hulls. Backtransformation icons represent the estimated shape of a species in that corresponding location in morphospace. Phenograms showing projection of the phylogenetic tree in a space defined by each PC axes. Branches are color coded according to substrate preference type.

given areas of morphospace occupied exclusively by species that prefer hard of soft substrate (Figure 4.7). We find evidence that hard substrate preference is associated with a significant increase in the rate of morphological diversification for the full body and forehead subsets, with a ratio of 6.8 and 15.9 higher when compared to their soft substrate counterparts, respectively (Figure 4.A.11 & Figure 4.A.12).

4.5 Discussion

The integration of geometric morphometrics with a robust time-calibrated phylogenetic hypothesis and ecological traits of foraging behavior and substrate preference yields novel insight into the ecomorphological radiation of the goatfishes. In this study, we present evidence for the association of substrate preference with various aspects of goatfish morphology. Although morphological diversity is tightly linked to their phylogenetic history, we discuss the evolution of hard substrate feeding within the genus *Parupeneus* and its potential influence on patterns of rapid morphological diversification. We conclude that several aspects of goatfish morphology, particularly the forehead shape, are associated with unique characteristics that allow for the utilization of hard substrates during foraging in marine habitats.

4.5.1 Forehead and body shape is associated with preferential foraging on different substrates

Transitions across ecological niches in association with morphological diversification can greatly impact a lineages ability to survive and diversify. In goatfishes, we find evidence that substrate preference impacts patterns of head and body shape diversification. We find that although species that prefer to feed on hard substrate make up a small proportion of the total species diversity, they tend to occupy a distinct set of morphological characteristics (Figure 4.4 & Figure 4.5). These characteristics include an elongated snout with a promi-

nent rostrum, shorter body lengths, greater body depths, longer barbels, and larger head sizes (Figure 4.4 & Figure 4.5). Interestingly, the evolution of hard substrate preference is restricted to several species within the genus *Parupeneus* (Figure 4.6). Despite this, forehead shapes associated with hard substrate feeding occupy comparable levels of overall disparity. The observed association between all morphometric subsets and substrate preference in the goatfishes provides evidence of robust ecomorphological relationships.

The influence of head and body shape on fish foraging behavior has been extensively studied across a wide variety of fish clades, including cichlids, damselfishes, wrasses, and morymrids (Liem 1973; Westneat 1995; Cooper and Westneat 2009; Parsons et al. 2011; López-Fernández et al. 2013; Evans et al. 2022; Peterson et al. 2022). However, fishes that interact with the benthic substrate require a specific set of morphological traits that allow them to manipulate the substrate to gain access to infaunal food sources or to burrow within. For example, many benthic foraging specialists have evolved ventral facing, protrusible mouths, that allow for feeding off the seafloor (McCormick 1995; Lukoschek and McCormick 2001; Ribeiro et al. 2018; Friedman et al. 2020).

Although all species of goatfish are specialized benthic foragers, we find evidence that aspects of their cranial morphology might constrain the diversity of behaviors they are able to perform. Species that are associated with preferential feeding on soft substrates lack elongated pre-orbital morphology, which may limit their ability to manipulate and maneuver around more complex substrate types, such as coral rubble (Mittelheiser et al. 2022). For example, individuals of the Spangled Emperor, *Lethrinus nebulosus*, have been observed using their elongated snouts to forage within the substrate to access food sources (Sweatman 1996). Notably, evidence suggests that individuals of *L. nebulosus* rarely exceed a foraging depth that exceeded their snout length. A similar observation has been made for *P. barberinus*, in which individuals rarely foraged deeper than their barbel or snout lengths, highlighting the potential association between foraging ability and morphological variables

(Lukoschek and McCormick 2001). Potential adaptation-based explanations include a lack of ability to detect prey at depths greater than the reach of their sensory barbels, as well as a reduction of time where their eyes are below the sediment surface to decrease the likelihood of a predation event (Lukoschek and McCormick 2001). Therefore, evolution towards elongated foreheads may allow for access to food in deeper and more complex substrate types.

Many benthic associated fish lineages have convergently evolved sensory projections that increase the ability to detect potential prey items within the substrate or environments with limited visibility (Katagiri 1966; Diogo and Chardon 2000; Kim et al. 2001; Kiyohara et al. 2002; Eakin et al. 2006; LeClair and Topczewski 2010). Notably, we find evidence that a longer barbel length is associated with preferential foraging on hard substrates, despite the loss of significance when controlling for body length. The presence of longer barbels in species that preferentially forage in more complex environments makes functional sense, as longer barbels represent an increase in the surface area of exposed tissue that is capable of chemosensation and mechanical potential (McCormick 1993; Lombarte and Aguirre 1997; Uiblein et al. 1998). Goatfish barbels are unique in their developmental patterns, structure, and functionality when comparted to the barbels of other fish clades (Gosline 1984; McCormick 1993; Lombarte and Aguirre 1997; Kim 2002; Kirino et al. 2006). Their pair of hyoid barbels are modified, rod-like branchiostegal rays that provide stability and a strengthening support (Gosline 1984). Modifications of branchiostegal rays in conjunction with sensory systems are rare, as they have only convergently evolved within the Mullidae and the genus *Polymixia* (Gosline 1984; Kim et al. 2001). Each barbel is able to move independently due to their own set of musculature, and can be used to probe for food and as an excavation tool (Gosline 1984).

Previous research has shown that there is evidence for plasticity in barbel development, which is dependent on habitat complexity and food availability during the time of post-larval settlement (McCormick 1993). Larval goatfishes have the typical teleost arrangement

of their branchiostegeal rays and only begin to utilize their barbels post settlement on a reef or associated habitat (Gosline 1984; McCormick 1993; Pavlov et al. 2011). During development, the anterior branchiostegeal ray separate from the other rays and move forward on the hyoid bar to become level with the tip of the glossohyal (Gosline 1984; McCormick 1993; Pavlov et al. 2011). The barbels show a period of rapid growth during the settlement phase, in which the individual transitions from its pelagic life phase to its benthic phase (McCormick 1993; McCormick and Milicich 1993; McCormick and Molony 1993). For example, the amount and depth of food availability during the settlement phase of individuals of *Upeneus tragula* led to significant variability in barbel development, with the food deprived treatment group developed the longer barbels and larger taste buds compared to their counterparts in a treatment of excess food (McCormick 1993). Although this phenomenon has not been examined in other species of goatfish, it highlights potential ecomorphological implications of barbel morphology in relation to available substrates. However, it also poses the question about the degree of plasticity across species and whether this relates to their propensity to utilize differing substrate types while foraging.

4.5.2 Recent evolution of preferential foraging on hard substrates

The evolution of preferential feeding on hard substrates is restricted to several species within the genus *Parupeneus*, indicating that this foraging mode likely derived within the past 5.5 Ma (Figure 4.4). When considering the evolution of forehead shape diversity, we find evidence that preferential feeding on soft substrate may act as a significant constraining force (Figure 4.4). Our relative rate analyses revealed that the foreheads of species that preferentially forage on hard substrate evolved 15.9 times faster than soft substrate species, which is corroborated by the statistical non-significance between the disparities associated with each foraging mode despite differences in species richness and origination time (Figure 4.4). The goatfishes underwent a pulse of phylogenetic and morphological diversification associ-

ated with the Miocene-Pliocene transition, which is marked by increase in the interchange of Indo-West Pacific fauna caused by the separation of the Arabian and African plates, despite the complete separation of the Mediterranean Sea from the Red Sea and the Indian Ocean due to terrain uplift (Briggs 1995; Nash et al. 2022).

Although this period of rapid morphological diversification was observed for the body shape across all species of goatfish, it remains unclear why this rapid evolution of forehead shape was restricted to species within *Parupeneus*. Given that *Parupeneus* partitions the morphospace with minimal overlap to other genera, this suggests that there was possibly a burst of morphological diversification associated with the initial divergence of *Parupeneus* or a pattern of gradual morphological diversification followed by extinction of the intermediate or overlapping forms (Sidlauskas 2008). Another possibility is that there are other phenotypic and biomechanical factors that constrain head shape to a particular range of morphologies within other genera (Sidlauskas 2008). For example, one of the diagnostic anatomical traits associated with *Parupeneus* is the presence of a crest on the dorsal midline of the frontal bone, which is the attachment point of the epaxial muscles (Gosline 1984; Kim 2002). The epaxial muscles within *Parupeneus* reach the anterior portion of the frontal near the nasal and mesethmoid bones, which is substantially farther than observed in other genera of Mullidae (Gosline 1984; Kim 2002). The anteriorly located epaxial muscles potentially allow the neurocranium to exert more force when manipulation substrate, such as shifting larger pieces of rubble, compared to other confamilial species.

Given that preferential foraging on hard substate is restricted to species within *Parupeneus*, but is not ubiquitous across the genus, this suggests that the morphological radiation associated with *Parupeneus* allows for expansion into multiple foraging substrates. Although we find evidence that there are general trends across morphological subsets that differentiate species that forage on either hard or soft substrate types, we find limited evidence that supports a significant different between these two foraging modes (Figure 4.7). This concept of

foraging flexibility is reinforced by observations of individuals of species that preferentially feed on hard substrate, such as *P. multifasciatus*, foraging on soft substrate sporadically (personal observations). Although this behavior was observed at a relatively low frequency, it highlights the potential that species within *Parupeneus* can use alternative foraging modes on a situational basis.

4.5.3 Ecological and evolutionary implications for diversity in foraging behaviors

Radiations of head and jaw morphology have been observed across a large variety of taxa and habitats, suggesting that there are multiple factors, such as environmental and developmental, that can influence trends of adaptive divergence. Two common mechanisms for variation in phenotype are genetic differentiation among phenotypic forms and environmentally induced modifications that result in discrete morphologies during development (Robinson and Parsons 2002). Adaptive variation among jaw morphologies is relatively common, and the specific genetic mechanisms that drive this adaptive variation has been examined across a variety of clades (Roberts et al. 2011; Reeck and Oxford 2022). Morphological variation can initially occur without genetic change. For example, new or unusual external inputs during development can facilitate the origination and evolution of novel morphological traits (West-Eberhard 2005). These new phenotypes can head start the process of adaptation as they may to lead to traits that will undergo favorable natural selection. For this to occur, the trait must be phenotypically plastic, which indicates that it is able to respond to environmental stimuli (Via et al. 1995; West-Eberhard 2005).

Phenotypic plasticity has the potential to enhance the chance of genetic divergence through the potential increase in the probability of survival and reproduction within differing environments. However, for morphological plasticity to be effective in reducing the probability of lineage extinction, individuals must respond behaviorally different in order to

a utilize a variety the available habitats (Abaad et al. 2016). Although we find evidence that species of goatfish have a preferential foraging substrate type, the specific behaviors that individuals use to interact with each substrate can differ drastically among species (personal observations, Gosline 1984; McCormick 1995; Krajewski et al. 2006). Additionally, it has been shown that despite variability in substrate preference and feeding behavior, the diets of most species are statistically indistinguishable from each other (Platell et al. 1998). This follows the paradigm known as Liem's Paradox, in which lineages with specialized morphologies can be ecological generalists and highlights the importance of evaluating the ability of a lineage to adapt in the presence of competition and ecological stress (Liem 1980). For example, differences in the craniofacial structure of Lake Malawi Cichlids vary in their ability to collect and process food items, despite overlapping feeding strategies (Albertson 2008). In goatfishes, it has been shown that species that co-occur can employ habitat partitioning in order to reduce competition over infaunal invertebrates in the substrate (Platell et al. 1998; Lombarte et al. 2000; Krajewski et al. 2006; Mittelheiser et al. 2022). The disparity in the efficacy of these feeding modes would be obscured during periods of plentiful food sources and would become evident during shortages of available resources (Albertson 2008).

In conclusion, there are significant associations between head and body shape with foraging substrate preference across the goatfishes. Although preferential foraging on hard substrates recently evolved within several lineages within the genus *Parupeneus*, we find evidence that it is associated with a dramatic increase in the rate of morphological change across the forehead shape of the goatfishes. We recognize that categorizing species into preferential feeding on hard or soft substrate potentially masks the underlying diversity of behavioral adaptations that occur within each of these foraging modes and across species. However, this examination of the relationship between morphological traits and foraging mode highlights the importance of examining the evolution of behavior using an integrative framework.

APPENDIX

4.A Supplementary Figures

Supplementary tables for this chapter can be viewed at `https://figshare.com/s/465545284ac987a7b372`

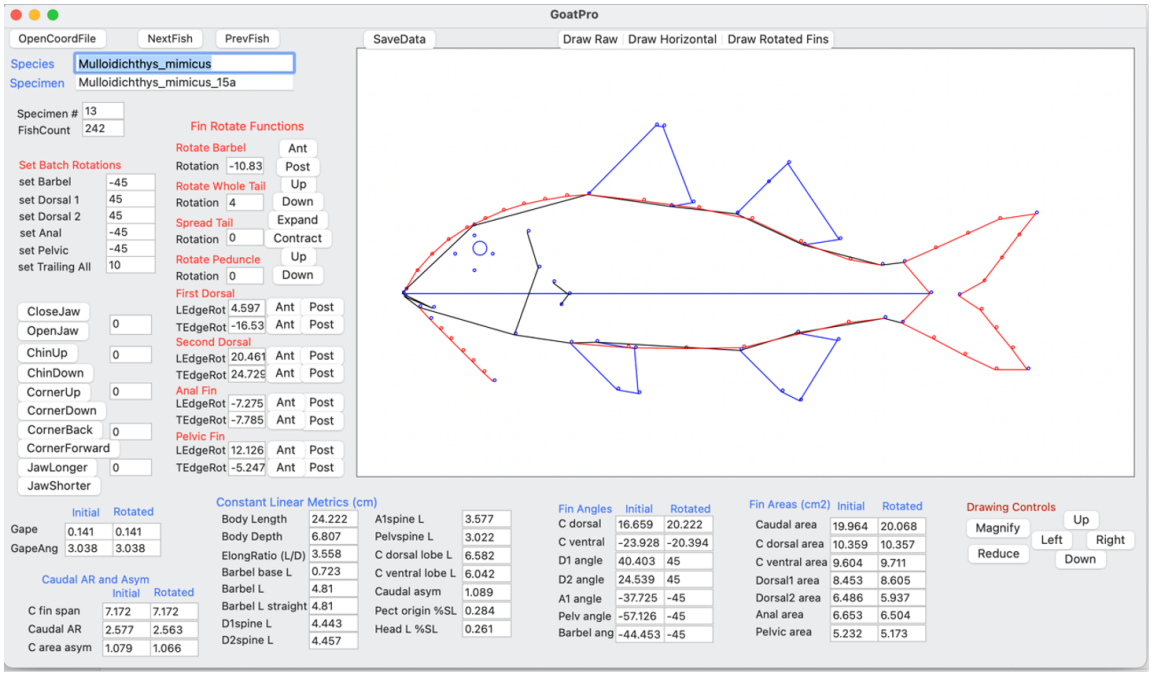

Figure 4.A.1: **GoatPro Interface.**The desktop app GoatPro, in which the rotational biomechanical flexibility of barbel, jaws, fins, and caudal peduncle enables the user to adjust morphometric coordinates for the leading and trailing edges of median fins to a standard angle, straighten the barbel and posterior body, and close the lower jaw.

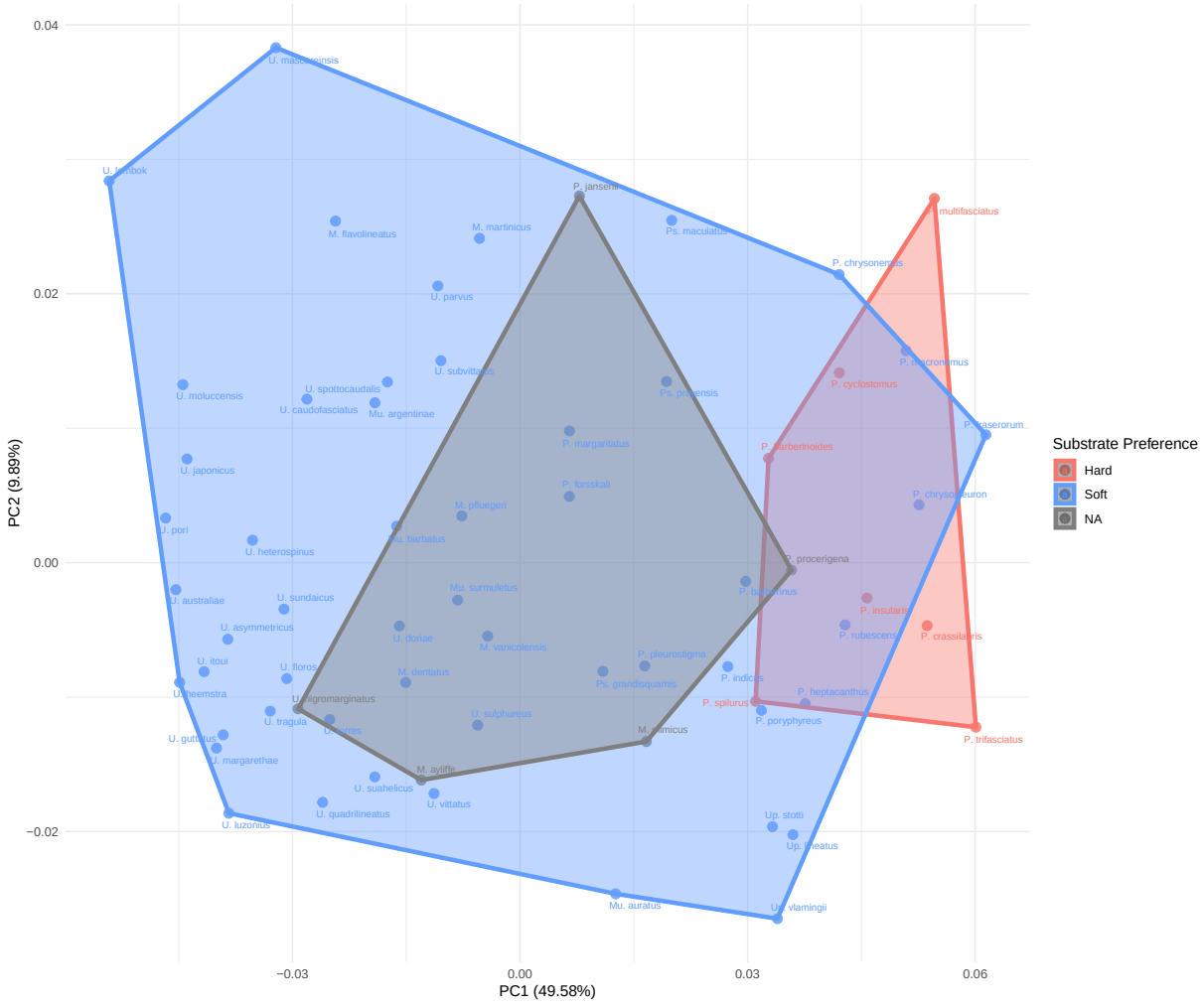

Figure 4.A.2: **Substrate preference morphospace of full body dataset of the Mullidae.** Substrate preference is indicated by convex hulls.

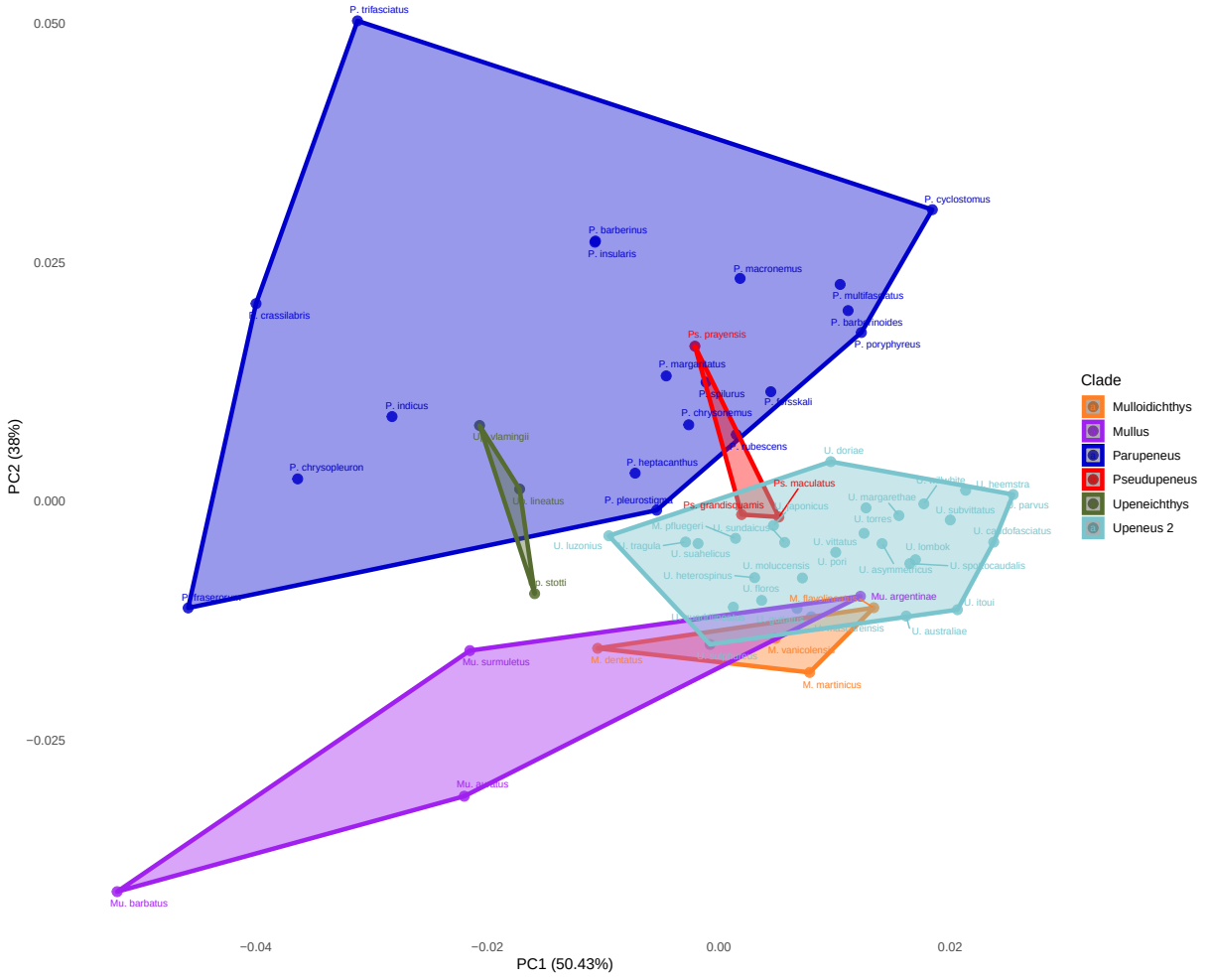

Figure 4.A.3: **Morphospace of forehead curve subset of the Mullidae.** Morphospace of forehead curve subset of the Mullidae. Genera are indicated by convex hulls.

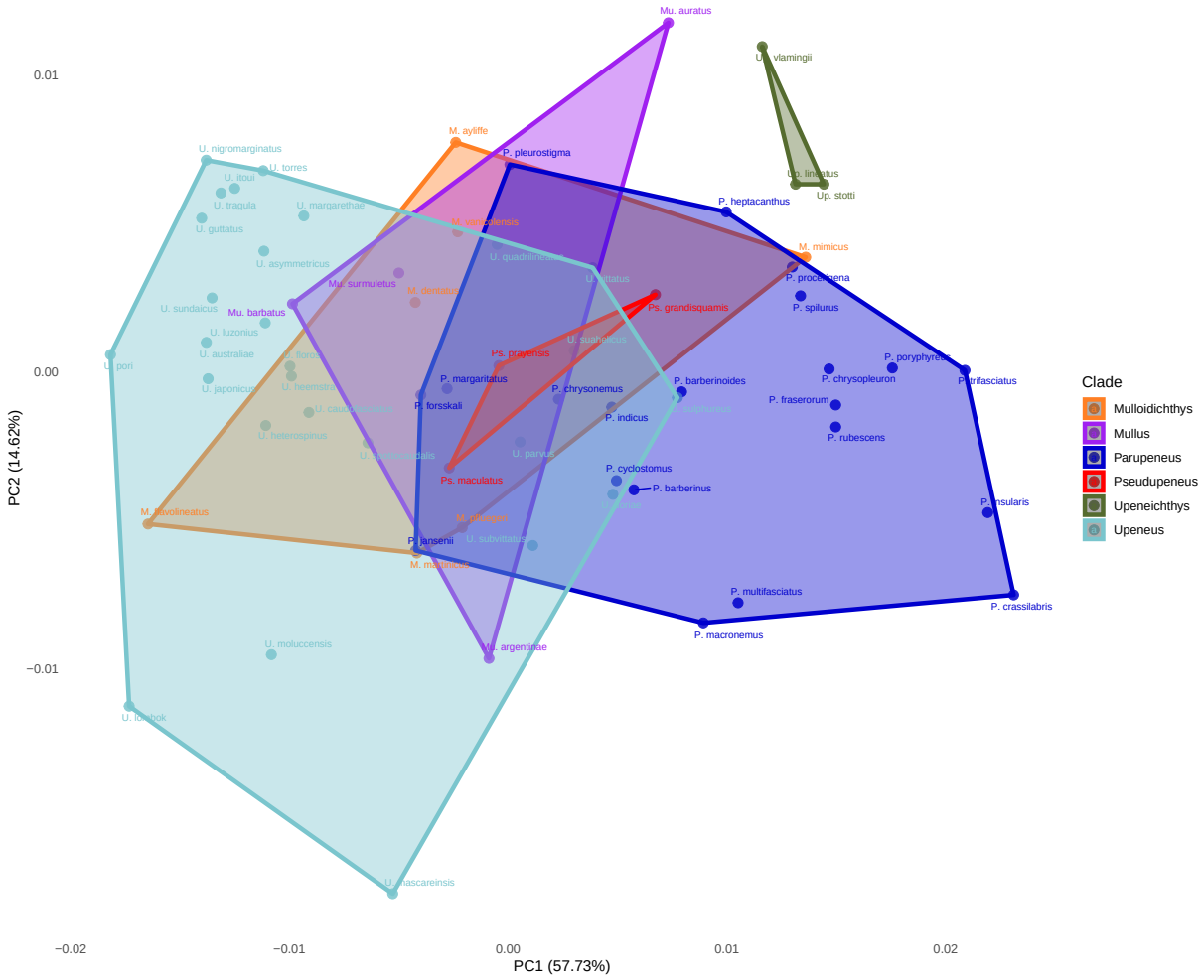

Figure 4.A.4: **Morphospace of Dorsal/Ventral curve subset of the Mullidae.** Genera are indicated by convex hulls.

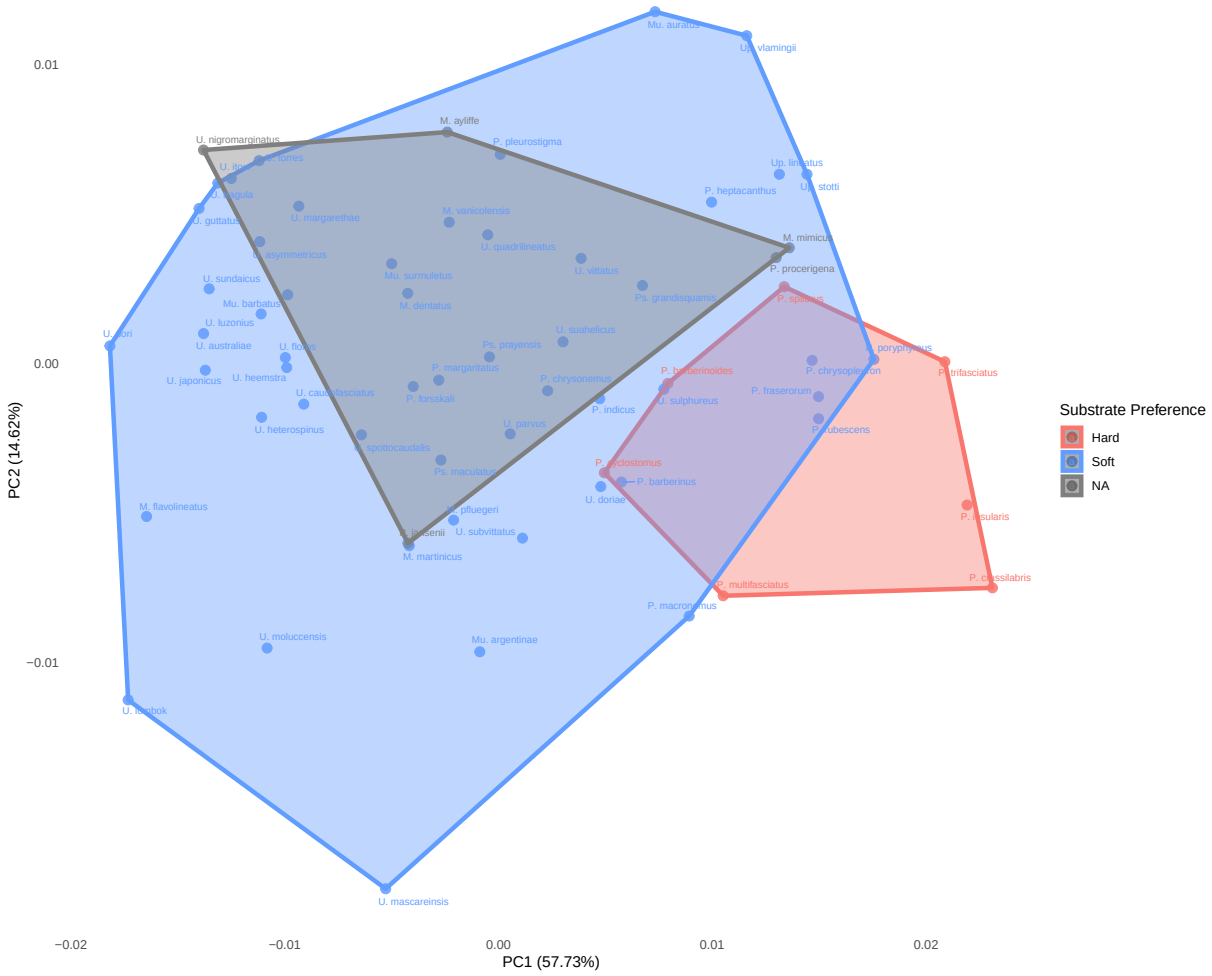

Figure 4.A.5: **Substrate preference morphospace of Dorsal/Ventral curve subset of the Mullidae.** Substrate preference morphospace of Dorsal/Ventral curve subset of the Mullidae. Substrate preference is indicated by convex hulls.

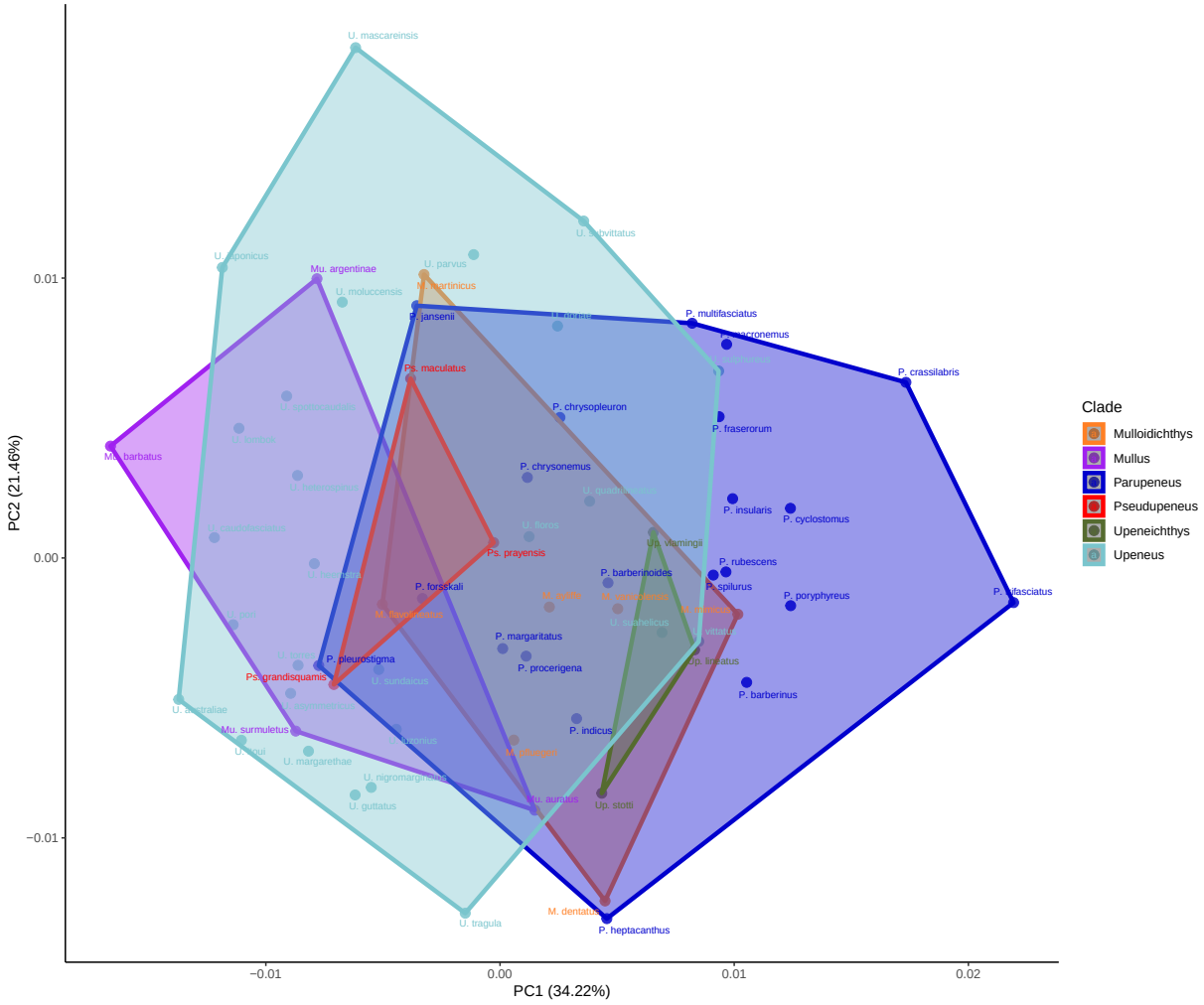

Figure 4.A.6: **Morphospace of caudal fin curve subset of the Mullidae.** Genera are indicated by convex hulls.

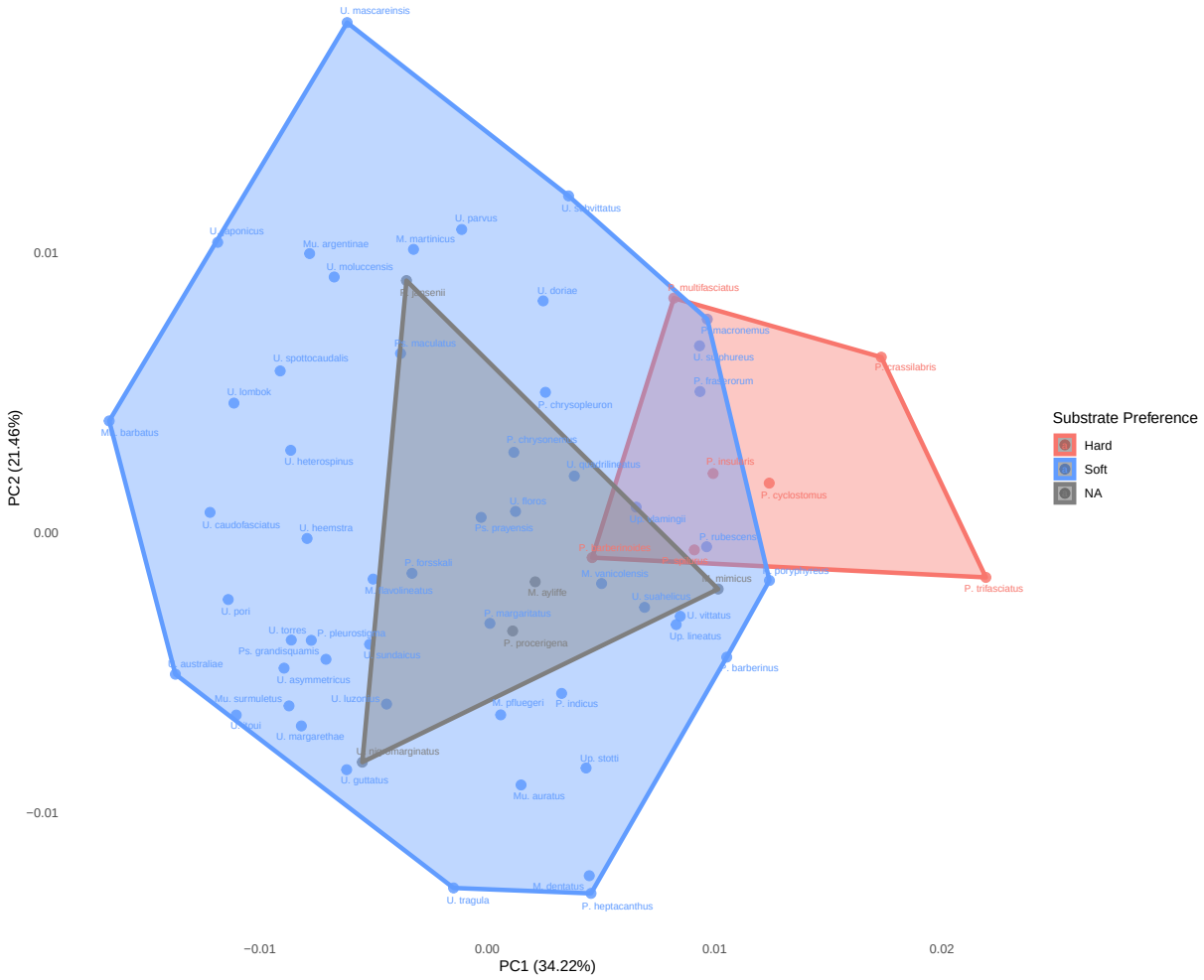

Figure 4.A.7: **Substrate preference morphospace of caudal fin curve subset of the Mullidae.** Substrate preference morphospace of caudal fin curve subset of the Mullidae. Genera are indicated by convex hulls.

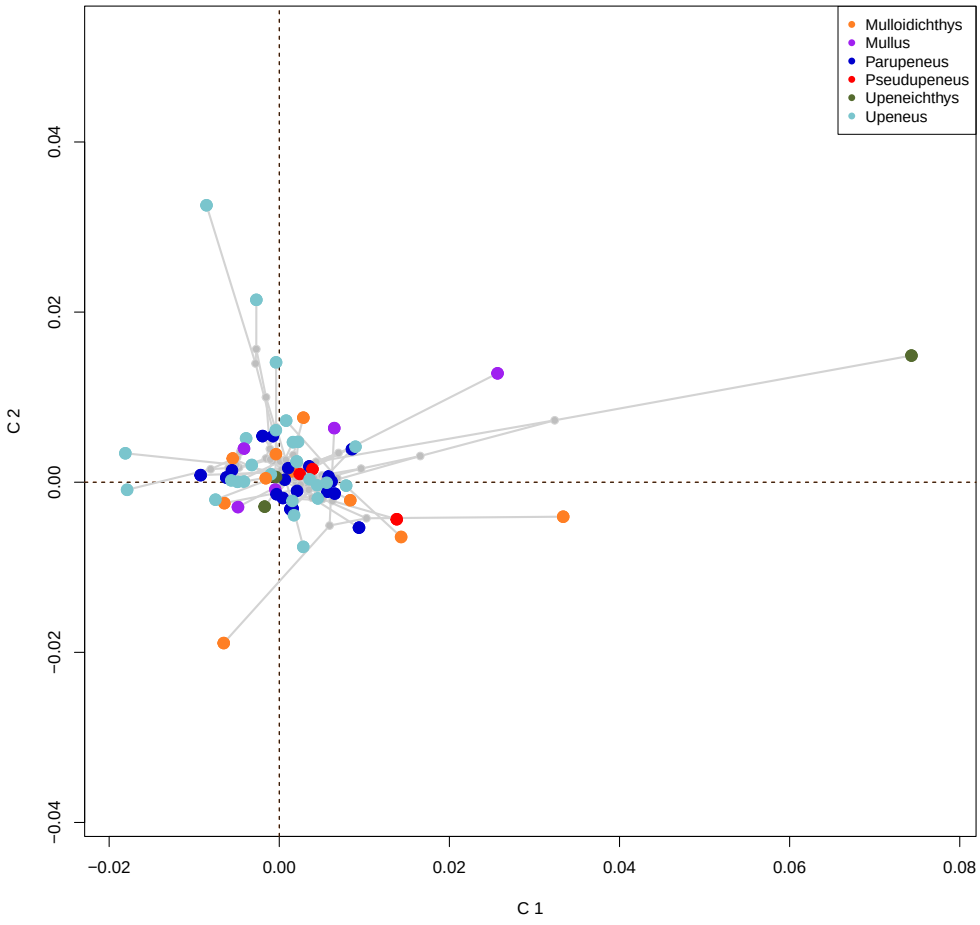

Figure 4.A.8: **Phylomorphospace of forehead curve subset of the Mullidae.**

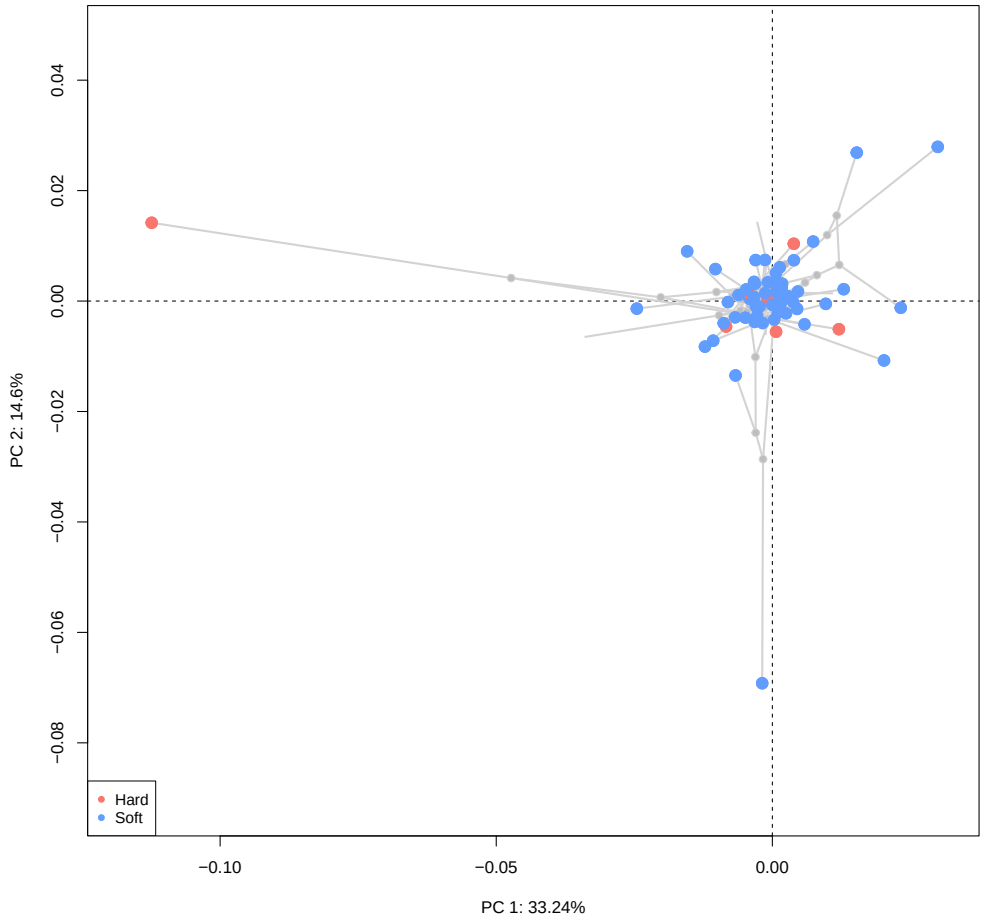

Figure 4.A.9: **PaCA of full body dataset for the Mullidae**PaCA of full body dataset for the Mullidae. Substrate preference is indicated by tip color on the phylogenetically aligned components using Procrustes shape coordinates.

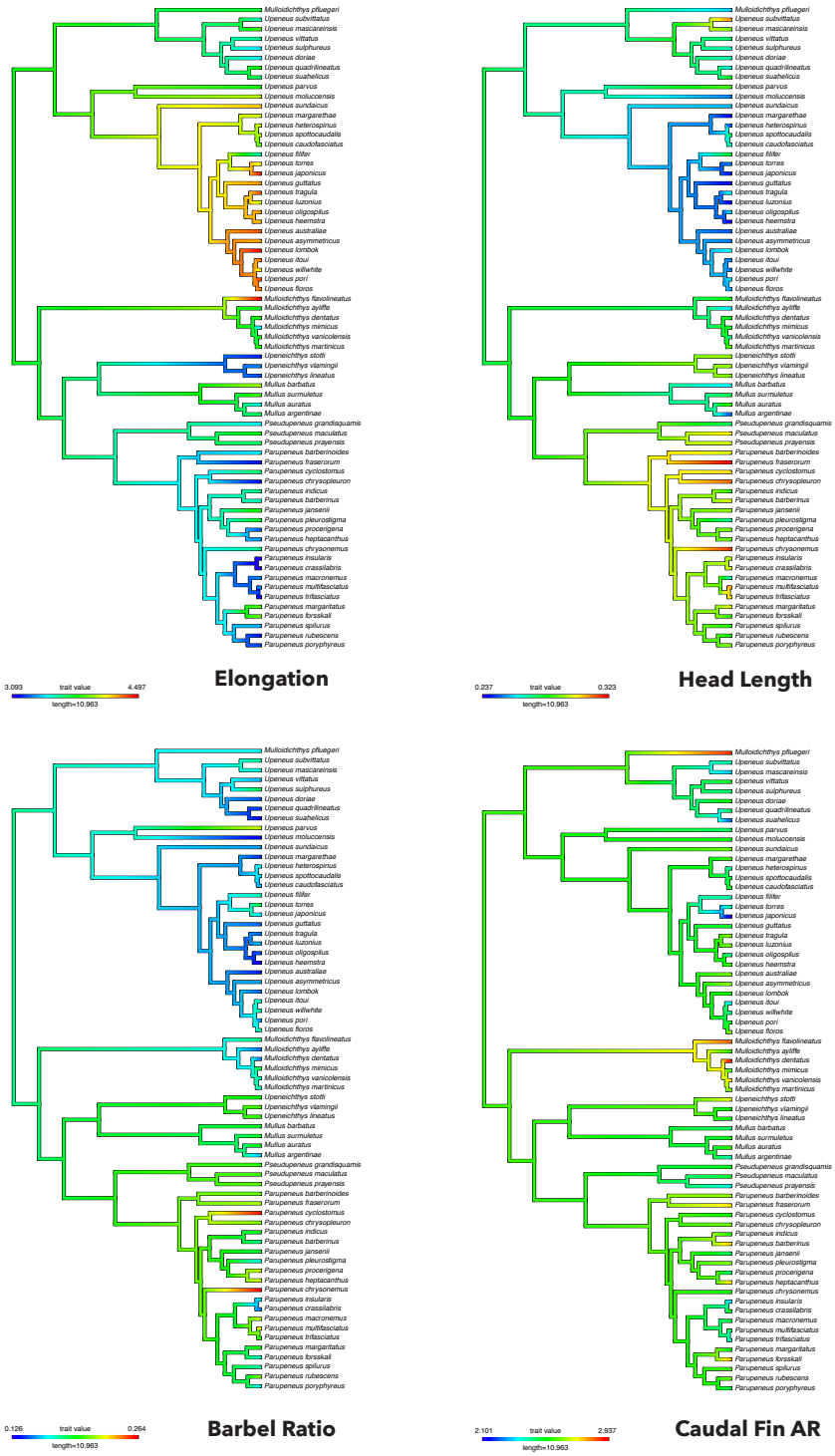

Figure 4.A.10: **Mapped ancestral estimations for morphological metrics.**Mapped ancestral estimations for morphological metrics. All were inferred using the ContMap from phytools, in which branch lengths are color coded in associated with the mapped character state: warmer colors indicate larger values and cooler colors represent smaller values.

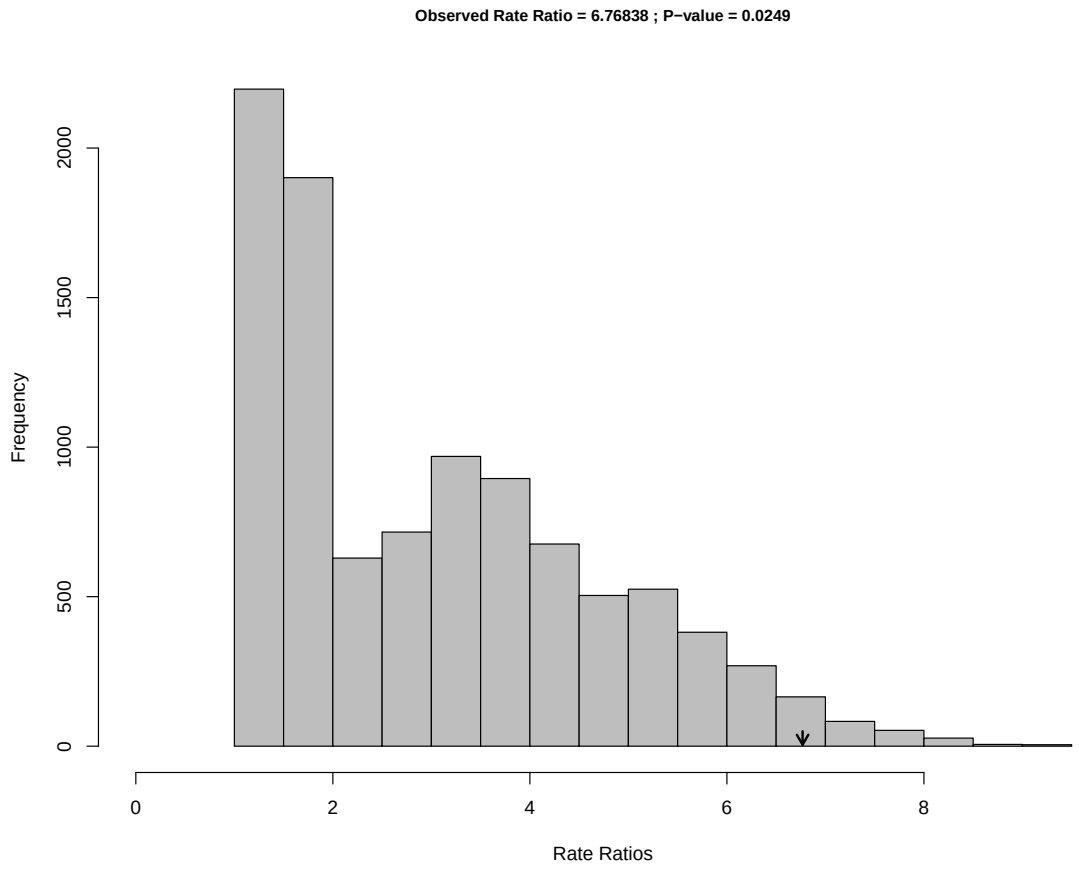

Figure 4.A.11: **Distribution of rate ratios of Full Body shape evolution associated with each substrate preference.**

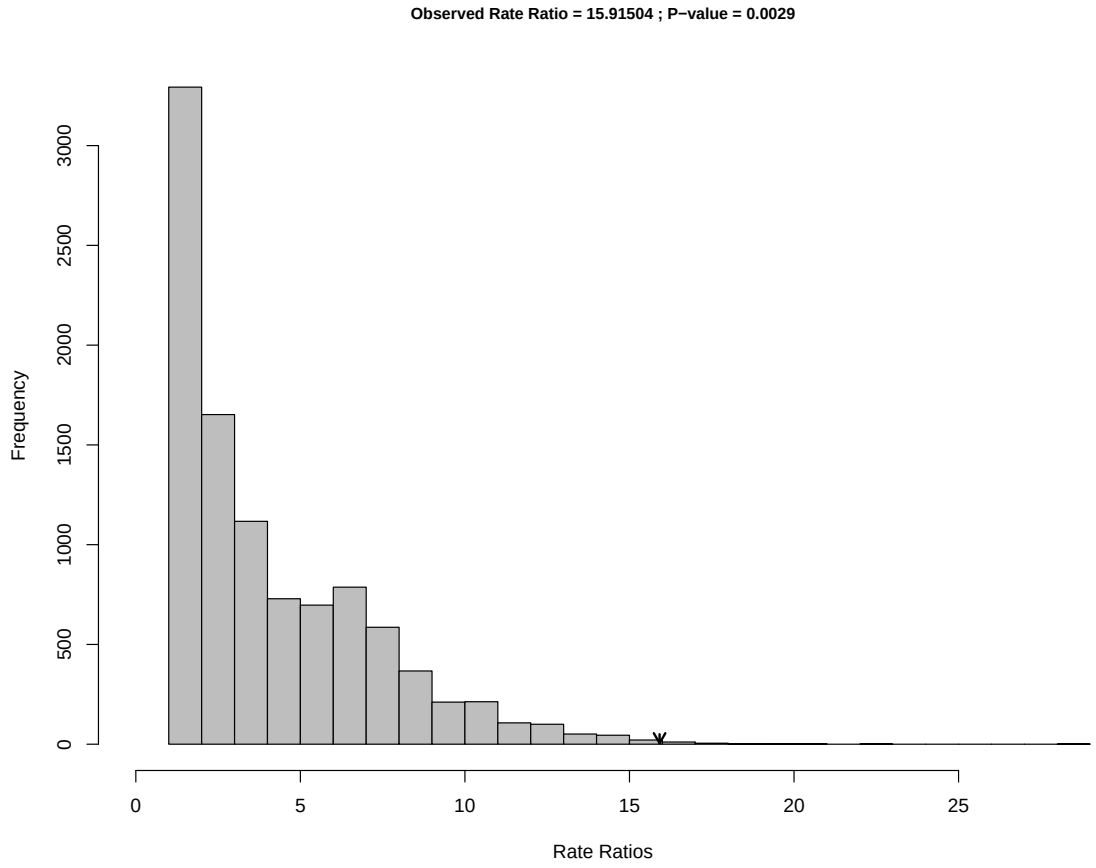

Figure 4.A.12: **Distribution of rate ratios of Forehead shape evolution associated with each substrate preference.**

CHAPTER 5

CONCLUSION AND FUTURE DIRECTIONS

5.1 Conclusion

The overarching goal of this thesis was to explore the phylogenetics, evolution and biogeography of the goatfishes, with a focus on the unique foraging behavior associated with substrate preferences and barbel use. Prior to this thesis, little was known about their evolutionary history, spatial distribution, morphological diversity, and ecological niches in comparison to other reef associated families, despite their common occurrence in these systems. This dissertation represents a major advance in the integrative examination of the evolutionary relationships, assemblage patterns, and ecomorphological associations across the Mullidae. The results of this thesis provide a comprehensive examination of the many unique characteristics associated with the goatfishes, provides the phylogenetic framework required to fully understand the biology of this family, and enables us to begin addressing questions about how future changes in reef ecosystem health will impact the biology of this clade.

One primary result of this thesis is the inference of the most species rich, time-calibrated phylogeny of the goatfishes to date. This phylogeny is used as the basis for all evolutionary analyses performed throughout this thesis and reveals novel aspects of goatfish evolution that were previously unknown. For example, the root age of the Mullidae is estimated to have occurred within the Miocene, which is far younger than most comparable families with similar species richness and biogeographic distribution. Additionally, we find evidence for unstable rates of diversification, which is based on the combination of long internal branch lengths paired with divergences within the past 5 Ma between most extant species. This variability in diversification and species accumulation is also associated with an increase in body shape diversification across the family.

Another main result of this dissertation is the delineation of goatfish assemblage groups using bioregionalization approaches. The boundaries between these bioregions are specific to the Mullidae, highlighting global patterns of species turnover and the locations of potential barriers to dispersal. These bioregions represent the complex influence of scale on biogeographic distributions, as the boundaries between regions represent the impact of both historic vicariance events and current ecological processes on extant species ranges. This varied history results in variability in species richness, phylogenetic diversity, and rates of endemism across regions, with the major distinction found between regions in the Atlantic/East Pacific compared to those in the Indo-Pacific oceanic basins. Additionally, we find limited evidence for regional patterns in lineage diversification, but rather we find that patterns of biodiversity are driven by areas of range overlap with relatively high levels of phylogenetic diversity.

The final major result of this dissertation is robust evidence for an ecomorphological relationship between multiple morphometric variables of body shape, fin shape and forehead shape, and preferential foraging on different substrate types across the goatfishes. Geometric morphometrics of head, fin, barbel and body shape morphology show that most patterns of shape dispersion are strongly associated with phylogenetic relationships, with major axes of variability including elongation ratio, head length, barbel length, eye size, and dorsal fin positioning. Despite this overall variability in morphology, hard substrate foraging is primarily associated with forehead shape, elongation ratio, and head width. Notably, preferential feeding on hard substrate, such as coral rubble, is restricted to several species with the genus *Parupeneus*. Although this transition occurred relatively recently, the rate of morphological change associated with forehead shape is significantly higher than in species that forage on soft substrates. Given the variability in behavior associated with foraging, these results suggest that species that preferentially forage on soft substrates may be limited in the types of behaviors they are capable of due to aspects of their morphology.

5.2 Future Directions

5.2.1 Comparative barbel and cranial anatomy

The way in which organisms receive information from their environment is through sense organs that detect and process a range of stimuli, such as chemical, electrical, mechanical, and optical. Due to the large variability of sensory stimuli within habitats and trophic ecology, an examination of the mechanisms that influence and regulate sensory systems is essential to understanding the distribution of taxa across space and time (Lombarte and Aguirre 1997). In fishes, sensory systems involved with feeding behaviors have evolved into a diverse array of morphological traits that result in both the specialization and loss of certain sensory capabilities. These physical adaptations highlight the strong connection between ontogenetic development and environmental stimuli, which have the potential to greatly impact sensory development, growth patterns, behavior, and mortality rates of individuals (McCormick 1993). The unique sensory morphology of goatfishes make this clade ideal for a robust examination of the comparative anatomy across species, in comparison with variability in foraging behaviors across a variety of benthic habitats.

The gustatory system in fishes has undergone rather remarkable adaptive evolution resulting in large amounts of variation based on the functional needs of the taxa. Gustation, in conjunction with olfaction, have evolved to become a functionally interconnected chemosensory complex that plays an important role in fish feeding (Devitsina 2006). These sensory systems have become highly specialized due to chemical stimuli being present within the water column and benthos, facilitating chemosensory food search behaviors (Kasumyan 2002). In particular, the evolution of the extraoral system of taste reception has resulted in morphological variability among actinopterygians due to the development of extraoral taste buds located within the epithelium, which increases the surface area that is in direct contact with the environment (Herrick 1903; Devitsina 2006; LeClair and Topczewski 2010). These

Figure 5.1: **Unique use of barbels in *Mulloidichthys flavolineatus*.** Source: C. Nash.

are most commonly expressed in the form of ventrally located projections, which range from facial barbels to extensions of pectoral fin rays, that contain sense organs (Gosline 1984). One hypothesis for the evolution of barbels is the need for a compensatory mechanism with increased sensory sensitivity in turbid waters where the visual field is reduced (Lombarte and Aguirre 1997; LeClair and Topczewski 2010). From an ecomorphological perspective, the evolution of barbels allows for the utilization of resources that are otherwise unavailable within the benthos.

In the simplest anatomical definition, a teleost barbel must consist of an outer epithelium, dermal connective tissue, blood vessels, and taste buds that are innervated by extensions of the facial nerve (LeClair and Topczewski 2010). However, the number, location, and structure of barbels is highly variable among and within fish species. Historically barbels were used as a phylogenetically informative character used within the taxonomy due to their variability (LeClair and Topczewski 2010). However, the usefulness of barbels as a phylogenetic signal has become relatively obsolete due to the strong body of evidence suggesting that barbels were the result of numerous independent evolutionary events (Kim

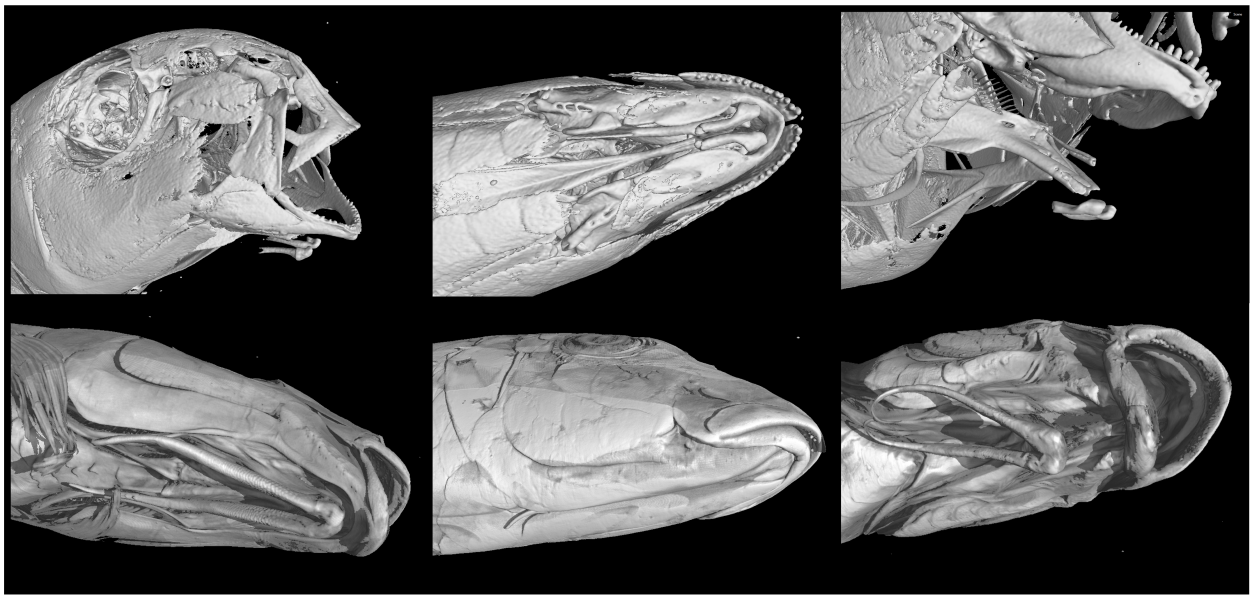

Figure 5.2: **Head and barbel CT scans for three species of goatfish.** A) *Mulloidichthys flavolineatus*, B) *Pseudupeneus maculatus*, and C) *Parupeneus bifasciatus.*

2002; LeClair and Topczewski 2010).

The most derived and specialized barbel structure is found within the Mullidae. The goatfish's hyoid barbels are modified, rod-like branchiostegal rays that provide stability and a strengthening support (Gosline 1984). Modifications of branchiostegal rays in conjunction with sensory systems are rare, having convergently evolved in only the Mullidae and the distantly related genus Polymixia (Gosline 1984). The barbel originates from a cap of fibrocartilage that forms a socket with the tip of the forward projection of the anterior ceratohyal of the hyoid bar (Gosline 1984; Kim et al. 2001). This articulation increases the freedom of movement of the barbel, which facilitates the barbel's role in probing for food and as an excavation tool (Gosline 1984; Kim et al. 2001). The indiviual barbel includes a large nerve ending encased within each barbel, with only the basal portion being completely ossified (Kim et al. 2001). Notably, there is variarion in the degree of ossification of the barbel bone across species (Kim et al. 2001; Kim 2002).

There are two mechanically separate components that facilitate barbel manipulation for species within the Mullidae. The first component is with regard to the movement of the barbel in relation to their articulation with the hyoid bars, and the second component controls the lowering and raising of the anterior ends of the pair of hyoid bars in relation to the head of the fish (Gosline 1984). Each barbel has their own musculature that allows for independent movement, with the following four muscles associated with each hyoid arch: the extensor tentaculi, retractor tentaculi, and two sections of the rotator tentaculi (Gosline 1984; Kim et al. 2001). Additionally, the taste buds on the barbels are found in high densities and are innervated in a strict orthogonal fashion by bundles of nerve fibers from the facial nerve (McCormick 1993; Kirino et al. 2006, 2013). In a comparison among two species, it was observed that although the epidermal area and the mean size of the taste buds are highly correlated with barbel length, taste bud density and cross-sectional area of the nerve bundle were not (McCormick 1993). The distribution of taste buds can also differ even among species, with the taste buds on some species forming clusters while others are solitary (Lombarte and Aguirre 1997). These differences in the distribution of taste buds among clades indicates that there is variability in the organization of the gustatory system in different lineages of goatfishes (Kirino et al. 2006). For example, an examination of the Mullidae genus *Parupeneus* revealed external taste buds on the barbels and the outer skin of the head, while the taste buds possessed by members of the genus *Upeneus* are limited to the barbel region (Kirino et al. 2006).

The primary aims of a developing future study are to quantify anatomical variation related to the cranium and barbel to examine the evolution of these traits across the Mullidae. This will yield a comprehensive dataset in which key anatomical traits can be analyzed to examine patterns of feeding capabilities and behavior. Additionally, this morphological trait data will be used to analyze the distribution of traits within and among different biogeographic regions to elucidate complex relationships among feeding capabilities, ecological

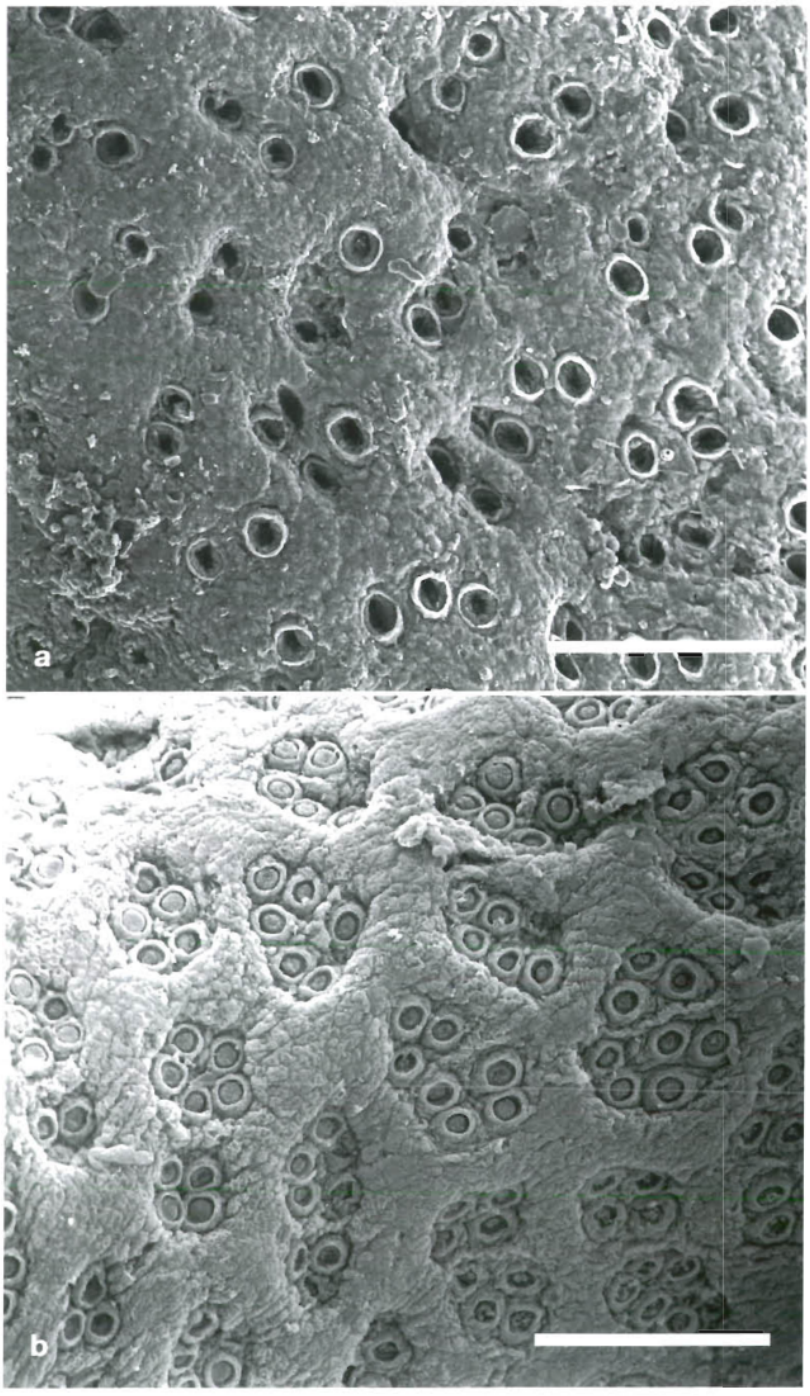

Figure 5.3: **Distribution of taste buds from two species of Mullidae.** A) *Mullus surmuletus* and B) *Mullus barbatus*. Source: Lombarte and Aguirre (1997).

niches, and range determinants. Variation in external facial and barbel morphology can be quantified through either standard morphometric landmarks or three-dimensional surface scanning. Data on the osteological and soft tissue anatomy of the barbel mechanism can be collected using micro-computed tomography (μCT) scans of the craniofacial region. μCT scanning is an ideal method because it is non-destructive to the specimens and can be used to more accurately quantify volume and spatial variation. Additionally, it would be of interest to quantify intraspecific variation and ontogenetic changes given evidence for plasticity in barbel development (McCormick 1993; Pavlov et al. 2011). Variation in taste bud structure and distribution would be collected using Scanning Electron Microscopy (SEM). Combined, these scans will be used to quantify key traits, such as 1) muscle attachment angle, 2) muscle volume, 3) barbel length, width, and volume, 4) ceratohyal length, 5) dentition, and 6) taste bud complexity. Quantification of this anatomical variation will be used to examine the evolution and biomechanical potential of the barbel apparatus across species, which can then be applied to increasing our understanding of how head and barbel structure influence the unique foraging behavior of the goatfishes.

5.2.2 Comparative kinematics of barbel use during foraging

A goatfish's barbels are the primary appendage used to detect food items within loose substrates through a large variety of behaviors that differ among lineages. Different species of goatfish are hypothesized to have a strong preference on the substrate type that they inhabit and feed on (Gosline 1984). For example, some species use their barbels as an excavation device to turn over soft substrate (i.e. sand and mud), while other species use their barbels to probe within crevices and dislodge small animals (Gosline 1984). Another interesting behavior exhibited by *Mulloidichthys flavolineatus* is the blowing of the substrate away once a food item is detected within the substrate. Goatfishes also can feed on plankton within the water column through jaw protrusion while their barbels are retracted, although no adult

Figure 5.4: **Frame of *P. multifasciatus* foraging.** Lateral tank view from kinematics experiment in Moorea, French Polynesia. Source: C. Nash.

goatfish has been shown to feed primarily on the plankton (Krajewski and Bonaldo 2006). An examination of the complex interrelationships among feeding morphology, behavior, and habitat use can be used to inform predictions on how effectively species of goatfish will be able to adapt to changing environments. These experiments and observations will provide an estimate of how well goatfishes can perform when feeding on non-preferential substrate types, revealing species that may be particularly vulnerable to changes in habitat availability.

To accomplish these aims, we have begun the collection of three-dimensional kinematic data through filming a controlled foraging event on multiple substrate types using high speed videography. In collaboration with UChicago undergraduate Olivia Grobmyer, we are digitizing and calculating the motion of the head and barbels during each feeding event. The data from these videos will be used to quantitatively calculate the intra- and interspecific variation in feeding behavior across different substrate types. Additionally, this data will be integrated with morphological measurements from micro-CT scans of museum specimens to create a 3D biomechanical model of how individuals use their barbels during a foraging event. This model will be used to infer feeding capabilities across the goatfish phylogeny

to predict which species are able to feed on non-preferential substrate types and adapt to multiple habitat types.

Currently we have data on *Parupeneus multifascitus*, and plan to expand our behavior observations across species. To examine the variability in feeding behavior and habitat specificity in situ, I will generate a robust observational database of foraging and feeding behaviors from different reef localities. One option is to use a time budget approach, in which one follows an individual for a fixed amount of time and estimates the percentage of time spent performing different behaviors, such as searching for food, digging, feeding, and interacting with other individuals. Another option is to use stationary camera rigs that will record video footage for a set amount of time. To stimulate foraging behavior, baited food sources can be placed in front of the camera rig. Ideally, several cameras would be placed in areas with divergent substrate types and reef zones to capture any variability in behavior across species. Additionally, this set up would be able to capture other behaviors, such as conspecific and interspecific social interactions. Environmental data, such as temperature, salinity, current speed, and turbidity will be recorded for each site. Additional data, such as diet composition, depth range, diurnal patterns, and spawning times, will be collected through a meta-analysis of the primary literature and personal observations.

5.2.3 Comparative plasticity in barbel development across the goatfishes

Due to the large variability within habitats and trophic ecology, an examination of the mechanisms that influence and regulate sensory systems is essential to understanding the distribution of taxa across space and time (Lombarte and Aguirre 1997). In fishes, sensory systems involved with feeding behaviors have evolved into a diverse array of morphological traits that result in both the specialization and loss of certain sensory capabilities. For example, cavefish have undergone a severe reduction in their optic capabilities due to a lack of exposure to light sources and have subsequently undergone an increase in the functionality

of their olfactory systems as a form of sensory compensation (Jeffery 2009; Eifert et al. 2015). These physical adaptations highlight the strong connection between ontogenetic development and environmental stimuli, which have the potential to greatly impact sensory development, growth patterns, behavior, and mortality rates of individuals (McCormick 1993). Larval goatfishes have the general teleost arrangement of their branchiostegeal rays and only begin to utilize their barbels post settlement (Gosline 1984). During development, the anterior branchiostegeal ray separate from the other rays and migrate forward on the hyoid bar to become level with the tip of the glossohyal (Gosline 1984; McCormick and Molony 1993). The barbels show a period rapid growth during the settlement phase, in which the individual transitions from its pelagic life phase to its benthic phase (McCormick and Molony 1993). This is represented by an over 50% increase in barbel length over a twelve-hour period (McCormick and Molony 1993). Additionally, it has been observed that the level of food availability during the settlement phase of *Upeneus tragula* in affects the size of the barbel and the size of the taste buds depending on the treatment (McCormick 1993). Interestingly, goatfishes deprived of food developed the longest barbels with the largest taste buds, while individuals that had access to an excess of food developed shorter barbels (McCormick and Molony 1993). This relationship, although not necessarily intuitive, does make sense functionally. A longer barbel with larger taste buds can increase the surface area available to detect and obtain food. This also indicates that, at least in the specific species studied, there is some degree of plasticity in the development of the barbels dependent on the environment at the time of settlement.

The ability for the sensory system to adapt based on the environment that an organism experiences as it develops greatly increases the potential fitness of that individual. Current research suggests that different elements of the teleost sensory system have the mechanisms required to respond to a variety of environmental and biotic coniditions (Holland 1978; McCormick and Molony 1993). For barbels in particular, the ability to regenerate taste buds

after being damaged is a recurrent trait found across many lineages. Catfishes, in particular, have been studied for over 100 years to investigate mechanisms of cell and sensory regeneration. Although the quality of the regenerated barbel is dependent on the species, an experiment by Katagiri (1966) demonstrated that the regenerated barbel of a juvenile Japanese catfish, *Parasilurus asotus*, expressed functional taste buds three days after amputation and eventually reached its initial length and size (Katagiri 1966). Sturgeons have also been observed to regenerate their barbels, as the effectiveness of using the amputation a barbels as a mark-and-recapture technique for individuals of short-nose sturgeon was found to be rather ineffective method for long-term study as the 100% of the individuals partially regenerated their barbel within two years (Collins et al. 1994). Barbel regeneration has also been tested in zebrafish, in which regenerated barbels were found to be grossly similar in terms of proportion and pigmentation to the original barbel (LeClair and Topczewski 2010). Additionally, the vasculature, nerves, and positioning of taste buds were restored by similar cell types in their original position (LeClair and Topczewski 2010). However, the central rod of connective tissue within the barbel was replaced by a scar tissue like assemblage of mesenchymal cells within extracellular matrix (LeClair and Topczewski 2010).

No prior work appears to have explicitly examined the regeneration of goatfish barbels. An evaluation of the plasticity of the sensory system begs the question of what precisely is causing the variation observed among taxa. Various fish clades have adapted and utilized their extra-oral taste buds in order to access a novel and underutilized food source (Kiyohara et al. 2002; Kirino et al. 2006). However, the mechanism by which these adaptations have occurred remains a mystery. For example, is the observed variation in barbel and brain structure within fishes caused by alterations in fish feeding behavior that resulted in changes in brain structure, or did the changes in brain structure precede and drive the changes in feeding behavior? To answer these questions requires an intimate knowledge of the variation within and among species, as well as the mechanisms that produce morphological change.

Although this is an arduous task, this knowledge is crucial in understanding the exact ways in which organisms interact with the environment and which ways the environment impacts the evolution of an organism. One place to start would be a detailed examination of the plasticity of barbel and brain development among goatfishes under different environmental conditions.

This study will provide insight on both the level of plastic capabilities in both barbel and brain structure, as well as the degree in which the environment has an impact on morphological development. Another important avenue of research is a detailed analysis of the genome and gene expression patterns in order to better understand the underlying genetic mechanisms and potential epigenetic effects. It is unquestionable that the goatfishes have evolved extraordinary morphological and sensory adaptations in order to best utilize their habitats. An understanding of how goatfishes interact and evolve with their environment is essential in order to inform accurate predictions on the resilience of goatfish species to impending environmental changes.

5.2.4 Concluding Remarks

The research described in this thesis will greatly contribute to our understanding of evolutionary and ecological patterns in a relatively poorly studied clade, will enhance our ability to analyze and predict marine biogeographic patterns, and will enable broader analyses of reef fishes across multiple families. Integrated with patterns of biodiversity, an examination of the evolutionary diversification of important functional traits is a powerful tool that can be used to understand the processes that drive current lineage distributions and dispersal patterns, to explore how shifts between regions have occurred historically, and to predict where lineages will persist in the future. The datasets generated as part of this thesis will be used to address further questions about the life history, comparative anatomy, and phylogeography across species within this clade and beyond. In conclusion, the research presented in this dissertation yields a substantial contribution to our understanding of the phylogenetic history, biogeographic distribution, comparative morphology, and ecomorphological associations of the goatfishes.

REFERENCES

Abaad, M., Tuset, V. M., Montero, D., Lombarte, A., Otero-Ferrer, J. L., and Haroun, R. (2016). Phenotypic plasticity in wild marine fishes associated with fish-cage aquaculture. *Hydrobiologia*, 765:343–358.

Ackerly, D. D., Schwilk, D. W., and Webb, C. O. (2006). Niche evolution and adaptive radiation: Testing the order of trait divergence. *Ecology*, 87:50–61.

Adams, D., Collyer, M., Kaliontzopoulou, A., and Baken, E. (2021). Geomorph: Software for geometric morphometric analyses. R package version 4.0.2.

Adams, D. C. (2014). Quantifying and comparing phylogenetic evolutionary rates for shape and other high-dimensional phenotypic data. *Systematic Biology*, 63:166–177.

Adams, D. C. and Collyer, M. L. (2018). Multivariate phylogenetic comparative methods: evaluations, comparisons, and recommendations. *Systematic biology*, 67:14–31.

Adams, D. C. and Collyer, M. L. (2019). Comparing the strength of modular signal, and evaluating alternative modular hypotheses, using covariance ratio effect sizes with morphometric data. *Evolution*, 73:2352–2367.

Adams, D. C. and Otárola-Castillo, E. (2013). geomorph: an r package for the collection and analysis of geometric morphometric shape data. *Methods in Ecology and Evolution*, 4:393–399.

Adams, D. C., Rohlf, F. J., and Slice, D. E. (2013). A field comes of age: Geometric morphometrics in the 21st century. *Hystrix*, 24:7–14.

Aiello, B. R., Hardy, A. R., Westneat, M. W., and Hale, M. E. (2018). Fins as mechanosensors for movement and touch-related behaviors. *Integrative and Comparative Biology*, 58:844–859.

Albertson, R. C. (2008). Morphological divergence predicts habitat partitioning in a Lake Malawi Cichlid species complex. *Copeia*, 2008:689–698.

Angelini, D. (2023). borealis: Tools for reproducible geometric morphometric analysis. R package version 2022.10.27.

Bakali, M. E., Talbaoui, M., and Bendriss, A. (2010). Régime alimentaire du rouget de roche (*Mullus surmuletus* l.)(téléostéen, Mullidae) de la côte nord-ouest méditerranéenne du maroc (région de m'diq). *Bulletin de l'Institut Scientifique, Rabat, section Sciences de la Vie*, 32:87.

Baken, E. K., Collyer, M. L., Kaliontzopoulou, A., and Adams, D. C. (2021). geomorph v4.0 and gmshiny: Enhanced analytics and a new graphical interface for a comprehensive morphometric experience. *Methods in Ecology and Evolution*, 12:2355–2363.

Bannikov, A. F. (1999). Unusual new percoid fish from the Eocene of Bolca. *Miscellanea Paleontologica. Studi e Ricerche sui Giacimenti Terziari di Bolca*, 8:117–128.

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

Bartlett, M. S. (1937). Properties of sufficiency and statistical tests. *Proceedings of the Royal Society of London. Series A - Mathematical and Physical Sciences*, 160:268–282.

Barton, D. E. and David, F. N. (1959). The dispersion of a number of species. *Journal of the Royal Statistical Society: Series B (Methodological)*, 21:190–194.

Baselga, A. and Orme, C. D. L. (2012). betapart: an r package for the study of beta diversity. *Methods in Ecology and Evolution*, 3:808–812.

Bauer, D. F. (1972). Constructing confidence sets using rank statistics. *Journal of the American Statistical Association*, 67:687–690.

Beddow, H. M., Liebrand, D., Sluijs, A., Wade, B. S., and Lourens, L. J. (2016). Global change across the Oligocene-Miocene transition: High-resolution stable isotope records from iodp site u1334 (equatorial Pacific ocean). *Paleoceanography*, 31:81–97.

Bellwood, D. R., Goatley, C. H. R., and Bellwood, O. (2016). The evolution of fishes and corals on reefs: Form, function and interdependence. *Biological Reviews*, 92:878–901.

Bellwood, D. R. and Wainwright, P. C. (2002). The history and biogeography of fishes on coral reefs. *Coral reef fishes: dynamics and diversity in a complex ecosystem*, 5:32.

Bender, M. G., Leprieur, F., Mouillot, D., Kulbicki, M., Parravicini, V., Pie, M. R., Barneche, D. R., Oliveira-Santos, L. G. R., and Floeter, S. R. (2017). Isolation drives taxonomic and functional nestedness in tropical reef fish faunas. *Ecography*, 40:425–435.

Bender, M. G., Pie, M. R., Rezende, E. L., Mouillot, D., Floeter, S. R., Macroecology, M., and Lab, B. (2013). Biogeographic, historical and environmental influences on the taxonomic and functional structure of Atlantic reef fish assemblages. *Global Ecology and Biogeography*, 22:1173–1182.

Betancur, R. R., Wiley, E. O., Arratia, G., Acero, A., Bailly, N., Miya, M., Lecointre, G., and Ortí, G. (2017). Phylogenetic classification of bony fishes. *BMC Evolutionary Biology 2017 17:1*, 17:1–40.

Betancur-R, R., Broughton, R. E., Wiley, E. O., Carpenter, K., López, J. A., Li, C., Holcroft, N. I., Arcila, D., Sanciangco, M., II, J. C. C., Zhang, F., Buser, T., Campbell, M. A., Ballesteros, J. A., Roa-Varon, A., Willis, S., Borden, W. C., Rowley, T., Reneau, P. C., Hough, D. J., Lu, G., Grande, T., Arratia, G., and Ortí, G. (2013). The tree of life and a new classification of bony fishes. *PLoS currents*, 5.

Bivand, R. and Rundel, C. (2021). rgeos: Interface to geometry engine - open source ('geos'). https://r-forge.r-project.org/projects/rgeos/.

Bivand, R. S., Pebesma, E., and Gomez-Rubio, V. (2013). *Applied spatial data analysis with R, Second edition*. Springer, NY.

Blomberg, S. P., Garland, T., and Ives, A. R. (2003). Testing for phylogenetic signal in comparative data: Behavioral traits are more labile. *Evolution*, 57:717–745.

Boettiger, C., Lang, D. T., and Wainwright, P. (2012). rfishbase: exploring, manipulating and visualizing fishbase data from r. *Journal of Fish Biology*.

Bolger, A. M., Lohse, M., and Usadel, B. (2014). Trimmomatic: a flexible trimmer for illumina sequence data. *Bioinformatics*, 30:2114–2120.

Bos, A. R. (2014). *Upeneus nigromarginatus*, a new species of goatfish (Perciformes: Mullidae) from the Philippines phylogenetic relationships among populations of Indo-Pacific fishes. *Bulletin of Zoology*.

Bouckaert, R., Heled, J., Kühnert, D., Vaughan, T., Wu, C. H., Xie, D., Suchard, M. A., Rambaut, A., and Drummond, A. J. (2014). Beast 2: A software platform for Bayesian evolutionary analysis. *PLOS Computational Biology*, 10:e1003537.

Bouton, N., Witte, F., and Alphen, J. J. M. V. (2002). Experimental evidence for adaptive phenotypic plasticity in a rock-dwelling Cichlid fish from Lake Victoria. *Biological Journal of the Linnean Society Biological Journal of the Linnean Society Biological Journal of the Linnean Society*, 77:185–192.

Bradbury, I. R., Laurel, B., Snelgrove, P. V., Bentzen, P., and Campana, S. E. (2008). Global patterns in marine dispersal estimates: the influence of geography, taxonomic category and life history. *Proceedings of the Royal Society B: Biological Sciences*, 275:1803–1809.

Braga, P. H. P., Kembel, S., and Peres-Neto, P. (2023). Historical and contemporary processes drive global phylogenetic structure across geographical scales: Insights from bat communities. *Global Ecology and Biogeography*.

Brandl, S. J., Robbins, W. D., Bellwood, D. R., and Brandl, S. J. (2015). Exploring the nature of ecological specialization in a coral reef fish community: morphology , diet and foraging microhabitat use. *Proceedings of the Royal Society B*, 282:20151147.

Bridge, T. C. L., Luiz, O. J., Coleman, R. R., Kane, C. N., and Kosaki, R. K. (2016). Ecological and morphological traits predict depth-generalist fishes on coral reefs. *Proceedings of the Royal Society B: Biological Sciences*, 283:20152332.

Briggs, J. C. (1961). The east Pacific barrier and the distribution of marine shore fishes. *Evolution*, 15:545–554.

Briggs, J. C. (1974). Operation of zoogeographic barriers. *Systematic Biology*, 23:248–256.

Briggs, J. C. (1995). *Global biogeography*.

Bryant, J. A., Lamanna, C., Morlon, H., Kerkhoff, A. J., Enquist, B. J., and Green, J. L. (2008). Microbes on mountainsides: Contrasting elevational patterns of bacterial and plant diversity. *Proceedings of the National Academy of Sciences*, 105:11505–11511.

Burns, M. D. and Sidlauskas, B. L. (2019). Ancient and contingent body shape diversification in a hyperdiverse continental fish radiation. *Evolution*, 73:569–587.

Butler, M. and King, A. (2004). Phylogenetic comparative analysis: A modeling approach for adaptive evolution. *The American Naturalist*, 164:683–695.

Cantalice, K. and Alvarado-Ortega, J. (2016). *Eekaulostomus cuevasae gen. and sp. nov.*, an ancient armored trumpetfish (Aulostomoidea) from Danian (Paleocene) marine deposits of Belisario Domínguez, Chiapas, southeastern Mexico. *Palaeontologia Electronica*.

Carnevale, G. and Bannikov, A. F. (2019). A dragonet (Teleostei, Callionymoidei) from the Eocene of Monte Bolca, Italy. *Bollettino della Societa Paleontologica Italiana*, 58:295–307.

Carnevale, G., Bannikov, A. F., Landini, W., and Sorbini, C. (2006). Volhynian (early Sarmatian sensu lato) fishes from Tsurevsky, north Caucasus, Russia. *Journal of Paleontology*, 80(4):684–699.

Ceccarelli, S., Justi, S. A., Rabinovich, J. E., Filho, J. A. F. D., and Villalobos, F. (2020). Phylogenetic structure of geographical co-occurrence among new world Triatominae species, vectors of chagas disease. *Journal of Biogeography*, 47:1218–1231.

Chamberlain, S., Barve, V., Mcglinn, D., Oldoni, D., Desmet, P., Geffert, L., and Ram, K. (2023). rgbif: Interface to the global biodiversity information facility api. R package version 3.7.5.

Chernomor, O., Haeseler, A. V., and Minh, B. Q. (2016). Terrace aware data structure for phylogenomic inference from supermatrices. *Systematic Biology*, 65:997–1008.

Choat, J. H. and Kingett, P. D. (1982). The influence of fish predation on the abundance cycles of an algal turf invertebrate fauna. *Oecologia*, 54:88–95.

Ciampaglio, C. N., Kemp, M., and McShea, D. W. (2001). Detecting changes in morphospace occupation patterns in the fossil record: characterization and analysis of measures of disparity. *Paleobiology*, 27:695–715.

Claverie, T. and Wainwright, P. C. (2014). A morphospace for reef fishes: elongation is the dominant axis of body shape evolution. *PloS one*, 9(11):e112732.

Collins, M. R., Smith, T. I. J., and Heyward, L. D. (1994). Effectiveness of six methods for marking juvenile shortnose sturgeons. *The Progressive Fish-Culturist*, 56:250–254.

Cooper, W. J. and Westneat, M. (2009). Form and function of damselfish skulls: Rapid and repeated evolution into a limited number of trophic niches. *BMC Evolutionary Biology*, 9:24.

Cowman, P. F. (2014). Historical factors that have shaped the evolution of tropical reef fishes: A review of phylogenies, biogeography, and remaining questions. *Frontiers in Genetics*, 5:1–15.

Cowman, P. F. and Bellwood, D. R. (2013). The historical biogeography of coral reef fishes: Global patterns of origination and dispersal. *Journal of Biogeography*, 40:209–224.

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

Cowman, P. F., Parravicini, V., Kulbicki, M., and Floeter, S. R. (2017). The biogeography of tropical reef fishes: Endemism and provinciality through time. *Biological Reviews*.

Daru, B. H., Elliott, T. L., Park, D. S., and Davies, T. J. (2017). Understanding the processes underpinning patterns of phylogenetic regionalization. *Trends in Ecology and Evolution*, 32:845–860.

Daru, B. H., van der Bank, M., Maurin, O., Yessoufou, K., Schaefer, H., Slingsby, J. A., and Davies, T. J. (2016). A novel phylogenetic regionalization of phytogeographical zones of southern Africa reveals their hidden evolutionary affinities. *Journal of Biogeography*, 43:155–166.

Denton, J. S. S. and Adams, D. C. (2015). A new phylogenetic test for comparing multiple high-dimensional evolutionary rates suggests interplay of evolutionary rates and modularity in lanternfishes (Myctophiformes; Myctophidae). *Evolution*, 69:2425–2440.

Devitsina, G. V. (2006). Adaptive variability of the gustatory system receptor part of the carp *Cyprinus carpio* (Cyprinidae, Teleostei) after chronic anosmia. *Journal of Evolutionary Biochemistry and Physiology*, 42:743–750.

Diogo, R. and Chardon, M. (2000). The structures associated with catfish (Teleostei: Siluriformes) mandibular barbels: Origin, anatomy, function, taxonomic distribution, nomenclature and synonymy. *Netherlands Journal of Zoology*, 50:455–478.

Drummond, A. J., Nicholls, G. K., Rodrigo, A. G., and Solomon, W. (2002). Estimating mutation parameters, population history and genealogy simultaneously from temporally spaced sequence data. *Genetics*, 161:1307–1320.

Eakin, R. R., Eastman, J. T., and Vacchi, M. (2006). Sexual dimorphism and mental barbel structure in the south Georgia plunderfish *Artedidraco mirus* (Perciformes: Notothenioidei: Artedidraconidae). *Polar Biology*, 30:45–52.

Eifert, C., Farnworth, M., Schulz-Mirbach, T., Riesch, R., Bierbach, D., Klaus, S., Wurster, A., Tobler, M., Streit, B., Indy, J. R., Arias-Rodriguez, L., and Plath, M. (2015). Brain size variation in extremophile fish: Local adaptation versus phenotypic plasticity. *Journal of Zoology*, 295:143–153.

Espiñeira, M. and Vieites, J. M. (2014). Genetic system for traceability of goatfishes by fins methodology and authentication of mullets (*Mullus barbatus* and *Mullus surmuletus*) by rt-pcr. *European Food Research and Technology*, 240:423–429.

Evans, K. M., Larouche, O., West, J. L., Gartner, S. M., and Westneat, M. W. (2023). Burrowing constrains patterns of skull shape evolution in wrasses. *Evolution & Development*, 25(1):73–84.

Evans, M. E., Smith, S. A., Flynn, R. S., and Donoghue, M. J. (2009). Climate, niche evolution, and diversification of the "bird-cage" evening primroses (*Oenothera*, sections *Anogra* and *Kleinia). American Naturalist*, 173:225–240.

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

Faircloth, B. C., McCormack, J. E., Crawford, N. G., Harvey, M. G., Brumfield, R. T., and Glenn, T. C. (2012). Ultraconserved elements anchor thousands of genetic markers spanning multiple evolutionary timescales. *Systematic Biology*, 61:717–726.

Faircloth, B. C., Sorenson, L., Santini, F., and Alfaro, M. E. (2013). A phylogenomic perspective on the radiation of ray-finned fishes based upon targeted sequencing of ultraconserved elements (UCEs). *PLOS ONE*, 8:e65923.

Felsenstein, J. (1973). Maximum-likelihood estimation of evolutionary trees from continuous characters. *American journal of human genetics*, 25:471.

Fernandez-Silva, I., Randall, J. E., Coleman, R. R., Dibattista, J. D., Rocha, L. A., Reimer, J. D., Meyer, C. G., and Bowen, B. W. (2015). Yellow tails in the Red Sea: Phylogeography of the Indo-Pacific goatfish *Mulloidichthys flavolineatus* reveals isolation in peripheral provinces and cryptic evolutionary lineages. *Journal of Biogeography*, 42:2402–2413.

Fernandez-Silva, I., Snelgrove, B. N., and Bowen, B. W. (2013). Twelve microsatellite dna markers to resolve population structure of the yellow-striped goatfish *Mulloidichthys flavolineatus* (family Mullidae). *Conservation Genetics Resources*, 5:565–568.

Fietz, K., Trofimenko, E., Guerin, P. E., Arnal, V., Torres-Oliva, M., Lobréaux, S., Pérez-Ruzafa, A., Manel, S., and Puebla, O. (2020). New genomic resources for three exploited Mediterranean fishes. *Genomics*, 112:4297–4303.

Floeter, S. R. and Gasparini, J. L. (2000). The southwestern Atlantic reef fish fauna: Composition and zoogeographic patterns. *Journal of Fish Biology*, 56:1099–1114.

Floeter, S. R., Rocha, L. A., Robertson, D. R., Joyeux, J. C., Smith-Vaniz, W. F., Wirtz, P., Edwards, A. J., Barreiros, J. P., Ferreira, C. E., Gasparini, J. L., Brito, A., Falcón, J. M., Bowen, B. W., and Bernardi, G. (2008). Atlantic reef fish biogeography and evolution. *Journal of Biogeography*, 35:22–47.

Foote, M. (1997). The evolution of morphological diversity. *Annual Review of Ecology and Systematics*, 28:129–152.

Freckleton, R. P., Harvey, P. H., and Pagel, M. (2002). Phylogenetic analysis and comparative data: a test and review of evidence. *The American Naturalist*.

Fricke, R., Eschmeyer, W. N., and der Laan, R. V. (2022). *Eschmeyer's Catalog of Fishes: Genera, Species, References*.

Friedman, S. T., Martinez, C. M., Price, S. A., and Wainwright, P. C. (2019). The influence of size on body shape diversification across Indo-Pacific shore fishes. *Evolution*, 73:1873–1884.

Friedman, S. T., Price, S. A., Corn, K. A., Larouche, O., Martinez, C. M., and Wainwright, P. C. (2020). Body shape diversification along the benthic–pelagic axis in marine fishes. *Proceedings of the Royal Society B: Biological Sciences*, 287:20201053.

Frontier, S. (1976). Étude de la décroissance des valeurs propres dans une analyse en composantes principales: Comparaison avec le moddle du bâton brisé. *Journal of Experimental Marine Biology and Ecology*, 25:67–75.

Gaither, M. R., Bowen, B. W., Rocha, L. A., and Briggs, J. C. (2016). Fishes that rule the world: circumtropical distributions revisited. *Fish and Fisheries*, 17:664–679.

Galarza, J. A., Turner, G. F., Macpherson, E., and Rico, C. (2009). Patterns of genetic differentiation between two co-occurring demersal species: The red mullet (*Mullus barbatus*) and the striped red mullet (*Mullus surmuletus*). *Canadian Journal of Fisheries and Aquatic Sciences*, 66:1478–1490.

Galil, B. S. (2007). Loss or gain? invasive aliens and biodiversity in the Mediterranean Sea. *Marine Pollution Bulletin*, 55:314–322.

Gans, C. and Northcutt, R. G. (1983). Neural crest and the origin of vertebrates: A new head. *Source: Science, New Series*, 220:268–274.

George, A. B. and Westneat, M. W. (2019). Functional morphology of endurance swimming performance and gait transition strategies in balistoid fishes. *Journal of Experimental Biology*, 222.

Golani, D. (1999). Genetic relationship in goatfishes (Mullidae: Perciformes) of the Red Sea and the Mediterranean, with remarks on Suez Canal migrants. *Scientia Marina*, 63:129–135.

Gosline, W. A. (1984). Structure, function, and ecology in the goatfishes (family Mullidae). *Pacific Science*, 38:312–323.

Gould, S. J. and Eldredge, N. (1977). Punctuated equilibria: the tempo and mode of evolution reconsidered. *Paleobiology*, 3:115–151.

Gower, J. C. (1975). Generalized Procrustes analysis. *Psychometrika 1975 40:1*, 40:33–51.

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., Palma, F. D., Birren, B. W., Nusbaum, C., Lindblad-Toh, K., Friedman, N., and Regev, A. (2011). Full-length transcriptome assembly from RNA-seq data without a reference genome. *Nature Biotechnology*, 29:644–652.

Grant, W. S. and Bowen, B. W. (1998). Shallow population histories in deep evolutionary lineages of marine fishes: Insights from sardines and anchovies and lessons for conservation. *Journal of Heredity*, 89:415–426.

Gregory, W. K. (1933). Fish skulls: a study of the evolution of natural mechanisms. *Trans Amer Philos Soc*, 23:75–481.

Grenié, M., Mouillot, D., Villéger, S., Denelle, P., Tucker, C. M., Munoz, F., and Violle, C. (2018). Functional rarity of coral reef fishes at the global scale: Hotspots and challenges for conservation. *Biological Conservation*, 226:288–299.

Guillemot, N., Kulbicki, M., Chabanet, P., and Vigliola, L. (2011). Functional redundancy patterns reveal non-random assembly rules in a species-rich marine assemblage. *PLoS ONE*, 6:e26735.

Guralnick, R. and Constable, H. (2010). Vertnet: creating a data-sharing community. *BioScience*, 60:258–259.

Hanel, R., Westneat, M. W., and Sturmbauer, C. (2002). Phylogenetic relationships, evolution of broodcare behavior, and geographic speciation in the wrasse tribe Labrini. *Journal of Molecular Evolution*, 55:776–789.

Harmon, L. J., Losos, J. B., Davies, T. J., Gillespie, R. G., Gittleman, J. L., Jennings, W. B., Kozak, K. H., McPeek, M. A., Moreno-Roark, F., Near, T. J., et al. (2010). Early bursts of body size and shape evolution are rare in comparative data. *Evolution: International Journal of Organic Evolution*, 64:2385–2396.

Harmon, L. J., Schulte, J. A., Larson, A., and Losos, J. B. (2003). Tempo and mode of evolutionary radiation in iguanian lizards. *Science*, 301:961–964.

Harmon, L. J., Weir, J. T., Brock, C. D., Glor, R. E., and Challenger, W. (2008). Geiger: investigating evolutionary radiations. *Bioinformatics*, 24:129–131.

Hedman, M. M. (2010). Constraints on clade ages from fossil outgroups. *Paleobiology*, 36:16–31.

Helmus, M., Bland, T., Williams, C., and Ives, A. (2007). Phylogenetic measures of biodiversity. *The American Naturalist*, 169:E68–E83.

Hench, K., Helmkampf, M., McMillan, W. O., and Puebla, O. (2022). Rapid radiation in a highly diverse marine environment. *Proceedings of the National Academy of Sciences of the United States of America*, 119.

Hernández, C. M., Paris, C. B., Vaz, A. C., Benjamin, Jones, T., Kellner, J. B., David, Richardson, E., Sponaugle, S., Cowen, R. K., and Llopiz, J. K. (2023). Diverse patterns of larval coral reef fish vertical distribution and consequences for dispersal and connectivity. *Coral Reefs 2023*, 1:1–13.

Herrick, C. J. (1903). *The Organ and Sense of Taste in Fishes*. U. S. Commission of Fish and Fisheries.

Hoang, D. T., Chernomor, O., Haeseler, A. V., Minh, B. Q., and Vinh, L. S. (2018). Ufboot2: Improving the ultrafast bootstrap approximation. *Molecular Biology and Evolution*, 35:518–522.

Hobson, E. S. and Chess, J. R. (2001). Influence of trophic relations on form and behavior among fishes and benthic invertebrates in some California marine communities. *Environmental Biology of Fishes*, 60:411–457.

Hoeksema, B. W. (2007). *Delineation of the Indo-Malayan Centre of Maximum Marine Biodiversity: The Coral Triangle*.

Holland, K. (1978). Chemosensory orientation to food by a Hawaiian goatfish (*Parupeneus porphyreus*, Mullidae). *Journal of Chemical Ecology*, 4:173–186.

Hollander, M., Wolfe, D. A., and Chicken, E. (2013). *Nonparametric statistical methods*. John Wiley & Sons.

Hughes, L. C., Ortí, G., Huang, Y., Sun, Y., Baldwin, C. C., Thompson, A. W., Arcila, D., Betancur, R., Li, C., Becker, L., Bellora, N., Zhao, X., Li, X., Wang, M., Fang, C., Xie, B., Zhoui, Z., Huang, H., Chen, S., Venkatesh, B., and Shi, Q. (2018). Comprehensive phylogeny of ray-finned fishes (Actinopterygii) based on transcriptomic and genomic data. *Proceedings of the National Academy of Sciences of the United States of America*, 115:6249–6254.

Hughes, L. C., Ortí, G., Saad, H., Li, C., White, W. T., Baldwin, C. C., Crandall, K. A., Arcila, D., and Betancur-R, R. (2021). Exon probe sets and bioinformatics pipelines for all levels of fish phylogenomics. *Molecular Ecology Resources*, 21:816–833.

Hughes, T. P., Bellwood, D. R., and Connolly, S. R. (2002). Biodiversity hotspots, centres of endemicity, and the conservation of coral reefs. *Ecology Letters*, 5:775–784.

Ives, A. R., Midford, P. E., and Garland, T. (2007). Within-species variation and measurement error in phylogenetic comparative methods. *Systematic Biology*, 56:252–270.

Jeffery, W. R. (2009). Regressive evolution in *Astyanax* cavefish. *Annual Review of Genetics*, 43:25–47.

Kalyaanamoorthy, S., Minh, B. Q., Wong, T. K., Haeseler, A. V., and Jermiin, L. S. (2017). Modelfinder: fast model selection for accurate phylogenetic estimates. *Nature Methods 2017 14:6*, 14:587–589.

Kasumyan, A. O. (2002). Sturgeon food searching behaviour evoked by chemical stimuli: A reliable sensory mechanism. *Journal of Applied Ichthyology*, 18:685–690.

Katagiri, Y. (1966). Regeneration of the mandibular barbels of the fry of the catfish, *Parasilurus asotus*, and comparison of histological structure of the mandibular barbels of three catfishes inhabiting Lake Biwa. *Japanese Journal of Ichthyology*, 13:169–175.

Katoh, K., Misawa, K., Kuma, K. I., and Miyata, T. (2002). Mafft: a novel method for rapid multiple sequence alignment based on fast fourier transform. *Nucleic Acids Research*, 30:3059–3066.

Katoh, K. and Standley, D. M. (2013). Mafft multiple sequence alignment software version 7: Improvements in performance and usability. *Molecular Biology and Evolution*, 30:772–780.

Kembel, S. W., Cowan, P. D., Helmus, M. R., Cornwell, W. K., Morlon, H., Ackerly, D. D., Blomberg, S. P., and Webb, C. O. (2010). Picante: R tools for integrating phylogenies and ecology. *Bioinformatics Applications Note*, 26:1463–1464.

Kilbourne, B. M. (2017). Selective regimes and functional anatomy in the mustelid forelimb: Diversification toward specializations for climbing, digging, and swimming. *Ecology and Evolution*, 7:8852–8863.

Kim, B. (2002). Comparative anatomy and phylogeny of the family Mullidae (Teleostei: Perciformes). *Mem. Grad. Sch. Fish. Sci. Hokkaido Univ.*, pages 1–75.

Kim, B., Yabe, M., and Nakaya, K. (2001). Barbels and related muscles in Mullidae (Perciformes) and Polymixiidae (Polymixiiformes). *Ichthyological Research*, 48:409–413.

Kim, B. J. and Amaoka, K. (2001). A new species, *Parupeneus procerigena*, from the Saya de Malha Bank in the western Indian Ocean (Perciformes: Mullidae). *Ichthyological Research*, 48:45–50.

Kirino, M., Ikenaga, T., Tsukahara, J., Lamb, C. F., and Kiyohara, S. (2006). Primary taste center in the goatfish of genus *Parupeneus*. *Fisheries Science*, 72:461–468.

Kirino, M., Parnes, J., Hansen, A., Kiyohara, S., and Finger, T. E. (2013). Evolutionary origins of taste buds: phylogenetic analysis of purinergic neurotransmission in epithelial chemosensors. *Open biology*, 3:130015.

Kisel, Y., Mcinnes, L., Toomey, N. H., and Orme, C. D. L. (2011). How diversification rates and diversity limits combine to create large-scale species-area relationships. *Philosophical Transactions of the Royal Society B: Biological Sciences*, 366:2514–2525.

Kiyohara, S., Sakata, Y., Yoshitomi, T., and Tsukahara, J. (2002). The 'goatee' of goatfish: innervation of taste buds in the barbels and their representation in the brain. *Proceedings of the Royal Society B: Biological Sciences*, 269:1773–1780.

Klingenberg, C. P. (2016). Size, shape, and form: concepts of allometry in geometric morphometrics. *Development Genes and Evolution*, 226:113.

Kolasinski, J., Frouin, P., Sallon, A., Rogers, K., Bruggemann, H. J., and Potier, M. (2009). Feeding ecology and ontogenetic dietary shift of yellowstripe goatfish *Mulloidichthys flavolineatus* (Mullidae) at Reunion Island, SW Indian Ocean. *Marine Ecology Progress Series*, 386:181–195.

Krajewski, J. P. and Bonaldo, R. M. (2006). Plankton-picking by the goatfish *Pseudupeneus maculatus* (Mullidae), a specialized bottom forager. *Journal of Fish Biology*, 68:925–930.

Krajewski, J. P., Bonaldo, R. M., Sazima, C., and Sazima, I. (2006). Foraging activity and behaviour of two goatfish species (Perciformes: Mullidae) at Fernando de Noronha Archipelago, tropical West Atlantic. *Environmental Biology of Fishes*, 77:1–8.

Kreft, H. and Jetz, W. (2010). A framework for delineating biogeographical regions based on species distributions. *Journal of Biogeography*, 37:2029–2053.

Kulbicki, M., Parravicini, V., Bellwood, D. R., Arias-Gonzàlez, E., Chabanet, P., Floeter, S. R., Friedlander, A., McPherson, J., Myers, R. E., Vigliola, L., and Mouillot, D. (2013). Global biogeography of reef fishes: A hierarchical quantitative delineation of regions. *PLoS ONE*, 8:e81847.

Landini, W. and Sorbini, C. (2005). Evolutionary dynamics in the fish faunas of the Mediterranean basin during the Plio-Pleistocene. *Quaternary International*, 140-141:64–89.

Larsson, A. (2014). Aliview: a fast and lightweight alignment viewer and editor for large datasets. *Bioinformatics*, 30:3276–3278.

LeClair, E. E. and Topczewski, J. (2010). Development and regeneration of the zebrafish maxillary barbel: A novel study system for vertebrate tissue growth and repair. *PLoS ONE*, 5.

Leis, J. M. (2007). Behaviour as input for modelling dispersal of fish larvae: Behaviour, biogeography, hydrodynamics, ontogeny, physiology and phylogeny meet hydrography. *Marine Ecology Progress Series*, 347:185–193.

Leprieur, F., Descombes, P., Gaboriau, T., Cowman, P. F., Parravicini, V., Kulbicki, M., Melián, C. J., de Santana, C. N., Heine, C., Mouillot, D., Bellwood, D. R., and Pellissier, L.

(2016). Plate tectonics drive tropical reef biodiversity dynamics. *Nature Communications*, 7:11461.

Leprieur, F., Pellissier, L., Mouillot, D., and Gaboriau, T. (2021). Influence of historical changes in tropical reef habitat on the diversification of coral reef fishes. *Scientific Reports*, 11:20731.

Lessios, H. A. and Robertson, D. R. (2006). Crossing the impassable: genetic connections in 20 reef fishes across the eastern Pacific barrier. *Proceedings. Biological sciences*, 273:2201–8.

Lessios, H. A. and Robertson, D. R. (2013). Speciation on a round planet: Phylogeography of the goatfish genus *Mulloidichthys*. *Journal of Biogeography*, 40:2373–2384.

Lester, S. E. and Ruttenberg, B. I. (2005). The relationship between pelagic larval duration and range size in tropical reef fishes: a synthetic analysis. *Proceedings. Biological sciences*, 272:585–91.

Li, J. and Asner, G. P. (2023). Global analysis of benthic complexity in shallow coral reefs. *Environmental Research Letters*, 18.

Liem, K. F. (1973). Evolutionary strategies and morphological innovations: Cichlid pharyngeal jaws. *Systematic Zoology*, 22:425.

Liem, K. F. (1980). Adaptive significance of intra-and interspecific differences in the feeding repertoires of Cichlid fishes. *American Zoologist*, 20:295–314.

Lombarte, A. and Aguirre, H. (1997). Quantitative differences in the chemoreceptor systems in the barbels of two species of Mullidae (*Mullus surmuletus* and *M. barbatus*) with different bottom habitats. *Marine Ecology Progress Series*, 150:57–64.

Lombarte, A., Recasens, L., González, M., and de Sola, L. G. (2000). Spatial segregation of two species of Mullidae in relation to habitat. *Marine Ecology Progress Series*, 206:239–249.

Longo, S., Faircloth, B., Meyer, A., Westneat, M., Alfaro, M., and Wainwright, P. (2017). Phylogenomic analysis of a rapid radiation of misfit fishes (Syngnathiformes) using ultraconserved elements. *Molecular Phylogenetics and Evolution*.

Louca, S. and Pennell, M. W. (2020). Extant timetrees are consistent with a myriad of diversification histories. *Nature*, 580:502–505.

Luiz, O. J., Allen, A. P., Robertson, D. R., Floeter, S. R., Kulbicki, M., Vigliola, L., Becheler, R., and Madin, J. S. (2013). Adult and larval traits as determinants of geographic range size among tropical reef fishes. *Proceedings of the National Academy of Sciences*, 110:16498–16502.

Luiz, O. J., Madin, J. S., Robertson, D. R., Rocha, L. A., Wirtz, P., and Floeter, S. R. (2012). Ecological traits influencing range expansion across large oceanic dispersal barriers: Insights from tropical Atlantic reef fishes. *Proceedings of the Royal Society B: Biological Sciences*, 279:1033–1040.

Lukoschek, V. and McCormick, M. I. (2000). A review of multi-species foraging associations in fishes and their ecological significance. *Proceedings of the Ninth International Coral Reef Symposium*, I:467–474.

Lukoschek, V. and McCormick, M. I. (2001). Ontogeny of diet changes in a tropical benthic carnivorous fish, *Parupeneus barberinus* (Mullidae): Relationship between foraging behaviour, habitat use, jaw size, and prey selection. *Marine Biology*, 138:1099–1113.

López-Fernández, H., Arbour, J. H., Winemiller, K. O., and Honeycutt, R. L. (2013). Testing for ancient adaptive radiations in neotropical Cichlid fishes. *Evolution*, 67:1321–1337.

Lübbers, J., Kuhnt, W., Holbourn, A. E., Bolton, C. T., Gray, E., Usui, Y., Kochhann, K. G., Beil, S., and Andersen, N. (2019). The middle to late Miocene "carbonate crash" in the equatorial Indian Ocean. *Paleoceanography and Paleoclimatology*, 34:813–832.

Maggio, T., Brutto, S. L., Garoia, F., Tinti, F., and Arculeo, M. (2009). Microsatellite analysis of red mullet *Mullus barbatus* (Perciformes, Mullidae) reveals the isolation of the adriatic basin in the Mediterranean Sea. *ICES Journal of Marine Science*, 66:1883–1891.

Martinez, C. M., McGee, M. D., Borstein, S. R., and Wainwright, P. C. (2018). Feeding ecology underlies the evolution of Cichlid jaw mobility. *Evolution*, 72:1645–1655.

McCord, C. L., Nash, C. M., Cooper, W. J., and Westneat, M. W. (2021). Phylogeny of the damselfishes (Pomacentridae) and patterns of asymmetrical diversification in body size and feeding ecology. *PLoS One*, 16(10):e0258889.

McCord, C. L. and Westneat, M. W. (2016). Evolutionary patterns of shape and functional diversification in the skull and jaw musculature of triggerfishes (Teleostei: Balistidae). *Journal of Morphology*, 277:737–752.

McCormick, M. I. (1993). Development and changes at settlement in the barbel structure of the reef fish, *Upeneus tragula* (Mullidae). *Environmental Biology of Fishes*, 37:269–282.

McCormick, M. I. (1995). Fish feeding on mobile benthic invertebrates: influence of spatial variability in habitat associations. *Marine Biology*, 121:627–637.

McCormick, M. I. and Milicich, M. J. (1993). Late pelagic-stage goatfishes: distribution patterns and inferences on schooling behaviour. *Journal of Experimental Marine Biology and Ecology*, 174:15–42.

McCormick, M. I. and Molony, B. W. (1993). Quality of the reef fish *Upeneus tragula* (Mullidae) at settlement: is size a good indicator of condition? *Marine Ecology Progress Series*, 98:45–54.

McCormick, M. I. and Molony, B. W. (1995). Influence of water temperature during the larval stage on size, age and body condition of a tropical reef fish at settlement. *Marine ecology progress series. Oldendorf*, 118:59–68.

McGee, M. D., Borstein, S. R., Meier, J. I., Marques, D. A., Mwaiko, S., Taabu, A., Kishe, M. A., O'Meara, B., Bruggmann, R., Excoffier, L., and Seehausen, O. (2020). The ecological and genomic basis of explosive adaptive radiation. *Nature*, 586:75–79.

McMahan, C. D., Chakrabarty, P., Sparks, J. S., Smith, W. M., and Davis, M. P. (2013). Temporal patterns of diversification across global Cichlid biodiversity (Acanthomorpha: Cichlidae). *PLOS ONE*, 8:e71162.

Melo, B. F., Sidlauskas, B. L., Near, T. J., Roxo, F. F., Ghezelayagh, A., Ochoa, L. E., Stiassny, M. L., Arroyave, J., Chang, J., Faircloth, B. C., MacGuigan, D. J., Harrington, R. C., Benine, R. C., Burns, M. D., Hoekzema, K., Sanches, N. C., Maldonado-Ocampo, J. A., Castro, R. M., Foresti, F., Alfaro, M. E., and Oliveira, C. (2021). Accelerated diversification explains the exceptional species richness of tropical characoid fishes. *Systematic Biology*, 71:78–92.

Meyer, C. G., Holland, K. N., Wetherbee, B. M., and Lowe, C. G. (2000). Movement patterns, habitat utilization, home range size and site fidelity of whitesaddle goatfish, *Parupeneus porphyreus*, in a marine reserve. *Environmental Biology of Fishes*, 59:235–242.

Miller, E. C., Hayashi, K. T., Song, D., and Wiens, J. J. (2018). Explaining the ocean's richest biodiversity hotspot and global patterns of fish diversity. *Proceedings of the Royal Society B*, 285.

Minh, B. Q., Schmidt, H. A., Chernomor, O., Schrempf, D., Woodhams, M. D., Haeseler, A. V., Lanfear, R., and Teeling, E. (2020). IQ-TREE 2: New models and efficient methods for phylogenetic inference in the genomic era. *Molecular Biology and Evolution*, 37:1530–1534.

Mitchell, A. F. and Krzanowski, W. J. (1985). The Mahalanobis distance and elliptic distributions. *Biometrika*, 72(2):464–467.

Mittelheiser, L., Lepoint, G., Gillet, A., and Frédérich, B. (2022). Ecomorphology of six goatfish species (Mullidae) from Toliara Reef, Madagascar. *Environmental Biology of Fishes*, 105:1015–1032.

Mora, C., Treml, E. A., Roberts, J., Crosby, K., Roy, D., and Tittensor, D. P. (2012). High connectivity among habitats precludes the relationship between dispersal and range size in tropical reef fishes. *Ecography*, 35:89–96.

Mouillot, D., Bortoli, J. D., Leprieur, F., Parravicini, V., Kulbicki, M., and Bellwood, D. R. (2013). The challenge of delineating biogeographical regions: Nestedness matters for Indo-Pacific coral reef fishes. *Journal of Biogeography*, 40:2228–2237.

Mouillot, D., Villeger, S., Parravicini, V., Kulbicki, M., Arias-Gonzalez, J. E., Bender, M., Chabanet, P., Floeter, S. R., Friedlander, A., Vigliola, L., and Bellwood, D. R. (2014). Functional over-redundancy and high functional vulnerability in global fish faunas on tropical reefs. *Proceedings of the National Academy of Sciences*, 111:13757–13762.

Munro, J. L. (1976). Aspects of the biology and ecology of Caribbean reef fishes: Mullidae (goat-fishes). *J. Fish B i d*, 9:79–97.

Nakamura, Y., Horinouchi, M., Nakai, T., and Sano, M. (2003). Food habits of fishes in a seagrass bed on a fringing coral reef at Iriomote Island, southern Japan. *Ichthyological Research*, 50:15–22.

Nakamura, Y. and Sano, M. (2004). Overlaps in habitat use of fishes between a seagrass bed and adjacent coral and sand areas at Amitori bay, Iriomote Island, Japan: Importance of the seagrass bed as juvenile habitat. *Fisheries Science*, 70:788–803.

Nash, C. M., Lungstrom, L. L., Hughes, L. C., and Westneat, M. W. (2022). Phylogenomics and body shape morphometrics reveal recent diversification in the goatfishes (syngnatharia: Mullidae). *Molecular Phylogenetics and Evolution*, 177:107616.

Nee, S., May, R. M., and Harvey, P. H. (1994). The reconstructed evolutionary process. *Philosophical Transactions of the Royal Society of London. Series B: Biological Sciences*, 344:305–311.

Olsen, A. M. (2017). Feeding ecology is the primary driver of beak shape diversification in waterfowl. *Functional Ecology*, 31:1985–1995.

Olsen, A. M. and Westneat, M. W. (2015). Stereomorph: an r package for the collection of 3D landmarks and curves using a stereo camera set-up. *Methods in Ecology and Evolution*, 6:351–356.

Pagel, J. and Schurr, F. M. (2012). Forecasting species ranges by statistical estimation of ecological niches and spatial population dynamics. *Global Ecology and Biogeography*, 21:293–304.

Pagel, M. (1999). Inferring the historical patterns of biological evolution. *Nature*, 401:877–884.

Paradis, E. and Schliep, K. (2019). ape 5.0: an environment for modern phylogenetics and evolutionary analyses in r. *Bioinformatics*, 35:526–528.

Parravicini, V., Kulbicki, M., Bellwood, D. R., Friedlander, A. M., Arias-Gonzalez, J. E., Chabanet, P., Floeter, S. R., Myers, R., Vigliola, L., D'Agata, S., and Mouillot, D. (2013). Global patterns and predictors of tropical reef fish species richness. *Ecography*, 36:1254–1262.

Parsons, K. J., Cooper, W. J., and Albertson, R. C. (2011). Modularity of the oral jaws is linked to repeated changes in the craniofacial shape of African Cichlids. *International Journal of Evolutionary Biology*, 2011:1–10.

Pavlov, D. A., Emel'yanova, N. G., Thuan, L. T. B., and Ha, V. T. (2011). Reproduction and initial development of manybar goatfish *Parupeneus multifasciatus* (Mullidae). *Journal of Ichthyology*, 51:604–617.

Pebesma, E. J. and Bivand, R. S. (2005). Classes and methods for spatial data in r. *R News*, 5:9–13.

Pennell, M. W., Sarver, B. A., and Harmon, L. J. (2012). Trees of unusual size: Biased inference of early bursts from large molecular phylogenies. *PLOS ONE*, 7:e43348.

Peterson, R. D., Sullivan, J. P., Hopkins, C. D., Santaquiteria, A., Dillman, C. B., Pirro, S., Betancur-R, R., Arcila, D., Hughes, L. C., and Ortí, G. (2022). Phylogenomics of bony-tongue fishes (Osteoglossomorpha) shed light on the craniofacial evolution and biogeography of the weakly electric clade (Mormyridae). *Systematic Biology*, 71:1032–1044.

Pinheiro, H. T., MacDonald, C., Quimbayo, J. P., Shepherd, B., Phelps, T. A., Loss, A. C., Teixeira, J. B., and Rocha, L. A. (2023). Assembly rules of coral reef fish communities along the depth gradient. *Current Biology*.

Pinto, G., Mahler, D. L., Harmon, L. J., and Losos, J. B. (2008). Testing the island effect in adaptive radiation: rates and patterns of morphological diversification in Caribbean and mainland Anolis lizards. *Proceedings of the Royal Society B: Biological Sciences*, 275:2749–2757.

Pittman, S. J. and Olds, A. D. (2015). NOAA seascape ecology of fishes on coral reefs. *Ecology of fishes on coral reefs*, page 274.

Platell, M. E., Potter, C., and Clarke, K. R. (1998). Do the habitats, mouth morphology and diets of the mullids *Upeneichthys stotti* and *U. lineatus* in coastal waters of south-western Australia differ? *Journal of Fish Biology*, 52:398–418.

Polly, P. D., Lawing, A. M., Fabre, A.-C., and Goswami, A. (2013). Phylogenetic principal components analysis and geometric morphometrics. *Hystrix*, 24:33.

Provoost, P. and Bosch, S. (2017). robis: R client to access data from the obis api. R package version 1.0.0.

Pulliam, H. (2000). On the relationship between niche and distribution. *Ecology Letters*, 3:349–361.

Rabiee, M., Sayyari, E., and Mirarab, S. (2019). Multi-allele species reconstruction using ASTRAL. *Molecular Phylogenetics and Evolution*, 130:286–296.

Rabosky, D. L., Chang, J., Title, P. O., Cowman, P. F., Sallan, L., Friedman, M., Kaschner, K., Garilao, C., Near, T. J., Coll, M., and Alfaro, M. E. (2018). An inverse latitudinal gradient in speciation rate for marine fishes. *Nature*, 559:392–395.

Rambaut, A., Drummond, A. J., Xie, D., Baele, G., and Suchard, M. A. (2018). Posterior summarization in Bayesian phylogenetics using tracer 1.7. *Systematic Biology*, 67:901–904.

Randall, J. E. and Heemstra, E. (2009). Three new goatfishes of the genus *Parupeneus* from the western Indian Ocean, with resurrection of *P. seychellensis*. *Smithiana Bulletin*, 10:37–50.

Randall, J. E. and Myers, R. F. (2002). *Parupeneus insularis*, a new central Pacific species of goatfish (Perciformes: Mullidae) of the *P. trifasciatus* complex. *Zoological Studies-Taipei*, 41:431–440.

Reeck, J. C. and Oxford, J. T. (2022). The shape of the jaw—zebrafish *Col11a1a* regulates Meckel's cartilage morphogenesis and mineralization. *Journal of Developmental Biology*, 10(4):40.

Renema, W., Bellwood, D. R., Braga, J. C., Bromfield, K., Hall, R., Johnson, K. G., Lunt, P., Meyer, C. P., McMonagle, L. B., Morley, R. J., O'Dea, A., Todd, J. A., Wesselingh, F. P., Wilson, M. E., and Pandolfi, J. M. (2008). Hopping hotspots: Global shifts in marine biodiversity. *Science*, 321:654–657.

Revell, L. J. (2009a). Size-correction and principal components for interspecific comparative studies. *Evolution*, 63:3258–3268.

Revell, L. J. (2009b). Size-correction and principal components for interspecific comparative studies. *Evolution; international journal of organic evolution*, 63:3258–68.

Revell, L. J. (2012). phytools: an r package for phylogenetic comparative biology (and other things). *Methods in Ecology and Evolution*, 3:217–223.

Revell, L. J., Johnson, M. A., Schulte, J. A., Kolbe, J. J., and Losos, J. B. (2007). A phylogenetic test for adaptive convergence in rock-dwelling lizards. *Evolution; international journal of organic evolution*, 61:2898–912.

Ribeiro, E., Davis, A. M., Rivero-Vega, R. A., Ortí, G., and Betancur, R. (2018). Post-Cretaceous bursts of evolution along the benthic-pelagic axis in marine fishes. *Proceedings of the Royal Society B*, 285.

Rincon-Sandoval, M., Duarte-Ribeiro, E., Davis, A. M., Santaquiteria, A., Hughes, L. C., Baldwin, C. C., Soto-Torres, L., Acero, A. P., Walker, H. J., Carpenter, K. E., Sheaves, M., Ortí, G., Arcila, D., and Betancur, R. R. (2020). Evolutionary determinism and convergence associated with water-column transitions in marine fishes. *Proceedings of the National Academy of Sciences of the United States of America*, 117:33396–33403.

Roberts, R. B., Hu, Y., Albertson, R. C., and Kocher, T. D. (2011). Craniofacial divergence and ongoing adaptation via the hedgehog pathway. *Proceedings of the National Academy of Sciences of the United States of America*, 108:13194–9.

Robertson, D. R., Grove, J. S., and McCosker, J. E. (2004). Tropical trans-Pacific shore fishes. *Pacific Science*, 58:507–565.

Robinson, B. and Parsons, K. (2002). Changing times, spaces, and faces: Tests and implications of adaptive morphological plasticity in the fishes of northern postglacial lakes. *Canadian Journal of Fisheries and Aquatic Sciences*, 59:1819–1833.

Rocha, L. A., Craig, M. T., and Bowen, B. W. (2007). Phylogeography and the conservation of coral reef fishes. *Coral Reefs*, 26:501–512.

Rohlf, F. J. and Slice, D. (1990). Extensions of the Procrustes method for the optimal superimposition of landmarks. *Systematic Biology*, 39:40–59.

Royston, J. P. (1982a). Algorithm AS 181: The w test for normality. *Applied Statistics*, 31:176.

Royston, J. P. (1982b). An extension of Shapiro and Wilk's w test for normality to large samples. *Applied Statistics*, 31:115.

Royston, P. (1995). Remark as r94: A remark on Algorithm AS 181: The W-test for normality. *Applied Statistics*, 44:547.

Russ, G. R., Bergseth, B. J., Rizzari, J. R., and Alcala, A. C. (2015). Decadal-scale effects of benthic habitat and marine reserve protection on Philippine goatfish (f: Mullidae). *Coral Reefs*, 34:773–787.

Saeed, A. A. R., Moharram, S. G., and Al-Abbasi, F. A. (2016). A comparative gametogenesis and biochemical differentiation of goatfish (family: Mullidae) from the Red Sea. *Russian Journal of Marine Biology*, 42:42–50.

Sambrook, K., Hoey, A. S., Andréfouët, S., Cumming, G. S., Duce, S., and Bonin, M. C. (2019). Beyond the reef: The widespread use of non-reef habitats by coral reef fishes. *Fish and Fisheries*, 20:903–920.

Santaquiteria, A., Siqueira, A. C., Duarte-Ribeiro, E., Carnevale, G., White, W. T., Pogonoski, J. J., Baldwin, C. C., Ortí, G., Arcila, D., and Ricardo, B. R. (2021). Phylogenomics and historical biogeography of seahorses, dragonets, goatfishes, and allies (Teleostei: Syngnatharia): Assessing factors driving uncertainty in biogeographic inferences. *Systematic Biology*, 70:1145–1162.

Sayyari, E. and Mirarab, S. (2016). Fast coalescent-based computation of local branch support from quartet frequencies. *Molecular Biology and Evolution*, 33:1654–1668.

Sazima, C., Krajewski, J. P., Bonaldo, R. M., and Guimarães, P. R. (2006). The goatfish *Pseudupeneus maculatus* and its follower fishes at an oceanic island in the tropical west Atlantic. *Journal of Fish Biology*, 69:883–891.

Shibuya, S. and Motomura, H. (2021). *Parupeneus williamsi*, a new species of goatfish (Perciformes: Mullidae) endemic to the Marquesas Islands, French Polynesia. *Ichthyological Research*, 68:11–20.

Sidlauskas, B. (2008). Continuous and arrested morphological diversification in sister clades of characiform fishes: A phylomorphospace approach. *Evolution*, 62:3135–3156.

Sierra, L. M., Claro, R., and Popova, O. A. (1994). Alimentación y relaciones tróficas. *Ecologia de los peces marinos de Cuba*, pages 263–320.

Simpson, G. G. (1945). Tempo and mode in evolution. *Transactions of the New York Academy of Sciences*, 8:45–60.

Siqueira, A. C., Morais, R. A., Bellwood, D. R., and Cowman, P. F. (2020). Trophic innovations fuel reef fish diversification. *Nature Communications 2020 11:1*, 11:1–11.

Slater, G. J. (2022). Topographically distinct adaptive landscapes for skulls, skeletons, and size explain the dietary adaptive radiation of Carnivora (Mammalia). *bioRxiv*, page 2022.04.01.486739.

Slater, G. J., Price, S. A., Santini, F., and Alfaro, M. E. (2010). Diversity versus disparity and the radiation of modern cetaceans. *Proceedings of the Royal Society B: Biological Sciences*, 277:3097–3104.

Smith, L. L., Fessler, J. L., Alfaro, M. E., Streelman, J. T., and Westneat, M. W. (2008). Phylogenetic relationships and the evolution of regulatory gene sequences in the parrotfishes. *Molecular Phylogenetics and Evolution*, 49:136–152.

Song, H. Y., Mabuchi, K., Satoh, T. P., Moore, J. A., Yamanoue, Y., Miya, M., and Nishida, M. (2014). Mitogenomic circumscription of a novel percomorph fish clade mainly comprising "Syngnathoidei" (Teleostei). *Gene*, 542:146–155.

Spalding, M. D., Fox, H. E., Allen, G. R., Davidson, N., Ferdaña, Z. A., Finlayson, M., Halpern, B. S., Jorge, M. A., Lombana, A., Lourie, S. A., Martin, K. D., McManus, E., Molnar, J., Recchia, C. A., and Robertson, J. (2007). Marine ecoregions of the world: A bioregionalization of coastal and shelf areas. *BioScience*, 57:573–583.

Stadler, T. (2013). How can we improve accuracy of macroevolutionary rate estimates? *Systematic Biology*, 62:321–329.

Strübin, C., Steinegger, M., and Bshary, R. (2011). On group living and collaborative hunting in the yellow saddle goatfish (*Parupeneus cyclostomus*). *Ethology*, 117:961–969.

Sweatman, H. (1996). Impact of tourist pontoons on fish assemblages on the Great Barrier Reef. *Tech. Rep. CRC Reef Res. Cent.*, page 54.

Szabó, Z., Snelgrove, B., Craig, M. T., Rocha, L. A., and Bowen, B. W. (2014). Phylogeography of the manybar goatfish, *Parupeneus multifasciatus*, reveals isolation of the Hawaiian Archipelago and a cryptic species in the Marquesas Islands. *Bulletin of Marine Science*, 90:493–512.

Tedesco, P. A., Paradis, E., Lévêque, C., and Hugueny, B. (2017). Explaining global-scale diversification patterns in Actinopterygian fishes. *Journal of Biogeography*, 44:773–783.

Westneat, M., Alfaro, M. E., Wainwright, P. C., Bellwood, D. R., Grubich, J. R., Fessler, J. L., Clements, K. D., and Smith, L. L. (2005). Local phylogenetic divergence and global evolutionary convergence of skull function in reef fishes of the family Labridae. *Proceedings of the Royal Society B: Biological Sciences*, 272:993–1000.

Thacker, C. E., Shelley, J. J., Tyler Mccraney, Unmack, P. J., and Mcgee, M. D. (2022). Delayed adaptive radiation among New Zealand stream fishes: Joint estimation of divergence time and trait evolution in a newly delineated island species flock. *Syst. Biol*, 71:13–23.

Thorndike, R. L. (1953). Who belongs in the family? *Psychometrika*, 18:267–276.

Tittensor, D. P., Mora, C., Jetz, W., Lotze, H. K., Ricard, D., Berghe, E. V., and Worm, B. (2010). Global patterns and predictors of marine biodiversity across taxa. *Nature*, 466:1098–1101.

Torfstein, A. and Steinberg, J. (2020). The Oligo–Miocene closure of the Tethys Ocean and evolution of the proto-Mediterranean Sea. *Scientific Reports 2020 10:1*, 10:1–10.

Tosetto, L., Williamson, J., White, T., and Hart, N. (2021). Can the dynamic colouration and patterning of bluelined goatfish (Mullidae; *Upeneichthys lineatus*) be perceived by conspecifics? *Brain, Behavior and Evolution*, 96:103–123.

Turan, C. (2006). Phylogenetic relationships of Mediterranean Mullidae species (Perciformes) inferred from genetic and morphologic data. *Scientia Marina*, 70:311–318.

Uiblein, F. (1991). Ontogenetic shifts in resource use and shoaling tendency related to body size in Red Sea goatfish (*Parupeneus forsskali*, Mullidae). *Marine Ecology*, 12:153–161.

Uiblein, F. (2007). Goatfishes (Mullidae) as indicators in tropical and temperate coastal habitat monitoring and management. *Marine Biology Research*, 3:275–288.

Uiblein, F. (2011). Taxonomic review of western Indian Ocean goatfishes of the genus *Mulloidichthys* (family Mullidae), with description of a new species and remarks on colour and body form variation in Indo-West Pacific species. *Smithiana Bulletin*, 13:51–73.

Uiblein, F. and Causse, R. (2013). A new deep-water goatfish of the genus *Upeneus* (Mullidae) from Vanuatu, South Pacific. *Zootaxa*, 3666:337–344.

Uiblein, F. and Gouws, G. (2014). A new goatfish species of the genus *Upeneus* (Mullidae) based on molecular and morphological screening and subsequent taxonomic analysis. *Marine Biology Research*, 10:655–681.

Uiblein, F. and Gouws, G. (2015). Distinction and relatedness – taxonomic and genetic studies reveal a new species group of goatfishes (*Upeneus*; Mullidae). *Marine Biology Research*, 11:1021–1042.

Uiblein, F. and Heemstra, P. C. (2010). A taxonomic review of the western Indian Ocean goatfishes of the genus *Upeneus* (family Mullidae), with descriptions of four new species. *Smithiana Bulletin*, 11:35–71.

Uiblein, F. and Heemstra, P. C. (2011a). Description of a new goatfish species, *Upeneus randalli* n. sp. (Mullidae), from the Persian Gulf, with remarks and identification keys for the genus *Upeneus*. *Scientia Marina*, 75:585–594.

Uiblein, F. and Heemstra, P. C. (2011b). A new goatfish, *Upeneus seychellensis* sp. nov. (Mullidae), from the Seychelles Bank, with remarks on *Upeneus guttatus* and a key to western Indian Ocean *Upeneus* species. *Marine Biology Research*, 7:637–650.

Uiblein, F. and Heemstra, P. C. (2011c). A new goatfish, *Upeneus seychellensis* sp. nov. (Mullidae), from the Seychelles Bank, with remarks on *Upeneus guttatus* and a key to western Indian Ocean *Upeneus* species. *Marine Biology Research*, 7:637–650.

Uiblein, F., Hoang, T. A., Alama, U., Causse, R., Chacate, O. E., Fahmi, G., and Matiku, P. (2018). A new species and new records of goatfishes of the genus *Parupeneus* (Mullidae) from the Indian Ocean, with updated occurrence information for *P. jansenii* in the western Pacific. *Cybium: International Journal of Ichthyology*, 42:229–256.

Uiblein, F., Hoang, T. A., and Gledhill, D. (2017). Redescription and new records of Jansen's goatfish, *Parupeneus jansenii* (Mullidae), from the western Pacific and eastern Indian Ocean. *Zootaxa*, 4344:541–559.

Uiblein, F., Köhler, C., and Tian, M. C. (1998). Quantitative examination of morphological variability among goatfishes of the genus *Upeneus* from the Malayan province (Pisces: Perciformes: Mullidae). *Senckenbergiana maritima*, 28:123–132.

Uiblein, F. and Lisher, M. (2013). A new goatfish of the genus *Upeneus* (Mullidae) from Angoche, northern Mozambique. *Zootaxa*, 3717(1):85–95.

Uiblein, F. and Mcgrouther, M. (2012). A new deep-water goatfish of the genus *Upeneus* (Mullidae) from northern Australia and the Philippines, with a taxonomic account of *U. subvittatus* and remarks on *U. mascareinsis*. *Zootaxa*, 3550:61–70.

Uiblein, F. and Motomura, H. (2021). Three new goatfishes of the genus *Upeneus* from the eastern Indian Ocean and western Pacific, with an updated taxonomic account for *U. itoui* (Mullidae: Japonicus-species group). *Zootaxa*, 4938:298–324.

Uiblein, F. and White, W. T. (2015). A new goatfish of the genus *Upeneus* (Mullidae) from Lombok, Indonesia and first verified record of *U. asymmetricus* for the Indian Ocean. *Zootaxa*, 3980:51–66.

Vasil'eva, E. D. (2011). Importance of morphological traits and coloration for diagnostics of species of the genus *Mullus* (Mullidae, Perciformes), assessment of the taxonomic status of populations, and revision of ranges. *Journal of Ichthyology*, 51:14–27.

Vermeij, G. J. (2001). Community assembly in the sea: geologic history of the living shore biota. *Marine community ecology*, pages 39–60.

Vermeij, G. J. (2005). One-way traffic in the western Atlantic: causes and consequences of Miocene to early Pleistocene molluscan invasions in Florida and the Caribbean. *Paleobiology*, 31:624–642.

Via, S., Gomulkiewicz, R., Jong, G. D., Scheiner, S. M., Schlichting, C. D., and Tienderen, P. H. V. (1995). Adaptive phenotypic plasticity: Consensus and controversy. *Trends in Ecology & Evolution*, 10:212–217.

Villalobos, F., Olalla-Tárraga, M., Cianciaruso, M. V., Rangel, T. F., and Diniz-Filho, J. A. F. (2017). Global patterns of mammalian co-occurrence: Phylogenetic and body size structure within species ranges. *Journal of Biogeography*, 44:136–146.

Villalobos, F., Rangel, T. F., and Diniz-Filho, J. A. F. (2013). Phylogenetic fields of species: Cross-species patterns of phylogenetic structure and geographical coexistence. *Proceedings of the Royal Society B: Biological Sciences*, 280.

Villéger, S., Miranda, J. R., Hernández, D. F., and Mouillot, D. (2010). Contrasting changes in taxonomic vs. functional diversity of tropical fish communities after habitat degradation. *Ecological Applications*, 20:1512–1522.

Waechter, L. S., Luiz, O. J., Leprieur, F., and Bender, M. G. (2021). Functional biogeography of marine vertebrates in Atlantic ocean reefs. *Diversity and Distributions*, 24:1–14.

Wainwright, P. C., Bellwood, D. R., Westneat, M. W., Grubich, J. R., and Hoey, A. S. (2004). A functional morphospace for the skull of labrid fishes: Patterns of diversity in a complex biomechanical system. *Biological Journal of the Linnean Society*, 82:1–25.

Wainwright, P. C. and Richard, B. A. (1995). Predicting patterns of prey use from morphology of fishes. *Environmental Biology of Fishes*, 44:97–113.

Wang, M., Lloyd, G. T., Zhang, C., and Zhou, Z. (2021). The patterns and modes of the evolution of disparity in Mesozoic birds. *Proceedings of the Royal Society B*, 288.

Ward, J. H. (1963). Hierarchical grouping to optimize an objective function. *Journal of the American Statistical Association*, 58:236–244.

Webb, C. O., Ackerly, D. D., and Kembel, S. W. (2008). Phylocom: software for the analysis of phylogenetic community structure and trait evolution. *Bioinformatics*, 24:2098–2100.

West-Eberhard, M. J. (2005). Phenotypic accommodation: Adaptive innovation due to developmental plasticity. *Journal of Experimental Zoology Part B: Molecular and Developmental Evolution*, 304:610–618.

Westneat, M. W. (1995). Phylogenetic systematics and biomechanics in ecomorphology. *Environmental Biology of Fishes*, 44:263–283.

White, A. E., Dey, K. K., Mohan, D., Stephens, M., and Price, T. D. (2019). Regional influences on community structure across the tropical-temperate divide. *Nature Communications*, 10.

Wiens, J. J. and Graham, C. H. (2005). Niche conservatism: Integrating evolution, ecology, and conservation biology. *Annual Review of Ecology, Evolution, and Systematics*, 36:519–539.

Winemiller, K. O., Fitzgerald, D. B., Bower, L. M., and Pianka, E. R. (2015). Functional traits, convergent evolution, and periodic tables of niches. *Ecology letters*, 18(8):737–751.

Wismer, S., Hoey, A. S., and Bellwood, D. R. (2009). Cross-shelf benthic community structure on the Great Barrier Reef: Relationships between macroalgal cover and herbivore biomass. *Marine Ecology Progress Series*, 376:45–54.

Zhang, C., Rabiee, M., Sayyari, E., and Mirarab, S. (2018). ASTRAL-iii: Polynomial time species tree reconstruction from partially resolved gene trees. *BMC Bioinformatics*, 19:15–30.

Žalohar, J., Hitij, T., and Križnar, M. (2009). Two new species of seahorses (Syngnathidae, *Hippocampus*) from the Middle Miocene (Sarmatian) coprolitic horizon in Tunjice Hills, Slovenia: The oldest fossil record of seahorses. *Annales de Paleontologie*, 95:71–96.